ISSN 2279-9362

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

Lorenzo Stanca

No. 693
July 2023

Carlo Alberto Notebooks
www.carloalberto.org/research/working-papers

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

Lorenzo Stanca†

March 30, 2023

Abstract

Models of recursive utility are of central importance in many economic applications. This paper investigates a new behavioral feature exhibited by these models: aversion to risks that exhibit persistence (positive autocorrelation) through time, referred to as correlation aversion. I introduce a formal notion of such a property and provide a characterization based on risk attitudes, and show that correlation averse preferences admit a specific variational representation. I discuss how these findings imply that attitudes toward correlation are a crucial behavioral aspect driving the applications of recursive utility in fields such as asset pricing, climate policy, and optimal fiscal policy.

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

JEL classification: C61, D81.

---

*This paper is a revised and extended version of Chapter 1 of my dissertation at Northwestern University; several results also appeared in my job market paper which shared the same title. I thank Nabil Al-Najjar, Peter Klibanoff, and Marciano Siniscalchi for their help and guidance. I also thank Larry Epstein for providing extremely valuable comments that significantly improved the paper.

†Collegio Carlo Alberto and University of Turin, Department of ESOMAS, Corso Unione Sovietica, 218 Bis, 10134, Turin, Italy. E-mail: lorenzomaria.stanca@unito.it

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 and Zin (1989), Epstein and Zin (1991)), precautionary savings (Weil (1989), Hansen et al. (1999)), business cycles (Tallarini (2000)), risk-sharing (Epstein (2001), Anderson (2005)) and more recently have been applied to climate change (Bansal et al. (2017), Cai and Lontzek (2019)), optimal fiscal policy (Karantounias (2018)), and repeated games Kochov and Song (2021), 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

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

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 each another. In contrast, recursive preferences allow for a more general recursive formulation

$$V_t = u(c_t) + \beta \phi^{-1} \left(\mathbb{E}_p \phi \left(V_{t+1} \right) \right), \tag{1}$$

where the curvature of u reflects intertemporal substitution and ϕ reflects attitudes toward risk, hence obtaining the desired separation between the two.[1]

The present paper introduces a new axiom called *correlation aversion*, as it requires aversion to persistent consumption shocks. In the dynamic setting I consider, this axiom imposes restrictions on an individual's willingness to pay for *non-instrumental information* about future consumption. I provide bounds—based on risk attitudes—on attitudes toward non-instrumental information that are necessary and sufficient for recursive preferences to satisfy correlation aversion. Recall that since the work of Kreps and Porteus (1978), recursive utility has been understood to entail a preference for non-instrumental information, also referred to as a preference for early resolution of uncertainty. To illustrate, consider a gamble in which consumption is fixed at 0 for every $t = 1, \dots, T-1$ 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

[1]See Cochrane (2009), Chapter 21.3 or Campbell (2017), Chapter 6.4 for a textbook treatment.

$t = 2$ 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 preference for useless information. 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)). For example, it would be concerning if the implications of asset pricing models or estimates of the social cost of carbon depended on the demand for irrelevant information. In contrast, correlation aversion can be illustrated with a different thought example. Consider two gambles: A and B.[2] 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 t, giving a level of consumption equal to 1 if heads and 0 otherwise. 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. Correlation aversion requires the hedging motive to dominate the preference for non-instrumental information so that B is preferred to A.

I consider a risk setting, where preferences are defined over temporal lotteries. A novel and general notion of increasing autocorrelation between consumption at two different time periods is introduced. As discussed when comparing gambles A and B, the total effect of increasing correlation is determined by the relative strength of the hedging motive described earlier and preferences for non-instrumental information. The first main result, Theorem 2, shows that the hedging motive is stronger if and only if ϕ satisfies increasing relative risk aversion (IRRA). I illustrate how increasing relative risk aversion limits preferences for non-instrumental information. Notably, IRRA is one of the most common classes of functions in applications (e.g., see Arrow (1971), p. 96), including the Epstein-Zin and Hansen-Sargent as special cases. This result implies that under reasonable conditions the value of non-instrumental information is dominated by the value of intertemporal hedging. When one further restricts

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

preferences for non-instrumental information, the second main result (Theorem 3) shows that recursive preferences admit a variational of representation

$$V_t = u(c_t) + \beta \left[\min_q \mathbb{E}_q V_{t+1} + I^t_{\phi,u,\beta}(q||p) \right], \tag{2}$$

where $I^t_{\phi,u,\beta}$ is a cost function. Such a representation provides a connection between correlation aversion and robustness to model misspecification. An interpretation of equation (2) is that the decision maker does not fully trust the distribution of future consumption given by the reference probability p. Instead, other possible distributions q are considered plausible depending on their dissimilarity from p as measured by $I^t_{\phi,u,\beta}$. It is known that when ϕ has an exponential form, the cost function $I^t_{\phi,u,\beta}$ is given by relative entropy; the result shows that this fear of model misspecification is true for a much larger class of models that satisfy correlation aversion, including Epstein-Zin preferences. In particular, in the Epstein-Zin case the cost function $I^t_{\phi,u,\beta}$ is defined in terms of the Rényi divergence, a common type of measure of divergence between probability measures that has application in several fields.

The restrictions on preferences for non-instrumental information I obtain are strong enough so that on certain domains of consumption there is no preference for non-instrumental information. In these domains, consumption programs can be ranked in terms of persistence, as is often the case in applications. Proposition 4 demonstrates that such domains can separate risk aversion from intertemporal substitution without implying preference for useless information. However, I show that recursive preferences in (1) cannot disentangle risk aversion from correlation aversion, emphasizing the need for more general recursive utility models.

Together, these results have important implications for applications of recursive utility. The literature on consumption-based asset pricing has considered consumption processes that involve persistence, such as the long-run risk model of Bansal and Yaron (2004) in which 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. 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 Epstein-Zin preferences satisfy IRRA, making correlation aversion a more dominant feature of preferences.[3] This point has important implications for Epstein et al. (2014)'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 supports Epstein et al.'s assertion that greater quantitative rigor is necessary for accurately modeling investors' preferences. Potential solutions are briefly discussed.

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 fear of model misspecification. Correlation aversion has a straightforward connection with aversion to model misspecification as exemplified by the robust representation in (2). To better understand this point, observe that an equivalent way of thinking about gamble A 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 misspecification. 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. In summary, correlation aversion enables us to connect various crucial properties—aversion to long-run risk, model uncertainty, and volatility in continuation utility—to observable consumption choice behavior.

[3] This point has to be contrasted with the common understanding of the long-run risk model, e.g. Bansal et al. (2016) who state "The long-run risks (LRR) asset pricing model emphasizes the role of low-frequency movements [...] along with investor preferences for early resolution of uncertainty, as an important economic-channel that determines asset prices".

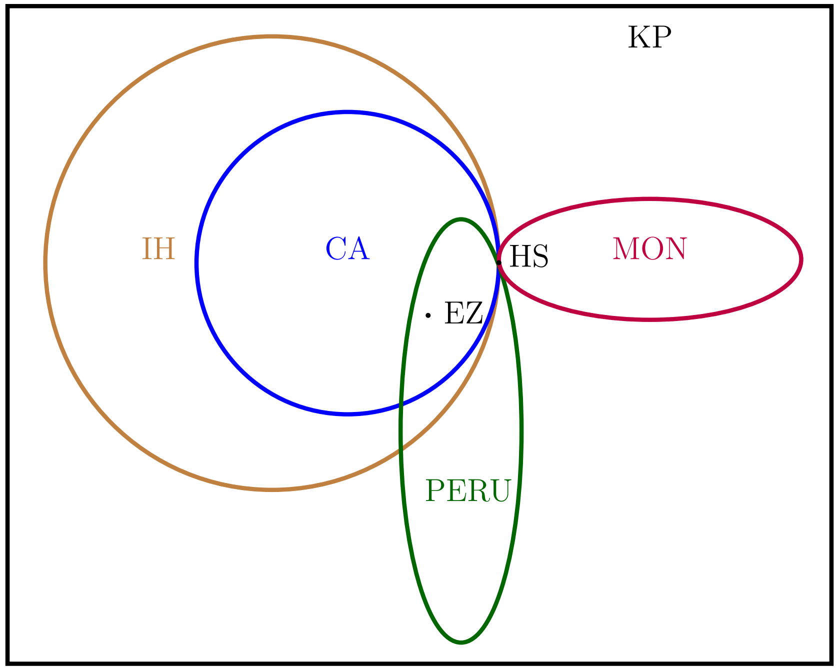

Figure 1: Relationship between correlation averse (CA) preferences and other recursive risk preferences: recursive preferences that satisfy intertemporal-hedging (IH), Epstein-Zin (EZ) preferences, multiplier-preferences (HS), monotone recursive preferences (MON), and preferences that exhibit a preference for early resolution of uncertainty (PERU).

1.1 Related literature

The literature on dynamic choice has considered a notion of correlation aversion derived from the literature on risk aversion with multiple commodities started by Kihlstrom and Mirman (1974) (see also Richard (1975) or Epstein and Tanny (1980)). In particular, Bommier (2007) considers a notion of correlation aversion based on the Kihlstrom and Mirman approach in a continuous time setting. The extension to a purely subjective setting with discrete time is studied in Kochov (2015) and Bommier et al. (2019), which they refer to as intertemporal hedging. Intertemporal hedging involves comparing intertemporal gambles that do not differ in terms of temporal resolution of uncertainty (see Section 3 for a discussion). Miao and Zhong (2015) and Andersen et al. (2018) relate Epstein-Zin utility to an analogous notion of intertemporal hedging and provide experimental evidence in its favor. Within the class of recursive preferences in (1)—which I refer to as Kreps-Porteus preferences—intertemporal hedging is equivalent to ϕ being concave, i.e. risk aversion. The notion of correlation aversion studied in the present paper involves a trade-off between non-instrumental information and intertemporal hedging. Therefore additional restrictions on risk aversion are required for correlation aversion to hold. Such restrictions are satisfied by Epstein and Zin's (1989) preferences and Hansen and Sargent's (2001) multiplier preferences. In particular an important consequence of the present paper is that within the Kreps-Porteus setting multiplier preferences are the ones to jointly satisfy correlation aversion and monotonicity as defined in Bommier et al. (2017). Figure 1 illustrates the relationship just discussed between correlation aversion and other prominent classes of recursive preferences. I discuss the relationship of correlation aversion with the work of DeJarnette et al. (2020) and Dillenberger et al. (2020) on preferences that satisfy stochastic impatience will be discussed more in depth in Section 5.4.

1.2 Organization of the paper

Section 2 introduces the notation and the main choice-theoretic objects used in the paper, and provides a novel treatment of preference for early resolution of uncertainty based on the Blackwell's order. This novel treatment is used in Section 3 for the main results related to correlation aversion. Section 4 examines the relationships between correlation aversion, risk aversion, and intertemporal substitution. The major implications of these results for the applied literature are examined in Section 5, and Section 6 offers a final summary of the paper.

2 Preliminaries

2.1 Choice setting

I assume that time is discrete and varies over a finite horizon $2 \leq T < \infty$. The Supplemental Appendix discusses the case of an infinite horizon $T = \infty$. The consumption set C is assumed to be an unbounded interval of non-negative real numbers

$C = [0,\infty)$.[4] Given a Polish space X, let $\Delta_s(X), \Delta_b(X)$ denote the space of simple and Borel probability measures with bounded support over X, respectively. Observe that $\Delta_s(X) \subseteq \Delta_b(X)$ and that $\Delta_b(X)$ is a mixture space. Given $\ell, m \in \Delta_b(X)$ such that ℓ is absolutely continuous with respect to m, indicated by $\ell \ll m$, $\frac{d\ell}{dm}$ denotes the Radon-Nikodym derivative. Endow $\Delta_b(X)$ with the weak* topology. Given $x \in X$, I denote with $\delta_x \in \Delta_b(X)$ the Dirac probability defined by $\delta_x(A) = 1$ when $x \in A$ and $\delta_x(A) = 0$ otherwise. I denote with $\bigoplus_{i=1}^n \pi_i m_i$ the mixture of n probabilities $(m_i)_{i=1}^n$ in $\Delta_b(X)$ with a probability vector $(\pi_i)_{1 \le i \le n}$. Further, note that every two-stage lottery $m \in \Delta_s(\Delta_s(X))$ can be (uniquely up to permutations) associated to a stochastic matrix $M[m]$ whose rows describe each probability $M[m](\cdot|i) \in \text{supp}m$ in the support of m for $i = 1, \ldots, |\text{supp}m|$.

Temporal lotteries $(D_t)_{t=0}^T$ are defined by $D_T := C$ and recursively,

$$D_t := C \times \Delta_b(D_{t+1}),$$

for every $t = 0, \ldots, T-1$. Likewise, simple temporal lotteries are defined by $D_{T,s} := C$, $D_{t,s} := C \times \Delta_s(D_{t+1,s})$ for every $t = 0, \ldots, T-1$. Simple temporal lotteries can be intuitively represented using a tree diagram, as illustrated in Figure 2. I write $(c_0, (c_1, m)) \in D_0$ for a temporal lottery that consists of two periods of deterministic consumption, c_0 and c_1, followed by the lottery $m \in \Delta_b(D_2)$. More generally, for any consumption vector $c^t = (c_0, \ldots, c_{t-1}) \in C^t$ and $m \in \Delta_b(D_t)$, the temporal lottery $(c_0, (c_1, (c_2, (\ldots, (c_{t-1}, m))))) \in D_0$ or (c^t, m) for brevity is one that consists of t periods of deterministic consumption followed by the lottery m. Given two Polish spaces X, Y and $m \in \Delta_b(X \times Y)$ I denote with $\text{marg}_X\, m$ the marginal probability over X, i.e., $\text{marg}_X\, m(A) = m(A \times Y)$ for every measurable $A \subseteq X$. A function $I : X \times X \to [0, \infty]$ is a generalized distance in the sense of Csiszár (1995) if it satisfies $I(m||\ell) = 0$ if and only if $m = \ell$ for every $m, \ell \in X$.

