Recursive Preferences, Correlation Aversion, and the Temporal Resolution of Uncertainty*

Lorenzo Stanca†

December 19, 2021

Abstract

Models of recursive utility are commonly associated with a preference for early resolution of uncertainty, often regarded as an important economic channel in applications. This paper provides a different understanding of recursive preferences based on attitudes toward correlation, and in particular aversion to intertemporally correlated risks. I formalize and investigate such a property. I show that an increase in correlation makes a decision maker that prefers early resolution worse off, even when increasing correlation also provides non-instrumental information about future consumption. Relatedly, I show that one can separate risk aversion from intertemporal substitution by considering a domain of choice in which pure preferences for early resolution of uncertainty play no role. Finally, I apply the insights of this paper to better understand the features possessed by existing models of recursive utility. I argue that attitudes toward correlation are the key behavioral feature driving the results of consumption-based asset pricing models.

Keywords: Intertemporal substitution, risk aversion, correlation aversion, recursive utility, preference for early resolution of uncertainty.

JEL classification: C61, D81.

---

*I visited the department of Decision Sciences at Bocconi University in the fall of 2019 and my thinking on the issues addressed in this paper began at that time. I thank Simone Cerreia-Vioglio, Fabio Maccheroni, and Massimo Marinacci for their hospitality and helpful discussions. I thank Nabil Al-Najjar, Peter Klibanoff, and Marciano Siniscalchi for their help and guidance. I also thank Matteo Camboni, Xiaoyu Cheng, Annie Liang, Matteo Magnaricotte, Michael Porcellacchia, Michael Powell and Kelly Gail Strada for providing very valuable comments.

†Department of Managerial Economics and Decision Sciences, Kellogg School of Management, Northwestern University, Evanston, IL 60208, USA. E-mail: lorenzo.stanca@kellogg.northwestern.edu

1 Introduction

Recursive preferences are of central importance in many economic settings. They play a key role in models of consumption-based asset pricing (Epstein & Zin (1989), Epstein & Zin (1991)), precautionary savings (Weil (1989), Hansen et al. (1999)), business cycles (Tallarini (2000)), optimal fiscal policy (Karantounias (2018)), and risk-sharing (Epstein (2001), Anderson (2005)), among many others. Part of their success is due to their ability to disentangle risk aversion from intertemporal substitution. This property is relevant in many settings to quantify the impact of these two different features on quantities of interest, such as asset prices or precautionary savings. Recall that the standard model of discounted expected utility in its recursive form can be written as

$$V_t = u(c_t) + \beta \mathbb{E}\left(V_{t+1}\right).$$

In this model, risk aversion and attitudes toward consumption smoothing are both captured by the curvature of u and therefore they cannot be separately identified from one another. In contrast, recursive preferences allow for a more general recursive formulation

$$V_t = W\left(c_t, I\left(V_{t+1}\right)\right),$$

where the so-called time aggregator W reflects intertemporal substitution and the certainty equivalent I reflects attitudes toward risk (or uncertainty), hence obtaining the desired separation between the two. Ever since the work of Kreps & Porteus (1978) it has been understood that separating these two important features entails a preference for non-instrumental information, also referred to as a preference for early resolution of uncertainty. For example, consider a gamble in which consumption is fixed at 0 for every $t = 1, \ldots, T-1$ and pays either 1 or 0 at $t = T$ depending on the outcome of a coin toss. A strict preference for tossing the coin at $t = 1$ over $t = 2$ indicates a preference for non-instrumental information. There is no planning advantage to tossing the coin early: in this sense choosing to toss the coin at $t = 1$ reveals a pure preference for information, even if such information is useless. The standard additive expected utility model is indifferent between tossing the coin at $t = 1$ and $t = 2$, while models of recursive utility typically prefer early resolution of uncertainty. A strict preference for information that is useless is seen as puzzling, and in this sense it is seen as a cost of separating risk aversion from intertemporal substitution (see for example Epstein et al. (2014)).

This paper provides a different understanding of recursive preferences based on attitudes toward correlation, and in particular aversion to *intertemporally* correlated risks. To illustrate, consider two gambles: A and B.[1] In gamble A a fair coin is tossed at $t = 1$. If the outcome is heads, then consumption is constant at the level 1 for every period $t = 1, \ldots, T$. Otherwise, it is constant at the level 0 at every period. In gamble B, consumption is determined by tossing a fair coin at every time period

[1]The example is a modified version of the example in Duffie & Epstein (1992), p. 355.

t, giving a level of consumption equal to 1 if heads and 0 otherwise. Introspection suggests that one should treat these gambles differently.[2] A hedging motive suggests that B should be preferred to A. But at the same time B resolves gradually while for A all uncertainty resolves at $t = 1$. In other words, A features early resolution of uncertainty, making the comparison between these two gambles non-obvious: A has the advantage of resolving all uncertainty at $t = 1$, while B is more desirable because of its hedging value. The standard discounted expected utility model is always indifferent between these two, while in general recursive preferences allow for non-indifference between A and B. In this case, it is the standard discounted expected utility model that presents an undesirable feature, being unable to distinguish between A and B.

I present a novel and general notion of increasing correlation when uncertainty resolves gradually. I start from a general notion of an i.i.d. process and introduce positive correlation between consumption at different time periods, and study the attitudes toward such an increase in correlation for the major models of recursive utility. Notably, such a notion does not require probabilistic sophistication allowing for the study of recursive models of ambiguity aversion. As discussed when comparing gambles A and B, more correlation means also having more predictive power about future consumption, i.e. non-instrumental information. Therefore, the total effect of increasing correlation is determined by the relative strength of the hedging motive described earlier and preferences for non-instrumental information. I show that the major models of recursive utility exhibit correlation aversion, even if they also have a preference for early resolution of uncertainty. In this sense, aversion to correlation is the more dominant property of recursive utility. At a technical level, correlation aversion is identified by the time aggregator W being submodular (i.e., having negative cross derivative). In contrast, from Kreps & Porteus (1978) it is known that convexity of W in its second argument characterizes a preference for early resolution of uncertainty. In words, submodularity means that the marginal value of continuation utility is lower when today's consumption value is higher. While for well-known preferences such as Epstein-Zin the two properties of the aggregator coincide, I show that this fact is not true for more general recursive preferences. In particular, I show that the effect of correlation aversion is stronger than that of preferences for early resolution of uncertainty under standard regularity assumptions. This point is especially important in applications such as asset pricing since it clarifies how certain modelling assumptions influence the results.[3]

As a consequence of this analysis, I then argue that in order to identify a pure preference for non-instrumental information one must observe choices over consumption processes in which correlation plays no role. To illustrate, consider the first

[2] I review related concepts of correlation aversion and the experimental evidence in favor of it in Section 6.

[3] For example, recursive maxmin expected utility is indifferent to the timing of resolution of uncertainty (Strzalecki (2013)), but non-indifferent to correlation.

example discussed above. The decision maker is comparing two gambles that feature deterministic consumption at every time period except when consumption is random, differing only for the time at which uncertainty is resolved. This is not mere chance, but rather a necessity: deterministic consumption is needed to remove the effect of attitudes toward correlation and therefore to identify pure preferences for non-instrumental information. I suggests that consumption processes of this type are not relevant for standard dynamic problems of consumption choice under uncertainty. For this reason, I introduce a new domain of choice that does not contain such consumption processes and study preferences over such a restricted domain, therefore weakening completeness of preferences.[4] I show that such a domain, nevertheless, is rich enough to axiomatize a general recursive representation of preferences and can also allow for disentangling risk aversion from intertemporal substitution. As a consequence, one can separate risk aversion from intertemporal substitution without implying a pure preference for early resolution of uncertainty, while at the same time having a rich domain in terms of applications.

Together, these results provide a unified understanding of the applications of recursive utility. The literature on consumption-based asset pricing has progressively considered consumption processes that involve more persistence. For example, in the long-run risk model of Bansal & Yaron (2004) consumption growth contains a small, persistent predictable component. Such persistence provides non-instrumental information: realizations of consumption growth today provide non-instrumental information about consumption growth for the long-run future. An investor with preferences for early resolution of uncertainty should enjoy such non-instrumental information, and hence demand a lower premium on equity if the persistence of consumption growth increases. From this perspective, the ability of consumption-based models to explain the observed premium on equity is hindered rather than helped. However, the persistent component also increases positive correlation between consumption growth at different time periods. Therefore, the equity premium in this model is higher relative to the discounted expected utility benchmark because correlation aversion is the more dominant feature of preferences, and not due to a preference for early resolution.[5]

[4] As discussed by Aumann (1962, p. 446), Schmeidler (1989, p. 576) and Wakker (1989), completeness of preferences is a questionable condition in decision theory. In particular, my point here is closely related to Wakker (1989, p. 42):

> Some people may object against completeness because they want the decision maker to have the right, for a pair of alternatives, simply not to choose between them. This is not the objection we have in mind. [...] Our objection against completeness is that many choice situations are not actual, but hypothetical, and that it is unrealistic to suppose that the decision maker is confronted with very many, some unrealistic, choice situations.

[5] This point has to be contrasted with the common understanding of the long-run risk model, e.g. Bansal et al. (2016) state "The long-run risks (LRR) asset pricing model emphasizes the role

A strand of the literature (e.g., Hansen et al. (1999)) has motivated the use of models of recursive utility with robustness concerns and in particular aversion to model uncertainty. Correlation aversion has a straightforward connection with model uncertainty. Consider again gamble A. An equivalent way of thinking about such a gamble is that a biased coin is tossed at every period, but there is uncertainty about the bias: with equal chance the coin always returns heads or always returns tails. In contrast, gamble B features no such uncertainty: the coin is known to be unbiased. In other words, a preference for B over A indicates aversion to model uncertainty.[6] In optimal fiscal policy and risk sharing problems, the key feature of recursive utility is aversion to volatility in future utility (see for example Karantounias (2018), p. 2284, or Anderson (2005), p. 94). Correlation aversion has a strong connection with such a property: in gamble B, at $t = 0$ future utility is constant and equal to $\frac{1}{2}$, while for gamble A future utility is volatile, being either 0 or 1. Thus preferring B to A indicates aversion to volatility in future utility. To sum up, thinking through the lenses of correlation aversion allows a variety of important properties — aversion to long-run risk, aversion to model uncertainty, and aversion to volatility in continuation utility — to be related to observable consumption choice behavior.[7] In contrast, as previously mentioned, preferences for early resolution of uncertainty play no role in any of these applications.

Finally, I examine the consequences of using correlation aversion for the analysis of recursive preferences. First, I study which parameters determine correlation aversion for the major recursive preferences. I also study the implications of this analysis for dynamic ordinal certainty equivalent (DOCE) preferences, an alternative approach to separating risk aversion from intertemporal substitution. Further, I revisit Epstein, Farhi & Strzalecki's result which suggests that timing premia for the long-run risk model seem implausibly high based on introspection. Following the analysis based on aversion to correlation, I ask a different question: "what fraction of your consumption stream would you give up to remove all persistence in consumption growth?" Under standard parameter specifications, a preliminary analysis suggests that an investor would be willing to give up a share of his wealth which is not consistent with the experimental evidence. This result reinforces Epstein, Farhi & Strzalecki's point that the quantitative discipline of the long-run risks model has been lax in modeling aspects of investors preferences.

The results are stated in a setting of *uncertainty* (unlike a setting of *risk* such as in Kreps & Porteus (1978) or Epstein & Zin (1989)). The main reason that I consider such a setting is that it allows us to consider recursive models of ambiguity aversion,

of low-frequency movements [...] along with investor preferences for early resolution of uncertainty, as an important economic-channel that determines asset prices.".

[6]This connection is made formal by Al-Najjar & Shmaya (2019). They obtain a representation of Epstein-Zin recursive utility which admits a direct interpretation in terms of model (or parameter) uncertainty.

[7]Beyond macroeconomics, Kochov & Song (2021) apply recursive preferences to repeated games. In their case, correlation aversion plays a key role in expanding the set of feasible payoffs.

which are well known to be relevant in the asset pricing literature (e.g see Ju & Miao (2012) or Collard et al. (2018)).

2 Preliminaries

2.1 Framework

Time is discrete and varies over a finite horizon $t \in \{0, 1, \ldots, T\} \equiv \mathcal{T}$. The information structure is described by a filtered space $(\Omega, \{\mathcal{G}_t\}_{t \in T})$ where Ω is an arbitrary set of states of the world and $\mathcal{G} = \{\mathcal{G}_t\}_{t \in T}$ is a sequence of σ-algebras such that $\mathcal{G}_0 = \{\Omega, \emptyset\}$ that satisfy $\mathcal{G}_t \subset \mathcal{G}_{t+1}$ for $t = 0, \ldots, T-1$.[8] For simplicity, assume that every algebra $\mathcal{G}_t, t \in T$, is generated by a finite partition, where $\mathcal{G}_t(\omega)$ denotes the element of the partition containing $\omega \in \Omega$.

Let X denote the set of outcomes, which is assumed to be a convex subset of $\mathbb{R}^n$. The main cases we are interested in are $X = \mathbb{R}_+$ and $X = [1, \infty)$. An act or process is an X-valued, $\mathcal{G}$-adapted process, that is, a sequence $(f_t)_{t \in \mathcal{T}}$ such that $f_t : \Omega \to X$ is $\mathcal{G}_t$ measurable for every $t \in \mathcal{T}$. $\mathcal{F}$ is the set of all processes or acts. Elements of $\mathcal{F}$ should be thought of as consumption processes. $\mathcal{D}$ is the set of deterministic processes, $d = (d_0, d_1, \ldots, d_n) \in \mathcal{D}$ if and only if d_t is measurable w.r.t. $\mathcal{G}_0$ for all t. Since each $\mathcal{G}_t$ is finitely generated, then set of all $\mathcal{G}_t$-measurable acts can be endowed with the product topology. It follows that we can endow $\mathcal{F}$ with the product topology.[9] The assumptions of finiteness are not necessary but avoid the need to formally establish the existence of each recursive utility model considered in this paper.[10]

Given a measurable space (S, Σ) and $K \subseteq \mathbb{R}$, let $B_0(\Sigma, K)$ denote the set of simple Σ measurable function with range contained in K. A function $I : B_0(\Sigma, K) \to \mathbb{R}$ (i) continuous if it continuous in the sup-norm topology (ii) monotone if $\xi(s) \geq \xi'(s)$ for every $s \in S$ implies $I(\xi) \geq I(\xi')$ (iii) strictly monotone if it is monotone and $\xi(s) \geq \xi'(s)$ for every $s \in S$ with one strict inequality implies $I(\xi) > I(\xi')$ (iv) normalized if $I(x) = x$ for every $x \in \mathbb{R}$ (where x denotes the constant function $x1_\Omega$) (v) concave if $I(\alpha\xi + (1-\alpha)\xi') \leq \alpha I(\xi) + (1-\alpha)I(\xi')$ for every $\xi, \xi' \in B_0(\Sigma, K)$ (vi) constant-additive if $I(\xi + k) = I(\xi) + k$ for every $k \in K$ (vii) positive homogeneous if $I(\beta\xi) = \beta I(\xi)$ for every $\beta > 0$. Given a probability measure P defined on (S, Σ), an expected utility functional is given by $\mathbb{E}_P \xi = \int \xi(s) \mathrm{d}P(s)$.

The primitives of interest are a family of $\mathcal{G}$-adapted weak orders (complete and

[8] See Stokey & Lucas (1989) for canonical interpretations of this setting in terms of shocks/observations.

[9] Denoting with $|\mathcal{G}_t|$ the number of elements of the partition that generates $\mathcal{G}_t$, the set

$$\{f : \Omega \to X : f \text{ is measurable w.r.t. } \mathcal{G}_t\},$$

can be identified with a subset of $\mathbb{R}^{|\mathcal{G}_t|}$, and therefore $\mathcal{F}$ can be endowed with the product topology.

[10] See Marinacci & Montrucchio (2010) for a thorough treatment of the topic.

transitive relations) $\{\succeq_{t,\omega}\}_{(t,\omega)\in\mathcal{T}\times\Omega}$ on $\mathcal{R}$ where $\mathcal{D} \subseteq \mathcal{R} \subseteq \mathcal{F}$.[11] Unless otherwise stated, assume that $\mathcal{R} = \mathcal{F}$. Let $\succeq_0$ denote the preference at time zero. For brevity, I typically denote the collection of preferences $\{\succeq_{t,\omega}\}_{(t,\omega)\in\mathcal{T}\times\Omega}$ with just $\succeq_{t,\omega}$.

2.2 Recursive preferences

I provide a definition of a general recursive representation of preferences.

Definition 1. $\succeq_{t,\omega}$ *admits a general recursive representation if and only if there exist* $(V_t(\omega,\cdot))_{t,\omega}$ *that represent* $\succeq_{t,\omega}$ *satisfying the recursive relation* $V_T(\omega,h) = u(h_T(\omega))$ *for some continuous* $u : X \to \mathbb{R}$ *that satisfies* $u(z) = 0$ *for some* $z \in X$ *and for* $t < T$,

$$V_t(\omega,h) = W\left(h_t(\omega), I_{t,\omega}\left(V_{t+1}(\cdot,h)\right)\right) \text{ for every } h \in \mathcal{R}, \tag{1}$$

where $V_t(\omega,\mathcal{R}) = V_t(\omega',\mathcal{R}) \equiv V_t$*, each*

$$I_{t,\omega} : \{\xi \in B_0(\mathcal{G}_{t+1}, V_{t+1}) : \xi = V_{t+1}(\cdot,f), f \in \mathcal{R}\} \to \mathbb{R},$$

is continuous, normalized, and strictly monotone with $I_{t,\omega} = I_{t,\omega^*}$ *when* $\mathcal{G}_t(\omega) = \mathcal{G}_t(\omega^*)$ *and* $W : X \times \cup_{\tau\geq t+1} V_\tau \to \mathbb{R}$ *is a time aggregator that is continuous and strictly increasing in the second variable that satisfies* $W(x, u(z)) = u(x)$.

A general recursive representation of $\succeq_{t,\omega}$ can be identified by its parameters $(W, u, (I_{t,\omega})_{t,\omega})$. Below I describe the most common types of specifications for $(W, u, (I_{t,\omega})_{t,\omega})$.[12] These can be divided into two cases: risk models and ambiguity models.

- Recursive discounted expected utility (RDEU) preferences, where $W(x,y) = u(x) + \beta y$, $\beta \in (0,1)$ and $I_{t,\omega}(\xi) = \mathbb{E}_{P_{t,\omega}}\xi$ with each $P_{t,\omega}$ being a probability on $(\Omega, \mathcal{G}_{t+1})$.

- Recursive second-order expected utility preferences, where $W(x,y) = u(x)+\beta y$, $\beta \in (0,1)$ and $I(\xi) = \phi^{-1}\left(\mathbb{E}_{P_{t,\omega}}\phi(\xi)\right)$ for some strictly increasing and concave function $\phi : u(X) \to \mathbb{R}$ and each $P_{t,\omega}$ is a probability on Ω. Such a class of recursive preferences offers a simple separation between risk aversion and intertemporal substitution. The most important to instances of such preferences are given by:

  (1) Epstein-Zin preferences (EZ) Epstein & Zin (1989), which are given by[13]

$$u(x) = \begin{cases} \frac{x^\rho}{\rho} & 0 \neq \rho < 1, \\ \log(x) & \rho = 0, \end{cases}$$

[11] By $\mathcal{G}$-adapted I mean that $\succeq_{t,\omega} = \succeq_{t,\omega^*}$ whenever $\mathcal{G}_t(\omega) = \mathcal{G}_t(\omega^*)$.

[12] For brevity, I omit the conditions required for strict monotonicity of the certainty equivalent.