The preferences of a decision maker over temporal lotteries are given by a collection $(\succeq_t)_{t=0}^T$ where each $\succeq_t$ is a weak order over D_t and $\succ_t$ denotes the asymmetric part of $\succeq_t$. To ease notation, I denote with $\succeq := (\succeq_t)_{t=0}^T$ the entire collection of preferences. I consider preferences that admit the following general recursive representation described in (1).

[4] Alternatively, one could consider an infinite horizon with a compact consumption set, say $C = [0,1]$. However, an unbounded consumption set is more germane to economic applications, where an unbounded consumption set is needed.

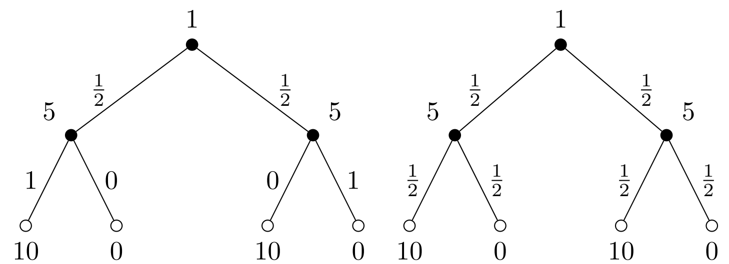

Figure 2: Probability tree representation of two temporal lotteries with $T = 2$

Definition 1. *Preferences $\succeq$ admits a Kreps-Porteus (KP) recursive representation (ϕ, u, β) if and only if, for $t = 0, \ldots, T$, each $\succeq_t$ is represented by $V_t : D_t \to \mathbb{R}$,*

$$V_t(c, m) = u(c) + \beta \phi^{-1} \left(\mathbb{E}_m \phi \left(V_{t+1}\right)\right) \quad \text{for } t = 0, \ldots, T-1,$$

and $V_T(c) = u(c)$ for every $c \in C$, where $\beta \in (0, 1]$ is the discount factor, $u : C \to \mathbb{R}$ is unbounded above, continuous, and strictly increasing, and $\phi : u(C) \to \mathbb{R}$ is a continuous and strictly increasing function.

This representation of preferences effectively separates risk aversion (as captured by the function ϕ) from intertemporal substitution (as modeled by the utility function u). The axiomatic foundation of this representation is well-known (see for example Proposition 4 in Sarver (2018)). The parameter β is unique, while u is cardinally unique and ϕ is cardinally unique given u. This class of preferences comprises many common cases used in applications. Two notable examples are Epstein-Zin preferences (EZ), given by $u(x) = \frac{x^\rho}{\rho}$ for every $x \in u(C)$ and $\phi(x) = \frac{\rho}{\alpha} x^{\frac{\alpha}{\rho}}$ for every $x \in u(C)$, where $0 \neq \alpha < 1, 0 \neq \rho < 1$ and $\alpha < \rho$; Hansen-Sargent multiplier preferences (HS) are given by $\phi(x) = -\exp\left(-\frac{x}{\theta}\right)$ with $0 < \theta < \infty$ for every $x \in u(C)$.[5] The results in the present paper will consider KP representations with ϕ that is concave and satisfies certain differentiability assumptions. Assumptions on the smoothness of ϕ are needed to employ standard tools from the theory of risk aversion. In particular, I will make ample use of the Arrow-Pratt index $A_\phi : \operatorname{int} u(C) \to \mathbb{R}$ defined by

$$A_\phi(x) = -\frac{\phi''(x)}{\phi'(x)} \quad \text{for every } x \in \operatorname{int} u(C),$$

[5] Under the present taxonomy, EZ preferences do not overlap with HS preferences, but they would if one allowed for $\rho = 0$, see for example Hansen et al. (2007), Example 2.3.

and the index of relative risk aversion defined by $R_\phi(x) = xA_\phi(x)$ for every $x \in \operatorname{int} u(C)$. A function ϕ is decreasing absolute risk averse (DARA) if A_ϕ is nonincreasing, it is increasing absolute risk averse (IARA) if its index A_ϕ is nondecreasing, and it is constant absolute risk averse (CARA) if it is both DARA and IARA. Decreasing (DRRA), increasing (IRRA), and constant (CRRA) relative risk averse functions are defined analogously by replacing the index A_ϕ with R_ϕ.

2.2 Preferences for (non-instrumental) information

I reframe the theory of preferences for early resolution of uncertainty using the language of information economics. Temporal lotteries are partially ordered by means of a version of Blackwell order, which allows comparing them in terms of their (non-instrumental) informativeness. In addition to its theoretical appeal and generality, this approach will permit building a formal link between correlation and information.

Definition 2. *Given $d, d' \in D_{0,s}$ say that d is more informative than d', denoted $d \geq_B d'$, if and only if for some $t \leq T-2$ and $c^t \in C^t$, $d = (c^t, m)$ and $d' = (c^t, m')$ with $m, m' \in \Delta_s(C \times \Delta_s(D_{t+1,s}))$ satisfying $\operatorname{marg}_C m' = \operatorname{marg}_C m$, and*

$$M\left[\operatorname{marg}_{\Delta_s(D_{t+1,s})} m'\right] = GM\left[\operatorname{marg}_{\Delta_s(D_{t+1,s})} m\right], \tag{3}$$

where G is a stochastic matrix, with each row forming a probability vector.

In words, the expression $d \geq_B d'$ means that the two lotteries, d and d', have the same distribution of consumption in the future period $t+1$. However, the actual realization of consumption in period $t+1$ provides more information about future values of consumption (from period $t+2$ onwards) for the lottery d compared to the lottery d'. Observe that $\geq_B$ is a partial order just like the standard Blackwell order. The following examples help to further clarify this notion of comparative information.

Example 1. Assume $T = 2$. Let $d = \left(1, \frac{1}{2}(5, 10) \oplus \frac{1}{2}(5, 0)\right)$ and $d' = (1, 5, (\frac{1}{2}10 \oplus \frac{1}{2}0))$. Figure 2 provides a graphical representation of these two temporal lotteries. We have

$$M\left[\operatorname{marg}_{\Delta_s(D_{2,s})} m'\right] = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} M\left[\operatorname{marg}_{\Delta_s(D_{2,s})} m\right],$$

so that $d \geq_B d'$. In words, the terminal value of consumption is fully revealed by a coin toss at $t = 1$ for d but only revealed at $t = 2$ for d'.

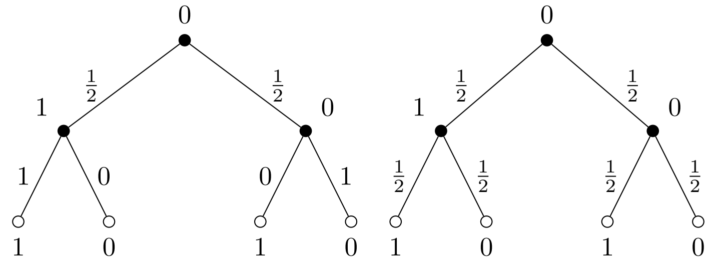

Figure 3: Probability tree representation of a temporal lottery

Example 2. Again assume $T = 2$. Consider d, d' given by $d = \left(1, \frac{1}{2}(1,1) \oplus \frac{1}{2}(0,0)\right)$ and $d' = \left(1, \frac{1}{2}\left(1, \left(\frac{1}{2}1 \oplus \frac{1}{2}0\right)\right) \oplus \frac{1}{2}\left(0, \left(\frac{1}{2}1 \oplus \frac{1}{2}0\right)\right)\right)$. Figure 3 provides a graphical representation of these two temporal lotteries. We have

$$M\left[\mathrm{marg}_{\Delta_s(D_{2,s})} m'\right] = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} M\left[\mathrm{marg}_{\Delta_s(D_{2,s})} m\right],$$

so that $d \geq_B d'$. In words, d' is an "iid" temporal lottery while d is perfectly correlated.

This notion of comparative information is extended to arbitrary temporal lotteries by means of the following standard procedure.

Definition 3. *For every $d, d' \in D_0$, write $d \geq_B d'$ if and only if there exist sequences $(d_n)_{n=0}^{\infty}, (d'_n)_{n=0}^{\infty}$ in $D_{0,s}$ such that $\lim_n d_n = d$, $\lim_n d'_n = d'$ and $d_n \geq_B d'_n$ for every $n \geq 0$.*

A preference for non-instrumental information or, equivalently, for early resolution of uncertainty over a set of temporal lotteries is defined as monotonicity with respect to the order $\geq_B$.

Axiom 1. $\succeq$ exhibits a preference for information over a set $B \subseteq D_0$ if and only if for every $d, d', \in B$

$$d \geq_B d' \implies d \succeq_0 d'.$$

The Kreps and Porteus's approach restricts attention to temporal lotteries such as those in Example 1 in which the draw at $t = 1$ provides information about consumption at $t = 2$ but at the same time consumption at $t = 1$ is deterministic. Examples with correlation are excluded and it will become clearer why. Formally, in this case the set B is given by[6]

$$B := \left\{ \left(c^t, m \right) : c^t \in C^t, m \in \Delta_s(D_{t+1,s}), \operatorname{marg}_C m = \delta(\bar{c}), \bar{c} \in C \right\}. \tag{4}$$

Observe that within the consumption domain B, consumption at time $t+1$ is consistently at level $\bar{c}$, which implies that it is uncorrelated with consumption in the following periods. Preferences valuing non-instrumental information over B exhibit the following characterization.

Theorem 1. *Assume $\succeq$ admits a KP representation (ϕ, u, β) with ϕ twice continuously differentiable. Then $\succeq$ exhibits a preference for information over B if and only if*

$$-\beta \frac{\phi''(\beta x + y)}{\phi'(\beta x + y)} \leq -\frac{\phi''(x)}{\phi'(x)}, \tag{5}$$

for every $x, y \in \operatorname{int} u(C)$.

Outline of the proof. The key idea behind the proof is to demonstrate, using established results from information economics (e.g., as outlined in Theorem 4 in Kihlstrom (1984)), that $\succeq$ has a preference for information if and only if all functions $U_t : \Delta_s(D_{t+1,s}) \to \mathbb{R}$ defined by

$$U_t(m) = \phi\left(u(\bar{c}) + \beta \phi^{-1} \left(\mathbb{E}_m \phi(V_{t+1}) \right) \right) \quad \text{for every } m \in \Delta_s(D_{t+1,s}), \tag{6}$$

are convex for every $\bar{c} \in C$ and $t = 1, \dots, T-2$. Straightforward calculations show that convexity of each U_t is equivalent to (5). $\square$

Condition (5) appears also in Strzalecki (2013) (p. 1051). Here I focus on risk attitudes that exhibit a preference for information regardless of the level of impatience or intertemporal substitution. Say ϕ satisfies uniform preference for information (UPI) if and only if every $\succeq$ with KP representation (ϕ, u, β) exhibits a preference for information. Observe that if ϕ satisfies UPI, then it also satisfies DARA by (5). This relationship provides a connection between classical risk attitudes and preference for information.

[6] For a more recent treatment, see, for example, Definition 2 in Bommier et al. (2017).

3 Attitudes toward correlation

3.1 The case $T = 2$

I introduce a general notion of an increase in positive correlation between consumption at two distinct periods. For ease of exposition, consider first the case in which there are two risky periods, i.e. $T = 2$. Later, I will show how it can be used to introduce persistence over time to study long-run risk, i.e. persistence over multiple periods. To this purpose, I introduce a class of temporal lotteries that can be defined by (i) the distribution of consumption at time $t = 1$ and (ii) the conditional distribution of consumption at time $t = 2$, given consumption in the prior period. Let

$$M_s^* := \{m \in \Delta_s(C \times \Delta_s(C)) : (c, \mu), (c, \mu') \in \operatorname{supp} m \implies \mu = \mu'\}.$$

Every such $m \in M_s^*$ can be (uniquely) associated with $m_1 \in \Delta_s(C)$ and $m_2(\cdot|\cdot) \in \Delta_s(C)^{\operatorname{supp} m_1}$, defined by $m_1 = \operatorname{marg}_C m$, and

$$m_2(\cdot|c) = \mu(\cdot),$$

where μ is the unique element of $\Delta_s(C)$ such that $(c, \mu) \in \operatorname{supp} m$. Conversely, given $m_1 \in \Delta_s(C)$ and $m_2(\cdot|\cdot) \in \Delta_s(C)^{\operatorname{supp} m_1}$, we can uniquely define $m \in M_s^*$ by

$$m(c, m_2(\cdot|c)) := m_1(c) \quad \text{for every } c \in \operatorname{supp} m_1.$$

In words, m_1 describes the distribution of time 1 consumption while $m_2(\cdot|c)$ is the conditional distribution of consumption at the final time period given a realization of $t = 1$ consumption. The set $D_{0,s}^* := \{(c, m) \in D_{0,s} : m \in M_s^*\}$ is the set of temporal lotteries that can be described in terms of a pair (m_1, m_2). Likewise, one can define the associated cumulative distributions $m_1(c_1 \leq c)$, $m_2(c_2 \leq c | c_1 \leq c')$.

Definition 4. *Consider $d = (c_0, m), d' = (c_0, m') \in D_{0,s}^*$. Say that d differs from d' by an **intertemporal elementary correlation increasing transformation** (IECIT) if and only if $m_1 = m_1'$ and there exist $\varepsilon \geq 0$, and a pair (c, c') such that $c \neq c', m_1(c), m_1(c') \neq 0$ and*

$$m_2(c|c) = m_2'(c|c) + \frac{\varepsilon}{m_1'(c)},$$

$$m_2(c'|c) = m_2'(c'|c) - \frac{\varepsilon}{m_1'(c)},$$

$$m_2(c'|c') = m_2'(c'|c') + \frac{\varepsilon}{m_1'(c')},$$

and

$$m_2(c|c') = m_2'(c|c') - \frac{\varepsilon}{m_1'(c')}.$$

and $m_2 = m_2'$ *otherwise.*

The following two examples serve as an illustration of this concept.