[13] Epstein & Zin (1989) consider a more general class of certainty equivalents, but for simplicity I focus on the case of expected utility.

and

$$\phi(x) = \begin{cases} \frac{\rho}{\alpha} x^{\frac{\alpha}{\rho}} & 0 \neq \alpha < 1, 0 \neq \rho < 1, \\ \frac{1}{\alpha} \exp \alpha x & 0 \neq \alpha < 1, \rho = 0. \end{cases}$$

(2) Recursive multiplier preferences (RM), see Hansen & Sargent (2001), where

$$I_{t,\omega}(\xi) = \min_{p \in \Delta(\Omega)} \mathbb{E}_p \xi + \theta R(p \| P_{t,\omega}),$$

where $R(p\|P_{t,\omega})$ is the relative entropy of p with respect to some fixed countably additive and nonatomic measure $P_{t,\omega}$ on $(\Omega, \mathcal{G}_{t+1})$, and $\theta \in (0, \infty]$ is a parameter. It is well-known (e.g., see Strzalecki (2011)) that such preferences admit the equivalent representation in terms of recursive second-order expected utility with $\phi(\cdot)$ given by

$$\phi(x) = \begin{cases} -\exp\left(-\frac{x}{\theta}\right) & \text{for } \theta < \infty, \\ x & \text{for } \theta = \infty. \end{cases}$$

- Recursive Epstein-Uzawa (REU) preferences see (Uzawa (1968) and Epstein (1983)), where u is strictly increasing and $W(x, y) = u(x) + b(x)y$ for some continuous function $b : X \to \mathbb{R}$ with $b(X) \subseteq (0, 1)$ and $I(\xi) = \mathbb{E}_{P_{t,\omega}}(\xi)$ for some $P_{t,\omega}$.

- Recursive discounted ambiguity averse preferences (RDAA), see Strzalecki (2013), where $u(X) = \mathbb{R}_+$ or $u(X) = \mathbb{R}$ and

$$V_t(\omega, h) = u\left(h_t(\omega)\right) + \beta I_{t,\omega}\left(V_{t+1}(\cdot, h)\right),$$

where $\beta \in (0, 1)$ and $I_{t,\omega}$ is a concave functional that is constant-additive or positive homogeneous. Two notable cases of such preferences are:

(1) Recursive maxmin expected utility (RMEU) preferences, see Epstein & Wang (1994), Epstein & Schneider (2003*b*) where $I(\xi) = \min_{p \in C_{t,\omega}} \mathbb{E}_p \xi$, with each set $C_{t,\omega}$ being convex and weak * -closed set of probabilities on $(\Omega, \mathcal{G}_{t+1})$.

(2) Recursive smooth ambiguity preferences (RSA), (see Klibanoff et al. (2005), Klibanoff et al. (2009)), where Ω is finite and

$$I_{t,\omega}(\xi) = \phi^{-1}\left(\mathbb{E}_{\mu_{t,\omega}} \phi\left(\mathbb{E}_P \xi\right)\right)$$

for some strictly increasing and concave function $\phi : u(X) \to \mathbb{R}$ and with each $\mu_{t,\omega}$ being a Borel probability measure on $\Delta(\Omega)$. In particular, I assume ϕ is $\phi(x) = -\exp(-\frac{x}{\alpha})$ or

$$\phi(x) = \begin{cases} \frac{x^\alpha}{\alpha} & 0 \neq \alpha < 1, \\ \log(x) & \alpha = 0. \end{cases}$$

2.3 IID and attitudes toward temporal resolution

As discussed by Strzalecki (2013), to study attitudes toward timing of resolution of uncertainty it is helpful to make an assumption of "constant beliefs," typically referred to in the literature as IID (Independently and Indistinguishably Distributed) ambiguity (see Epstein & Schneider (2003*a*)). Specifically, such an assumption requires that that $\Omega = S^T$ with $T \geq 2$, where (S, Σ) is a finite measurable space. Moreover, $\Sigma = 2^S$ and let $\mathcal{G}_t = \Sigma^t \times \{\emptyset, S\}^{T-t}$. In words, this means that at time t one knows the realization of $(s_1, \ldots, s_t)$, but is ignorant about the future. More precisely, observe that in this case we have $\mathcal{G}_t((s_1, \ldots, s_T)) = \{s_1\} \times \ldots \times \{s_t\} \times \{\emptyset, S\}^{T-t}$.

For every act $f = (f_0, \ldots, f_T)$ each element f_t, $t \geq 1$ can be written as a function of the first t elements, $(s_1, \ldots, s_t) \equiv s^t$. Hence we can write each node (t, ω) equivalently as s^t, $\succeq_{t,\omega} = \succeq_{s^t}$ and $I_{t,\omega} = I_{s^t}$. I assume that $I_{s^t} = I_{\bar{s}^t} \equiv I$ for every $s^t, \bar{s}^t$. This assumption is required to avoid assuming that attitudes toward timing of resolution are influenced by changing beliefs. To illustrate this point, in the case of risk models with beliefs given by $(P(\cdot|s_1, \ldots, s_t))_{s^t}$ it implies that for some probability P over S it holds that $P(s_{t+1}|s_1, \ldots, s_t) = P(s_{t+1})$ for every s^{t+1}.

In such a setting, a preference for early resolution of uncertainty can be defined as follows.

Definition 2. *Fix $t \leq T-2$. Say that $h \in \mathcal{F}$ resolves earlier than $h' \in \mathcal{F}$ whenever there exist $f_{t+2}, \ldots, f_T \in X^S$ and $x_0, \ldots, x_{t+1} \in X$ such that $h_j = h'_j = x_j$ for all $j \leq t+1, h_j(s_1, \ldots, s_j) = f_j(s_{t+1})$ for $j \geq t+2$, and $h'_j(s_1, \ldots, s_j) = f_j(s_{t+2})$.*

$\succeq_{t,\omega}$ exhibits a preference for earlier resolution of uncertainty if and only if for all $h, h' \in \mathcal{F}$ and $t \leq T-2$,

$$h \succeq_{t',\omega} h'$$

for all $t' \leq t$ and $\omega \in \Omega$. The notion of indifference is defined analogously.[14]

Figure 1 contains an example of acts that resolve early (bottom) and late (top).[15] The next result summarizes what we know from the literature about attitudes toward timing of resolution for the most used recursive models in the literature.

Theorem 0 (Chew & Epstein (1991), Strzalecki (2013))**.** *Suppose that $\succeq_{t,\omega}$ admits a recursive representation that is either EZ with $\alpha \leq \rho$, RM with $\theta \geq 0$ or RSA. Then $\succeq_{t,\omega}$ exhibits a preference for early resolution of uncertainty.*

Suppose that $\succeq_{t,\omega}$ admits an RDEU, MEU, or REU representation. Then $\succeq_{t,\omega}$ is indifferent to the timing of resolution of uncertainty.

[14] My definition is slightly more general than the one in Strzalecki (2013). It is needed later for Proposition 4. However, it is equivalent for RDAA preferences defined on $\mathcal{F}$.

[15] In particular, the act f that resolves early is defined by $f_t = z$ for $t = 0, 1$ and $f_2(s_1, s) = x$, $f'_2(s_2, s) = y$ for $s = s_1, s_2$ while the act f' that resolves late is defined by $f'_t = z$ for $t = 0, 1$ and $f'_2(s, s_1) = x$, $f'_2(s, s_2) = y$ for $s = s_1, s_2$.

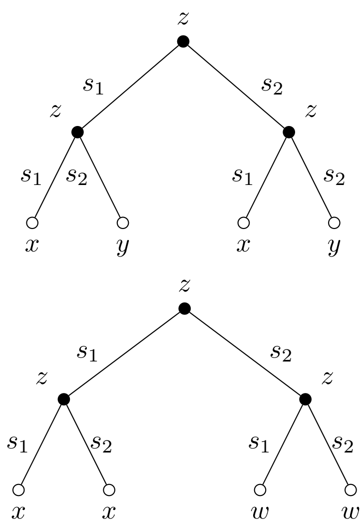

Figure 1: Early resolution vs late resolution.

3 Attitudes toward correlation

3.1 Increasing correlation

As discussed in the introduction, I introduce a general notion of an increase in positive correlation when uncertainty resolves gradually, and study the implications of such an increase in correlation for the major models of recursive utility. In applications, uncertainty typically resolves gradually: the value of consumption at time t is never fully known until time t itself is reached. For this reason, understanding the implications of an increase in correlation in such a setting is of fundamental importance. Notably, such a notion is "belief-free," which allows me to study preferences that are not necessarily probabilistically sophisticated. I do so by taking the equivalent in our setting of an i.i.d. process and introducing correlation.

A generalized i.i.d. process f^{iid} is defined by taking $f : S \to X$ and letting $f_t^{iid}(s^t) = f(s_t)$ for every $t \geq 1$ and $f_0 = x$ for some $x \in X$. The maintained IID assumption says that states are independently distributed. If f_{t+1} depended on the entire history $(s_1, \ldots, s_t)$, the consumption process would not be independently distributed. The function f removes any dependence of f_{t+1} on past realizations of the states. Now, to introduce correlation one could, for example, change f^{iid} to f^{corr} so that for fixed $s, s' \in S$ it holds that $f_t^{corr}((s_1, \ldots, s_{t-2}, s, s')) = f(s)$. In words, this means that if s is realized at time $t-1$ then $f(s)$ will be paid when s' is realized at time t. A specific example is represented in Figure 2 in the case of $S = \{s_1, s_2, s_3\}$, $\Sigma = 2^S$, $T = 2$ and $X = \mathbb{R}_+$. It can immediately be seen that such an increase in (positive) correlation will make f^{corr} strictly better than f^{iid} (see Figure 2). For this reason, I focus on increases in correlation that are "symmetric" in a precise sense: if

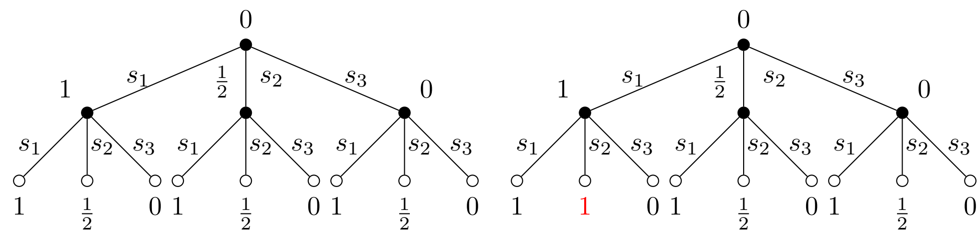

Figure 2: Asymmetric increase in correlation

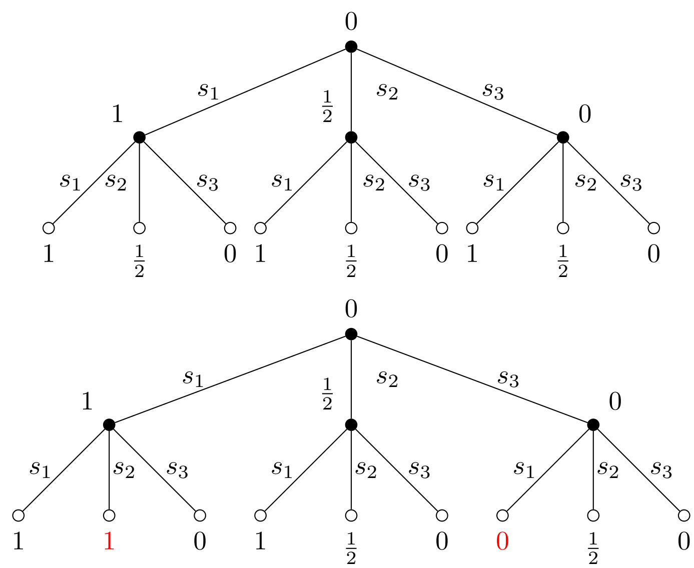

Figure 3: Symmetric increase in correlation.

$f(s_1 \ldots, s, s') = f(s)$ then also $f(s_1 \ldots, s', s) = f(s')$, as depicted in Figure 3.

Definition 3. *Given $f : S \to X$, denote with G_f the set of functions $g_f : S \times S \to X$ that satisfy*

$$g_f(s_j, s_i) = \begin{cases} f(s_i) \text{ or} \\ f(s_j), g_f(s_i, s_j) = f(s_i), \end{cases}$$

for every (s_j, s_i). For any $f : S \to X$ define the class of correlated acts as[16]

$$F_f^{corr} = \{f^{corr} \in \mathcal{F} : f_t^{corr}(s_1, \ldots, s_t) = g_f(s_{t-1}, s_t) \text{ for some } t \geq 2, g_f \in G_f,$$
$$\text{and } f_j^{corr}(s_1, \ldots, s_j) = f(s_j) \text{ for } j \neq t\}.$$

[16] I write $f_t^{corr}(s_1, \ldots, s_t) = g_f(s_{t-1}, s_t)$ as short for $f_t^{corr}(s_1, \ldots, s_t) = g_f(s_{t-1}, s_t)$ for every $(s_{t-1}, s_t) \in S \times S$.

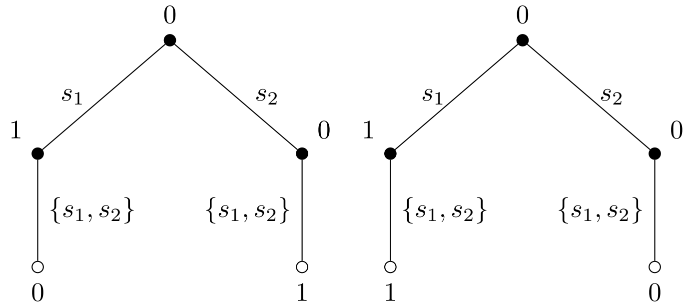

Figure 4: Non-correlated (left) and correlated (right) processes

The set of all correlated acts is given by

$$CORR = \bigcup_{f \in X^S} F_f^{corr}.$$

Remark 1. To study the effect on preferences of increasing positive correlation, it is enough to consider an increase in correlation between consumption at some time t and the subsequent period $t+1$. However, it is straightforward to extend the definition to allow for correlation over multiple periods. Section 6.2 presents a more detailed discussion.

Going back to the interpretation discussed in the introduction, the function g_f can be thought of as introducing model (or parameter) uncertainty into the consumption process: the law that determines consumption at time t is uncertain and depends on a state realized at time $t-1$. A symmetric increase in correlation is represented in Figure 3. This figure illustrates the interpretation mentioned in the introduction, that increasing correlation also increases the volatility of future utility: if at time $t=1$ the state s_1 is realized consumption is more likely to be high also at $t=2$, but the opposite is true in case the state s_3 is realized. It is important to contrast such a notion of an increase in positive correlation with the standard way of comparing a non-correlated process with a correlated one described in Figure 4. Such a comparison involves two processes that differ only in their pattern of correlation, but that have the same timing of resolution of uncertainty. On the contrary, when uncertainty resolves gradually, increasing correlation has a twofold effect: (i) not only does it increase the likelihood that consumption at time t matches consumption at time $t+1$ of future utility, but also (ii) it provides non-instrumental information about future consumption. Preferences exhibit correlation aversion when the aversion to (i) is stronger than the positive value provided by (ii).

Definition 4. $\succeq_{t,\omega}$ *exhibit correlation aversion if and only if*

$$f^{iid} \succeq_0 f^{corr}$$

for every $f \in X^S$ *and* $f^{corr} \in F_f^{corr}$.

Intuitively, recursive discounted expected utility should be indifferent to both effects, and indeed such preferences are always indifferent between f^{iid} and f^{corr} (see Lemma 2 in the appendix). From Theorem 0 we know that for RMEU and REU, preferences are indifferent to the timing of resolution; therefore (ii) should be irrelevant and only (i) should have an effect on such preferences. For other types of recursive preferences, it is not obvious what effect should prevail. In general, one would expect EZ, RM, and RSA preferences with standard parameter specifications to be averse to correlation. Yet, by Theorem 0 we also know that such recursive models prefer early resolution of uncertainty. The final effect on utility will depend on the relative strength of these two effects. Studying which effect prevails is a standard practice in economics, such as when one tries to asses the relative strength of income and substitution effects. The next two sections study attitudes toward correlation for the most common types of recursive preferences.

3.2 Attitudes toward correlation: risk

When a unique probability P defined on S is given, each correlated process f^{corr} can be equivalently described by means of probabilistic transformations of the process $(s^t)_t$ of states, in the spirit of Epstein & Tanny (1980). For simplicity, I illustrate this point for the case in which every state is equally likely.

Definition 5. *Write $S = \{s_1, \dots, s_n\}$. Consider $P \in \Delta(S)$ such that $P(s_1) = \dots = P(s_n) = \frac{1}{n}$. An elementary correlation-increasing transformation of P is given by taking $0 \le \varepsilon \le \frac{1}{n}$, $i, j \in \{1, \dots, n\}$, and defining the conditional probabilities $(P^\varepsilon(\cdot|s))_{s\in S}$ on S as follows: $P^\varepsilon(\cdot|s_k) = P(\cdot)$ for every $k \neq i, j$, $P(s_k|s_i) = P(s_k|s_i) = P(s_k)$ for every $k \neq i, j$ $P^\varepsilon(s_j|s_j) = P(s_j) + \varepsilon$, $P^\varepsilon(s_i|s_i) = P(s_i) + \varepsilon$, $P^\varepsilon(s_i|s_j) = P(s_j) - \varepsilon$, and $P^\varepsilon(s_j|s_i) = P(s_i) - \varepsilon$.*

Such correlation-increasing transformations add memory to the stochastic process of states. It is then easy to see that any f^{corr} can be obtained equivalently by performing finitely many of such correlation-increasing transformations of the process of states. The following example illustrates this point. Moreover, it suggests that in the case of RM preferences, aversion to correlation is the more dominant effect.

Example 1. $S = \{s_1, s_2, s_3\}$, $\Sigma = 2^S$, $T = 2$ and $X = \mathbb{R}_+$. Assume multiplier preferences, i.e. $W(x, y) = u(x) + \beta y$ and $I(\xi) = -\log(\mathbb{E}_P e^{-\xi})$ with $P(s_1) = P(s_2) = P(s_3) = \frac{1}{3}$. Let $f : S \to X$ be defined by $f(s_1) = \frac{3}{5}$, $f(s_1) = \frac{1}{2}$, $f(s_3) = 0$. Let $g(s_2, s_3) = f(s_2) = \frac{1}{2}$, $g(s_3, s_2) = f(s_3) = 0$ and $g(s_j, s_i) = f(s_i)$ otherwise. Equivalently, such an increase in correlation can be expressed by means of the following correlation increasing transformation: $P^{\frac{1}{3}}(s_2|s_2) = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$, $P^{\frac{1}{3}}(s_3|s_2) = \frac{1}{3} - \frac{1}{3} = 0$, $P^{\frac{1}{3}}(s_3|s_2) = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$, and $P^{\frac{1}{3}}(s_2|s_3) = \frac{1}{3} - \frac{1}{3} = 0$. Figure 5 and 6 represent the two acts f^{iid} and f^{corr}. It is easy to check that $f^{iid} \succ_0 f^{corr}$ for any $\beta \in (0, 1)$. $\triangle$

I prove that such a result holds in general: preferences for early resolution of uncertainty are dominated by correlation aversion.

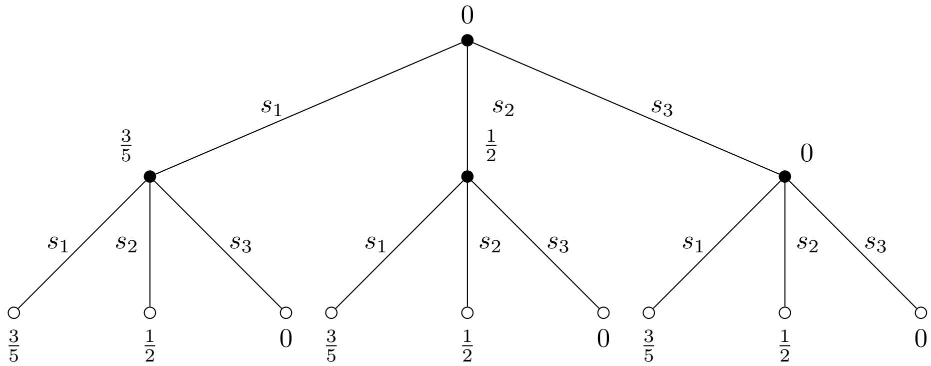

Figure 5: "i.i.d" process

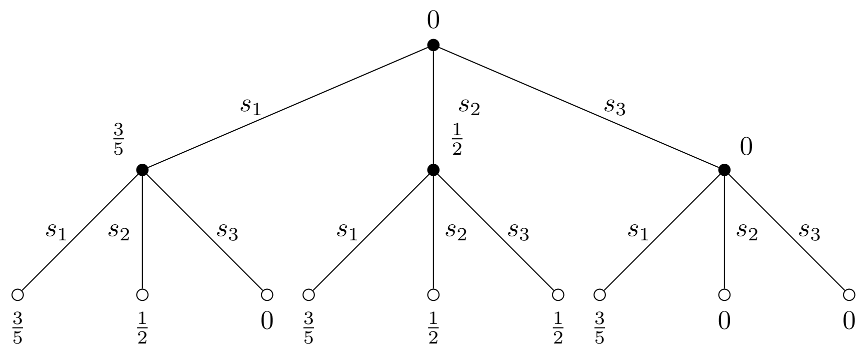

Figure 6: Introducing positive correlation may eliminate ambiguity

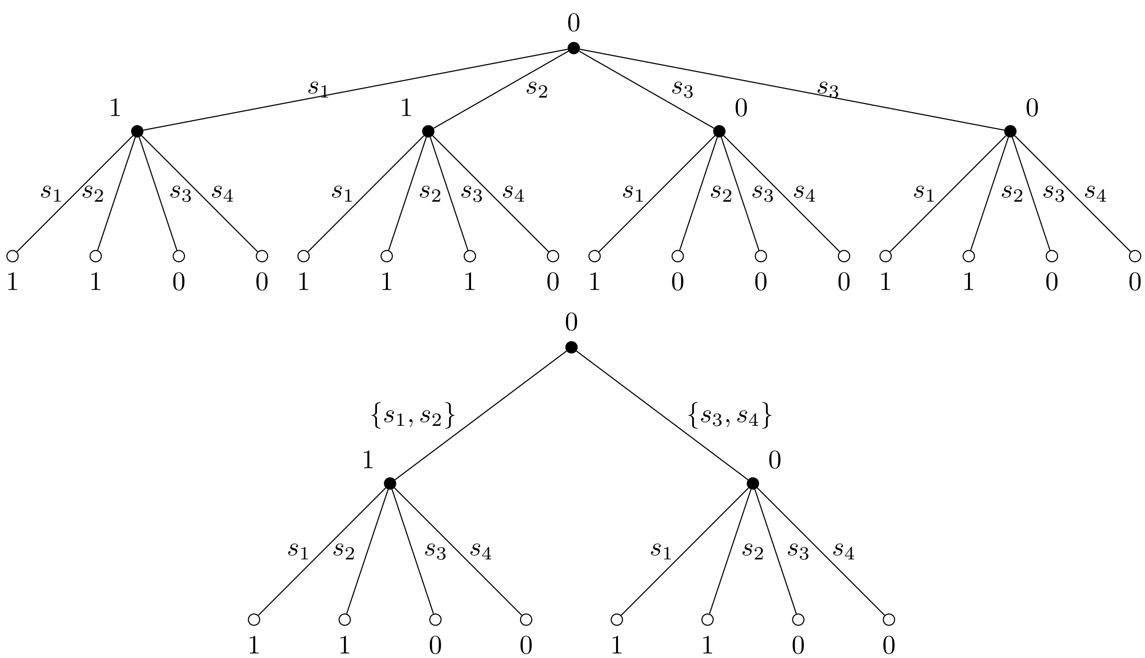

Figure 7: Introducing positive correlation creates ambiguity.

Theorem 1. *Suppose that* $\succeq_{t,\omega}$ *has a recursive representation that is either*

(i) EZ with $\alpha \leq \rho$*,*

(ii) RM with $\theta \geq 0$*,*

(iii) REU with $b(\cdot)$ *decreasing,*[17]

then it exhibits aversion to correlation.

Proof. See the appendix. □

In particular, as shown in the appendix, one will have $f^{iid} \succ_0 f^{corr}$ whenever $f^{iid} \neq f^{corr}$ and (i) $\alpha < \rho$ (ii) $\theta > 0$ (iii) $b(\cdot)$ is strictly decreasing. In other words, for EZ and RM preferences attitudes toward correlation are dominated by aversion to correlation. Only the latter matters. As for REU, the result confirms the intuition that only attitudes toward correlation should matter.

To gain a better understanding of this result, observe that both EZ and RM preferences belong to the class of recursive second-order expected utility preferences. Such preferences can be equivalently written with a time aggregator

$$W(x, y) = \phi(u(x) + \beta\phi^{-1}(y)),$$

and an expected utility certainty equivalent. Whenever ϕ is concave and twice differentiable, as is the case of EZ and RM preferences, such an aggregator is submodular: an increase in today's utility lowers the marginal value of continuation utility. Correlation aversion is implied by such a property of the time aggregator. It is important to contrast this result with what we know from Kreps & Porteus (1978): a preference for early resolution of uncertainty coincides with convexity of the aggregator in y. As shown by Strzalecki (2013) (Lemma 3), when ϕ is twice differentiable such a condition is equivalent to

$$\beta \left[-\frac{\phi''(\beta y + x)}{\phi'(\beta y + x)} \right] \leq \left[-\frac{\phi''(y)}{\phi'(y)} \right], \tag{2}$$

for every $y, x \in u(X)$. Hence, in general, attitudes toward correlation differ from attitudes toward temporal resolution.[18] The proof shows that when ϕ satisfies (2) for

[17] Bommier et al. (2019) provide a result for REU preferences related to Kochov's 2015 notion of intertemporal hedging. Such a notion involves comparing stochastic processes that differ in terms of correlation but not in terms of temporal resolution, such as those in Figure 4.

[18] For example, when ϕ is given by

$$\phi(x) = \int_0^x e^{-t^2} dt,$$

for every $x \in \mathbb{R}_+$, it will in general fail to satisfy condition (2). In such cases the effect of correlation aversion will be even stronger. More in general, the Koopmans, Diamond & Williamson's aggregator

$$W(x, y) = (1/\theta) \log\left(1 + \beta x^\gamma + \delta y\right), \quad 0 \leq x \leq 1, 0 \leq y \leq 1,$$

is concave in y and at the same time submodular under standard parameter specifications.

every $\beta \in (0,1)$, if it also exhibits IRRA (increasing relative risk aversion) the effect of correlation aversion will be stronger than that of preference for early resolution of uncertainty. IRRA is one of the most important classes of utility functions (e.g., see Arrow (1971), p. 96), which in turn contains as a special case the CRRA and CARA cases represented by EZ and RM preferences. Therefore, such a result holds for a very important class of aggregators. More generally, I show that correlation aversion dominates preferences for early resolution of uncertainty under the condition that the Arrow-Pratt index of risk aversion $\frac{\phi''}{\phi'}$ is convex. As for REU preferences, they posses a similar submodularity property, which can be seen immediately whenever $b(\cdot)$ is differentiable and decreasing. Indeed, the cross derivative of the aggregator will take the form

$$W_{xy} = b'(\cdot) \leq 0,$$

while such an aggregator is linear in y, and therefore indifferent to timing of resolution.

3.3 Attitudes toward correlation: ambiguity

Recursive ambiguity averse preferences have more complex attitudes toward correlation. The next example shows that this fact need not be true for recursive ambiguity averse preferences.

Example 2. Assume the same setting of the previous example, i.e. $S = \{s_1, s_2, s_3\}$, $\Sigma = 2^S$, $T = 2$ and $X = \mathbb{R}_+$ and same acts. Suppose this setting models the following situation: there is one urn containing 90 balls: 30 balls are red, while the remaining 60 balls are either black or yellow in unknown proportions. The state s_1 represents the event that a black ball is drawn, s_2 a yellow ball, and s_3 a red one. Consider recursive smooth ambiguity preferences, with $p^1(s_1) = 1 - p^1(s_2) = \frac{2}{9}$, $p^1(s_1) = 1 - p^1(s_2) = \frac{4}{9}$, $p^1(s_3) = p^2(s_3) = \frac{1}{3}$ and $\mu(p^1) = \mu(p^2) = \frac{1}{2}$. Assume $\phi(x) = \log(x)$, $u(x) = x$ and $\beta = \frac{9}{7\sqrt{47} - 8\sqrt{17}} \approx 0.6$. Under these assumptions,

$$V_0(f^{corr}) = \phi^{-1}\mathbb{E}_\mu \phi(\mathbb{E}_p V_1(\cdot, f^{corr})) = \mathbb{E}_{\bar{p}} V_1(\cdot, f^{corr}),$$

where $\bar{p} = \frac{1}{2}p^1 + \frac{1}{2}p^2$. Introducing positive correlation removes ambiguity from the perspective of time 0, in the sense that $V_1(\cdot, f^{corr})$ is unambiguous while $V_1(\cdot, f^{iid})$ is ambiguous. We have

$$V_0(f^{corr}) \approx 0.584 > 0.537 \approx V_0(f^{iid}).$$

$\triangle$

In other words, for recursive ambiguity preferences, an increase in correlation might reduce the ambiguity about future utility, thus making an increase in correlation desirable. The key is that we added correlation to an "ambiguous" consumption process. The next example suggests that if we add correlation to an "unambiguous" consumption process then this will produce (weakly) more volatility of future utility.

Example 3. Suppose now there is one urn containing 100 balls: 50 balls are red or blue with at least 20 of each color, while the remaining 50 balls are either black or yellow, again with at least 20 of each color. We can represent this as follows $S = \{s_1, s_2, s_3, s_4\}$, $\Sigma = 2^S$, $T = 2$ and $X = \mathbb{R}_+$. Let $f(s_1) = f(s_2) = 1$, $f(s_3) = f(s_4) = 0$, $g(s_2, s_3) = 1$ and $g(s_3, s_2) = 0$. The acts f^{iid} and f^{corr} are represented in Figure 7. Preferences are given by RMEU, with $I(\xi) = \min_{P \in C} \mathbb{E}_P \xi$, that satisfies $P(s_1, s_2) = \frac{1}{2}$, $P(s_3, s_4) = \frac{1}{2}$, $P(s_i) \geq \frac{1}{5}$ for $P \in C$ and $i = 1, 2, 3, 4$. It is easy to check that $f^{iid} \succ_0 f^{corr}$ for any $\beta \in (0, 1)$. In this case $f : S \to \mathbb{R}_+$ was chosen to be unambiguous (I make this notion more precise later). Positive correlation creates ambiguity, in the sense that at time 0 the future utility $V_1(\cdot, f^{iid})$ is unambiguous while $V_1(\cdot, f^{corr})$ is ambiguous. $\triangle$

I generalize this idea in the next theorem. First, I define formally an ambiguous act in this setting.

Definition 6. *A certainty equivalent $I : B_0(\Sigma, u(X)) \to \mathbb{R}$ admits a global benchmark if the set*

$$E_I = \{P \in \Delta : \mathbb{E}_P \xi \geqslant I(\xi) \text{ for all } \xi \in B_0(\Sigma, u(X))\},$$

is non-empty. Call $f : S \to X$ unambiguous if for some $P \in E_I$ it holds $I(u(f)) = \mathbb{E}_P u(f) = \mathbb{E}_Q u(f)$ for every $P, Q \in E_I$. Let $U_I \subseteq X^S$ denote the class of unambiguous acts.[19]

Remark 2. I identified ambiguity neutrality with expected utility. Such an assumption may be too restrictive in some cases. In the appendix (see subsection 7) I generalize the above definition by identifying ambiguity neutrality with probabilistic sophistication.

The next result extends Theorem 1 to recursive ambiguity averse preferences.

Theorem 2. *Assume that $\succeq_{t,\omega}$ has an RDAA representation where I has a global benchmark. Then $f^{iid} \succeq_0 f^{corr}$ for every $f \in U_I$.*

Proof. See the appendix. $\square$

Remark 3. RMEU and RSA preferences with ϕ CARA or CRRA satisfy the assumption of the theorem. The global benchmark for RMEU is given by any $P^* \in C$, and for RSA by $P^* = \mathbb{E}_\mu P$.

In the appendix I discuss in detail when the inequality will be strict, the main condition being that f^{corr} is not unambiguous in a precise sense (i.e increasing correlation effectively leads to ambiguity in future utility). I also discuss how (see subsection 7.4.4) Theorem 1 and 2 can be extended to an even more general class of recursive preferences.

[19]This notion is derived from Ghirardato & Marinacci (2002).

3.4 When do attitudes toward timing of resolution matter?

As a consequence of Theorems 1 and 2, to identify pure preferences for non-instrumental information one needs to remove correlation in the consumption process. This can be achieved by fixing consumption to a certain level independently of what state is realized before the resolution of uncertainty, as in Figure 8. Indeed Definition 2 considers processes in which consumption is fixed to a constant before uncertainty resolves.

A different setting in which attitudes toward non-instrumental information matter is the following. Consider enlarging the state space so as to consider processes that share the same pattern of correlation but whose uncertainty resolves at different times. For example, suppose that $S = A \times A$ for a finite set A and assume $T = 2$. Let $f = A \to X$ and define $h, h' : \Omega \to X$ by $h_1((a_1, a_2)) = f(a_1)$, $h_2((a_1, a_2)(a'_1, a'_2)) = f(a'_1)$, $h'_1((a_1, a_2), (a_1, a_2)) = f(a_1)$ and $h'_1((a_1, a_2), (a_1, a_2)) = f(a_2)$. Figure 9 gives an example for the case $A = \{a_1, a_2\}$. Observe that both h and h are "i.i.d." but h' resolves gradually while h resolves immediately. A preference for h' over h indicates a preference for one-shot resolution of uncertainty over gradual resolution.

However, in the standard dynamic consumption problem under uncertainty, consumption processes of the type described above do not play a role. First, consumption at every time t is never fully deterministic. Indeed, the consumption processes I described in the introduction do not allow for consumption to be constant at any time $t \geq 1$. Second, uncertainty resolves always gradually and therefore processes that feature one-shot resolution of uncertainty such as h are excluded. To illustrate, in the consumption-savings applications, consumption c_t at every period t is a non-trivial function of income y_t, and uncertainty about income resolves gradually. In consumption-based asset pricing models, in equilibrium one has $c_t = d_t$ where $(d_t)_t$ is the dividend process, whose uncertainty resolves gradually and which is usually assumed to be non-deterministic at every period t.

In other words, processes such as those in Figure 2 or h in Figure 9 that are used to identify pure preferences for non-instrumental information are not relevant in standard applications. For this reason, I suggest one should not take $\mathcal{F}$ as a domain of choice, but rather a strict subset of it, a relevant domain. Consider general information structures from section 2.1. Say that $f \in \mathcal{F}$ involves early resolution if for some $t \geq 1$, f_t is measurable w.r.t. $\mathcal{G}_\tau$ for some $\tau < t$. In words, this means that time t consumption is known at the earlier period τ.

Definition 7 (Relevant Domain)**.** *For every $t \in T$, let $\mathcal{F}_t$ denote the set given by*

$$\mathcal{F}_t = \{f \in \mathcal{F} : f_t \text{ is } \mathcal{G}_\tau\text{-measurable for some } \tau < t \implies f \in \mathcal{D}\}.$$

Define the relevant domain to be $\mathcal{F}^r = \cap_{t=1}^T \mathcal{F}_t$, which I endow with the relative topology.

In words, this means that an element of $\mathcal{F}^r$ either involves no early resolution or it is deterministic. Notably, it excludes processes as in Figures 8 and 9. At the same

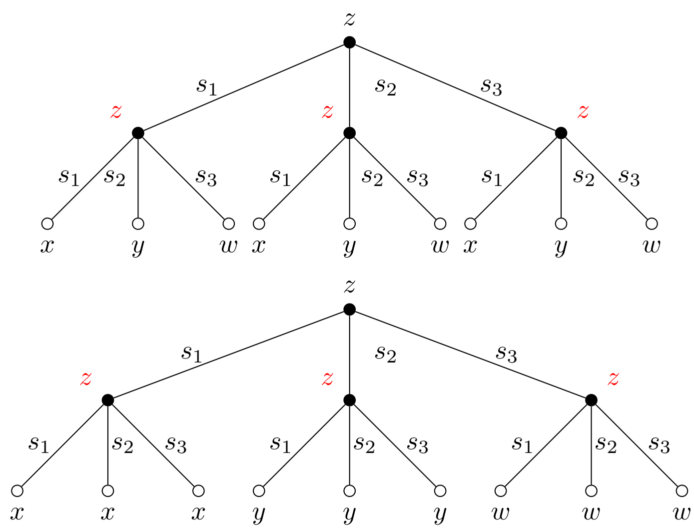

Figure 8: Early resolution vs late resolution

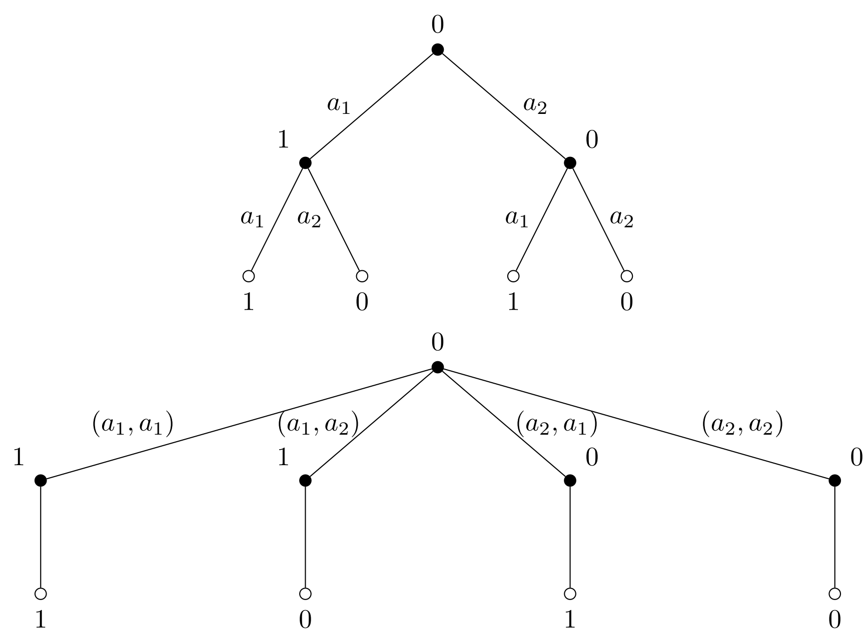

Figure 9: h and h'

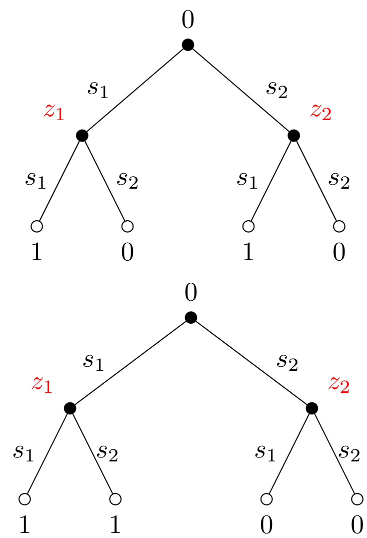

Figure 10

time, observe that $\mathcal{F}^r$ is rich enough to contain the consumption processes used in applications such as those I described in the introduction. In such a domain, pure attitudes toward timing of resolution do not play a role, in the sense that the ranking of $f, g \in \mathcal{F}^r$ will be determined by the interaction of attitudes toward correlation and attitudes toward non-instrumental information. For example, the ranking of the processes in Figure 10 will depend on how the values of z_1, z_2 affect the correlation between consumption at $t = 1$ and $t = 2$.

In the next section I assume that one can observe choices only over a subset $\mathcal{R}$ of $\mathcal{F}^r$. I show that one can axiomatize a general recursive representation and disentangle a general notion of risk aversion from intertemporal substitution on a choice domain in which preferences for non-instrumental information play no role.

4 Recursive preferences on the relevant domain $\mathcal{F}^r$

4.1 Representation theorem

Consider again the general information structures from section 2.1. Assume one observes choices only over a subset $\mathcal{R}$ of the relevant domain with $\mathcal{D} \subseteq \mathcal{R} \subseteq \mathcal{F}^r$. I therefore take $\succeq_{t,\omega}$ on $\mathcal{R}$ to be the primitive. As discussed in the introduction, such an assumption amounts to a weakening of the axiom of completeness of preferences.