Example 3 (Example 2 continued)**.** In this case we have $m_1 = m_1'$, $m_2(1|1) = 1 = m_2'(1|1) + \frac{1}{1/2}\frac{1}{4} = \frac{1}{2} + \frac{1}{2}$, $m_2(1|1) = 0 = m_2'(1|1) - \frac{1}{1/2}\frac{1}{4} = \frac{1}{2} - \frac{1}{2}$, $m_2(1|0) = 0 = m_2'(1|0) - \frac{1}{1/2}\frac{1}{4} = \frac{1}{2} - \frac{1}{2}$ and $m_2(0|0) = 1 = m_2'(0|0) + \frac{1}{1/2}\frac{1}{4} = \frac{1}{2} + \frac{1}{2}$. It follows that d differs from d' by an IECIT with $\varepsilon = \frac{1}{4}$. Therefore, the perfectly correlated temporal lottery can be obtained from the "iid" lottery by means of an IECIT. In this case, an IECIT increases the informativeness of a temporal lottery.

Example 4. Consider the temporal lotteries $d = (c_0, m), d' = (c_0, m') \in D_{0,s}$ where $m_1(x)' = m_1(x) = \frac{1}{2}$ and $m_2(1|0) = m_2(0|1) = 1$, $m_2'(1|1) = m_2'(0|0) = 1$. Figure 4 provides a graphical representation of these two lotteries. The lottery d is obtained by applying an IECIT with $\varepsilon = \frac{1}{2}$. The lotteries d and d' have perfect positive and negative correlation, respectively. We can immediately see that $d \geq_B d'$ and $d' \geq_B d$, meaning that d and d' are equally informative. The strict preference for d' over d, is referred to as correlation aversion by Bommier (2007) and *intertemporal hedging* by Kochov (2015). I adopt the latter terminology as it reflects the fact that only hedging considerations affect the evaluations of these two lotteries. The Supplemental Appendix demonstrates that intertemporal hedging is equivalent to the concavity of ϕ.

The concept of an IECIT correlation is an application of Epstein and Tanny's (1980) idea of generalized increasing correlation, applied in a dynamic setting. With the notion of an IECIT, it is possible to establish an ordering $\geq_C$ that can be used to rank temporal lotteries based on their positive autocorrelation.

Definition 5. *Given* $d, d' \in D_{0,s}^*$ *say that* d *is more informative than* d'*, denoted* $d \geq_C d'$*, if and only if* d *differs from* d' *by a finite amount of IECITs.*

I provide a necessary condition of when two temporal lotteries differ by a finite amount of IECITs.

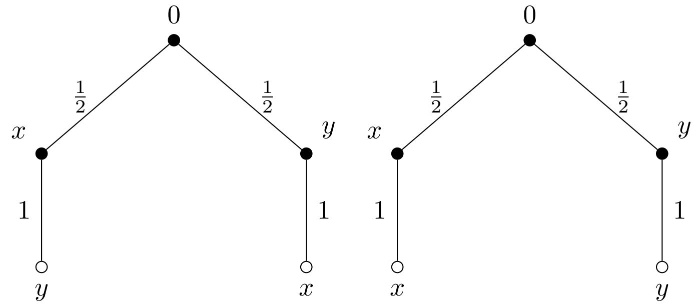

Figure 4: Negative vs positive correlation

Proposition 1. *If $d \geq_C d'$ then it holds that*

$$m_2'\left(c_2 \leq c \mid c_1 \leq c'\right) \leq m_2\left(c_2 \leq c \mid c_1 \leq c'\right) \text{ for every } (c, c') \in C \times C.$$

Notably, Proposition 1 implies that $\geq_C$ is transitive and thus a partial order. Finally, denote with D_0^* the weak* closure of $D_{0,s}^*$. It is possible to extend $\geq_C$ to D_0^* as follows.

Definition 6. *Given $d = (c, m), d' = (c, m') \in D_0^*$, write $d \geq_C d'$, if and only if $\operatorname{marg}_C m = \operatorname{marg}_C m'$ and there exist sequences $(d_n)_{n=0}^\infty, (d_n')_{n=0}^\infty$ in $D_{0,s}^*$ such that $\lim_n d_n = d$, $\lim_n d_n' = d'$ and $d_n \geq_C d_n'$ for every $n \geq 0$.*

The following result establishes a formal connection between IECITs and non-instrumental information by showing that increasing the correlation of "iid" temporal lotteries makes them more informative. To this end, define the "iid" temporal lottery for each $\ell \in \Delta_b(C)$ by $d^{iid}(\ell) = (c, m)$ where $m(A \times B) = \ell(A)$ if $\ell \in B$ and $m(A \times B) = 0$ otherwise.

Proposition 2. *Consider $\ell \in \Delta_b(C)$ and $d, d' \in D_0^*$. Then it holds that*

$$d \geq_C \geq d' \geq_C d^{iid}(\ell) \implies d \geq_B d' \geq_B d^{iid}(\ell).$$

Proof. See the Appendix. □

This proposition establishes formally the main trade-off described in the introduction: increasing persistence in consumption risks to an "iid" lottery provides more information about future consumption. Having fully established a theory of increasing correlation, we can define correlation aversion as a form of monotonicity with respect to the order $\geq_C$.

Axiom 2. $\succeq$ exhibits correlation aversion if and only if for every $d, d' \in D_0^*$ and $\ell \in \Delta_b(C)$

$$d \geq_C d' \geq_C d^{iid}(\ell) \implies d^{iid}(\ell) \succeq_0 d' \succeq_0 d.$$

The next result characterizes correlation averse preferences in terms of risk attitudes, under the assumption of UPI, i.e. when the trade-off between intertemporal hedging and non-instrumental information is relevant.

Theorem 2. *Consider ϕ that is twice continuously differentiable and satisfies UPI. Then every $\succeq$ with KP representation (ϕ, u, β) exhibits correlation aversion if and only if ϕ satisfies IRRA.*

Proof. See the Appendix. □

As shown by Proposition 2, increasing positive correlation entails more non-instrumental information. This result implies that under the restriction that such non-instrumental information is valuable, the effect of correlation aversion will be stronger than that of preference for early resolution of uncertainty. Conversely, if IRRA fails then correlation aversion will not be satisfied. 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 HS preferences. Moreover, empirical findings support DARA and IRRA (Wakker (2010), p. 83). To gain intuition, observe that IRRA means that R_ϕ is non-decreasing. Therefore, we have:

$$R'_\phi(x) \geq 0 \implies A'_\phi(x) \geq -\frac{A_\phi(x)}{x},$$

for every $x \neq 0$. Since under UPI it holds that $A'_\phi \leq 0$, IRRA results in a constraint on the reduction of absolute risk aversion for a given increase in consumption. Since a uniform preference for information implies DARA, this inequality can be seen as a constraint on the preference for information. Finally, observe that as a consequence of Theorem 2, indifference to correlation occurs only under risk neutrality, i.e. $\phi(x) = x$.

The following condition requires a mild strengthening of the notion of increasing relative risk aversion by adding the requirement that R_ϕ be convex, in addition to being non-decreasing. We refer to this condition as strong correlation aversion, given the connection between increasing relative risk aversion and correlation, as established by Theorem 2.

Definition 7. *We say that ϕ satisfies strong correlation aversion (SCA) if and only if R_ϕ is non-decreasing and convex.*

The next key result formalizes the connection between robustness to model misspecification and correlation aversion.

Theorem 3. *Assume that $\succeq$ admits a KP representation (ϕ, u, β) with ϕ four times continuously differentiable and satisfying UPI. If ϕ satisfies SCA, then $\succeq$ admits the recursive representation $(V_t)_{t=0}^2$ given by $V_2(c) = u(c)$ and*

$$V_t(c, m) = u(c) + \beta \min_{\ell \in \Delta_b(D_{t+1})} \left\{ \mathbb{E}_\ell V_{t+1} + I^t_{(u,\beta,\phi)}(\ell \| m) \right\} \quad \text{for } t = 0, 1,$$

where $I^t_{(u,\beta,\phi)}(\cdot, \cdot) : \Delta_b(D_{t+1}) \times \Delta_b(D_{t+1}) \to [0, \infty]$ is a generalized distance.

Outline of the proof. Using a general result due to Hardy et al. (1952) on certainty equivalents, I show that SCA implies that the certainty equivalent $\phi^{-1}\left(\mathbb{E}_m \phi\left(V_{t+1}\right)\right)$ is concave in utilities.[7] This result allows us to utilize the Fenchel-Moreau duality theorem, revealing that the certainty equivalent can be represented dually as $\phi^{-1}\left(\mathbb{E}_m \phi\left(V_{t+1}\right)\right) = \min_\ell \mathbb{E}\ell V t + 1 + I^t_{\phi,u,\beta}(\ell \| m)$, where each $I^t_{\phi,u,\beta}(\ell \| m)$ is a cost function whose general formulation is discussed in the Appendix. The cost functions $(I^t_{\phi,u,\beta})_{t=0}^1$ have therefore the same uniqueness properties as a KP representation (ϕ, u, β). The details are in the Appendix. □

It is well-known that when $\phi(x) = -e^{-\frac{x}{\theta}}$, the cost function $I^t_{\phi,u,\beta}$ is given by Relative Entropy:

$$I^t_{\phi,u,\beta}(\ell \| m) = I^t_\theta(\ell \| m) = \theta \left(\mathbb{E}_m \left[\frac{d\ell}{dm} \log \left(\frac{d\ell}{dm} \right) \right] \right),$$

when $\ell \ll m$ and $I^t_{\phi,u,\beta}(\ell \| m) = \infty$ otherwise.[8] The interpretation is that the decision-maker is concerned about misspecification of the distribution of future consumption. Therefore, alternative distributions are considered based on their distance from m, as measured by the cost function $I^t_{\phi,u,\beta}$. Theorem 3 implies that this interpretation in

[7] Cerreia-Vioglio et al. (2011) provide a similar representation under the assumption that ϕ is strictly increasing and concave. However, their result significantly differs from this one because they assume that $u(C) = (-\infty, \infty)$. This assumption is typically not satisfied in applications, such as the standard Epstein-Zin case.

[8] See for example Strzalecki (2011).

terms of model misspecification applies to all preferences satisfying strong correlation aversion. Similar to the variational preferences in Maccheroni et al. (2006), these cost functions can be interpretated as measure of aversion to model misspecification, or equivalently as an index of correlation aversion. A lower value of each $I^t_{\phi,u,\beta}$ indicates a higher degree of correlation aversion exhibited by the decision-maker, meaning that considering alternative distributions of future consumption becomes less costly. To illustrate, consider the common parametrization of Epstein-Zin used in asset pricing with intertemporal rate of substitution greater than unity $\frac{1}{1-\rho} > 1$ and $\alpha < 0$ (see Bansal and Yaron (2004)). As shown in the proof of the theorem, by setting $q = \frac{\alpha}{\alpha-\rho} > 0$ in this case we have the cost function

$$I^t_{\phi,u,\beta}(\ell \| m) = [\mathbb{E}_\ell V_{t+1}] \left[e^{\frac{1-q}{q} R_q(\ell\|m)} - 1 \right] \text{ if } \ell \ll m,$$

where $R_q(\ell\|m) = \frac{1}{q-1} \log \left(\mathbb{E}_m \left[\left(\frac{d\ell}{dm} \right)^q \right] \right)$ is the Rényi divergence. The Rényi divergence has applications in a variety of fields, including information theory, statistics, and machine learning (see Sason (2022) for a review). As the level of risk aversion $1 - \alpha$ increases the cost function correspondingly decreases. While the cost function for Hansen-Sargent preferences depends solely on ϕ through θ, the cost function for Epstein-Zin preferences also depends on the continuation utility, thus allowing for more complex patterns of correlation aversion.

3.2 The case $T < \infty$

The previous results are easily extended to an arbitrary horizon $T < \infty$. The case of an infinite horizon (that is, $T = \infty$) is discussed in the Supplemental Appendix (see Section 9.1). In this setting, it is possible to extend the previous analysis as follows. One can define the *present equivalent* $PE_{\succeq_t}(d)$ of each lottery $d \in D_t$ as the unique single period consumption level $c \in C$ such that $d \sim_t (c, 0, \dots, 0)$. Now observe that every $m \in \Delta_b(C \times \Delta_b(D_{t+1}))$ and $\succeq$ with KP representation (ϕ, u, β) induce the probability $m_\succeq$ over $\Delta_b(C \times \Delta_b(C))$ defined as follows:

$$m_\succeq(A \times B) = m(A \times B_\succeq) \quad \text{for every closed } A \times B \subseteq C \times \Delta_b(C),$$

where $B_\succeq = \{\ell \in \Delta_b(D_{t+1}) : \ell_\succeq \in B\}$ and $\ell_\succeq \in \Delta_b(C)$ is defined by $\ell_\succeq(A) = \ell(\{d \in D_{t+1} : PE_{\succeq_t}(d) \in A\})$.[9] In words, $m_\succeq$ describes the joint distribution between

[9]The present equivalent and consequently the lottery $m_\succeq$ are both well defined since preferences are continuous and u is unbounded above.

consumption at time $t+1$ and the continuation temporal lottery, where each temporal lottery is expressed in terms of one-period consumption. In this way, it is possible to extend the order $\geq_C$ and the correlation aversion axiom as follows.

Definition 8. *Consider $d = (c, m), d' = (c, m') \in D_0$ such that $(c, m_{\succeq}), (c, m'_{\succeq}) \in D_0^*$. Write $d \geq_C d'$ if and only if $(c, m_{\succeq}) \geq_C (c, m_{\succeq})$.*

Given $\ell \in \Delta_b(C)$, the "i.i.d." lottery is given by $d^{iid}(\ell) := (c, m)$ where $m_{\succeq}$ is such that $m(A \times B) = \ell(A)$ whenever $\ell \in B$ and $m(A \times B) = 0$ otherwise.

Axiom 3. $\succeq$ exhibits correlation aversion if and only for every $\ell \in \Delta_b(C)$ and $d = (c, m), d' = (c, m') \in D_0$ such that $(c, m_{\succeq}), (c, m'_{\succeq}) \in D_0^*$

$$d \geq_C d' \geq_C d^{iid}(\ell) \implies d^{iid}(\ell) \succeq_0 d' \succeq_0 d.$$

We can generalize Theorems 2 and 3 to this setting with an arbitrary finite horizon.