I show that four standard axioms restricted to $\mathcal{R}$ are enough to characterize a general recursive representation. The first axiom is a standard continuity requirement.

Axiom 1 (Continuity)**.** For every $h \in \mathcal{R}$ the sets

$$\{f \in \mathcal{R} : f \succeq_{t,\omega} h\},$$

and

$$\{f \in \mathcal{R} : h \succeq_{t,\omega} f\},$$

are closed.[20]

Given $\tau \in T$, $x, y, z \in X$, and $d \in \mathcal{D}$, $(d_{-t-1}, y, x_{T-t'}, z_{t-t'})$ denotes the deterministic consumption stream that pays d_τ at times $\tau = 0, \ldots, t-1$, y at time t, x at times $\tau = t+1, \ldots, T+t-t'$ and z at times $\tau = T+t-t'+1, \ldots, T$. The next axiom, stationarity, requires preference over deterministic acts to be independent of a time delay.

Axiom 2 (Stationarity)**.** There exist $z \in X$ such that for every $t \leq t', \omega, \omega' \in \Omega$, $d \in \mathcal{D}$, $y, \bar{y}, x, \bar{x} \in X$

$$(d_{-t-1}, y, x_{T-t'}, z_{t-t'}) \succeq_{t,\omega} (d_{-t-1}, \bar{y}, \bar{x}_{T-t'}, z_{t-t'}) \iff (d_{-t'-1}, y, x_{T-t'}) \succeq_{t',\omega'} (d_{-t'-1}, \bar{y}, \bar{x}_{T-t'}).$$

The next axiom, which I refer to as consequentialism, requires that the decision maker at a note (t, ω) does not care about (i) what an act pays on unrealized events nor (ii) what it pays at earlier time periods.

Axiom 3 (Consequentialism)**.** For all $t \in T$ and $\omega \in \Omega$, and all acts $f, g \in \mathcal{R}$, if $f_k(\omega') = g_k(\omega')$ for all $k \geq t$ and all $\omega' \in \mathcal{G}_t(\omega)$, then $f \sim_{t,\omega} g$.

Observe that the above axiom implies that the ranking of an act $f \in \mathcal{R}$ by $\succeq_{t,\omega}$ depends only on $(f_t(\omega), f_{t+1}, \ldots, f_T)$.

Finally, the last axiom excludes preference reversals as new information arrives.

Axiom 4 (Dynamic Consistency)**.** For all $t \in T$, and $\omega \in \Omega$, and acts $f, g \in \mathcal{R}$ that yield identical outcomes up to and including period t, if $f \succeq_{t+1,\omega'} g$ for all $\omega' \in \mathcal{G}_t(\omega)$, then $f \succeq_{t,\omega} g$ and if $f \succ_{t+1,\omega'} g$ for some $\omega' \in \mathcal{G}_t(\omega)$, then $f \succ_{t,\omega} g$.[21]

[20]Recall that $\mathcal{F}$ is endowed by the product topology, and that therefore $\mathcal{R}$ can be endowed with the relative topology

[21]It is possible to consider a weaker axiom notion of dynamic incosistency, which would result in a certainty equivalent that $I_{t,\omega}$ need not be strictly monotone. This could be done by defining appropriately the notion of a $\succeq_{t,\omega}$-nonnull event. Then one can require that if $f \succ_{t+1,\omega'} g$ for every ω' in a $\succeq_{t,\omega}$ -nonnull event, then $f \succ_{t,\omega} g$. I chose to present the stronger representation as DC is easier to state.

Our representation theorem characterizes recursive utility under very general conditions (cf. Kreps & Porteus (1978), Johnsen & Donaldson (1985), Chew & Epstein (1991), Skiadas (1998), Wang (2003), Hayashi (2005), Bommier et al. (2017)), allowing for both changing beliefs and ambiguity sensitive preferences. The only loss of generality is constituted by the exclusion of an infinite horizon, which however can be overcome by means of appropriate technical conditions.

Theorem 3 (Recursive representation)**.** *Assume that $\mathcal{R}$ is connected. $\succeq_{t,\omega}$ satisfy axioms 1-4 if and only if admits a general recursive representation.*

Proof. See the appendix. □

Remark 4. One could wonder whether given a recursive representation $(W, u, (I_{t,\omega})_{t,\omega})$ on $\mathcal{R}$ the only "reasonable" extension to $\mathcal{F}$ is given by the straightforward extension of $(W, u, (I_{t,\omega})_{t,\omega})$ to $\mathcal{F}$. In Section 7.2, I show that one can extend preferences in a different fashion. Specifically, I introduce preferences that have an Epstein-Zin representation on $\mathcal{R}$ but on $\mathcal{F} \setminus \mathcal{R}$ admit the representation introduced by Selden & Stux (1978) and Selden (1978). Notably, such a "hybrid" model features indifference to timing of resolution of uncertainty.

Remark 5. The theorem makes no reference to uniqueness of the representation. Uniqueness can be achieved by adding further conditions that imply uniqueness of $u : X \to \mathbb{R}$. For example, one can assume that X is the set of lotteries over a finite set Z and obtain uniqueness of u by means of specific axioms such as independence.

At a technical level, the main difficulty introduced by weakening the completeness axiom is related to showing that $\mathcal{R} \subseteq \mathcal{F}^r$ is rich enough to construct a representation (more precisely, showing that $\mathcal{R}$ is connected). In the appendix (see Lemma and 1 and Remark 7) I show that $\mathcal{F}^r$ is connected. An example of great interest of a subset of $\mathcal{F}^r$ that is connected is given by

$$IND = \{h \in \mathcal{F} : \text{ there exist } (f_t)_t \text{ with } f_t \in X^S \text{ such that } h_t(s_1, \ldots, s_{t-1}, \cdot) = f_t(\cdot), \\ \text{and if for some } t', f_t' \text{ is constant } \implies h \in \mathcal{D}\}.$$

In words, this set contains processes that are "independent" (h_t does not depend on $(s_1, \ldots, s_{t-1})$) but not necessarily "identically distributed" (h_t depends on a function $f_t : S \to X$ which is not identical over time). Hence, in such a domain attitudes toward temporal resolution or correlation play no role.

4.2 Separating intertemporal substitution from attitudes toward uncertainty

A simple yet important implication of Theorem 3 is that to separate risk aversion from the intertemporal rate substitution it is enough to observe only choices over a subset $\mathcal{D} \subseteq \mathcal{R} \subseteq \mathcal{F}^r$, and therefore there are no implications for attitudes toward timing of resolution.

Consider preferences $\succeq^i_{t,\omega}$, $i = 1, 2$ that admit the representation in (1). Comparative risk aversion can be defined in a similar fashion as in Epstein & Zin (1989) (pp. 949-950) and Chew & Epstein (1991) (Theorem 3.2). For any $f \in \mathcal{R}$, (t, ω) and $d \in \mathcal{D}$, denote with $(f_t(\omega), d_{T-t})$ the consumption stream that pays $f_t(\omega)$ at time t and d_τ at times $\tau = t+1, \dots, T$.

Definition 8. *$\succeq^1_{t,\omega}$ is more risk averse than $\succeq^2_{t,\omega}$ if for every $f \in \mathcal{R}$, $d \in \mathcal{D}$ and (t, ω) with $t < T$*

$$(f_t(\omega), d_{T-t}) \succeq^2_{t,\omega} (f_t(\omega), f_{t+1}, \dots, f_T) \implies (f_t(\omega), d_{T-t}) \succeq^1_{t,\omega} (f_t(\omega), f_{t+1}, \dots, f_T),$$

and

$$(f_t(\omega), d_{T-t}) \succ^2_{t,\omega} (f_t(\omega), f_{t+1}, \dots, f_T) \implies (f_t(\omega), d_{T-t}) \succ^1_{t,\omega} (f_t(\omega), f_{t+1}, \dots, f_T).$$

Then the following result is immediate.

Proposition 1. *$\succeq^1_{t,\omega}$ is more risk averse than $\succeq^2_{t,\omega}$ if and only if they admit recursive representations $(W^i, u^i, (I^i_{t,\omega})_{t,\omega})$, $i = 1, 2$ such that $u^1 = u^2$, $W^1 = W^2$ and $I^1_{t,\omega}(\xi) \leq I^2_{t,\omega}(\xi)$ for every $\xi \in \{\xi \in B_0(\mathcal{G}_{t+1}, V_{t+1}) : \xi = V_{t+1}(\cdot, f), f \in \mathcal{R}\}$ and every (t, ω).*

Proof. First observe that if $W_1 = W_2$, then

$$(f_t(\omega), d_{T-t}) \succeq^i_{t,\omega} (f_t(\omega), f_{t+1}, \dots, f_T) \iff V_{t+1}(d) \geq I^i_{t,\omega}(V^i_{t+1}(\cdot, f)), \tag{3}$$

Now if $\succeq^1_{t,\omega}$ is more risk averse than $\succeq^2_{t,\omega}$ then it is straightforward to check that they rank prospects in $\mathcal{D}$ in the same way. It follows that they must admit recursive representations $(W^i, u^i, (I^i_{t,\omega})_{t,\omega})$, $i = 1, 2$ such that $u^1 = u^2$ and $W^1 = W^2$. By (3) it follows that $I^1_{t,\omega}(\xi) \leq I^2_{t,\omega}(\xi)$ for every $\xi \in \{\xi \in B_0(\mathcal{G}_{t+1}, V_{t+1}) : \xi = V_{t+1}(\cdot, f), f \in \mathcal{R}\}$. The converse follows immediately by (3). □

Remark 6. Observe that if $\mathcal{R} = \mathcal{D}$ then the "only if" part of the statement is trivially true since $I^i_{t,\omega}$ are defined on deterministic prospects so that $I^1_{t,\omega} = I^2_{t,\omega}$. More in general, this will be true whenever

$$\{\xi \in B_0(\mathcal{G}_{t+1}, V_{t+1}) : \xi = V_{t+1}(\cdot, f), f \in \mathcal{R}\} = \{\xi \in B_0(\mathcal{G}_{t+1}, V_{t+1}) : \xi = V_{t+1}(d), d \in \mathcal{D}\}. \tag{4}$$

Now consider the special case of EZ preferences. Assume that $\succeq^i_{t,\omega}$ are represented by

$$V^i_t(\omega, h) = \frac{h_t(\omega)^{\rho_i}}{\rho_i} + \beta_i(\mathbb{E}_{P_{t,\omega}}(V^i_{t+1}(\cdot, h)^{\frac{\alpha_i}{\rho_i}})^{\frac{\rho_i}{\alpha_i}}, \tag{5}$$

for $0 \neq \rho < 1$ and that (4) does not hold. In this case we obtain the following.

Corollary 1. $(\succeq_{t,\omega})^1_{t,\omega}$ *is more risk averse than* $(\succeq_{t,\omega})^2_{t,\omega}$ *if and only if* $\beta_1 = \beta_2$, $\rho_1 = \rho_2$ *and* $\alpha_1 \leq \alpha_2$.

Proof. The result follows immediately by the previous proposition upon observing that $I(\xi) = (\mathbb{E}_P \xi^\alpha)^{\frac{1}{\alpha}}$ is increasing in α (see Theorem 16, Hardy et al. (1952)). □

A separation exists between risk aversion and intertemporal substitution, but the domain is restricted to a domain of choice in which pure preferences for non-instrumental information play no role. In particular, one can take

$$\mathcal{R} = IND \cup (CORR \cap \mathcal{F}^r).$$

This domain is rich enough to disentangle risk aversion from intertemporal substitution, but as a consequence of Theorems 1 and 2, only attitudes toward temporal resolution play no role. Rather, only attitudes toward correlation matter in such a domain.

5 Important consequences

5.1 What determines aversion to correlation?

If aversion to correlation is the relevant behavioral property, then it is of great importance to understand what drives it. Theorem 1 implies that for EZ preferences, attitudes toward timing of resolution of uncertainty and attitudes are modeled by the same parameters. This is not true in general. From the literature we know that both RMEU and REU preferences are neutral toward timing of resolution, while from Theorems 1 and 2 we know that they are not indifferent to an increase in correlation. In other words, this means that in general attitudes toward timing of resolution and attitudes toward correlation need not be modeled by the same parameters.

In the case of REU preferences the degree of time non-separability modeled by $b(\cdot)$ drives aversion to correlation. However, for such preferences, risk aversion is tied to intertemporal substitution. It follows that risk aversion plays no role in determining aversion to correlation. In contrast, for EZ, RM and RDAA, correlation aversion is driven by static attitudes toward risk or ambiguity. An important implication for the LRR model is that if only purely static attitudes toward risk drive aversion to correlation, one is going to need a lot of persistence for correlation aversion to make a difference. I make this statement more precise in the next section.

It is natural to ask whether one can achieve a separation between risk aversion/ambiguity aversion, intertemporal substitution and attitudes toward correlation. Such a question will be pursued in future research.

σ	0.0078
φ	0.044
a	0 and 0.9790
β	0.998
$1-\alpha$	7.5
ρ	0
x_0	0

Table 1: Parameters of the LRR model (see Epstein et al. (2014), Table 1)

5.2 How much would you pay to eliminate persistence?

In light of the previous analysis, I re-examine Epstein, Farhi & Strzalecki's (2014) result that common parameter specifications lead to implausibly high timing premia. I ask a different question: "what fraction of your wealth would you give up to remove all persistence in consumption?" Consider the consumption process (case I) studied in Bansal & Yaron (2004)

$$\begin{aligned} \log\left(\frac{c_{t+1}}{c_t}\right) &= m + x_{t+1} + \sigma\epsilon_{c,t+1}, \\ x_{t+1} &= ax_t + \varphi\sigma\epsilon_{x,t+1}, \\ \epsilon_{c,t+1}, \epsilon_{x,t+1} &\sim \text{ i.i.d. } N(0,1). \end{aligned} \tag{6}$$

f^{iid} is given by $f_t^{iid} = \log(\frac{c_t}{c_{t-1}})$ for $a = 0$ (no persistence) and f^{corr} is given by $f_t^{corr} = \log(\frac{c_t}{c_{t-1}})$ for $a = 0.9790$ (a standard specification for persistence in the literature, see Bansal & Yaron (2004)). Table 1 summarizes the parameters of the model.

Using the same approach of Epstein et al. (2014) we can compute the utility associated to both f^{corr}, f^{iid}:

$$\log V_0(f^{corr}) = \log c_0 + \frac{\beta}{1-\beta a}x_0 + \frac{\beta}{1-\beta}m + \frac{\alpha}{2}\frac{\beta\sigma^2}{1-\beta}\left(1 + \frac{\varphi^2\beta^2}{(1-\beta a)^2}\right),$$

and

$$\log V_0(f^{iid}) = \log c_0 + \beta x_0 + \frac{\beta}{1-\beta}m + \frac{\alpha}{2}\frac{\beta\sigma^2}{1-\beta}\left(1 + \varphi^2\beta^2\right).$$

Define the persistence premium by

$$\pi = 1 - \frac{V(f^{corr})}{V(f^{iid})} = 1 - e^{\frac{\beta}{1-\beta a}x_0 - \beta x_0 + \frac{\alpha}{2}\frac{\beta\sigma^2}{1-\beta}\left(\frac{\varphi^2\beta^2}{(1-\beta a)^2} - \varphi^2\beta^2\right)}.$$

Under these assumptions, I obtain the persistence premium:

$$\pi = 1 - \exp\left(-6.5 \times 0.998 \times \frac{0.0078^2}{2(1-0.998)}\left(0.044^2 \times \frac{0.998^2}{(1-0.998\times 0.979)^2} - 0.044^2 \times 0.998^2\right)\right)$$
$$\approx 1 - e^{-8} \approx 0.3028.$$

When $1-\alpha = 10$, one obtains $\pi \approx 40\%$. In other words, an investor with such preferences would be willing to give up either 30% or 40% of his wealth to get rid of persistence. The experimental evidence from Andersen et al. (2018) suggests that one should have at most $\pi \approx 20\%$ (see section 7.4.5). This result suggests that persistence is either implausibly high or that risk aversion is too high relative to intertemporal substitution (recall from the previous section that the relationship between ρ and α determines attitudes toward correlation for EZ preferences).

5.3 Dynamic ordinal certainty equivalent models

Dynamic ordinal certainty equivalent (DOCE) preferences offer an alternative approach to disentangle risk aversion from intertemporal substitution. Axiomatized in Selden (1978) and Selden & Stux (1978), such preferences replace risky consumption in each period by certainty equivalents with respect to a utility function $v(\cdot)$ and evaluate the resulting sequence of certainty equivalents with discounted utility with respect to a utility function $u(\cdot)$. More precisely, Selden-Stux preferences $\succeq^{SS}_{s^t}$ over $\mathcal{F}$ are represented by

$$V_t(s^t, h) = u(h_t(s_1, \ldots, s_t)) + \sum_{j=1}^{T-t} \beta^j u\left(v^{-1}\left[\mathbb{E}_{\prod_{\tau=1}^{j} P(s_{t+\tau})} v\left(h_{t+j}(s^t, \cdot)\right)\right]\right).$$

Hall (1985), Zin et al. (1987), Attanasio & Weber (1989) and Kubler et al. (2019) have studied applications of such preferences.

In contrast with recursive preferences, DOCE preferences are neutral to the timing of resolution of uncertainty (see Selden & Stux (1978) and Kubler et al. (2019)). However, such preferences are also indifferent to correlation, as I now show.

Proposition 2. *$f^{iid} \sim^{SS}_0 f^{corr}$ for every $f \in X^S$.*

Proof. See the appendix. □

Because of the above result, DOCE models have less freedom than recursive preferences in addressing existing asset pricing puzzles. The main leeway is related to generalizing the formulation of the certainty equivalent to non-expected utility, similarly to how Epstein & Zin (1990) showed that one can obtain a partial resolution of the equity premium puzzle by considering Yaari's dual theory of choice under uncertainty (in the context of recursive preferences).

6 Discussion

6.1 Intertemporal hedging and experimental evidence

The connection between attitudes toward correlation and recursive preferences is not completely new. Kochov (2015) introduced an axiom called intertemporal hedging which has a similar content. Bommier et al. (2019) study preferences for intertemporal hedging in the case of recursive models of ambiguity aversion. Figure 11 has a visual description of the intertemporal hedging axiom. As one can see, his definition involves comparing processes that have the same timing of resolution but differ only in terms of correlation, and therefore are never part of the relevant domain $\mathcal{F}^r$. Since in applications uncertainty resolves gradually, it is necessary to understand the implications of correlation aversion in such a domain. Moreover, in this way it becomes possible to compare the relative strength of correlation aversion with preferences for non-instrumental information.

The notion of intertemporal hedging is based on the literature on correlation aversion which in turn is based on the framework of risk aversion with multiple commodities introduced by Kihlstrom & Mirman (1974), see for example Richard (1975) and Epstein & Tanny (1980). Miao & Zhong (2015) relate Epstein-Zin utility to an analogous notion of intertemporal hedging and provide experimental evidence in its favor. Bommier (2007) studies intertemporal hedging for continuous time models, providing a formula that relates a measure of intertemporal hedging to intertemporal substitution and risk aversion. Andersen et al. (2018) provide evidence in favor of intertemporal hedging.

6.2 Aversion to long-run risk

Strzalecki (2013) introduced the notion of aversion to long-run risk. Such a notion is related to Duffie & Epstein's (1992) example discussed in the introduction. Such a notion considers only maximally correlated processes. Therefore, it excludes reasonable patterns of correlation in a consumption process. For this reason, several important conceptual issues are ignored. For example, the fact that recursive preferences that feature ambiguity aversion can prefer more correlation is consumption are ignored. Notably, RMEU preferences are always indifferent to long-run risk, while introspection suggests they should be sensitive to correlation. In my setting, the notion of a correlated process can be extended in a straightforward manner as follows

$$F_f^{corr} = \{f^{corr} \in \mathcal{F} : \text{ for some } g_f \in G_f \text{ and}$$
$$J \subseteq \{2, \ldots, T\}, f_t^{corr}(s_1, \ldots, s_t) = g_f(s_{t-1}, s_t) \text{ for every } t \in J\}.$$

Under such a notion, it is immediately observable that correlation aversion implies aversion to long-run risk, but not the other way around.