Theorem 4. *Consider ϕ that is twice continuously differentiable and satisfies UPI. Then every $\succeq$ with KP representation (ϕ, u, β) exhibit correlation aversion if and only if ϕ satisfies IRRA. Further, if $\succeq$ that admits a KP representation (ϕ, u, β) with ϕ that additionally satisfies SCA and is four times continuously differentiable then $\succeq$ admits the recursive representation $(V_t)_{t=0}^T$ given by $V_T(c) = u(c)$ and*

$$V_t(c, m) = u(c) + \beta \min_{\ell \in \Delta_b(D_{t+1})} \left\{ \mathbb{E}_\ell V_{t+1} + I^t_{(u,\beta,\phi)}(\ell || m) \right\} \quad \text{for } t = 0, \ldots, T-1$$

where $I^t_{(u,\beta,\phi)}(\cdot, \cdot) : \Delta_b(D_{t+1}) \times \Delta_b(D_{t+1}) \to [0, \infty]$ is a generalized distance.

Proof. The proof follows the same steps as the proof of Theorems 2 and 3 and is therefore omitted for brevity. $\square$

The theory presented thus far has focused on studying attitudes towards the correlation between consumption at two separate periods. However, it is also possible to consider more complex patterns of correlation, such as correlation between multiple periods. To explore this, I introduce a class of "Markov" temporal lotteries, in which the persistence of consumption between periods is determined by a single parameter, ε, which ranges from 0 to 1. This parameter is similar to the long-run risk concept introduced by Bansal and Yaron (2004). When $\varepsilon = 0$, the lottery outcomes are independent across periods, while for $\varepsilon = 1$ one has perfect positive correlation. Given

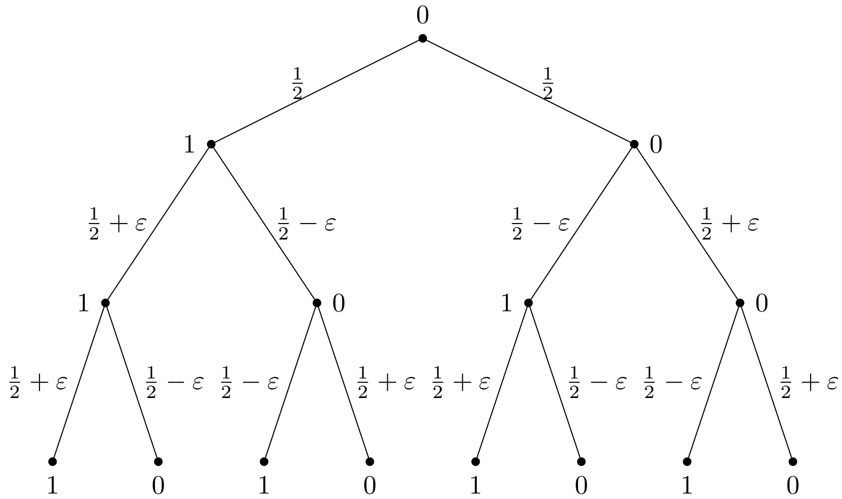

Figure 5: Example of $d_\varepsilon(\ell)$ with $\ell(1) = \ell(0) = \frac{1}{2}$ and $\varepsilon \in [0, 1]$

$\ell \in \Delta_s(C)$ and $\varepsilon \in [0, 1]$, define $d_\varepsilon(\ell)$ recursively as follows: $d_{T-1,c,\varepsilon}(\ell) = (c, m_{T-1,\varepsilon}(\ell))$ where

$$m_{T-1,\varepsilon}(\ell)(x) = \begin{cases} \ell(c) + (1 - \ell(c))\varepsilon & \text{if } x = c \\ \ell(x) - \ell(x)\varepsilon & x \neq c \end{cases}$$

and recursively for $2 \leq t \leq T - 2$ define $d_{t-1,c,\varepsilon}(\ell) = \left(c, m^\varepsilon_{t-1}(\ell)\right)$ by

$$m_{t-1,\varepsilon}(\ell)(d_{t,x,\varepsilon}(\ell)) = \begin{cases} \ell(c) + (1 - \ell(c))\varepsilon & \text{if } x = c \\ \ell(x) - \ell(x)\varepsilon & x \neq c \end{cases}$$

and finally set $d_\varepsilon(\ell) = (c_0, m_{1,\varepsilon})(\ell)$ where $m_{1,\varepsilon}(\ell) := \ell(c)$. Figure 5 provides a graphical example of a temporal lottery of this type. The following result demonstrates that, under the assumption of correlation aversion, a higher value of ε corresponds to a lower level of utility.

Proposition 3. *Consider $\succeq$ satisfying correlation aversion. Then for every $\ell \in \Delta_s(C)$ and $\varepsilon, \varepsilon' \in [0, 1]$*

$$\varepsilon \geq \varepsilon' \implies d^{\varepsilon'}(\ell) \succeq_0 d^\varepsilon(\ell).$$

Outline of the proof. The proof is a straightforward consequence of Theorem 4. See the Appendix. $\square$

Increasing the value of ε involves again a trade-off between non-instrumental information aversion and intertemporal hedging.[10] Under the interpretation that $\varepsilon \in [0,1]$ models the persistent component in consumption, the above result implies that correlation aversion is a sufficient condition for disliking long-run risks.

4 Substitution, risk aversion, and correlation aversion

A key motivation for the study of recursive preferences is that the two distinct aspects of preference—intertemporal substitutability and risk aversion—are not intertwined (see Epstein and Zin (1989) pp. 949-950 and Chew and Epstein (1991), Theorem 3.2). However, a potential drawback of this fact is that it leads to preferences for non-instrumental information in certain domains such as the one defined in (4). In this section I show that *under the assumption* of correlation aversion, one can distinguish between risk aversion and intertemporal substitution by means of a consumption domain in which preferences do not exhibit a preference for non-instrumental information. To this end, let

$$\mathcal{R} = \{d_\varepsilon(\ell) : \ell \in \Delta_s(C), \varepsilon \in [0,1]\} \cup C^T.$$

The consumption domain $\mathcal{R}$ contains all possible "Markov" temporal lotteries introduced in the previous section along with deterministic consumption streams. Consider preferences $\succeq_i$, $i = 1, 2$. Comparative risk aversion can be defined in a similar fashion as in Chew and Epstein (1991), but in a smaller domain and not the entire class of temporal lotteries D_0.

Definition 9. *Say that $\succeq^1$ is more risk averse than $\succeq^2$ if and only if for every $(c_0, m) \in \mathcal{R}$*

$$(c_0, m) \succeq_0^2 (c_0, c, \ldots, c) \implies (c_0, m) \succeq_0^1 (c_0, c, \ldots, c),$$

and

$$(c_0, m) \succ_0^2 (c_0, c, \ldots, c) \implies (c_0, m) \succ_0^1 (c_0, c, \ldots, c).$$

[10] Again, it holds that $d^\varepsilon(\ell) \geq_B d^{\varepsilon'}(\ell)$ just like in Proposition 2. The proof is omitted for brevity.

The next result shows the domain $\mathcal{R}$ is enough to distinguish risk aversion from intertemporal substitution.

Proposition 4. *Consider $\succeq_1$, $\succeq_2$ that both admit a KP representation. Then $\succeq_1$ is more risk averse than $\succeq_2$ if and only if they admit KP representations (ϕ_1, u_1, β_1) and (ϕ_2, u_2, β_2) such that $u_1 = u_2$, $\beta_1 = \beta_2$ and $A_{\phi_1} \geq A_{\phi_2}$.*

Outline of the proof. Since $\mathcal{R}$ contains all deterministic consumption streams it follows that $\succeq_1$ and $\succeq_2$ admit KP representations (ϕ_1, u_1, β_1) and (ϕ_2, u_2, β_2) such that $u_1 = u_2$, $\beta_1 = \beta_2$. Then one can apply the standard results on comparative risk aversion by means of lotteries of the type $\bigoplus_{i=1}^{n} \pi_i(c_i, \ldots, c_i)$ for every probability vector $(\pi_i)_{i=1}^{n}$ and constant consumption streams $(c_i, \ldots, c_i) \in C^T$ to obtain that $A_{\phi_1} \geq A_{\phi_2}$. $\square$

Under correlation aversion, a decision maker prefers less informative temporal lotteries to more informative ones. Therefore, preferences under this assumption do not exhibit a preference for information over the set $\mathcal{R}$. In fact, the preference for less informative lotteries, as represented by $d_{\varepsilon'}(\ell) \succeq_0 d_{\varepsilon}(\ell)$ when $\varepsilon \geq \varepsilon'$, aligns with preferences that are correlation averse but indifferent to non-instrumental information. By Theorem 1, preferences that are indifferent to information and not indifferent to correlation are characterized by $\beta = 1$ and $\phi(x) = -\exp\left(-\frac{x}{\theta}\right)$. As a result, when considering an empirically relevant restriction on risk attitudes, the consumption domain $\mathcal{R}$ is able to distinguish between risk aversion and intertemporal substitution, but attitudes towards non-instrumental information do not play a role.

However, even when considering the entire set of temporal lotteries D_0, KP preferences cannot differentiate between risk aversion and correlation aversion. To illustrate this point, as established in the proof of Theorem 3, if $\succeq_1$ is more risk averse than $\succeq_2$, then it follows that

$$I^t_{(\phi_1,u,\beta)}(\cdot||m) \leq I^t_{(\phi_2,u,\beta)}(\cdot||m),$$

for every $t = 0, \ldots, T-1$. Because $I^t_{\phi,u,\beta}$ is a measure of correlation aversion, it follows that for KP preferences, risk aversion and correlation aversion cannot be disentangled. For example, for EZ and HS preferences, correlation aversion and risk aversion are both determined by the parameters α and θ, respectively. However, this point does not hold for other classes of recursive preferences, such as Epstein-Uzawa. These preferences admit the representation

$$V_t(c, m) = u(c) + b(c)\mathbb{E}_m V_{t+1},$$

and $V_T(c) = u(c)$ for some continuous functions $u : X \to \mathbb{R}$ and $b : X \to (0, 1)$. These preferences are indifferent to non-instrumental information and value intertemporal hedging when b is non-increasing (Chew and Epstein (1991), Bommier et al. (2019)); but on the other hand cannot distinguish risk aversion from intertemporal substitution (Chew and Epstein (1991), p. 361). As I discuss next, not being able to distinguish between risk aversion and correlation aversion has important implications in asset pricing.

5 Implications

5.1 Long-run risk

Long-run risk models are a cornerstone in the consumption-based asset pricing literature for their ability to account for a wide range of asset pricing puzzles. A key model is that of Bansal and Yaron (2004), which relies on Epstein and Zin's preferences and a consumption process (case I) given by

$$\begin{aligned} \log\left(\frac{c_{t+1}}{c_t}\right) &= m + x_{t+1} + \sigma\epsilon_{c,t+1}, \\ x_{t+1} &= a x_t + \varphi\sigma\epsilon_{x,t+1}, \\ \epsilon_{c,t+1}, \epsilon_{x,t+1} &\sim \text{ iid } N(0, 1). \end{aligned} \tag{7}$$

In asset pricing, asset prices are determined by the claim on consumption growth process $d_t := \log\left(\frac{c_{t+1}}{c_t}\right)$.[11] Such a model faces the trade-off discussed in previous sections. An investor with recursive preferences values both early resolution of uncertainty and intertemporal hedging, with intertemporal hedging being more valuable

[11] While the present paper has focused on consumption levels, when u is isoelastic (such as is the case in applications), the same considerations apply to consumption growth. For example, when $u(x) = \log(x)$ we have the identity

$$(1 - \beta)\sum_{t=0}^{\infty} \beta^t \log(c_t) = \log(c_0) + \sum_{t=1}^{\infty} \beta^t \log\left(\frac{c_t}{c_{t-1}}\right).$$

due to Epstein-Zin preferences satisfying IRRA. My findings imply that the persistent component of consumption inflates the equity premium because of correlation aversion, thus making the ability of the model to match the risk premium independent of preferences for early resolution of uncertainty. Given the widespread use of long-run risk models in climate models to estimate the social cost of carbon (see Cai and Lontzek (2019)), this is a significant finding. Importantly, my research indicates that such estimates do not rely on preferences for irrelevant information, which is a crucial consideration for the credibility of climate models. However, some limitations of the model stem from the fact that the key feature of preferences, correlation aversion, is based solely on risk attitudes, which will be discussed further in the next section.

5.2 How much would you pay to remove long-run risk?

Epstein et al.'s (2014) have suggested that the long-run risk model entails implausibly high levels of preferences for early resolution of uncertainty. They introduce the concept of a "timing premium" to reflect, among other things, preferences for early resolution of uncertainty. However, when calculated using the standard parameters of the model as found in the literature, they note that the resulting timing premium seems excessively high compared to introspective assessments. I revisit their result that common parameter specifications lead to implausibly high timing premia in light of the theory on correlation aversion developed in the previous section. I ask a different question: "What fraction of your wealth would you give up to remove all persistence in consumption?" Formally, define the persistence premium by

$$\pi = 1 - \frac{V_0(d^{corr})}{V_0(d^{iid})},$$

where d^{iid} and d^{corr} are given by (7) with $a = 0$ (no persistence) and $a = 0.9790$, respectively. Table 1 summarizes the parameters of the model.[12] Under the level of risk aversion of $1 - \alpha = 7.5$, I obtain the persistence premium: $\pi \approx 0.3028$, while we have $\pi \approx 40\%$ when $1 - \alpha = 10$. 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 of consumption. Using the limited existing experimental evidence from Andersen et al. (2018), my calibration suggests that we should have at most $\pi \approx 20\%$ (see Section 7.2.2). This finding implies that either the level of persistence is

[12]This a standard specification for persistence in the literature, see Bansal and Yaron (2004).

σ	φ	a	β	$1-\alpha$	ρ	x_0	π
0.0078	0.044	0	0.998	7.5	0	0	0
0.0078	0.044	0.9790	0.998	7.5	0	0	30%
0.0078	0.044	0.9790	0.998	10	0	0	40%

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

unrealistic or the degree of risk aversion is too elevated. As previously discussed, one potential solution to this problem is to incorporate other sources of correlation aversion, such as endogenous discounting, into common recursive preferences like Epstein-Zin. However, it is important to note that the calibration exercise is based on limited experimental evidence and requires further empirical validation.

5.3 Utility smoothing

Karantounias (2018, 2022) demonstrates that standard Ramsey tax-smoothing prescriptions for optimal fiscal policy are significantly altered when the decision-maker has Epstein-Zin recursive preferences. The planner adopts a fiscal hedging policy: taxing less during unfavorable conditions and more during favorable conditions to mitigate income shocks. A key driver of this result is that with recursive preferences the planner is averse to volatility in future utilities (see Karantounias (2018), p. 2284). However, an important implication of the previous results is that such a feature of preferences emerges *in spite* of the fact that recursive preferences value early resolution of uncertainty, and is tightly connected with correlation aversion. As shown by Theorems 2 and 3, aversion to volatility in future utilities—mathematically reflected by concavity of the certainty equivalent—is characterized by bounds on preferences for early resolution of uncertainty. The findings of my paper demonstrate that the same implications for optimal fiscal policy may not hold when using recursive preferences that do not satisfy correlation aversion, as is the case with preferences that exhibit DRRA.

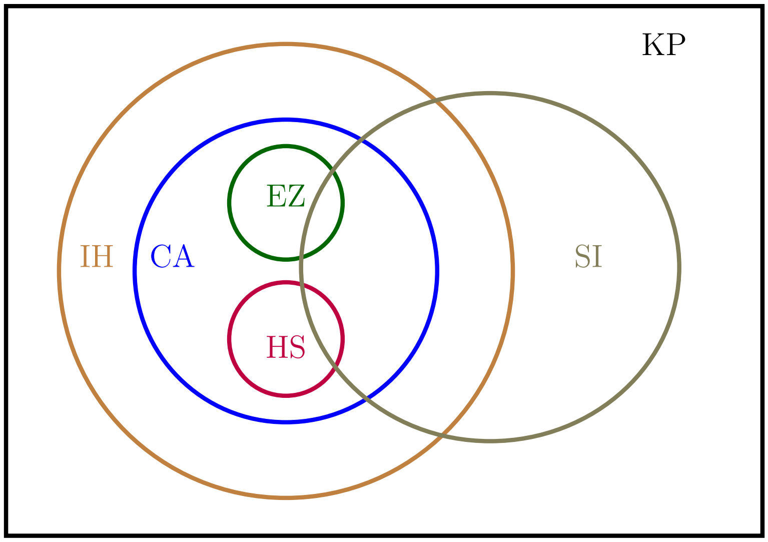

Figure 6: Relationship between correlation averse (CA) preferences and recursive preferences that satisfy intertemporal-hedging (IH), Epstein-Zin (EZ) preferences, and stochastic impatience (SI)

5.4 Non-EU and Stochastic impatience

DeJarnette et al. (2020) and Dillenberger et al. (2020) study stochastic impatience, a property that extends impatience to uncertain environments. Like correlation aversion, stochastic impatience is a normatively desirable behavioral postulate. They find that EZ and HS models exhibit stochastic impatience, provided that the level of risk aversion is not excessively high relative to the inverse elasticity of intertemporal substitution. The relationship between correlation aversion and stochastic impatience is represented in Figure 6. In particular, correlation aversion can be compatible with stochastic impatience. Similar to my findings, their results also advocate for a more comprehensive specification of preferences in order to reduce the level of risk aversion used in applications.

6 Concluding remarks

This paper has explored the relationship between non-instrumental information and intertemporal hedging in the context of recursive preferences. I have shown that under reasonable restrictions on risk attitudes intertemporal hedging is valued more than non-instrumental information. In other words, decision makers will exhibit an aversion to positive autocorrelation in consumption even when it provides information about future consumption. I have discussed the importance of this novel trade-off in various economic applications: asset pricing, optimal taxation, and model misspecification. Note that this trade-off may not be driven solely by risk aversion, as other features of preferences may also be at play. However, standard models affect correlation aversion only through risk aversion. Further research is necessary to develop models of decision making that enable a greater disentangling. This paper has suggested a potential solution by integrating the Kreps-Porteus recursive framework with time non-separable preferences, such as those found in the Epstein-Uzawa model.

7 Appendix

7.1 Proofs

7.1.1 Proof of Theorem 1

Lemma 1. *Each U_t defined in (6) is convex if and only if (5) holds.*

Proof. First I claim that each U_t defined in (6) is convex if and only if the function $\Phi : \phi(u(C)) \to \mathbb{R}$ defined by $x \mapsto \phi(y+\beta\phi^{-1}(x))$ is convex. To see this point, observe that for every $\bar{c} \in C$ we have that

$$U_t(\alpha m + (1-\alpha)m') \leq \alpha U(m) + (1-\alpha)U(m') \iff$$
$$\phi\left(\bar{c} + \beta\phi^{-1}\left(\alpha\mathbb{E}_m\phi\left(V_{t+1}\right) + (1-\alpha)\mathbb{E}_m\phi(V_{t+1})\right)\right) \leq$$
$$\alpha\phi\left(\bar{c} + \beta\phi^{-1}\left(\mathbb{E}_m\phi\left(V_{t+1}\right)\right)\right) + (1-\alpha)\phi\left(\bar{c} + \beta\phi^{-1}\left(\mathbb{E}_{m'}\phi\left(V_{t+1}\right)\right)\right).$$

Since $u(C)$ is unbounded above and the statement above has to hold for every $m, m' \in \Delta_s(D_{t+1,s})$ it follows that convexity of U_t is equivalent to

$$\phi\left(\bar{c} + \beta\phi^{-1}\left(\alpha x + (1-\alpha)y\right)\right) \leq \alpha\phi\left(\bar{c} + \beta\phi^{-1}\left(x\right)\right) + (1-\alpha)\phi\left(\bar{c} + \beta\phi^{-1}\left(y\right)\right),$$

for every $x, y \in \phi(u(C))$ which is equivalent to convexity of Φ for every $\bar{c} \in u(C)$. Finally, the claim follows by using Lemma 3 in Strzalecki (2013). □

Proof of Theorem 1. Now if ϕ satisfies (5), then U_t is convex by Lemma 1. Take $d, d' \in B$ such that $d = (c, m), d' = (c, m')$ and $d \geq_B d'$. By Theorem 4 in Kihlstrom (1984),

$W_t(\operatorname{marg}_{\Delta_s(D_{t+1,s})} m) \geq W_t(\operatorname{marg}_{\Delta_s(D_{t+1,s})} m')$ for every real-valued convex function $W_t : \Delta_s(D_{t+1,s}) \to \mathbb{R}$. By convexity of U_t, it follows that $U_t(\operatorname{marg}_{\Delta_s(D_{t+1,s})} m) \geq U_t(\operatorname{marg}_{\Delta_s(D_{t+1,s})} m')$, and therefore that $d \succeq_0 d'$.

Conversely, consider $d, d' \in B$ given by

$$d = (c_0, \alpha\,(\bar{c}, m_1) \oplus (1-\alpha)\,(\bar{c}, m_2))\,,$$

and

$$d' = (c_0, \bar{c}, \alpha m_1 \oplus (1-\alpha) m_2)\,,$$

where $\alpha \in [0,1]$ and $V_2(m_1) = x$, $V_2(m_2) = y$. We have that $d \succeq_0 d'$ if and only if

$$\alpha\phi(\bar{c} + \beta\phi^{-1}(x)) + (1-\alpha)\phi(\bar{c} + \beta\phi^{-1}(y)) \geq \phi(\bar{c} + \beta\phi^{-1}(\alpha x + (1-\alpha y))).$$

Since the statement has to hold for arbitrary $x, y \in u(C)$ (recall that u is unbounded above) and $\alpha \in [0,1]$, it follows that the mapping $x \mapsto \phi(\bar{c} + \beta\phi^{-1}(x))$ must be convex. Hence an immediate application of Lemma 1 concludes the proof. $\square$

7.1.2 Proof of Proposition 1

Proof. Take $d, d' \in D^*_{0,s}$. Without loss of generality we can assume m_1, m'_1 and have common support $\{c_1, \ldots, c_n\} \subseteq [0,1]$, and $(m_2(\cdot|c_i))_{i=1}^n, (m'_2(\cdot|c_i))_{i=1}^n$ have common support over $\{c_1, \ldots, c_m\} \subseteq [0,1]$, with $c_1 < \ldots < c_n$ and $c_1 < \ldots < c_m$. Let $g_{ij} = m_2(c_i|c_j)m_2(c_i)$, $f_{ij} = m'_2(c_i|c_j)m'_1(c_i)$ for $i = 1, \ldots, n$ and $j = 1, \ldots, m$. Likewise, let $G(c,c') = \sum\sum_{j:c_j \leq c}\sum_{i:c_i \leq c'} g_{ij}$ and $F(c,c') = \sum\sum_{j:c_j \leq c}\sum_{i:c_i \leq c'} f_{ij}$. Observe that if d differs from d' by an IECIT then G differs from F by an elementary correlation-increasing transformation as defined by Epstein and Tanny (1980) (see their Definition 1). By Theorem 1 in Epstein and Tanny (1980) it follows that if $d \geq_C d'$, we have $G \geq F$, which by using the fact that $m_1 = m'_1$ one obtains

$$\begin{aligned}
& G \geq F \\
\iff & m'\,(c_2 \leq c, c_1 \leq c') \leq m\,(c_2 \leq c, c_1 \leq c') \\
\iff & m'_2\,(c_2 \leq c \mid c_1 \leq c') \leq m_2\,(c_2 \leq c \mid c_1 \leq c')\,,
\end{aligned}$$

for every $(c, c') \in C \times C$ as desired. $\square$

7.1.3 Proof of Proposition 2

Proof. Denote with $\{c, c', \dots, c_N\}$ the support of $\ell \in \Delta_s(C)$. It suffices to show that if $d' \in D^*_{0,s}$ differs from some $d^{iid}(\ell) \in D^*_{0,s}$ by an IECIT and $d \in D^*_{0,s}$ differs from d' by an IECIT then $d \geq_B d' \geq_B d^{iid}(\ell)$. Suppose that d' that differs from $d^{iid}(\ell)$ by an IECIT. Then for some $\varepsilon \geq 0$ and (c, c') it holds that

$$\begin{bmatrix} \ell(c) & \ell(c') & \dots & \ell(c_N) \\ \ell(c) & \ell(c') & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \ell(c) & \ell(c') & \dots & \ell(c_N) \end{bmatrix} = \begin{bmatrix} x & 1-x & 0 & \dots & 0 \\ x & 1-x & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \\ 0 & \vdots & \vdots & 0 & 1 \end{bmatrix} \begin{bmatrix} \ell(c)+\varepsilon & \ell(c')-\varepsilon & \dots & \ell(c_N) \\ \ell(c)-\varepsilon & \ell(c')-\varepsilon & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \ell(c) & \dots & \dots & \ell(c_N) \end{bmatrix},$$

where $x \in [0,1]$, so that $d' \geq_B d^{iid}(\ell)$. Using the same reasoning it is immediate to show that $d \geq_B d'$. $\square$

7.1.4 Proof of Theorem 2

It is enough to prove that for every $d, d' \in D^*_{0,s}$ and $\ell \in \Delta_s(C)$,

$$d \geq_C d' \geq_C d^{iid}(\ell) \implies d \succeq_0 d' \succeq_0 d^{iid}(\ell).$$

The statement can be extended to arbitrary elements of means of continuity of preferences.[13] I prove first the following preliminary result.

Lemma 2. *Consider d, d' such that d' differs from some $d^{iid}(\ell)$ by an IECIT and d differs from d' by an IECIT. Then there exists a differentiable function $U : [0,1] \to \mathbb{R}$ such that*

1. *$U(0) = V_0(d)$ and $U(1) = V(d')$;*
2. *$\lim_{\varepsilon \to 0} U'(\varepsilon) \leq 0$ whenever $d = d^{iid}(\ell)$;*
3. *$U''(\varepsilon) \geq 0$ for every $\varepsilon \in (0,1)$.*

Proof. See the Supplemental Appendix. $\square$

[13] To see this, assume that $d \geq_C d' \geq_C d^{iid}(\ell)$. Then there exist sequences $(d_n)_{n=0}^\infty$, $(d'_n)_{n=0}^\infty$ and $(d^{iid}(\ell_n))_{n=0}^\infty$, such that $\lim_n d_n = d$, $\lim_n d'_n = d'$ and $\lim_n d^{iid}(\ell_n) = d^{iid}(\ell)$ and $d_n \geq_C d' \geq_C d^{iid}_n(\ell_n)$. Then $d^{iid}(\ell) \succeq_0 d' \succeq_0 d$ follows by continuity.

It is now possible to prove Theorem 2. To this end, given $\ell \in \Delta_s(C)$, denote with $d^{corr}(\ell) = (c, m) \in D^*_{0,s}$ defined by $m_1 = \ell$ and $m_2(c|c) = 1$ for every $c \in \operatorname{supp} \ell$.