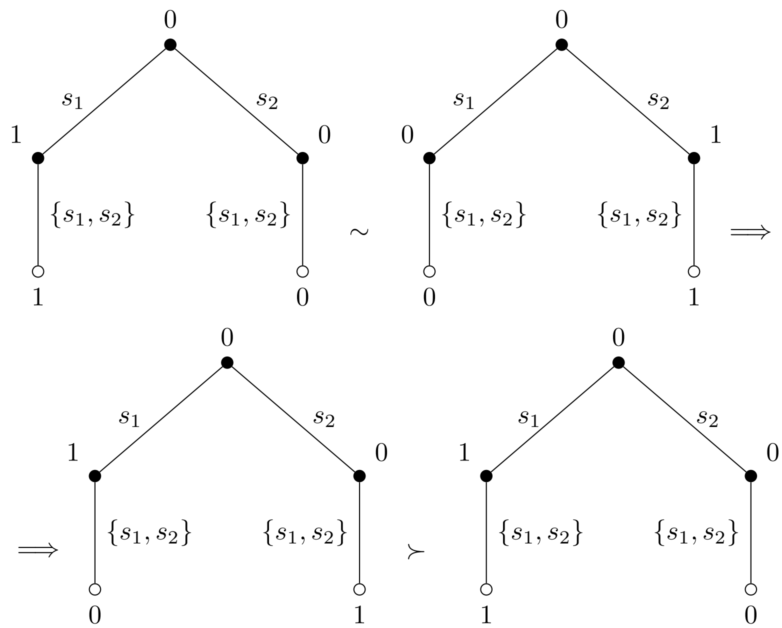

Figure 11: Intertemporal hedging

7 Appendix

7.1 Preferences for one-shot resolution of uncertainty

In this section I provide further results concerning attitudes toward temporal resolution of uncertainty. Specifically, I elaborate on the notion of timing premium introduced in Epstein et al. (2014). This notion quantifies how much would one pay to resolve all the uncertainty at time $t = 1$, as opposed to having gradual resolution of uncertainty. In other words, it quantifies preferences for one-shot resolution of uncertainty compared to gradual resolution. I translate their definition into an abstract framework.

Consider an IID setting. I impose more structure on S. Specifically, assume that $S = A_1 \times A_2 \ldots \times A_T$ where $A_1 = A_2 = \ldots = A_T$ and that each $(A_i, \mathcal{A}_i)$ is a finite measurable space with $\mathcal{A}_i = 2^{A_i}$. An arbitrary element $(s_1, \ldots, s_t)$ can be written as $((a_1^1, \ldots, a_T^1), \ldots, (a_1^t, \ldots, a_T^t))$.

Definition 9. *For any $f : A \to X$ measurable with respect to $\mathcal{A}_1$, let $\bar{f}_t(s_1, \ldots, s_t) = f(a_1^t)$ for every t, and let $\hat{f}$ be defined by $\hat{f}_t(s_1, \ldots, s_t) = f(a_t^1)$. The timing premium is defined by*

$$\pi^*(f) = 1 - \frac{V(\bar{f})}{V(\hat{f})}.$$

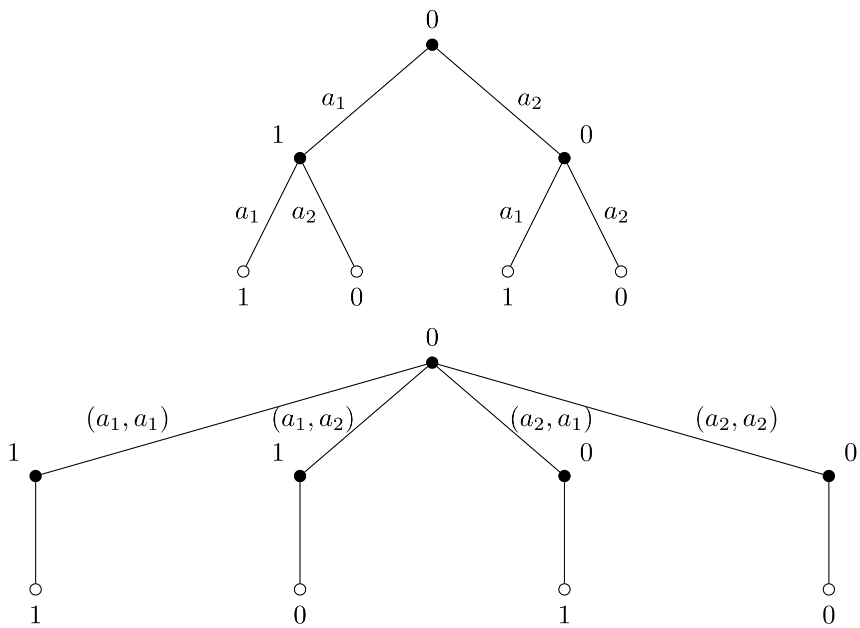

Figure 12: $\bar{f}$ and $\hat{f}$

Therefore $\succeq_0$ *has a non-negative time premium if and only if*

$$\hat{f} \succeq_0 \bar{f}.$$

An example of $\hat{f}$ with its associated $\bar{f}$ is represented in Figure 12. Observe that for simplicity, I am considering acts $\bar{f}$ that involve no correlation.

In general, the notions of preference for timing, timing premium, and attitudes toward timing are different. However, the three coincide for EZ preferences over $\mathcal{F}$. Recall that these are defined by

$$V_t(\omega, h) = \frac{h_t^\rho(\omega)}{\rho} + \beta(\mathbb{E}_P(V_{t+1}(\cdot, h)^{\frac{\alpha}{\rho}})^{\frac{\rho}{\alpha}}, \tag{7}$$

for $0 \neq \rho < 1$ and for $\rho = 0$

$$V_t(\omega, h) = \log(h_t(\omega)) + \beta \frac{1}{\alpha} \log\left(\mathbb{E}_P(\exp\left(\alpha V_{t+1}(\cdot, h)\right)\right). \tag{8}$$

Observe that I am further requiring that $P(a_1^1, \dots, a_T^1) = P(a_1^1, \dots, , a_T^t)$. Denote such preferences over $\mathcal{F}$ by $\succeq_{t,\omega}^{EZ}$.

Proposition 3. *Consider* $\succeq_{t,\omega}^{EZ}$ *over* $\mathcal{F}$*. Then the following are equivalent*

1. $\succeq_{t,\omega}$ *exhibits a preference for early resolution of uncertainty.*

2. *$\succeq_{t,\omega}$ is averse to correlation.*

3. *$\succeq_{t,\omega}$ has a non-negative time premium.*

In particular, 1-3 are equivalent to $\alpha \le \rho$.

Proof. See the appendix. □

In other words, for EZ preferences all such feature of preferences are modeled by the same parameter specifications. It should be no surprise that these three (different) properties of preferences have been conflated in the literature. In the next section I delve deeper into the relationship between preferences for early resolution of uncertainty and the timing premium.

7.2 A hybrid representation on $\mathcal{F}$

Proposition 3 established the equivalence between a non-negative timing premium and a preference for early resolution of uncertainty in the case of EZ preferences. I present an example of preferences $\mathcal{F}$ that are indifferent to timing of resolution but can have a positive timing premium.

Consider preferences $\succeq_{t,\omega}$ that have the representation (7) or (8) on $\mathcal{F}^r$.

I extend the representation to $\mathcal{F} \setminus \mathcal{F}^r$ by means of the DOCE representation introduced by Selden & Stux (1978), Selden (1978). Consider the following axioms on $\mathcal{F}$.

Axiom 5 (Indifference to timing)**.** Consider h, h' such that for some $\bar{s}^t = (\bar{s}_1, \dots, \bar{s}_t)$ with $1 \le t \le T-2$ the act

$$(h_0, h_1(\bar{s}_1), h_2(\bar{s}_1, \bar{s}_2), \dots, h_t(\bar{s}^t), \dots, h_T(\bar{s}^t, \cdot)),$$

resolves earlier than

$$(h'_0, h'_1(\bar{s}_1), h'_2(\bar{s}_1, \bar{s}_2), \dots, h'_t(\bar{s}^t), \dots, h'_T(\bar{s}^t, \cdot)),$$

and $h_\tau(s_1, \dots, s_\tau) = h'_\tau(s_1, \dots s_\tau)$ for every $(s_1, \dots, s_\tau)$ with $\tau < t$ and $h_\tau(s_1, \dots, s_\tau) = h'_\tau(s_1, \dots s_\tau)$ for every $(s_1, \dots s_\tau)$ such that $\tau \ge t$ and $(s_1, \dots, s_t) \ne (\bar{s}_1, \dots, \bar{s}_t)$. Then it holds that $h \sim_{t',\omega} h'$ for every $t' \le t$.

Axiom 6 (Consequentialism)**.** For all $t \in T$ and $\omega \in \Omega$, and all acts $f, g \in \mathcal{F}$ if $f_k(\omega') = g_k(\omega')$ for all $k \ge t$ and all $\omega' \in \mathcal{G}_t(\omega)$, then $f \sim_{t,\omega} g$.

Let $z \in X$ denote 0 when $\rho \ne 0$ and 1 when $\rho = 0$.

Axiom 7 (Consistency with Epstein-Zin)**.** Let h, h' be such that there exist, t, $f, f' : S \to X$ and $(s_1, \dots, s_{t-1})$ such that $h_t(s_1, \dots, s_{t-1}, \cdot) = f(\cdot)$, $h'_t(s_1, \dots, s_{t-1}, \cdot) = f'(\cdot)$, $h_\tau = h'_\tau = z$ for all $\tau \ne t$ and $h_t(\bar{s}^t) = h'_\tau(\bar{s}^t) = z$ whenever $(\bar{s}_1, \dots, \bar{s}_{t-1}) \ne (s_1, \dots, s_{t-1})$. Then

$$h \succeq_{t,\omega} h' \iff h \succeq^{EZ}_{t,\omega} h'.$$

Axiom 8 (Risk Independence)**.** Given any pair $h, h' \in \mathcal{F} \setminus \mathcal{F}^r$ which are identical except at the node $(s_1, \ldots, s_{t-1})$, then letting $\bar{h}_t(s_1 \ldots, s_{t-1}, \cdot) = h_t(s_1 \ldots, s_{t-1}, \cdot)$, $\bar{h}'_t(s_1 \ldots, s_{t-1}, \cdot) = h'_t(s_1 \ldots, s_{t-1}, \cdot)$ and $\bar{h}_t = \bar{h}'_t = z$ otherwise,

$$\bar{h} \sim_{t,\omega} \bar{h}' \implies h \sim_{t,\omega} h'.$$

Proposition 4. *$\succeq_{t,\omega}$ satisfies axioms 5-7 if and only if it is represented on $\mathcal{F} \setminus \mathcal{F}^r$ by*

$$V_t((s_1, \ldots, s_t), h) = u(h_t(s_1, \ldots, s_t)) + \sum_{j=1}^{T-t} \beta^j u\Bigg(\left[\mathbb{E}_{\prod_{\tau=1}^{j} P(s_{t+j})} h^{\alpha}_{t+j}(s^t, \cdot) \right]^{\frac{1}{\alpha}} \Bigg), \tag{9}$$

with

$$u(x) = \begin{cases} \frac{x^\rho}{\rho} & 0 \neq \rho < 1, \\ \log(x) & \rho = 0. \end{cases}$$

Moreover, $\succeq_{t,\omega}$ is represented on $\mathcal{F}$ by $V_t(\omega, \cdot)$ defined in (7)-(8) and (9).

Denote such preferences with $\succeq^{SS}_{t,\omega}$. Observe that $\succeq^{SS}_{t,\omega}$ satisfy dynamic consistency on $\mathcal{F}^r$ but at the same time are indifferent to timing of resolution of uncertainty. In general, the timing premium will be non-zero. However, in the specific case of i.i.d. processes, I can prove that it will always be smaller.

Proposition 5 (Timing premium inequality)**.** *Denote with π^{EZ} and π^{SS} the timing premium associated with $\succeq^{EZ}_{t,\omega}$ and $\succeq^{SS}_{t,\omega}$ EZ preferences and for the Selden-Stux representation, respectively. Then if $\alpha \leq \rho$ it holds that*

$$\pi^{SS}(f) \leq \pi^{EZ}(f),$$

for every $f : A \to X$.

Proof. See the appendix. □

In general, preferences for one-shot resolution of uncertainty reflect more than a pure attitude toward timing of resolution. To illustrate this fact, I compare below π^{EZ} and π^{SS} for the LRR consumption process in (6) (with the parameters in Table 1). As one can see from Table 2, π^{SS} can be also quite high even if such preferences are neutral to timing of resolution. Indeed, in this case, the high timing premium is driven by aversion to correlation.

7.3 Dynamic inconsistency

Epstein et al. (2014) make an important point

Table 2: Timing premium in the LRR model as a varies

	$a = 0$	$a = 0.8$	$a = 0.9$	$a = 0.972$	$a = 0.979$
Global EZ preferences	4.12%	4.3%	4.86%	12.39%	17.8%
EZ-SS preferences	0.2%	0.75%	1.36%	16.48%	26.1%

> At a psychic level, early resolution of risk may reduce anxiety. However, anxiety is plausibly more important when risk must be endured for a long time. Therefore, the risk premium required for bearing a lottery is greater the longer is the time that the individual has to live with the anxiety of not knowing how the lottery will be resolved. In other words, the willingness to bear a given risk declines as the date of resolution approaches, a form of dynamic inconsistency. However, such dynamic inconsistency is precluded when utility is recursive and thus anxiety cannot be a rationale for a timing premium given the utility functions considered here.

My approach permits addressing this point. Indeed, I can construct preferences that (i) are dynamically consistent on $\mathcal{F}^r$, (ii) have a preference for earlier resolution of uncertainty and (iii) the non-neutral attitudes toward timing stem from dynamic inconsistency.

Consider preferences $\succeq_{t,\omega}$ on $\mathcal{F}$ that are represented by

$$V_t(s^t, h) = u(h_t(s^t)) + \sum_{j=1}^{T} \beta^j [\mathbb{E}_{P(s^{t+j})} u(h_{t+j+1}(s^{t+j+1})) - \alpha(h)\beta \mathbb{E}_{P(s^{t+j})} \mathbb{VAR}_P(u(h_{t+j+1}(s^{t+j}, \cdot))], \tag{10}$$

with $\alpha(h) = 0$ for $h \in \mathcal{F}^r$ and for some $\alpha > 0$, $\alpha(h) = \alpha$ for every $h \in \mathcal{F} \setminus \mathcal{F}^r$.

In words, such preferences are RDEU on $\mathcal{F}^r$ but otherwise evaluate an act h by looking at its discounted expected value minus the expected discounted variance of h multiplied by a term $\alpha > 0$. It is easy to see that such preferences in general will not satisfy dynamic consistency. For example, suppose that $T = 2$ and $h = (z, z, f)$ for $f : S \to X$. Then $V_0(h) = \beta^2 \mathbb{E}_{P(s)} u(f(s)) - \alpha\beta^2 \mathbb{VAR}_P u(f(s))$. Notice that the functional $I(\xi) = \mathbb{E}_P \xi - \alpha \mathbb{VAR}_P \xi$ is not monotone. The intuition is simple: even if ξ might be better than ξ' in every state, ξ might be more volatile. Therefore, by Theorem 3 such preferences will not be dynamically consistent. Concerning attitudes toward timing, we have the following.

Proposition 6. *Assume preferences $\succeq_{t,\omega}$ have the representation in 10. Then $\succeq_{t,\omega}$ exhibits a preference for earlier resolution of uncertainty.*

Proof. See the appendix. □

Therefore, the parameter α can be thought of as measuring how anxious the decision maker is about not knowing how a process h will resolve: the later uncertainty resolves, the higher the future expected variance term will be.

7.4 Proofs

7.4.1 Proof of Theorem 3

I first start with a key lemma. Observe that since X is a subset of $\mathbb{R}^n$ and each $\mathcal{G}_t$ is generated by a finite partition, $\mathcal{F}$ is a metric space and so is $\mathcal{F}^r$ under the relative topology.

Lemma 1. *$\mathcal{F}^r$ is a separable and connected metric space.*

Proof. First observe that $\mathcal{F}$ is separable and therefore $\mathcal{F}^r$ is separable since any subset of a separable metric space is separable. I now show that $\mathcal{F}^r$ is path-connected. Take $h \in \mathcal{F}^r$ and $d \in \mathcal{D}$. Clearly if $h \in \mathcal{D}$ then the result follows by convexity of $\mathcal{D}$ (recall that X is convex). Assume $h \in \mathcal{F}^r \setminus \mathcal{D}$. Let $\{P_1^t, \ldots, P_{n^t}^t\}$ denote the partition of Ω that generates $\mathcal{G}_t$. I construct a continuous path $\iota : [0,1] \to \mathcal{F}^r$ that connects h to d. Since X is convex, for every t we just let $\iota_t(\alpha) = (1-\alpha)h_t + \alpha d_t$. Fix $t \geq 1$. Without loss of generality, assume that $n^t - n^{t-1} = 1$ and $P_{n^{t-1}}^{t-1} = P_{n^t}^t \cup P_{n^t-1}^t$. Let $\omega \in P_{n^t}^t$ and $\omega' \in P_{n^t-1}^t$. If $(1-\alpha)h_t(\omega)+\alpha d_t = (1-\alpha)h_t(\omega')+\alpha d_t$, we obtain a contradiction since $h_t(\omega) = h_t(\omega')$ but $h \in \mathcal{F}^r \setminus \mathcal{D}$. Therefore, $\iota_t(\alpha) \in \mathcal{F}^r$ for every α. It follows that we can connect via a path any $f \in \mathcal{F}^r$ to $d \in \mathcal{D}$. Hence, we can connect any $h, h' \in \mathcal{F}^r$ by a path. We conclude that $\mathcal{F}^r$ is path-connected and therefore connected.

□

Remark 7. It is easy to construct examples of strict subsets of $\mathcal{F}^r$ that are also connected. For example, consider the case in which $\Omega = S^T$ with $T \geq 1$ where (S, Σ) is a finite measurable space with $\Sigma = 2^S$ and $\mathcal{G}_t = \Sigma^t \times \{\emptyset, S\}^{T-t}$. Then the set

$$IND = \{h \in \mathcal{F} : \text{ there exist } (f_t)_t \text{ with } f_t \in X^S \text{ such that } h_t(s_1, \ldots, s_{t-1}, \cdot) = f_t(\cdot),$$
$$\text{and if for some } t', f_t' \text{ is constant } \implies h \in \mathcal{D}\},$$

is easily seen to be connected. Observe that such a domain is the natural extension to T periods of "certain × uncertain" consumption plans (e.g., see Selden (1978), Johnsen & Donaldson (1985)). Indeed, the two coincide when $T = 1$.

I turn now to the proof of the Theorem 3. See also Johnsen & Donaldson (1985), Proposition 2.

Proof of Theorem 3. I first prove sufficiency of the axioms. First by continuity, consequentialism and since $\mathcal{R}$ is connected and separable by Lemma 1, one can apply well known results from Debreu (1954) to show that there exist (sequentially) continuous functions $(V_t(\omega, \cdot))_{t,\omega}$ such that

$$V_t(\omega, h) = V_t(h_t(\omega), h_t, \ldots, h_T) \quad \text{for every } h \in \mathcal{R}.$$

Observe that by stationarity there exists a (sequentially) continuous function $u : X \to \mathbb{R}$ such that $V_T(\omega, h) = u(h_T(\omega))$ and $V_t(\omega, (x, z_{T-t-1})) = u(x)$ for every $\omega \in \Omega$ and $t < T$. Moreover, we can normalize $u(\cdot)$ from the stationarity axiom so that $u(z) = 0$.

I construct $I_{t,\omega} : B_0(\mathcal{G}_{t+1}, V_{t+1}(\omega, \mathcal{R})) \to \mathbb{R}$ as follows: for every h, by continuity, dynamic consistency, consequentialism, and stationarity we can construct $d_{\omega,t} = (d_{t+1}, \ldots, d_T) \in X^{T-t}$ such that for any $\bar{d} \in \mathcal{D}$

$$h \sim_{t,\omega} (\bar{d}_{-t-1}, h_t(\omega), d_{\omega,t}) \in \mathcal{D}. \tag{11}$$

Observe that all acts in (11) belong to $\mathcal{R}$. In particular, $d_{\omega,t}$ can be constructed recursively as follows. Starting from $t = T - 1$, observe that for any $\omega \in \Omega$, there exist $x, y \in X$ such that

$$V_{T-1}(h_{T-1}(\omega), x) \geq V_{T-1}(h_{T-1}(\omega), h_T) \geq V_{T-1}(h_{T-1}, y).$$

To see this, let $x = h_T(\bar{\omega})$ and $y = h_T(\underline{\omega})$, where $\bar{\omega} = \arg\max_\omega u(h_T(\omega))$ and $\underline{\omega} = \arg\min_\omega u(h_T(\omega))$. The statement follows by applying dynamic consistency. Therefore, by continuity and connectedness X we can find $d_{T-1,\omega} \in X$ such that $h \sim_{T-1,\omega} (\bar{d}_{-t-1}, h_T(\omega), d_{T-1,\omega})$. Now for any $t < T - 1$ and ω, assume one has constructed $d_{t+1,\omega'}$ for every $\omega' \in \mathcal{G}_t(\omega)$. Let $\bar{d}_{t,\omega} = (h_{t+1}(\bar{\omega}), d_{t+1,\bar{\omega}})$ and $\underline{d}_{t,\omega} = (h_{t+1}(\underline{\omega}), d_{t+1,\underline{\omega}})$ where

$$\bar{\omega} = \arg\max_{\omega'} V(h_{t+1}(\omega), d_{t+1,\omega'}),$$

and

$$\underline{\omega} = \arg\min_{\omega'} V(h_{t+1}(\omega), d_{t+1,\omega'}),$$

Then by dynamic consistency and stationarity we have

$$V_t(h_t(\omega), \bar{d}_{t,\omega}) \geq V_t(\omega, h) \geq V_t(h_t(\omega), \underline{d}_{t,\omega}).$$

Again, by connectedness of X and continuity we can find $d_{t,\omega}$ such that (11) is verified.

Now observe that this implies that for each (t, ω), $t = 0, \ldots, T$ and $\omega \in \Omega$ we have $V_t(\omega, \mathcal{R}) = V_t(\omega', \mathcal{R}) \equiv V_t$ (observe that $V_t \subseteq V_{t'}$ whenever $t' \leq t$). Define

$$I_{t,\omega} : B_0(\mathcal{G}_{t+1}, V_{t+1}) \to \mathbb{R},$$

by $I_{t,\omega}(\xi) = V_{t+1}(d_{\omega,t})$ and where $\xi(\omega) = V_{t+1}(\omega, h)$. Observe that $I_{t,\omega}$ is well defined by dynamic consistency.

I now claim that $I_{t,\omega}$ is strictly monotone, normalized and continuous. That $I_{t,\omega}$ is normalized follows by definition. Strict monotonicity follows by dynamic consistency. To prove continuity, assume that $\xi_n \to \xi$. Let h_n and h satisfy $\xi_n = V_{t+1}(\cdot, h_n)$, $\xi = V_{t+1}(\cdot, h)$ and $\lim h_n = h$. By contradiction, suppose that $I_{t,\omega}(\xi_n) \not\to I_{t,\omega}(\xi_n)$. It follows that $V_{t+1}(d^n_{t,\omega}) \not\to V_{t+1}(d_{t,\omega})$. Hence, there exists $\varepsilon \geq 0$ such that for every $N \in \mathbb{N}$ there exists $n \geq N$ such that

$$|V_{t+1}(d_{t,\omega}) - V_{t+1}(d^n_{t,\omega})| \geq \varepsilon > 0.$$

By dynamic consistency it follows that there exists $\epsilon > 0$ such that for every $N \in \mathbb{N}$ there exists $n \geq N$ such that

$$|V_t\left(h_t(\omega), V_{t+1}(d^n_{t,\omega})\right) - V_t\left(h_t(\omega), V_{t+1}(d_{t,\omega})\right)| \geq \epsilon > 0.$$

Observe that by continuity we have $V_t(\omega, h_n) \to V_t(\omega, h)$. Hence, we have arrived at a contradiction. Therefore $I_{t,\omega}(\xi_n) \to I_{t,\omega}(\xi_n)$ as desired.

Now assume that $h_t(\omega) = h'_t(\omega)$ and $I_{t,\omega}(V_{t+1}(\cdot, h)) = I_{t,\omega}(V_{t+1}(\cdot, h'))$. By dynamic consistency, it follows that $h \sim_{t,\omega} h'$. Moreover, if $I_{t,\omega}(V_{t+1}(\cdot, h)) > I_{t,\omega}(V_{t+1}(\cdot, h'))$ then $h_{t,\omega} \succ_{t,\omega} h'$. By Lemma 1 in Gorman (1968) it follows that there exists a continuous function $W_t : X \times V_{t+1} \to \mathbb{R}$ strictly increasing in its second argument such that

$$V_t(\omega, h) = W_t\left(h_t(\omega), I_{t,\omega}\left(V_{t+1}(\cdot, h)\right)\right).$$

Finally observe that by stationarity it holds that $W_t(x, y) = W_{t'}(x, y)$ for every t, t', $x \in X$ and $y \in V_{\max\{t,t'\}+1}$. Therefore, we can set $W \equiv W_0$, which delivers the representation.

I now turn to the necessity of the axioms. It is immediate to check that the recursive representation satisfies axiom 3. To show that the representation satisfies continuity, take $h \in \mathcal{R}$ and a sequence $(f_n)_n$ in $\mathcal{R}$ such that $f_n \succeq_{t,\omega} h$ and $\lim f_n = f$. This means that $V_t(\omega, f_n) \geq V_t(h)$ for every n so that by sequential continuity of $V_t(\omega, \cdot)$ we obtain that the set

$$\{f \in \mathcal{R} : f \succeq_{t,\omega} h\},$$

is closed. Showing that the set

$$\{f \in \mathcal{R} : h \succeq_{t,\omega} f\},$$

is closed can be done in the same way. Turn now to axiom 2. Let $z \in X$ be such that $u(z) = 0$ and $W(x, u(z)) = u(x)$ (we know z exists by assumption). Now for every $t \leq t', \omega, \omega'$, $d \in \mathcal{D}$, $y, \bar{y}, x, \bar{x} \in X$ it holds that $V_{t'+1}(x_{T-t'}) = V_{t+1}((x_{T-t'}, z_{t-t'}))$. It follows that

$$\begin{aligned}
V_t(\omega, (d_{-t-1}, y, x_{T-t'}, z_{t-t'})) &= W(y, V_{t+1}(x_{T-t'})) \\
\geq V_t\left(\omega, (d_{-t-1}, \bar{y}, \bar{x}_{T-t'}, z_{t-t'})\right) &= W(\bar{y}, V_{t+1}(\bar{x}_{T-t'})) \\
\iff V_{t'}(\omega, (d_{-t'-1}, y, x_{T-t'})) &= W(y, V_{t'+1}(x_{T-t'})) \\
\geq V_{t'}\left(\omega, (d_{-t'-1}, \bar{y}, \bar{x}_{T-t'})\right) &= W(\bar{y}, V_{t'+1}(\bar{x}_{T-t'})),
\end{aligned}$$

which implies that axiom 2 is satisfied. Finally, take $h, h' \in \mathcal{R}$ and (t, ω) with $h_t(\omega) = h'_t(\omega)$. If $h \succeq_{t+1,\omega'} h'$ for every $\omega' \in \mathcal{G}_t(\omega)$ then $V_{t+1}(\omega', h) \succeq_{t+1,\omega'} V_{t+1}(\omega', h')$ which by monotonicity of $I_{t,\omega}$ implies $I_{t,\omega}(V_{t+1}(\cdot, h)) \geq I_{t,\omega}(V_{t+1}(\cdot, h'))$. Since W is strictly increasing in its second variable, it follows that $V_t(\omega, h) \succeq_{t,\omega} V_t(\omega, h')$ as desired. Moreover, if for some $\omega' \in \mathcal{G}_t(\omega)$ the inequality is strict, then by strict monotonicity of $I_{t,\omega}$ we get $I_{t,\omega}(V_{t+1}(\cdot, h)) > I_{t,\omega}(V_{t+1}(\cdot, h'))$ as desired. □

7.4.2 Proof of Theorem 1

I first provide an important result that will be useful later.

Lemma 2. *Consider a probability p on $(S, 2^S)$. Then for every $f : S \to X$ and $g : S \times S \to X$ as in Definition 3 it holds that*

$$\sum_{s' \in S} p(s') \sum_{s \in S} p(s) u(g(s, s')) = \sum_{s \in S} p(s) u(f(s)).$$

Proof. Observe that

$$\sum_{s \in S} p(s) u(g(s, s')) = u(f(s')) + \varepsilon(s'), \tag{12}$$

for every $s' \in S$, where

$$\varepsilon(s') = \sum_{s \notin B(s')} p(s) u(f(s)) - u(f(s')) \sum_{s \notin B(s')} p(s),$$

and $B(s') = \{s \in S : g(s, s') = f(s')\}$.

Now the statement follows by observing that $\sum'_s p(s')\varepsilon(s') = 0$. Indeed, we have

$$\sum_{s' \in S} p(s')\varepsilon(s') = \sum_{s \in S'} p(s') \sum_{s \notin B(s')} p(s) u(f(s)) - \sum_{s' \in S} p(s') u(f(s')) \sum_{s \notin B(s')} p(s).$$

Observe that by definition of $g(\cdot, \cdot)$ it holds $s \in B(s') \iff s' \in B(s)$. Therefore

$$\begin{aligned}\sum_{s \in S'} p(s') \sum_{s \notin B(s')} p(s) u(f(s)) &= \sum_{s \in S'} \sum_{s \notin B(s')} p(s') p(s) u(f(s)) \\ &= \frac{1}{2} \sum_{s \in S'} \sum_{s \notin B(s')} p(s') p(s) (u(f(s)) + u(f(s'))),\end{aligned}$$

and

$$\begin{aligned}\sum_{s' \in S} p(s') u(f(s')) \sum_{s \notin B(s')} p(s) &= \sum_{s \in S'} \sum_{s \notin B(s')} p(s') p(s) u(f(s')) \\ &= \frac{1}{2} \sum_{s \in S'} \sum_{s \notin B(s')} p(s') p(s) (u(f(s)) + u(f(s'))).\end{aligned}$$

Hence

$$\sum_{s' \in S} p(s')\varepsilon(s') = 0,$$

as desired. □

Observe that Lemma 2 implies that $f^{iid} \sim_0^{RDEU} f^{corr}$ where $\succeq_{t,\omega}^{RDEU}$ is any RDEU preference.

Turning to the proof of Theorem 1, I first introduce formally important notions related to quasi-arithmetic means defined by

$$M_{\phi,P}(\xi) = \phi^{-1}\left(\mathbb{E}_P\left(\phi\left(\xi\right)\right)\right) = \phi^{-1}\left(\sum_{i=1}^{n} \phi(\xi(s_i))P(s_i)\right), \tag{13}$$

for every $\xi \in B_0(S, \mathbb{R}_+)$.

Definition 10. *Consider $\phi : \mathbb{R}_+ \to \mathbb{R}$ increasing, concave and twice differentiable. Say that ϕ satisfies*

1. *PERU (preference for early resolution of uncertainty) if for every $\beta \in (0, 1)$*

$$\beta\left[-\frac{\phi''(\beta y + x)}{\phi'(\beta y + x)}\right] \le \left[-\frac{\phi''(y)}{\phi'(y)}\right], \tag{14}$$

2. *DARA (decreasing absolute risk aversion) if $x \ge y$ implies*

$$-\frac{\phi''(x)}{\phi'(x)} \le -\frac{\phi''(y)}{\phi'(y)},$$

   and satisfies IARA (increasing absolute risk aversion) if the above inequality is true with the opposite sign;

3. *IRRA (increasing relative risk aversion) if $x \ge y$ implies*

$$-x\frac{\phi''(x)}{\phi'(x)} \le -y\frac{\phi''(y)}{\phi'(y)}.$$

The following result, whose proof I omit, is immediate.

Proposition 7. *Consider $\phi : \mathbb{R}_+ \to \mathbb{R}$ increasing, concave and twice differentiable. If ϕ satisfies PERU, then it satisfies DARA.*

I provide an important result concerning the concavity of the quasi-arithmetic means defined in (13).

Theorem 4. *Assume that $\phi : \mathbb{R}_+ \to \mathbb{R}$ strictly increasing, strictly concave, and twice differentiable. The following are equivalent:*

1. *The quasi-arithmetic mean $M_{\phi,P}$ is concave.*

2. *The Arrow-Pratt-de Finetti coefficient of absolute risk aversion $-\frac{\phi''}{\phi'}$ is convex.*

Proof. First observe that $-\frac{\phi''}{\phi'}$ is convex if and only if $\frac{\phi'}{\phi''}$ is convex. It suffices to prove that $\frac{\phi'}{\phi''}$ is convex if and only if $\frac{\phi''}{\phi'}$ is concave. Let $A(x) = \frac{\phi''(x)}{\phi'(x)}$ for every $x \in u(X)$. Take $x, y \in u(X)$. It is enough to prove that

$$\frac{1}{A\left(\frac{x+y}{2}\right)} \leq \frac{1}{2}\left(\frac{1}{A(x)} + \frac{1}{A(y)}\right). \tag{15}$$

Recall that in general it holds that

$$\frac{x+y}{2} \geq \frac{1}{\frac{1}{2}(\frac{1}{x} + \frac{1}{y})}.$$

Hence the claim follows by rewriting (15) as

$$\frac{1}{\frac{1}{2}\left(\frac{1}{A(x)} + \frac{1}{A(y)}\right)} \leq A\left(\frac{x+y}{2}\right).$$

Now the result follows by an application of Theorem 1 and Theorem 5 in Chudziak et al. (2019). □

Thanks to Theorem 4, we obtain the following powerful result, which shows that the conjunction of DARA and IRRA on ϕ imply the concavity of the quasi-arithmetic mean $M_{\phi,P}$.

Corollary 2. *Assume that ϕ satisfies PERU. Then $M_{\phi,P}$ is concave if and only if ϕ satisfies IRRA.*

Proof. First observe that by Theorem 4, if $M_{\phi,P}$ is concave then $\frac{\phi''}{\phi'}$ is convex, which is easily seen to imply that the mapping $x \mapsto -x\frac{\phi''(x)}{\phi'(x)}$ is increasing. This reasoning proves the "only if" part of the proof. To prove the "if" part, observe that

$$(xA(x))' = xA''(x) + 2A'(x).$$

If ϕ satisfies PERU, then $A'(x) \leq 0$ by Proposition 7. Hence if ϕ satisfies IRRA, it has to be the case that

$$A''(x) \geq 0.$$

The result therefore follows by Theorem 4. □

Such a result in a way completes Theorem 12 and Corollary 1 in Marinacci & Montrucchio (2010), which characterize when quasi-arithmetic means are constant subadditive and subhomogeneous. Indeed, one way to think about Corollary 2 is that it implies that for quasi-arithmetic means to be concave it is enough to assume constant superadditivity (DARA) and subhomogeneity (IRRA).

Corollary 3. *Assume that ϕ is given by*

$$\phi(x) = \int_0^x e^{-t^2} dt,$$

for every $x \in \mathbb{R}_+$. Then $M_{\phi,P}$ is concave.

Proof. The Arrow-Pratt-de Finetti index in this case is given by $-\frac{\phi''(x)}{\phi'(x)} = 2x$ and the result therefore follows by Theorem 4. □

Corollary 4. *Assume that ϕ is given by $\phi(x) = \frac{x^\lambda}{\lambda}$ for $\lambda < 1$ or $\phi(x) = -e^{-\theta x}$ with $\theta \geq 0$ for every $x \in \mathbb{R}_+$. Then $M_{\phi,P}$ is concave.*

Proof. In both cases, the proof follows by Theorem 4, upon observing that . □

Proof of Theorem 1. Let $\xi = u(f)$ and ξ' by $\xi'(s, s') = u(f_t^{corr}(s, s'))$ for every $s, s' \in S$. For (i) and (ii), we reason as follows. Now as mentioned, the functional

$$M_{\phi,P}(\cdot) = \phi^{-1}(\mathbb{E}_P \phi(\cdot)),$$

is concave in the EZ and RM case. Now by Lemma 2 we have that

$$\mathbb{E}_P(\xi'(s, \cdot)) = \xi + \varepsilon,$$

and

$$\mathbb{E}_P \varepsilon = 0.$$

Now let

$$V_{t-1}(\omega, f^{corr}) = \xi(\omega) + \beta\phi^{-1}\mathbb{E}_P\phi(\xi'(s, \cdot) + k),$$

and

$$V_{t-1}(\omega, f^{iid}) = \xi(\omega) + \beta\phi^{-1}\mathbb{E}_P\phi(\xi(\cdot) + k),$$

where $k = \beta\phi^{-1}\mathbb{E}_\mathbb{P}\phi(V_{t+1})$. By Corollary 4, $M_{\phi,P}$ is concave and therefore by further applying Lemma 2 we obtain

$$\mathbb{E}_P\phi^{-1}\mathbb{E}_P\phi(\xi'(s, \cdot) + k) \leq \phi^{-1}\mathbb{E}_P\phi(\mathbb{E}_P(\xi(\cdot) + k).$$

Hence the random variable

$$V_{t-1}(\omega, f^{corr}) = \xi(\omega) + \beta\phi^{-1}\mathbb{E}_P\phi(\xi'(s, \cdot) + k),$$

is dominated according to second order stochastic dominance by the random variable

$$V_{t-1}(\omega, f^{iid}) = \xi(\omega) + \beta\phi^{-1}\mathbb{E}_P\phi(\xi(\cdot) + k).$$

Since ϕ is concave, it follows that

$$V_{t-2}(\omega, f^{iid}) \geq V_{t-2}(\omega, f^{corr})$$

for every ω, whence $V_0(f^{corr}) - V_0(f^{iid}) \leq 0$. Moreover, observe that the inequality will be strict whenever g is non-trivial since ϕ is strictly concave.

As for (iii), first observe that since S is finite we can write it as $S = \{s_1, \ldots, s_n\}$. Therefore, $\mathbb{E}_P \xi = \sum_{i=1}^n P(s_i)\xi(s_i)$. Without loss of generality, assume that $f_2^{corr} = g$. Now we have

$$V_0(f^{corr}) - V_0(f^{iid}) =$$
$$\sum_{i=1}^n b(f(s_i)) \sum_{i=1}^n p(s_i) b(f(s_i)) \left(\sum_{j:s_j \notin B(s_i)} p(s_j) \left(u(g(s_i, s_j) - u(f(s_j))) \right) \right) +$$
$$+ \sum_{i=1}^n p(s_i) u(f(s_i)) \sum_{i=1}^n b(f(s_i)) p(s_i) \sum_{j:s_j \notin B(s_i)} p(s_j) \left[b(g(s_i, s_j) - b((f(s_j)))) \right]$$

Note that
$$\sum_{i=1}^n b(f(s_i)) p(s_i) \sum_{j:s_j \notin B(s_i)} p(s_j) \left[b(g(s_i, s_j) - bf(s_j)) \right] = 0.$$

Moreover,

$$\sum_{i=1}^n p(s_i) b(f(s_i)) \left(\sum_{j:s_j \in B(s_i)} p(s_j) \left[u(g(s_i, s_j) - u(f(s_j)) \right] \right) =$$
$$\sum_{i=1}^n \left(\sum_{j:s_j \notin B(s_i)} b(f(s_i)) p(s_i) p(s_j) \left[u\left(f(s_i)\right) - u(f(s_j)) \right] \right) =$$
$$\frac{1}{2} \sum_{i=1}^n \sum_{j:s_j \notin B(s_i)} p(s_i) p(s_j) \left[u(f(s_i)) - u(f(s_j)) \right] \left(b(f(s_j)) - b(f(s_i)) \right).$$

Now observe that

$$u(f(s_i)) - u(f(s)) \leq 0 \iff b(f(s)) - b(f(s_i)) \geq 0.$$

Therefore $V_0(f^{corr}) - V_0(f^{iid}) \leq 0$. Moreover, the inequality will be strict whenever b is strictly decreasing, f is not constant and $g \neq f$.