Proof of Theorem 2. By Lemma 2, there exists $U : [0,1] \to \mathbb{R}$ such that for some $q_1, q_2 \in [0,1]$ with $q_1 < q_2$ it holds that $U(0) = V_0(d^{iid}(\ell))$, $U(q_1) = V_0(d')$, $U(q_2) = V_0(d)$, $U(1) = V_0(d^{corr}(\ell))$, $\lim_{\varepsilon\to 0} U'(\varepsilon) \le 0$, and $U'(\varepsilon) \ge 0$ for every $\varepsilon \in (0,1)$ (where derivatives are intended in a weak sense, see Section 8.2 in Brezis (2010)).[14] I claim that it also holds that

$$\lim_{\varepsilon\to 1} U'(\varepsilon) \le 0.$$

Indeed, we have for some $p, q \in (0,1)$ and x, y such that $x \ge y$

$$\begin{aligned}
\lim_{\varepsilon\to 1} U'(\varepsilon) =& \lim_{\varepsilon\to 1} \frac{\partial}{\partial\varepsilon}\Bigg[p\phi\left(x + \beta\phi^{-1}\left(\phi(x)\left(p + q\varepsilon\right) + \phi(y)\left(q - q\varepsilon\right)\right)\right) + \\
& q\phi\left(y + \beta\phi^{-1}\left(\phi(x)\left(p - p\varepsilon\right) + \phi(y)\left(q + p\varepsilon\right)\right)\right)\Bigg] \\
& \le (\phi(x) - \phi(y))\left(\frac{\phi'((1+\beta)x)}{\phi'(x)} - \frac{\phi'((1+\beta)y)}{\phi'(y)}\right) \\
& = (\phi(x) - \phi(y))\int_y^x \frac{(1+\beta)\frac{\phi''(z(1+\beta))}{\phi'(z(1+\beta))} - \frac{\phi''(z)}{\phi'(z)}}{\left(\phi'(z)\right)^2} dz \le 0,
\end{aligned}$$

where the last inequality follows by the fact that ϕ satisfies IRRA, upon observing that

$$(1+\beta)\frac{\phi''(z(1+\beta))}{\phi'(z(1+\beta))} - \frac{\phi''(z)}{\phi'(z)} \le 0 \iff -z(1+\beta)\frac{\phi''(z(1+\beta))}{\phi'(z(1+\beta))} \ge -z\frac{\phi''(z)}{\phi'(z)}.$$

Applying the fundamental theorem of calculus for weak derivatives (see Theorem 8.2 in Brezis (2010)), it follows that

$$V(d') - V(d^{iid}) = \int_0^{q_1} U'(\tilde\varepsilon)d\tilde\varepsilon \le 0,$$

and

$$V(d) - V(d') = \int_{q_1}^{q_2} U'(\tilde\varepsilon)d\tilde\varepsilon \le 0.$$

[14] By applying Lemma 2, if there is a sequence $(d_i)_{i=0}^N$ such that each d_i differs from d_{i-1} by an IECIT, then one can construct $U : [0,1] \to \mathbb{R}$ that is continuous and weakly differentiable by setting $(U_i)_{i=1}^N$ using Lemma 2 and setting $U(x) = U_i\left(\frac{Nx}{i}\right)$ for $x \in \left[\frac{i-1}{N}, \frac{i}{N}\right)$, $i = 1, \dots, N-1$ and $x \in \left[\frac{N-1}{N}, 1\right]$.

Hence we obtain $d' \succeq_0 d \succeq_0 d^{iid}$ for every $\succeq$ with KP representation (ϕ, u, β) as desired.

Conversely, assume that ϕ does not satisfy IRRA. Then there exists $\underline{z} < \bar{z}$ such that R_ϕ is non-increasing over the interval $[\underline{z}, \bar{z}]$ and $R_\phi(\bar{z}) < R_\phi(\underline{z})$. Pick $\beta \in (0,1]$ such that $\frac{\bar{z}}{1+\beta} > \underline{z}$, and let $x = \frac{\bar{z}}{1+\beta}$, $y = \underline{z}$. Consider $d^{iid}(\ell)$ where $\ell(x) = \ell(y) = \frac{1}{2}$. Let $d^\varepsilon(\ell) = (c_0, m)$ where $m_2(x|x) = \ell(x) + \frac{1}{2}\varepsilon$, and $m_2(y|y) = \ell(y) + \frac{1}{2}\varepsilon$. Then $d^\varepsilon(\ell) \geq_C d^{\varepsilon'}(\ell) \geq_C d^{iid}(\ell)$ for $\varepsilon \geq \varepsilon'$. Now define $U : [0,1] \to \mathbb{R}$ such that $U(\varepsilon) = V_0(d^\varepsilon(\ell))$. Applying the same reasoning as in Lemma 2, we obtain $U''(\varepsilon) \geq 0$ for $\varepsilon \in (0,1)$ since ϕ satisfies UPI. Therefore by analogous calculations as before we have that since IRRA is not satisfied

$$\lim_{\varepsilon \to 1} U'(\varepsilon) \propto (\phi(x) - \phi(y)) \int_y^x \frac{(1+\beta)\frac{\phi''(z(1+\beta))}{\phi'(z(1+\beta))} - \frac{\phi''(z)}{\phi'(z)}}{(\phi'(z))^2} dz > 0,$$

which implies that for some $\bar{\varepsilon} < 1$ it most hold that $U'(\tilde{\varepsilon}) > 0$ for every $\tilde{\varepsilon} \in [\bar{\varepsilon}, 1)$. It follows that

$$V(d^1(\ell)) - V(d^{\bar{\varepsilon}}(\ell)) = \int_{\bar{\varepsilon}}^1 U'(\varepsilon) d\varepsilon > 0,$$

which implies that $d^1(\ell) \geq_C d^{\bar{\varepsilon}}(\ell) \geq_C d^{iid}(\ell)$ but $d^1(\ell) \succ_0 d^{\bar{\varepsilon}}(\ell)$. We can therefore conclude that ϕ must satisfy IRRA as desired. $\square$

7.1.5 Proof of Theorem 3

I introduce first some important notation: given a measurable space (S, Σ), $ca(\Sigma)$ is the set of all countably additive elements of the set of charges $ba(\Sigma)$, while $ca_+(\Sigma) = ca(\Sigma) \cap ba_+(\Sigma)$ is its positive cone and $\Delta(\Sigma)$ is the set of countably additive probability measures. Given $p \in ba(\Sigma)$, let $ba(\Sigma, p) = \{v \in ba(\Sigma) : B \in \Sigma \text{ and } p(B) = 0$ implies $v(B) = 0\}$. Observe that $ba(\Sigma, p)$ is isometrically isomorphic to (see **?**, Theorem IV.8.16) the dual of $L^\infty(p) := L^\infty(S, \Sigma, \mu)$ and $ca(\Sigma, p) = ca(\Sigma) \cap ba(\Sigma, p)$ is (isometrically isomorphic to) $L^1(p)$ (via the Radon-Nikodym derivative $\nu \mapsto d\nu/dp$). Turning to the proof of Theorem 2, I first introduce important notions related to quasi-arithmetic certainty equivalent functionals: given $p \in \Delta(\Sigma)$, let $M_{\phi,p} : L^\infty(p) \to \mathbb{R}$ be defined by

$$\phi^{-1}\left(\int \phi(\xi) dp\right) \text{ for every } \xi \in L^\infty(p).$$

The functional $M_{\phi,p}$ is well-defined whenever ϕ is continuous and non-decreasing. I provide an important result concerning the convex conjugate $M^*_{\phi,p}$ of the quasi-arithmetic mean.

Lemma 3. *Assume that $M_{\phi,p}$ satisfies $M_{\phi,p}(\xi+k) \geq M_{\phi,p}(\xi)+k$ for every $\xi \in L^\infty(p)$ and $k \in \mathbb{R}$. Then the Fenchel-Moreau conjugate satisfies $M^*_{\phi,p}(q) = -\infty$ when $q \notin \Delta(\Sigma)$.*

Proof. Observe first that by the aforementioned isometry between the dual of $L^\infty(p)$ and $ba(\Sigma)$, the convex conjugate $M^*_{\phi,p}$ can be seen as a mapping $ba(\Sigma, p) \to [-\infty, 0]$ defined by

$$M^*_{\phi,p}(q) = \inf_{\xi \in L^\infty(p)} \int \xi dq - M_{\phi,p}(\xi).$$

Now by assumption,

$$M^*_{\phi,p}(q) = \inf_{\xi \in L^\infty(p)} \int \xi dp - M_{\phi,q}(\xi) \leq \phi^{-1}\left(\inf_{\xi} \int \xi dp - \int \phi(\xi) dp\right).$$

Therefore, Corollary 2A in Rockafellar (1971) implies that $M^*_{\phi,p}(q) = -\infty$ whenever $q \notin ca(\Sigma, p)$. Further, assume that $q(S) \neq 1$. Again by assumption on $M_{\phi,p}$

$$\int (\xi + b)\, dq - M_{\phi,p}(\xi + b) \leq \int \xi dq - M_{\phi,p}(\xi) + b(q(S) - 1),$$

for all $b \in \mathbb{R}$ and so $M^*_{\phi,p}(q) = -\infty$ as desired. □

Denote with $L^\infty_+(p)$ be the non-negative orthant of $L^\infty(p)$.

Theorem 5 (See Hardy et al. (1952) Theorem 106, Chudziak et al. (2019) or Gollier (2001))**.** *Consider $\phi : \mathbb{R} \to \mathbb{R}$ strictly increasing, strictly concave, and twice differentiable over $(0, \infty)$. Then $M_{\phi,p}|L^\infty_+(p)$ is concave if and only if $A_{\phi|(0,\infty)}$ is convex.*

Proof. First observe that $A_{\phi|(0,\infty)}$ is convex if and only if $\frac{1}{A_{\phi|(0,\infty)}}$ is concave. To see this point, take $x, y \in (0, \infty)$. It is enough to prove that

$$\frac{1}{A_\phi\left(\frac{x+y}{2}\right)} \leq \frac{1}{2}\left(\frac{1}{A_\phi(x)} + \frac{1}{A_\phi(y)}\right). \tag{8}$$

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 (8) as

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

Now if A_ϕ is convex, by it follows that by setting $L^\infty_{s,+}(p) := \{\xi \in L^\infty_{s,+}(p) : \xi = \sum_{k=1}^n a_k \mathbf{1}_{A_k}, (a_k)_{k=1}^n \in \mathbb{R}^n_+\}$, one can apply the above claim combined with Theorem 1 and Theorem 5 in Chudziak et al. (2019) to show that $M_{\phi,p}|L^\infty_{s,+}(p)$ is concave. Convexity of $M_{\phi,p}|L^\infty_+(p)$ follows by the fact that $L^\infty_{s,+}(p)$ is dense in $L^\infty_+(p)$. Conversely, if $M_{\phi,p}|L^\infty_+(p)$ is concave then $M_{\phi,p}|L^\infty_{s,+}(p)$ is also concave, which by Theorem 1 and Theorem 5 in Chudziak et al. (2019) implies that $A_{\phi|(0,\infty)}$ must be convex. □

Thanks to Theorem 5, we obtain the following powerful result, which shows that the conjunction of DARA and SCA on ϕ implies the concavity of the quasi-arithmetic mean $M_{\phi,p}|L^\infty_+(p)$.

Corollary 1. *Assume that ϕ is four times continuously differentiable and satisfies UPI over $(0,\infty)$. Then if R_ϕ is convex $M_{\phi,p}|L^\infty_+(p)$ is concave.*

Proof. Observe that

$$R''_\phi(x) = (xA_\phi(x))'' = xA''_\phi(x) + 2A'_\phi(x).$$

If ϕ satisfies UPI, then $A'_\phi \leq 0$. Hence if ϕ satisfies SCA so that $R''_\phi \geq 0$, it has to be the case that

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

The result therefore follows by Theorem 5. □

Such a result in a way completes Theorem 12 and Corollary 1 in Marinacci and Montrucchio (2010), who characterize quasi-arithmetic certainty equivalents that are constant subadditive and subhomogeneous. Indeed, one way to think about Corollary 1 is that it implies that for quasi-arithmetic means to be concave it is enough to assume constant superadditivity (DARA) and SCA, where the latter property implies subhomogeneity (IRRA) of $M_{\phi,p}$. It is important to observe that both EZ and HS preferences satisfy such an assumption.

Corollary 2. *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. Immediate from Theorem 5. □

Now consider $\succeq$ with KP representation (ϕ, u, β). Without loss of generality, assume $u(C) = [0, \infty)$. I now show that letting

$$\hat{\phi}(x) = \begin{cases} \phi(x) & x \geq 0 \\ -\infty & x < 0, \end{cases}$$

then $M_{\hat{\phi},p}$ is concave if ϕ satisfies SCA.

Lemma 4. *If $\phi : [0, \infty) \to \mathbb{R}$ satisfies SCA, then $M_{\hat{\phi},p}$ is concave.*

Proof. By Corollary 1, $M_{\phi,p}|L^\infty_+(p)$ is concave. Now given $\xi, \xi' \in L^\infty(p)$ and $\alpha \in [0, 1]$, if $M_{\phi,p}(\alpha\xi + (1-\alpha)\xi') = -\infty$ then it must be the case that $M_{\phi,p}(\xi) = -\infty$ or $M_{\phi,p}(\xi') = -\infty$, so that $M_{\phi,p}(\alpha\xi + (1-\alpha)\xi') \geq \alpha M_{\phi,p}(\xi) + (1-\alpha)M_{\phi,p}(\xi')$. If $M_{\phi,p}(\alpha\xi + (1-\alpha)\xi') > -\infty$ and $M_{\phi,p}(\xi) = -\infty$ or $M_{\phi,p}(\xi') = -\infty$, then $M_{\phi,p}(\alpha\xi + (1-\alpha)\xi') \geq \alpha M_{\phi,p}(\xi) + (1-\alpha)M_{\phi,p}(\xi')$. Finally, if $M_{\phi,p}(\xi), M_{\phi,p}(\xi') > -\infty$ then it must be that $\xi, \xi' \in L^\infty_+(p)$ so that $M_{\phi,p}(\alpha\xi + (1-\alpha)\xi') \geq \alpha M_{\phi,p}(\xi) + (1-\alpha)M_{\phi,p}(\xi')$ as desired. □

It is now possible to deliver a proof of Theorem 3.

Proof of Theorem 3. Given $(V_t)_{t=1}^T$ from the KP representation, observe that for every $m_t \in \Delta_b(D_t)$, where $\mathcal{D}_t$ is the Borel σ-algebra of D_t, since each $V_t : D_t \to \mathbb{R}$, $t = 1, \ldots, T$ is continuous we have $V_t \in L^\infty_+(D_t, \mathcal{D}_t, m_t) := L^\infty_+(m_t)$. If ϕ satisfies SCA, then by Lemma $M_{\hat{\phi},m_t}$ is concave for each $t = 1, \ldots, T-1$. By applying the Fenchel-Moreau Theorem (see Phelps (2009), p. 42) and Lemma 4 it follows that

$$M_{\hat{\phi},m_t}(\xi) = \inf_{q \in \Delta(\mathcal{D}_t, m_t)} \mathbb{E}_q\xi - M^*_{\hat{\phi},m_t}(q) \quad \text{for all } \xi \in L^\infty(m_t).$$

Now using the isometry between $ca(\mathcal{D}_t, m_t)$ and $L^1(m_t)$, we can write

$$M^*_{\hat{\phi},m_t}(q) = M^*_{\phi,m_t}(q) = \inf_{\xi \in L^\infty_+(m_t) : \mathbb{E}_{m_t} \frac{dq}{dm_t}\xi = \mathbb{E}_q\xi} \left\{ \mathbb{E}_{m_t}\xi - \phi^{-1}\left(\mathbb{E}_{m_t}\phi\left(\xi\right)\right) \right\}.$$

By applying Proposition 1 in Frittelli and Bellini (1997) one obtains

$$\begin{aligned} M^*_{\phi,m_t}(q) &= \sup_{\xi \in L^\infty_+(m_t) : \mathbb{E}_{m_t}\left[\frac{dq}{dm_t}\xi\right] = \mathbb{E}_q\xi} \left\{ \phi^{-1}\mathbb{E}_{m_t}[\phi(\xi)] - \mathbb{E}_{m_t}\left[\frac{dq}{dm_t}\xi\right] \right\} \\ &= \phi^{-1}\left(\mathbb{E}_m(\phi \circ \psi)\left(k(q)\frac{dq}{dm_t}\right)\right) - \mathbb{E}_q V_{t+1}, \end{aligned}$$

where $\psi = (\phi')^{-1}$ and $k(q) \in (0, \infty)$ is the only solution to the equation

$$\mathbb{E}\psi\left(k(q)\frac{dq}{dm_t}\right)dm_t = \mathbb{E}_q\xi.$$

Hence if for $t = 0, \ldots, T-1$ we set

$$I^t_{\phi,u,\beta}(\ell||m_t) := \begin{cases} \mathbb{E}_\ell V_{t+1} - \phi^{-1}\left(\mathbb{E}_{m_t}(\phi \circ \psi)\left(k(\ell)\frac{d\ell}{dm_t}\right)\right) & \ell \ll m_t, \\ +\infty & \text{otherwise,} \end{cases}$$

then one obtains

$$V_t(c, m_t) = u(c) + \beta \min_{\ell \ll m_t}\left\{\mathbb{E}_\ell V_{t+1} + I^t_{\phi,u,\beta}(\ell\|m_t)\right\}$$

where the infimum is attained because $\{\ell \in \Delta_b(D_{t+1}) : \ell \ll m_t\}$ is a closed subset of the compact metric space $\Delta_b(D_{t+1})$ (see Epstein and Zin (1989), p. 962) for $t = 0, \ldots, T-1$. Observe that $\mathbb{E}_\ell V_{t+1}$ is entirely determined by (ϕ, u, β) and ℓ. Finally, observe that each $I^t_{\phi,u,\beta}$ is a premetric generalized distance in the sense of Csiszár (1995). Indeed, one can show that $I^t(\ell||m) = 0$ if and only if $m = \ell$ by adapting the same arguments as in Remark 8 in Frittelli and Bellini (1997). Further, Proposition 16 in Cerreia-Vioglio et al. (2011) can be used to show that

$$I^t_{(u,\beta,\phi_1)}(\cdot||m) \leq I^t_{(u,\beta,\phi_2)}(\cdot||m),$$

whenever $A_{\phi_1} \geq A_{\phi_2}$.

Further, observe that in the Epstein-Zin case we have (see Section 5.2 in Frittelli and Bellini (1997)) by setting $q = \frac{\alpha}{\alpha-\rho}$,

$$I^t_{\phi,u,\beta}(\ell\|m) = \mathbb{E}_\ell V_{t+1}\left\{\left(\mathbb{E}_m\left[\frac{d\ell}{dm}^q\right]\right)^{-\frac{1}{q}} - 1\right\},$$

so that upon noticing that the Rényi divergence is given for any $q > 0$, $q \neq 1$ (see Van Erven and Harremos (2014)) by

$$R_q(\ell\|m) = \frac{1}{q-1}\log\left(\mathbb{E}_m\left[\frac{d\ell}{dm}^q\right]\right),$$

we obtain that whenever $\alpha < 0$ and $\frac{1}{1-\rho} > 1$ it holds that $q > 0$, $q \neq 1$ so that

$$I_{\phi,u,\beta}(\ell, m) = \mathbb{E}_\ell V_{t+1}\left[e^{\frac{1-q}{q}R_q(\ell\|m)} - 1\right],$$

as desired. □

7.1.6 Proof of Proposition 3

Define again $U : [0,1] \to \mathbb{R}$ by $U(\varepsilon) = V_0(d^\varepsilon(\ell))$. Similarly to the proof of Theorem 2, for every $\varepsilon \geq \varepsilon'$ we have

$$\int_{\varepsilon'}^{\varepsilon} U'(\varepsilon)d\varepsilon = U(\varepsilon) - U(\varepsilon') \leq 0,$$

which implies $d^{\varepsilon'}(\ell) \succeq_0 d^\varepsilon(\ell)$ as desired.

7.1.7 Proof of Proposition 4

Suppose that for each $i \in \{1,2\}$, the preference relation $\succeq_i$ admits a KP representation of the form $(\tilde{\phi}_i, \tilde{u}_i, \tilde{\beta}_i)$. Observe first that if $\succeq^1$ is more risk averse than $\succeq^2$ then it must be that for every $c^T, \hat{c}^T \in C^T$

$$c^T \succeq_0^1 \hat{c}^T \iff c^T \succeq_0^2 \hat{c}^T,$$

which implies that $\tilde{\beta}_1 = \tilde{\beta}_2$ and $\tilde{u}_1 = a\tilde{u}_2 + b$ for some $a > 0$ and $b \in \mathbb{R}$. Therefore setting $u_2 = \frac{\tilde{u}_2}{a} - \frac{b}{a}$ and $\phi_2(x) = \tilde{\phi}_2(ax+b)$ for every $x \in u_2(C)$, the statement is satisfied with KP representations $(\tilde{\phi}_1, \tilde{u}_1, \tilde{\beta}_1)$ and $(\phi_2, u_2, \tilde{\beta}_2)$. Normalize $\tilde{u}_1(0) = \tilde{u}_2(0) = 0$. Now define $V : C \to u(C)$ by $V(c) := \sum_{t=1}^T \beta^t u(c)$. For each $\ell \in \Delta_s(u(C))$, there exists $(\pi)_{i=1}^n$ and $(\bar{c}_i)_{i=1}^n \in \prod_{i=1}^n C$ such that $\ell = V_\# \bigoplus_{i=1}^n \pi_i \bar{c}_i$.[15] Since $\succeq_1$ is more risk averse than $\succeq_2$, it follows that since $\left(c_0, \bigoplus_{i=1}^n \pi_i(\bar{c}_i, \ldots, \bar{c}_i)\right), (c_0, \mathbf{0}) \in \mathcal{R}$

$$\left(c_0, \bigoplus_{i=1}^n \pi_i(\bar{c}_i, \ldots, \bar{c}_i)\right) \succeq_0^2 (c_0, \mathbf{0}) \implies \left(c_0, \bigoplus_{i=1}^n \pi_i(\bar{c}_i, \ldots, \bar{c}_i)\right) \succeq_0^1 (c_0, \mathbf{0}).$$

for every $\ell \in \Delta_s(u(C))$, if $\mathbb{E}_\ell \phi_2(x) \leq \phi(0)$ then $\mathbb{E}_\ell \phi_1(x) \leq \phi(0)$. Hence the result follows by applying Proposition 2 in Gollier (2001). The converse follows by a straightforward application of Jensen's inequality.

7.2 The persistence premium

7.2.1 Long-run risk

We have that (see Epstein et al. (2014), pp. 2684-2685)

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

[15] Here $V_\# \bigoplus_{i=1}^n \pi_i \bar{c}_i$ denotes the pushforward of $(c_0, \bigoplus_{i=1}^n \pi_i(\bar{c}_i, \ldots, \bar{c}_i))$ by V on $u(C)$.

and

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

Therefore we obtain

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

$$\pi = 1 - \exp\left(-6.5 \times 0.998 \times \frac{0.0078^2}{2(1-0.998)} \left(0.044^2 \times \frac{0.998^2}{(1-0.998 \times 0.979)^2} - 0.044^2 \times 0.998^2\right)\right) \approx 0.302.$$

$$\pi = 1 - \exp\left(-9 \times 0.998 \times \frac{0.0078^2}{2(1-0.998)} \left(0.044^2 \times \frac{0.998^2}{(1-0.998 \times 0.979)^2} - 0.044^2 \times 0.998^2\right)\right) \approx 0.393.$$

Therefore we have that $\pi \approx 30\%$ with $\alpha = 7.50$ and $\pi \approx 40\%$ with $\alpha = 10$.

7.2.2 An upper bound on the persistence premium

There is no independent study that quantifies the persistence premium in the literature. To have a sense of a potential calibrated value, I conduct a thought experiment that provides an upper bound for the persistence. The thought experiment is based on the comparison between an "i.i.d." lottery and a maximally correlated lottery in the sense that there is no other temporal lottery more correlated than it. Andersen et al. (2018) estimate an intertemporal utility function under uncertainty which can be written as

$$V(f) = u^{-1}\phi^{-1}\mathbb{E}_P\left[\varphi\left(\sum_{t=1}^{2} \beta^t u(f_t)\right)\right],$$

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

Given $x > 0$ and $n = 2$, let d^{iid} be the lottery 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, d^{corr} is maximally correlated in the sense that there is no lottery d such that $d \geq_C d^{corr}$ and $d^{corr} \not\geq_C d$. In this case the persistence premium is given for every $x > 0$ by

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

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

8 Bibliography

References

Andersen, S., Harrison, G. W., Lau, M. I., and 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 Publishing.

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

Bansal, R., Ochoa, M., and Kiku, D. (2017). Climate change and growth risks. Technical report, National Bureau of Economic Research.

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

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

Brezis, H. (2010). *Functional analysis, Sobolev spaces and partial differential equations*. Springer.

Cai, Y. and Lontzek, T. S. (2019). The social cost of carbon with economic and climate risks. *Journal of Political Economy*, 127(6):2684–2734.

Campbell, J. Y. (2017). *Financial decisions and markets: a course in asset pricing*. Princeton University Press.

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

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

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

Cochrane, J. (2009). *Asset pricing: Revised edition*. Princeton university press.

Csiszár, I. (1995). Generalized projections for non-negative functions. In *Proceedings of 1995 IEEE International Symposium on Information Theory*, page 6. IEEE.

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

DeJarnette, P., Dillenberger, D., Gottlieb, D., and Ortoleva, P. (2020). Time lotteries and stochastic impatience. *Econometrica*, 88(2):619–656.

Dillenberger, D., Gottlieb, D., and Ortoleva, P. (2020). Stochastic impatience and the separation of time and risk preferences. *Available at SSRN*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3645071.

Duffie, D. and Epstein, L. G. (1992). Stochastic differential utility. *Econometrica: Journal of the Econometric Society*, 60(2):353–394.

Epstein, L. G. (1992). *Behaviour under risk: recent developments in theory and applications, in "Advances in Economic Theory: Sixth World Congress"(JJ Laffont, Ed.), Econometric Society Monograph*. Cambridge University Press, Cambridge.

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

Epstein, L. G., Farhi, E., and Strzalecki, T. (2014). How much would you pay to resolve long-run risk? *American Economic Review*, 104(9):2680–97.

Epstein, L. G. and Tanny, S. M. (1980). Increasing generalized correlation: a definition and some economic consequences. *Canadian Journal of Economics*, 13(1):16–34.

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

Epstein, L. G. and 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.

Frittelli, M. and Bellini, F. (1997). Certainty equivalent and no arbitrage: A reconciliation via duality theory. Working Paper, Università degli Studi di Milano.

Gollier, C. (2001). *The economics of risk and time*. MIT press.

Hansen, L. and Sargent, T. J. (2001). Robust control and model uncertainty. *American Economic Review*, 91(2):60–66.

Hansen, L. P., Heaton, J., Lee, J., and Roussanov, N. (2007). Intertemporal substitution and risk aversion. *Handbook of econometrics*, 6:3967–4056.

Hansen, L. P., Sargent, T. J., and Tallarini, T. D. (1999). Robust permanent income and pricing. *Review of Economic studies*, 66(4):873–907.

Hardy, G. H., Littlewood, J. E., and Pólya, G. (1952). *Inequalities*. Cambridge University Press.

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

Karantounias, A. G. (2022). A general theory of tax-smoothing. Working Paper, University of Surrey.

Kihlstrom, R. E. (1984). *A "Bayesian" exposition of Blackwell's theorem on the comparison of experiments, in "Bayesian models in economic theory"(Marcel Boyer, and Richard Kihlstrom (Eds.))*. North-Holland Amsterdam.

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

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

Kochov, A. and Song, Y. (2021). Intertemporal hedging and trade in repeated games with recursive utility. *Available at SSRN*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2714337.

Kreps, D. M. and Porteus, E. L. (1978). Temporal resolution of uncertainty and dynamic choice theory. *Econometrica*, 46(1):185–200.

Maccheroni, F., Marinacci, M., and Rustichini, A. (2006). Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica*, 74(6):1447–1498.

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

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

Phelps, R. R. (2009). *Convex functions, monotone operators and differentiability.* Springer.

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

Rockafellar, R. (1971). Integrals which are convex functionals. ii. *Pacific journal of mathematics*, 39(2):439–469.

Rothschild, M. and Stiglitz, J. E. (1970). Increasing risk: A definition. *Journal of Economic Theory*, 2:225–243.

Sarver, T. (2018). Dynamic mixture-averse preferences. *Econometrica*, 86(4):1347–1382.

Sason, I. (2022). Divergence measures: Mathematical foundations and applications in information-theoretic and statistical problems. *Entropy*, 24(5):712.