□

Remark 8. It is important to repeat that the result holds for a general class of recursive second-order expected utility preferences, i.e. those for which the function ϕ satisfies the condition that $-\frac{\phi''}{\phi'}$ be convex, even when ϕ does not satisfy PERU. For example, if ϕ is given by
$$\phi(x) = \int_0^x e^{-t^2} dt,$$
by Corollary 3 correlation aversion still applies. At the same time, for such a ϕ PERU in general will not be satisfied.

7.4.3 Proof of Theorem 2

Lemma 3. *Suppose $I : B_0(\Sigma, u(X))$ is concave with benchmark given by $\xi \mapsto \mathbb{E}_P\xi$. Then if $I(\xi) = \mathbb{E}_P\xi$ for any $x \in u(X)$ it holds that $I(\xi + \beta x) = \mathbb{E}_P(\xi + \beta x)$ for every $\beta \in (0, 1)$.*

Proof. If I is constant-additive it follows that $I(\xi + \beta x) = I(\xi) + \beta x = \mathbb{E}_P(\xi) + \beta x = \mathbb{E}_P(\xi + \beta x)$ as desired. If I is homogeneous, then $\mathbb{E}_P(\xi + \beta x) \geq I(\xi + \beta x) = I\left(\frac{1+\beta}{1+\beta}(\xi + \beta x)\right) = (1 + \beta)I\left(\frac{1}{1+\beta}\xi + \frac{1}{1+\beta}x\right) \geq I(\xi) + \beta I(x) = \mathbb{E}_P(\xi + \beta x)$. □

Theorem 5. *Consider the second-order quasi-arithmetic mean given by*

$$M_{\phi,\mu}(\xi) = \phi^{-1}\left(\int \phi\left(\sum_{i=1}^{n} \xi(s_i)P(s_i)\right) \mathrm{d}\mu(P)\right),$$

where μ is a probability measure over $\Delta(S)$. If $-\frac{\phi''}{\phi'}$ is convex, then $M_{\phi,\mu}$ is concave.

Proof. Take a sequence of discrete measures $(\mu_n)_n$ that converges to μ in the weak*-topology. Then by applying Theorem 4 we have for every ξ, η

$$M_{\phi,\mu_n}\left(\frac{1}{2}\xi + \frac{1}{2}\eta\right) \geq \frac{1}{2}M_{\phi,\mu_n}(\xi) + \frac{1}{2}M_{\phi,\mu_n}(\eta).$$

The result follows by taking limits. □

Remark 9. The previous result implies that RSA preferences feature a concave certainty equivalent, and therefore satisfy the assumptions of Theorem 2.

Proof of Theorem 2. Consider any unambiguous act f and associated f^{iid} and f^{corr}. Without loss of generality, assume that $f_2^{corr}(s_1, \ldots, s_t) = g(s_{t-1}, s_t)$. Let $\xi = u(f)$ and $\xi'(s, s') = u(g(s, s'))$. Let $B = \sum_{t=1}^{T-1} \beta^t$. Observe that by Lemma 3

$$V_0(f^{iid}) = u(f_0) + \beta I(\xi + BI(\xi)),$$

and

$$V_0(f^{corr}) = u(f_0) + \beta I\left(\xi + \beta I\left(\xi' + \frac{B - \beta}{\beta}I(\xi)\right)\right).$$

Denote with $\mathbb{E}_P(\cdot)$ the benchmark for I. Recall that I satisfies

$$\mathbb{E}_P\xi \geq \mathbb{E}_P\xi' \implies I(\xi) \geq I(\xi'), \tag{16}$$

whenever ξ is such that $u(f) = \xi$ and f is unambiguous.

Let us evaluate $\xi + \beta I\left(\xi' + \frac{B-\beta}{\beta}I(\xi)\right)$ according to $\mathbb{E}_P(\cdot)$,

$$\begin{aligned}\mathbb{E}_P\left[\xi + \beta I\left(\xi' + \frac{B - \beta}{\beta}I(\xi)\right)\right] &\leq \mathbb{E}_P\big(\xi\big) + \beta I\left(\mathbb{E}_P\xi' + \frac{B - \beta}{\beta}I(\xi)\right) \\ &\leq \mathbb{E}_P\big(\xi\big) + \beta I\left(\mathbb{E}_P\xi + \frac{B - \beta}{\beta}I(\xi)\right) \\ &= \mathbb{E}_P(\xi)(1 + B),\end{aligned}$$

where the first inequality follows from concavity of I, while the second equality from (16) and Lemma 2.

By (16) it follows that

$$I(\xi + BI(\xi)) \geq I\left(\xi + \beta I\left(\xi' + \frac{B-\beta}{\beta} I(\xi)\right)\right),$$

and hence

$$V_0(f^{iid}) \geq V_0(f^{corr}),$$

as desired.

In particular, the inequality will be strict whenever for some $s, s' \in S$ $\xi'(s,\cdot)) + \frac{B-\beta}{\beta} I(\xi)$ and $\xi'(s',\cdot)) + \frac{B-\beta}{\beta} I(\xi)$ are not unambiguous and I satisfies the following strict concavity property: for every $(\xi)_{i=1}^n$ such that for some $i \neq j$, ξ_i, ξ_j are not unambiguous and $(\alpha)_{i=1}^n$ positive weights summing to 1 it holds

$$\sum_{i=1}^n \alpha_i I(\xi_i) < I(\sum_{i=1}^n \alpha_i \xi_i).$$

□

7.4.4 Generalizing Theorem 1 and 2

Denote with Δ the set of probability simple probability measures on $\mathbb{R}$. A functional $J : \Delta \to \mathbb{R}$ is

1. monotone if it is monotone w.r.t. first order stochastic dominance;
2. risk averse if whenever μ is a mean preserving spread of λ then $I(\lambda) \geq I(\mu)$;
3. concave if for every $\mu, \lambda \in \Delta$ and $\alpha \in [0,1]$, $J(\alpha\mu + (1-\alpha)\lambda) \geq \alpha J(\mu) + (1-\alpha)J(\lambda)$.

Theorem 6. *Suppose that $\succeq_{t,\omega}$ admits a recursive representation given by*

$$V_t(\omega, h) = u(h_t(\omega)) + \beta J(P \circ V_{t+1}^{-1}(\cdot, h)),$$

for some probability P, where $J : \Delta \to \mathbb{R}$ is monotone, risk averse and concave. Then $\succeq_{t,\omega}$ exhibits aversion to correlation.

Proof. Let $\xi = u(f)$ and ξ' by $\xi'(s,s') = u(f_t^{corr}(s,s'))$ for every $s, s' \in S$. We have

$$\mathbb{E}_P \xi + \beta \mathbb{E}_P J(P \circ V_{t+1}^{-1}(\cdot, f^{corr})) \leq \mathbb{E}_P \xi + \beta J\left(\mathbb{E}_P\left[P \circ V_{t+1}^{-1}(\cdot, f^{corr})\right]\right).$$

Observe that

$$\mathbb{E}_P\left[P \circ V_{t+1}^{-1}(\cdot, f^{corr})\right] = \sum_{s' \in S} P(s') P \circ \xi^{-1'}(\cdot, s') + k,$$

where k is the discounted continuation utility and $\sum_{s'\in S} P(s')P \circ \xi^{-1'}(\cdot, s')$ is the mixture in Δ of the collection of probabilities $(P \circ \xi^{-1}(\cdot, s'))_{s\in S}$ with weights given by $(P(s'))_{s'\in S}$. By Lemma 2, $\mu = \sum_{s'\in S} P(s')P\circ\xi^{-1'}(\cdot, s')$ is a mean preserving spread of $P \circ \xi^{-1}(\cdot)$ so that by risk aversion of J,

$$\mathbb{E}_P\xi + \beta\mathbb{E}_P J(P \circ V_{t+1}^{-1}(\cdot, f^{corr})) \leq \mathbb{E}_P\xi + \beta\mathbb{E}_P J(P \circ V_{t+1}^{-1}(\cdot, f^{iid})).$$

By applying risk aversion once again it follows that

$$J(\xi + \beta J(P \circ V_{t+1}^{-1}(\cdot, f^{corr}))) \leq J(u(h_t(\omega)) + \beta J(P \circ V_{t+1}^{-1}(\cdot, f^{iid}))),$$

from which we obtain $f^{iid} \succeq_0 f^{corr}$ for every $f : S \to X$ which concludes the proof. $\square$

For every probability measure on S and $\xi \in B_0(\Sigma, u(X))$, let $P \circ \xi^{-1}$ denote the element of Δ given by $P \circ \xi^{-1}(x) = P(\xi^{-1}(x))$ for every $x \in \mathbb{R}$.

Example 4. An example of such recursive preferences is given by the recursive utility model studied in Epstein & Zin (1990) that uses Yaari's 1987 rank-dependent model. In this case we have that $X = \mathbb{R}_{++}$ and for every μ with support given by $(x_i)_{i=1}^n$, where $x_i < x_{i+1}$, $i = 1, \ldots, n-1$

$$J(\mu) = \phi^{-1}\left(\sum_{i=1}^{n}\left[g\left(\sum_{j=1}^{i}\mu_j\right) - g\left(\sum_{j=1}^{i-1}\mu_j\right)\right]\phi(x_i)\right),$$

and ϕ is CARA.

Definition 11. *A certainty equivalent $I : B_0(\Sigma, u(X)) \to \mathbb{R}$ admits a global benchmark w.r.t. a monotone and risk averse $J : \Delta \to \mathbb{R}$ if the set*

$$E_I = \left\{P \in \Delta : J\left(P \circ \xi^{-1}\right) \geqslant I(\xi) \text{ for all } \xi \in B_0(\Sigma, u(X))\right\},$$

is non-empty. Call $f : S \to X$ unambiguous if for some $P \in E_I$ it holds $I(u(f)) = J(P \circ \xi^{-1})$. Let $U_I \subseteq X^S$ denote the class of unambiguous acts.

Example 5. Hansen & Sargent (2020) and Cerreia-Vioglio et al. (2020) consider the following criterion

$$V_t(s^t, h) = u(h_t(s^t)) + \beta \min_{P\in\Delta(S)}\left\{\int V_{t+1}(\cdot, h)\mathrm{d}P + \alpha \min_{q\in Q} R(P\|q)\right\},$$

where $\alpha > 0$ and Q is a convex compact subset of $\Delta(S)$. It is easy to show that these preferences admit the equivalent representation

$$V_t(s^t, h) = u(h_t(s^t)) + \beta \min_{q\in Q}\phi^{-1}\int\phi\left(V_{t+1}(\cdot, h)\right)\mathrm{d}q,$$

where $\phi(x) = -\exp(-\frac{x}{\alpha})$. The benchmark in this case is given by $J(\mu) = \phi^{-1}\mathbb{E}_\mu\phi(x)$ and P is any element of Q. Then U_I will be non-empty whenever the set $AG = \{E \subseteq S : P(E) = P'(E) \text{ for all } P, P' \in Q\}$ is non-empty.

Theorem 7. *Assume that $\succeq_{t,\omega}$ has an RDAA representation where I has a global benchmark w.r.t J. Then it exhibits correlation aversion for the class U_I.*

Proof of Theorem 7. The proof works in the same way as the proof of Theorem 2. Using the fact that J is risk-averse we have:

$$\mathbb{E}_P\xi \geq \mathbb{E}_P\xi' \implies J(P\circ\xi) \geq J(P\circ\xi') \implies I(\xi) \geq I(\xi'), \tag{17}$$

whenever ξ is such that $u(f) = \xi$ and f is unambiguous. □

7.4.5 An upper bound on the persistence premium

Andersen et al. (2018) estimate an intertemporal utility function under uncertainty which can be written as

$$V(f) = u^{-1}\phi^{-1}\mathbb{E}_P\left[\varphi\left(\sum_{t=1}^{n}\beta^t u(f_t)\right)\right],$$

where $\beta \approx 0.998$, $\phi(x) = x^{0.68}$ and $u(x) = x^{0.65}$.

Given $x > 0$ and $n = 2$, let f^{iid} be the process that pays x and 0 with probability $\frac{1}{2}$ each and f^{corr} the process that pay $\{x, x\}$ and $\{0, 0\}$ with probability $\frac{1}{2}$ each. Therefore, f^{corr} is maximally correlated in the sense that the probability that consumption at $t = 1$ matches consumption at $t = 2$ is 1. In this case the persistence premium is given by

$$\pi = 1 - \frac{\left(0.5\left(x^{1-0.35} + \frac{x^{1-0.35}}{1+0.114}\right)^{1-0.32}\right)^{1/(1-0.32)(1-0.35)}}{\left(\left(x^{1-0.35}\right)^{1-0.32}\times 0.5 + \left(\frac{x^{1-0.35}}{1+0.114}\right)^{1-0.32}(1-0.5)\right)^{1/(1-0.32)(1-0.35)}} \approx 1 - 0.8 \approx 0.2.$$

Hence $\pi \approx 20\%$ provides an upper bound for the persistence premium.

7.4.6 Proof of Proposition 3

Definition 12 (Definition 2 in Strzalecki (2013)). *$\succeq_{t,\omega}$ exhibits a preference for earlier resolution of uncertainty if and only if for all $h, h' \in \mathcal{F}$ and $t \leq T-2$ such that there exist $x_0, \ldots, x_{t+1}, x_{t+3}, \ldots, x_T$ and $f : S \to X$ such that $h_j = h'_j = x_j$ for all $j \neq t+1$ and for $j = t+2$, $h_j(s_1, \ldots, s_j) = f(s_{t+1})$ and $h'_j(s_1, \ldots, s_j) = f(s_{t+2})$ it holds that*

$$h \succeq_{t',\omega} h'$$

for all $t' \leq t$ and $\omega \in \Omega$.

Lemma 4. *Assume that $\succeq_{t,\omega}$ has an RDAA representation on $\mathcal{F}$. Then it exhibits a preference for early resolution if and only if it exhibits a preference for early resolution according to definition 12.*

Proof. Take $x_0, \dots, x_{t+1} \in X$ and $f_{t+2}, \dots, f_T \in X^S$ and define h, h' appropriately. We have

$$\begin{aligned} &V_t\left(\omega, \left(x_0, \dots, x_{t+1}, h_{t+2}, h_{t+3}, \dots, h_T\right)\right) \\ &\quad = u\left(x_t\right) + \beta I\left(u\left(x_{t+1}\right) + \beta\left(\sum_{j=0}^{T-t-1} \beta^j u\left(f_{t+j+1}\left(s_{t+1}\right)\right)\right)\right). \end{aligned}$$

In contrast,

$$\begin{aligned} &V_t\left(\omega, \left(x_0, \dots, x_{t+1}, h'_{t+2}, h'_{t+3}, \dots, h'_T\right)\right) \\ &\quad = u\left(x_t\right) + \beta I\left(u\left(x_{t+1}\right) + \beta I\left(\sum_{j=0}^{T-t-1} \beta^j u\left(f_{t+j+1}\left(s_{t+2}\right)\right)\right)\right). \end{aligned}$$

If we let $u(x_{t+1}) = k$ and $u(f_i) = \xi_i$ for $i = t+1, \dots, T$, we obtain that a preference for early resolution is equivalent to

$$I\left(\beta \sum_{i=0}^{T-t-1} \beta^i \xi_{t+1+i} + k\right) \geq \beta \sum_{i=0}^{T-t-1} \beta^i I(\xi_{t+1+i}) + k.$$

Letting $\sum_{i=0}^{T-t-1} \beta^i \xi_i = \eta$, since by assumption $u(X)$ is unbounded for every $\zeta \in B_0(\Sigma, u(X))$, there exists $(\xi_i)_i$ such that $\sum_{i=0}^{T-t-1} \beta^i \xi_i = \zeta$ we obtain

$$I\left(\beta\eta + k\right) \geq \beta I(\eta) + k,$$

for all $\eta \in B_0(\Sigma, u(X))$ and $k \in u(X)$. Now by Lemma 1 in Strzalecki (2013), $\succeq_{t,\omega}$ displays a preference toward earlier resolution of uncertainty if and only if

$$I(\beta\xi + k) \geq \beta I(\xi) + k,$$

for all $\xi \in B_0(\Sigma, u(X))$ and all $k \in u(X)$. Therefore, the result follows. □

Lemma 5. *Let $(A_i, \mathcal{A}_i)_{i=1}^n$ be a collection of measurable spaces and let $(A, \mathcal{A})$ be the product of such measurable spaces. For every $\xi, \xi' \in B_0(\mathcal{A}, \mathbb{R}_+)$ such that ξ is measurable w.r.t. $\mathcal{A}_i$ and ξ' measurable w.r.t. $\mathcal{A}_j$ with $i \neq j$ it holds that*

$$\int\int \phi(\beta\xi' + \xi) \mathrm{d}P(a_j) \mathrm{d}P(a_i) \leq \int \phi\left(\beta\phi^{-1}\left(\int \phi(\xi') \mathrm{d}P(a_j)\right) + \xi\right) \mathrm{d}P(a_i).$$

where $\phi(x) = x^{\frac{\alpha}{\rho}}$ with $\rho < \alpha$, $\rho, \alpha < 1$. Moreover, the inequality is strict whenever ξ' is not constant.

Proof. By an application of Theorem 4 in Strzalecki (2013), we have

$$\int \phi(\beta\xi' + \xi(a_i)) \mathrm{d}P(a_j) \leq \phi\left(\beta\phi^{-1}\left(\int \phi(\xi') \mathrm{d}P(a_j)\right) + \xi(a_i)\right),$$

for every $a_i \in A$. It follows that

$$\int \int \phi(\beta\xi' + \xi)\mathrm{d}P(a_j)\mathrm{d}P(a_i) \leq \int \phi\left(\beta\phi^{-1}\left(\int \phi(\xi')\mathrm{d}P(a_j)\right) + \xi\right)\mathrm{d}P(a_i),$$

as desired. Since Theorem 4 in Strzalecki (2013) uses Jensen's inequality and ϕ is strictly convex, the inequality is strict whenever ξ' is not constant. □

Proof of Proposition 3. $(i) \implies (ii)$ By Theorem 4 (Remark 1) in Strzalecki (2013) and Lemma 2, it follows that $\alpha \leq \rho$. Therefore, the result follows by Theorem 1.

$(ii) \implies (i)$ Applying Theorem 1, if $\rho < \alpha$, then preferences are not averse to correlation. Therefore, if preferences are averse to correlation it has to be that $\alpha \leq \rho$.

$\implies (iii) \implies (i)$. Suppose $\rho < \alpha$. I claim that $V_0(\hat{f}) < V_0(\bar{f})$ for every non-constant f.

Let $\xi_t(a_1^t) = u(f(a_1^t))$ whenever $a_1^t = a_t^1$ (observe that $\xi_t(a_1^t) = u(f(a_t^1)))$ and let

$$\eta_T = \xi_{T-1} + \beta\phi^{-1}\int \phi(\xi_T)\mathrm{d}P(a_1^T),$$

and for $t < T$

$$\eta_t = \xi_{t-1} + \beta\phi^{-1}\int \phi(\eta_{t+1})\mathrm{d}P(a_1^t).$$

Then we have

$$V_0(\bar{f}) = u(x) + \beta\phi^{-1}\int \phi\left(\eta_1\right)\mathrm{d}P(a_1^1),$$

and

$$V_0(\hat{f}) = u(x) + \beta\phi^{-1}\int \phi\left(\sum_{t=0}^{T-1}\beta^t\xi_{t+1}\right).$$

By repeated applications of Lemma 5, we obtain

$$V_0(\hat{f}) = \int \ldots \int \phi\left(\sum_t \beta^t\xi_t + \xi_1\right)\mathrm{d}P(a_1^1)\ldots\mathrm{d}P(a_T^1) \leq \beta\phi^{-1}\int \phi(\eta_1)\mathrm{d}P(a_1^1).$$

Specifically, the inequality will be strict for any non constant f. Therefore, it follows that the timing premium is negative. Hence, if the timing premium is always non-negative it must be the case that $\alpha \leq \rho$.

□

7.4.7 Proof of Proposition 2

Proof of Proposition 2. We have

$$
\begin{aligned}
&V_0(f^{iid}) - V_0(f^{corr}) = \\
&= \beta^t \left(u \left(v^{-1} \left[\mathbb{E}_{\prod_{j=1}^t P(s_j)} v\left(f_t^{iid}(\cdot) \right) \right] \right] - u\left(v^{-1} \left[\mathbb{E}_{\prod_{t=1}^T P(s_t)} v\left(f_t^{corr}(\cdot) \right) \right] \right) \right].
\end{aligned}
$$

By Lemma 2 it follows that

$$
u \left(v^{-1} \left[\mathbb{E}_{\prod_{j=1}^t P(s_j)} v\left(f_t^{iid}(\cdot) \right) \right] \right) - u\left(v^{-1} \left[\mathbb{E}_{\prod_{t=1}^T P(s_t)} v\left(f_t^{corr}(\cdot) \right) \right] \right) = 0.
$$

whence $V_0(f^{iid}) = V_0(f^{corr})$ so that $f^{iid} \sim_0^{SS} f^{corr}$. □

7.4.8 Proof of Proposition 4

Proof of Proposition 4. The proof parallels that of Selden & Stux (1978) (proof of Lemma 1). I prove sufficiency of the axioms in the case of $\succeq_0$, and using consequentialism the result follows for $\succeq_{t,\omega}$ analogously. I claim that for every $h \in \mathcal{F} \setminus \mathcal{F}^r$, there exists $\bar{c} = (h_0, c_1, \ldots, c_T) \in \mathcal{D}$ such that $\bar{c} \sim_0 h$ and

$$
c_t = \left[\mathbb{E}_{\prod_{\tau=1}^t P(s_\tau)} h_t^\alpha \right]^{\frac{1}{\alpha}},
$$

which establishes the representation since $\succeq_0$ has an EZ representation so that

$$
V_0((h_0, c_1, \ldots, c_T))) = u(h_0) + \sum_{t=1}^T \beta^j u(c_j),
$$

as desired. First, for every $s^{t-1} = (s_1, \ldots, s_{t-1})$, let

$$
c_t^1(s^{t-1}) = \left[\mathbb{E}_{P(s_t)} h_t^\alpha \right]^{\frac{1}{\alpha}}.
$$

Observe that axioms 7 and 8, we have

$$
h \sim_0 (h_0, \ldots, h_{T-1}, c_T^1).
$$

Now by further applying axioms 7 and 8 we get

$$
(h_0, \ldots, h_{T-1}, c_T^1) \sim_0 (h_0, \ldots, c_{T-1}^1, c_T^1).
$$

By axiom 5,

$$
(h_0, \ldots, c_{T-1}^1, c_T^1) \sim_0 (h_0, \ldots, c_{T-1}^1, \hat{c}_T^1),
$$

where $\hat{c}_T^1(s_1, s_{T-2} \ldots, \cdot, s_T)$ is constant and

$$
\hat{c}_T^1(s_1, \ldots, \cdot) = c^1(s_1, \ldots, \cdot).
$$

By another application of axioms 7 and 8 we obtain:

$$(h_0, \dots, c^1_{T-1}, \hat{c}^1_T) \sim_0 (h_0, \dots, c^1_{T-1}, c^2_T),$$

where

$$c^2_T = \left[\mathbb{E}_{P(s_{T-1})P(s_T)} h^\alpha_T\right]^{\frac{1}{\alpha}}.$$

Proceeding as in the previous to steps, we obtain at step t

$$h \sim_0 (h_0, \dots, c^1_{T-t+1}, \dots, c^{t-1}_{T-1}, c^t_T),$$

where

$$c^t_j = \left[\mathbb{E}_{\prod_{\tau=T-t+1}^{j} P(s_\tau)} h^\alpha_j\right]^{\frac{1}{\alpha}}.$$

Specifically, after T steps we get

$$h \sim_0 (h_0, c_1, \dots, c_T),$$

as desired.

I turn to the necessity of the axioms. It is immediately verified that axioms 6-8 are satisfied. I prove that the representation satisfies indifference to timing.

Take h, h' such that for some $\bar{s}^t = (\bar{s}_1, \dots, \bar{s}_t)$ with $1 \leq t \leq T-2$ the act

$$(h_0, h_1(\bar{s}_1), h_2(\bar{s}_1, \bar{s}_2), \dots, h_t(\bar{s}^t), \dots, h_T(\bar{s}^t, \cdot)),$$

resolves earlier than

$$(h'_0, h'_1(\bar{s}_1), h'_2(\bar{s}_1, \bar{s}_2), \dots, h'_t(\bar{s}^t), \dots, h'_T(\bar{s}^t, \cdot)),$$

and $h_\tau(s_1, \dots, s_\tau) = h'_\tau(s_1, \dots s_\tau)$ for every $(s_1, \dots, s_\tau)$ with $\tau < t$ and $h_\tau(s_1, \dots, s_\tau) = h'_\tau(s_1, \dots s_\tau)$ for every $(s_1, \dots s_\tau)$ such that $\tau \geq t$ and $(s_1, \dots, s_t) \neq (\bar{s}_1, \dots, \bar{s}_t)$. Then we have for $t' \leq t$

$$V_{t'}((s_1, \dots, s_{t'}), h) - V_{t'}((s_1, \dots, s_{t'}), h') \propto$$
$$\sum_{j=0}^{T-t} \beta^j u\left(\left[\mathbb{E}_{\prod_{\tau=1}^{j} P(s_{t+\tau})} h^\alpha_{t+j}\right]^{\frac{1}{\alpha}}\right) - \sum_{j=0}^{T-t} \beta^j u\left(\left[\mathbb{E}_{\prod_{\tau=1}^{j} P(s_{t+\tau})} h'^\alpha_{t+j}\right]^{\frac{1}{\alpha}}\right).$$

Observe that by assumption on h, h' we have

$$\mathbb{E}_{\prod_{\tau=1}^{j} P(s_{t+\tau})} h^\alpha_{t+j} = \mathbb{E}_{\prod_{\tau=1}^{j} P(s_{t+\tau})} h'^\alpha_{t+j},$$

for $j = 0, \dots, T-t$.

Therefore $V_{t'}((s_1, \dots, s_{t'}), h) - V_{t'}((s_1, \dots, s_{t'}), h') = 0$ whence the result follows.

□

7.4.9 Proof of Proposition 5

Proof of Proposition 5. For simplicity, I focus on the case $\rho \neq 0$. The proof for the case $\rho = 0$ follows by analogous arguments. Let

$$B = \sum_{t=0}^{T-1} \beta^t = \frac{\beta^T - 1}{(\beta - 1)\beta}.$$

Since $\alpha \leq \rho$ we have

$$\begin{aligned}
V_0^{EZ}(\hat{f}) &= u(f_0) + \beta \left[\mathbb{E}_{P((a_1^1,\ldots,a_T^1))} \left(V_{t+1}((a_1^1,\ldots,a_T^1),\hat{f}) \right)^{\frac{\alpha}{\rho}} \right]^{\frac{\rho}{\alpha}} \\
&= u(f_0) + \beta \left[\mathbb{E}_{P((a_1^1,\ldots,a_T^1))} \left(\frac{B}{B} \sum_{t=1}^{T-1} \beta^{t-1} u(f_t(a_t^1)) \right)^{\frac{\alpha}{\rho}} \right]^{\frac{\rho}{\alpha}} \\
&= u(f_0) + \beta B \left[\mathbb{E}_{P((a_1^1,\ldots,a_T^1))} \left(\sum_{t=1}^{T-1} \frac{\beta^{t-1}}{B} u(f_t(a_t^1)) \right)^{\frac{\alpha}{\rho}} \right]^{\frac{\rho}{\alpha}} \\
&\geq u(f_0) + \beta B \left[\sum_{t=0}^{T-1} \frac{\beta^t}{B} \mathbb{E}_{P((a_1^1,\ldots,a_T^1))} \frac{1}{\rho^\alpha} \left(f_t^\rho(a_t^1) \right)^{\frac{\alpha}{\rho}} \right]^{\frac{\rho}{\alpha}} = \\
&= u(f_0) + \beta B \frac{1}{\rho} \left(\left[\mathbb{E}_P(a_t^1) f_t^\alpha(a_t^1) \right]^{\frac{1}{\alpha}} \right)^\rho = V_0^{SS}(\hat{f}).
\end{aligned}$$

Hence the result follows by the fact that $V_0^{EZ}(\bar{f}) = V_0^{SS}(\bar{f})$ so that

$$\pi^{EZ}(f) \geq \pi^{SS}(f) \iff 1 - \frac{V_0^{EZ}(\bar{f})}{V_0^{EZ}(\hat{f})} \geq 1 - \frac{V_0^{SS}(\bar{f})}{V_0^{SS}(\hat{f})} \iff V_0^{EZ}(\hat{f}) \geq V_0^{SS}(\hat{f}),$$

as wanted. □

7.4.10 Proof of Proposition 6

Proof of Proposition 6. Take $t \leq T-2$ and $h, h' \in \mathcal{F}$ such that there exist $f_{t+2}, \ldots, f_T \in \mathcal{F}$ and $x_0, \ldots, x_{t+1} \in X$ such that $h_j = h'_j = x_j$ for all $j \leq t+1, h_j(s_1, \ldots, s_j) = f_j(s_{t+1})$ for $j \geq t+2$, and $h'_{t+2} = f_j(s_{t+2})$. Observe that

$$\sum_{j=0} \beta^j \mathbb{E}_{P(s^{t+j})} \mathbb{VAR}_P(u(h_{t+j+1}(s^{t+j}, \cdot))] = 0,$$

and

$$\sum_{j=0} \beta^j \mathbb{E}_{P(s^{t+j})} \mathbb{VAR}_P(u(h'_{t+j+1}(s^{t+j}, \cdot))] \geq 0.$$

Since it holds that

$$u(h_t(s^t)) + \sum_{j=1}^{T} \beta^j [\mathbb{E}_{P(s^{t+j})} u(h_{t+j+1}(s^{t+j+1})) = u(h'_t(s^t)) + \sum_{j=1}^{T} \beta^j [\mathbb{E}_{P(s^{t+j})} u(h'_{t+j+1}(s^{t+j+1})),$$

we obtain

$$h \succeq_{t',\omega} h',$$

for all $t' \leq t$ as desired. □

8 Bibliography

References

Al-Najjar, N. I. & Shmaya, E. (2019), 'Recursive utility and parameter uncertainty', *Journal of Economic Theory* **181**, 274–288.

Andersen, S., Harrison, G. W., Lau, M. I. & Rutström, E. E. (2018), 'Multiattribute utility theory, intertemporal utility, and correlation aversion', *International Economic Review* **59**(2), 537–555.

Anderson, E. W. (2005), 'The dynamics of risk-sensitive allocations', *Journal of Economic theory* **125**(2), 93–150.

Arrow, K. J. (1971), 'Essays in the theory of risk-bearing (markham pub. co., chicago, il)'.

Attanasio, O. P. & Weber, G. (1989), 'Intertemporal substitution, risk aversion and the euler equation for consumption', *The Economic Journal* **99**(395), 59–73.

Aumann, R. J. (1962), 'Utility theory without the completeness axiom', *Econometrica: Journal of the Econometric Society* pp. 445–462.

Bansal, R., Kiku, D. & Yaron, A. (2016), 'Risks for the long run: Estimation with time aggregation', *Journal of Monetary Economics* **82**, 52–69.

Bansal, R. & Yaron, A. (2004), 'Risks for the long run: A potential resolution of asset pricing puzzles', *The journal of Finance* **59**(4), 1481–1509.

Bommier, A. (2007), 'Risk aversion, intertemporal elasticity of substitution and correlation aversion', *Economics Bulletin* **4**(29), 1–8.

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

Bommier, A., Kochov, A. & Le Grand, F. (2019), 'Ambiguity and endogenous discounting', *Journal of Mathematical Economics* **83**, 48–62.

Cerreia-Vioglio, S., Hansen, L. P., Maccheroni, F. & Marinacci, M. (2020), 'Making decisions under model misspecification', *University of Chicago, Becker Friedman Institute for Economics Working Paper* (2020-103).

Chew, S. H. & Epstein, L. G. (1991), Recursive utility under uncertainty, *in* 'Equilibrium theory in infinite dimensional spaces', Springer, pp. 352–369.

Chudziak, J., GŁAZOWSKA, D., Jarczyk, J. & Jarczyk, W. (2019), 'On weighted quasi-arithmetic means which are convex', *Mathematical Inequalities & Applications* **22**(4), 1123–1136.

Collard, F., Mukerji, S., Sheppard, K. & Tallon, J.-M. (2018), 'Ambiguity and the historical equity premium', *Quantitative Economics* **9**(2), 945–993.

De Finetti, B. & Savage, L. J. (1962), 'Sul modo di scegliere le probabilità iniziali', *Biblioteca del Metron* **1**, 81–147.

Debreu, G. (1954), 'Representation of a preference ordering by a numerical function', *Decision processes* **3**, 159–165.

Duffie, D. & Epstein, L. G. (1992), 'Stochastic differential utility', *Econometrica: Journal of the Econometric Society* pp. 353–394.

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

Epstein, L. G. (1990), *Behaviour under risk: recent developments in theory and applications*, Department of Economics and Institute for Policy Analysis, University of Toronto.

Epstein, L. G. (2001), 'Sharing ambiguity', *American Economic Review* **91**(2), 45–50.

Epstein, L. G., Farhi, E. & Strzalecki, T. (2014), 'How much would you pay to resolve long-run risk?', *American Economic Review* **104**(9), 2680–97.

Epstein, L. G. & Schneider, M. (2003*a*), 'Iid: independently and indistinguishably distributed', *Journal of Economic Theory* **113**(1), 32–50.

Epstein, L. G. & Schneider, M. (2003*b*), 'Recursive multiple-priors', *Journal of Economic Theory* **113**(1), 1–31.

Epstein, L. G. & Tanny, S. M. (1980), 'Increasing generalized correlation: a definition and some economic consequences', *Canadian Journal of Economics* pp. 16–34.

Epstein, L. G. & Zin, S. E. (1989), 'Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework', *Econometrica: Journal of the Econometric Society* pp. 937–969.

Epstein, L. G. & Zin, S. E. (1990), "first-order'risk aversion and the equity premium puzzle', *Journal of monetary Economics* **26**(3), 387–407.

Epstein, L. G. & Zin, S. E. (1991), 'Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis', *Journal of political Economy* **99**(2), 263–286.

Epstein, L. & Wang, T. (1994), 'Intertemporal asset pricing under knightian uncertainty', *Econometrica* **62**(2), 283–322.

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

Gorman, W. M. (1968), 'The structure of utility functions', *The Review of Economic Studies* **35**(4), 367–390.

Hall, R. E. (1985), 'Real interest and consumption', *NBER working paper* (w1694).

Hansen, L. P. & Sargent, T. J. (2020), 'Structured ambiguity and model misspecification', *Journal of Economic Theory* p. 105165.

Hansen, L. P., Sargent, T. J. & Tallarini, T. D. (1999), 'Robust permanent income and pricing', *Review of Economic studies* pp. 873–907.

Hansen, L. & Sargent, T. J. (2001), 'Robust control and model uncertainty', *American Economic Review* **91**(2), 60–66.

Hardy, G. H., Littlewood, J. E., Pólya, G., Pólya, G., Littlewood, D. et al. (1952), *Inequalities*, Cambridge university press.

Hayashi, T. (2005), 'Intertemporal substitution, risk aversion and ambiguity aversion', *Economic Theory* **25**(4), 933–956.

Johnsen, T. H. & Donaldson, J. B. (1985), 'The structure of intertemporal preferences under uncertainty and time consistent plans', *Econometrica: Journal of the Econometric Society* pp. 1451–1458.

Ju, N. & Miao, J. (2012), 'Ambiguity, learning, and asset returns', *Econometrica* **80**(2), 559–591.

Karantounias, A. G. (2018), 'Optimal fiscal policy with recursive preferences', *The Review of Economic Studies* **85**(4), 2283–2317.

Kihlstrom, R. E. & Mirman, L. J. (1974), 'Risk aversion with many commodities', *Journal of Economic Theory* **8**(3), 361–388.

Klibanoff, P., Marinacci, M. & Mukerji, S. (2005), 'A smooth model of decision making under ambiguity', *Econometrica* **73**(6), 1849–1892.

Klibanoff, P., Marinacci, M. & Mukerji, S. (2009), 'Recursive smooth ambiguity preferences', *Journal of Economic Theory* **144**(3), 930–976.

Kochov, A. (2015), 'Time and no lotteries: An axiomatization of maxmin expected utility', *Econometrica* **83**(1), 239–262.

Kochov, A. & Song, Y. (2021), 'Intertemporal hedging and trade in repeated games with recursive utility', *Available at SSRN 2714337* .

Koopmans, T. C., Diamond, P. A. & Williamson, R. E. (1964), 'Stationary utility and time perspective', *Econometrica: Journal of the Econometric Society* pp. 82–100.

Kreps, D. M. & Porteus, E. L. (1978), 'Temporal resolution of uncertainty and dynamic choice theory', *Econometrica: journal of the Econometric Society* pp. 185–200.

Kubler, F., Selden, L. & Wei, X. (2019), 'Time consistency, temporal resolution indi§erence and the separation of time and risk', *Columbia Business School Research Paper* .

Lucas, R. E. (1978), 'Asset prices in an exchange economy', *Econometrica: journal of the Econometric Society* pp. 1429–1445.

Marinacci, M. & Montrucchio, L. (2010), 'Unique solutions for stochastic recursive utilities', *Journal of Economic Theory* **145**(5), 1776–1804.

Miao, B. & Zhong, S. (2015), 'Risk preferences are not time preferences: Separating risk and time preference: Comment', *American Economic Review* **105**(7), 2272–86.

Richard, S. F. (1975), 'Multivariate risk aversion, utility independence and separable utility functions', *Management Science* **22**(1), 12–21.

Savage, L. (1954), 'Foundations of statistics.'.

Schmeidler, D. (1989), 'Subjective probability and expected utility without additivity', *Econometrica: Journal of the Econometric Society* pp. 571–587.

Selden, L. (1978), 'A new representation of preferences over "certain x uncertain" consumption pairs: The "ordinal certainty equivalent" hypothesis', *Econometrica: Journal of the Econometric Society* pp. 1045–1060.

Selden, L. & Stux, I. E. (1978), *Consumption Trees, OCE Utility and the Consumption/Savings Decision*, Columbia University, Graduate School of Business.

Skiadas, C. (1998), 'Recursive utility and preferences for information', *Economic Theory* **12**(2), 293–312.

Stokey, N. L. & Lucas, R. E. (1989), *Recursive Methods in Economic Dynamics*, Harvard University Press.

Strzalecki, T. (2011), 'Axiomatic foundations of multiplier preferences', *Econometrica* **79**(1), 47–73.

Strzalecki, T. (2013), 'Temporal resolution of uncertainty and recursive models of ambiguity aversion', *Econometrica* **81**(3), 1039–1074.

Tallarini, T. D. (2000), 'Risk-sensitive real business cycles', *Journal of monetary Economics* **45**(3), 507–532.

Uzawa, H. (1968), Time preference, the consumption function, and optimum asset holdings, *in* 'Value, Capita & Growth', Routledge, pp. 485–504.

Wakker, P. P. (1989), *Additive representations of preferences: A new foundation of decision analysis*, Vol. 4, Springer Science & Business Media.

Wang, T. (2003), 'Conditional preferences and updating', *Journal of Economic Theory* **108**(2), 286–321.

Weil, P. (1989), 'The equity premium puzzle and the risk-free rate puzzle', *Journal of monetary economics* **24**(3), 401–421.

Yaari, M. E. (1987), 'The dual theory of choice under risk', *Econometrica: Journal of the Econometric Society* pp. 95–115.

Zin, S. E. et al. (1987), Intertemporal substitution, risk and the time series behaviour of consumption and asset returns, Technical report.