Savage, L. (1954). *Foundations of statistics.* Wiley.

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.

Van Erven, T. and Harremos, P. (2014). Rényi divergence and kullback-leibler divergence. *IEEE Transactions on Information Theory*, 60(7):3797–3820.

Wakker, P. P. (2010). *Prospect theory: For risk and ambiguity*. Cambridge university press.

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

9 Supplemental Appendix

This supplemental material contains two parts. Section 9.1 extends the analysis to an infinite horizon while Section 9.2 provides proofs of the claims made in the main text and the Appendix.

9.1 The case $T = \infty$

As the consumption set $C = [0, \infty) = \mathbb{R}_+$ is identical to that of Epstein and Zin (1989), I follow their approach in introducing the set of temporal lotteries for the case of an infinite horizon, with specific reference to their discussion on pages 940-944. The only deviation in my approach is the use of $\Delta(X)$ to denote the set of Borel probabilities defined on a metric space X. The set of temporal lotteries, denoted by $D(b)$, is defined in equation 2.3 of their paper and is characterized by the expressions given in equations 2.2-2.11, which define all the relevant objects. I also make use of their characterization of temporal lotteries in $D(b)$.

Theorem 6 (Theorem 2.2 in Epstein and Zin (1989))**.**

$$D(b) \text{ is homeomorphic to } C \times \hat{\Delta}(D(b)),$$

where

$$\hat{\Delta}(D(b)) := \left\{ m \in \hat{\Delta}(D(b)) : f\left(m_2\right) \in \bigcup_{l>0} \Delta(Y(b;l)), \quad m_2 = P_2 m \right\}.$$

Because of this result, each $d \in D(b)$ can be identified with $(c, m) \in C \times \hat{\Delta}(D(b))$. Further, each $m \in \hat{\Delta}(D(b))$ can be equivalently identified with an element $\hat{\Delta}(C \times \hat{\Delta}(D(b)))$. Preferences are given by a weak order $\succeq$ over $D(b)$. The utility function $V : D(b) \to \mathbb{R}$ is called recursive if it satisfies the following equation for every $(c, m) \in C \times \hat{\Delta}\left(D(b)\right)$,

$$V\left(c_0, m\right) = \left[c^{\rho} + \beta \phi^{-1}\left[\left(\mathbb{E}_m \phi\left(V\right)\right)\right]^{\rho}\right]^{1/\rho}, \quad 0 \neq \rho < 1, \quad 0 < \beta < 1, \tag{9}$$

where $\phi : [0, \infty) \to \mathbb{R}$. The next result shows that (9) has always a solution, thus making recursive utility well defined in this context.

Theorem 7. *Suppose that ϕ is concave. Then there exists a continuous $V : D(b) \to \mathbb{R}$ that satisfies* (9)*.*

Proof. Denote by $S^{+}(D(b))$ the set of functions that map from $D(b)$ into positive real numbers. Let $h \in S^{+}(D(b))$ be defined as in p. 963 of Appendix 3 in Epstein and Zin (1989). Further, define $S_h^{+}(D(b))$ as follows

$$S_h^{+}(D(b)) \equiv \left\{ X \in S^{+}(D(b)) : \|X\|_h \equiv \sup_{d \in D(b)} \frac{X(d)}{h(d)} < \infty \right\}.$$

Define $T : S_h^{+}(D(b)) \to S_h^{+}(D(b))$ by

$$T(X) = \left[c^{\rho} + \beta \phi^{-1}\left[\left(\mathbb{E}_m \phi\left(X\right)\right)\right]^{\rho}\right]^{1/\rho} \qquad \text{for every } X \in S_h^{+}(D(b)).$$

Let V^* be a continuous function such that

$$V^*\left(c_0, m\right) = \left[c^{\rho} + \beta\left[\mathbb{E}_m\left(V^*\right)\right]^{\rho}\right]^{1/\rho}, \quad 0 \neq \rho < 1, \quad 0 < \beta < 1,$$

which exists uniquely by Theorem 3.1 in Epstein and Zin (1989). Let $T^0(V^*) = T(V^*)$ and $T^n(V^*) = T(T^{n-1}(V^*))$. By Jensen inequality

$$\phi^{-1}\left(\mathbb{E}\phi(X)\right) \leq \mathbb{E}X \text{ for all } X \in S_h^{+}(D(b)) \in \implies T(V^*) \leq V^*.$$

Further, it holds that $T(V^*) \geqslant 0$. By induction, one obtains that the sequence $(T^n(V^*))_{n=0}^{\infty}$ is non-increasing and bounded below. Therefore we can define $V \in S_h^+(D(b))$ $S_h^+(D(b))$ as follows

$$V := \lim_{n\to\infty} T^n V^*,$$

which is continuous by continuity of V^* and by the fact that T maps continuous functions into continuous functions. I now claim that V solves (9). Since

$$T^n V^*(c_0, m) = \left[c^\rho + \beta\phi^{-1}\left[\left(\mathbb{E}_m \phi\left(T^{n-1}V^*\right)\right)\right]^\rho\right]^{1/\rho} \qquad \text{for every } m \in D(b),$$

the statement follows by the fact that

$$\begin{aligned}\lim_{n\to\infty} \phi^{-1}\left[\left(\mathbb{E}_m \phi\left(T^{n-1}V^*(m)\right)\right)\right]^\rho &= \phi^{-1}\left[\left(\mathbb{E}_m \phi\left(\lim_{n\to\infty} T^{n-1}V^*(m)\right)\right)\right]^\rho \\ &= \phi^{-1}\left[\left(\mathbb{E}_m \phi\left(V\right)\right)\right]^\rho.\end{aligned}$$

□

Consider now a weak order $\succeq$ over $D(b)$. Say that $\succeq$ admits a KP representation (ϕ, ρ, β) if there exists $V : D(b) \to \mathbb{R}$ that satisfies (9) and that represents $\succeq$. For every $d \in D(b)$ one can define the present equivalent $PE_{\succeq}(d)$ as the unique single period consumption level $c \in C$ such that $d \sim (c, \mathbf{0})$, where $\mathbf{0} \in D(b)$ is the temporal lottery that pays the constant zero level of consumption at every time period. Now observe that every $m \in \hat{\Delta}(C \times \hat{\Delta}(D(b)))$ and $\succeq$ with KP representation (ϕ, ρ, β) induce the probability $m_{\succeq}$ over $\hat{\Delta}(C \times \hat{\Delta}(C))$ defined as follows:

$$m_{\succeq}(A \times B) = m(A \times B_{\succeq}) \quad \text{ for every closed } A \times B \subseteq C \times \Delta_b(C),$$

where $B_{\succeq} = \{\ell \in \hat{\Delta}(D(b)) : \ell_{\succeq} \in B\}$ and $\ell_{\succeq} \in \Delta_b(C)$ is defined by $\ell_{\succeq}(A) = \ell(\{d \in D(b) : PE_{\succeq}(d) \in A\})$.[16] In words, $m_{\succeq}$ describes the joint distribution between consumption at time $t+1$ and the continuation temporal lottery, where each temporal lottery is expressed in terms of one-period consumption. In this way, it is possible to extend the order $\geq_C$ and the correlation aversion axiom as follows.

Definition 10. *Consider $d = (c, m), d' = (c, m') \in D(b)$ such that $(c, m_{\succeq})), \left(c, m'_{\succeq}\right) \in D_0^*$. Write that $d \geq_C d'$ if and only if $(c, m_{\succeq}) \geq_C (c, m_{\succeq})$.*

[16]The lottery $m_{\succeq}$ is well defined since preferences are continuous, $u(x) = x^\rho$ and each $m \in \hat{\Delta}(D(b))$ has compact support.

Correlation aversion can be then defined as in the main text.

Definition 11. *$\succeq$ exhibits correlation aversion if and only for every $d, d' \in D(b)$*

$$d \geq_C d' \geq_C d^{iid}(\ell) \implies d^{iid}(\ell) \succeq d' \succeq d.$$

The main results of the paper carry out in the same way.

Theorem 8. *Consider ϕ that is twice continuously differentiable and satisfies UPI. Then every $\succeq$ with KP representation (ϕ, ρ, β) exhibit correlation aversion if and only if ϕ satisfies IRRA. Further, if $\succeq$ that admits a KP representation (ϕ, ρ, β) with ϕ that additionally satisfies SCA and is four times continuously differentiable then $\succeq$ admits the recursive representation*

$$V(c, m) = \left[c^\rho + \beta \left(\min_{\ell \in \hat{\Delta}(D(b))} \left\{ \mathbb{E}_\ell V + I_{(u,\beta,\phi)}(\ell || m) \right\} \right)^\rho \right]^{\frac{1}{\rho}},$$

where $I_{(u,\beta,\phi)}(\cdot, \cdot) : \Delta(\hat{D}(b)) \times \Delta(\hat{D}(b)) \to [0, \infty]$ is a generalized distance.

Proof. The proof follows the same steps as the proof of Theorems 2 and 3. $\square$

9.2 Additional proofs

9.2.1 Proof of Lemma 2

Write the support of m_1 as $\{c_1, \ldots, c_N\}$ and $p_i = m_1(c_i)$ for every $i = 1, \ldots, N$. Let

$$U(\varepsilon) = \sum_{i=1}^{N} p_i \phi \left(x_i + \beta \phi^{-1} \left(\sum_{j=1}^{N} p^\varepsilon_{ji} \phi(x_i) \right) \right) \qquad \text{for every } \varepsilon \in [0, 1],$$

where for some $\underline{i}, \underline{j}$ it holds that $p^\varepsilon_{\underline{j}\underline{i}} = p_{\underline{j}\underline{i}} - p_{\underline{i}\underline{j}}\varepsilon$, $p^\varepsilon_{\underline{i}\underline{i}} = p_{\underline{i}\underline{i}} + p_{\underline{i}\underline{j}}\varepsilon$, $p^\varepsilon_{\underline{i}\underline{j}} = p_{\underline{i}\underline{j}} - p_{\underline{j}\underline{i}}\varepsilon$, $p^\varepsilon_{\underline{j}\underline{j}} = p_{\underline{j}\underline{j}} + p_{\underline{j}\underline{i}}\varepsilon$, and otherwise $p_{ji} = m(c_j|c_i)$ for every other j, i. Clearly U defined in such a way satisfies point (1) in the statement. To prove point (2), observe that in this case we have that $p_{ji} = m(c_j)$ for every $j, i = 1, \ldots, N$ and $p^\varepsilon_{ji} = p_j + \varepsilon$ for some

$\underline{i}, \underline{j}$. Now we have that for some $p, q \in (0,1)$ and $x > y$

$$\begin{aligned}
\lim_{\varepsilon\to 0} U'(\varepsilon) = &\lim_{\varepsilon\to 0} \frac{\partial}{\partial \varepsilon}\Big[p\phi\left(x+\beta\phi^{-1}\left(\phi(x)\left(p+q\varepsilon\right)+\phi(y)\left(q-p\varepsilon\right)\right)\right)+\\
&q\phi\left(y+\beta\phi^{-1}\left(\phi(x)\left(p-q\varepsilon\right)+\phi(y)\left(q+p\varepsilon\right)\right)\right)\Big]\\
&\leq (\phi(x)-\phi(y))\lim_{\varepsilon\to 0}\left[\frac{\phi'(x+\beta\phi^{-1}\left(\phi(x)\left(p+q\varepsilon\right)+\phi(y)\left(q-p\varepsilon\right)+k\right))}{\phi'\left(\phi^{-1}(\phi(x)\left(p+q\varepsilon\right)+\phi(y)\left(q-p\varepsilon\right)+k\right))}-\right.\\
&\left.\frac{\phi'(y+\beta\phi^{-1}\left(\phi(x)\left(q+p\varepsilon\right)+\phi(y)\left(p-p\varepsilon\right)+k\right))}{\phi'\left(\phi^{-1}(\phi(x)\left(q+p\varepsilon\right)+\phi(y)\left(p-p\varepsilon\right)+k\right))}\right]\\
&=(\phi(x)-\phi(y))\left[\frac{\phi'(x+\beta\phi^{-1}\left(\phi(x)p+\phi(y)q+k\right))}{\phi'\left(\phi^{-1}(\phi(x)p+\phi(y)q+k\right))}-\right.\\
&\left.\frac{\phi'(y+\beta\phi^{-1}\left(\phi(x)q+\phi(y)p+k\right))}{\phi'\left(\phi^{-1}(\phi(x)q+\phi(y)p+k\right))}\right]\\
&=\frac{(\phi(x)-\phi(y))}{\phi'\left(\phi^{-1}(\phi(x)q+\phi(y)p+k\right))}\int_y^x \phi'(z)\phi''\left(z+\beta\phi^{-1}\left(\phi(z)+k\right)\right)dz \leq 0,
\end{aligned}$$

where the last inequality follows by the fact that ϕ is strictly increasing and concave and that $x > y$. Now to prove point (3), observe that letting

$$g_1(\varepsilon) := p_{\underline{i}}\phi\left(x_i+\beta\phi^{-1}\left(\sum_j p^{\varepsilon}_{j\underline{i}}\phi(x_i)\right)\right),$$

and

$$g_2(\varepsilon) := p_{\underline{j}}\phi\left(x_i+\beta\phi^{-1}\left(\sum_j p^{\varepsilon}_{j\underline{j}}\phi(x_i)\right)\right),$$

which are convex by Lemma 1. Then we obtain

$$\begin{aligned}
U''(\varepsilon) =&\frac{\partial^2}{\partial\varepsilon^2}\left[p_{\underline{i}}\phi\left(x_i+\beta\phi^{-1}\left(\sum_j p^{\varepsilon}_{j\underline{i}}\phi(x_i)\right)\right)+p_{\underline{j}}\phi\left(x_i+\beta\phi^{-1}\left(\sum_j p^{\varepsilon}_{j\underline{j}}\phi(x_i)\right)\right)\right]\\
=&g_1''(\varepsilon)+g_2''(\varepsilon)\geq 0,
\end{aligned}$$

for every $\varepsilon \in (0,1)$ as desired.

9.2.2 Intertemporal hedging

Proposition 5. *$\succeq$ with KP representation (ϕ, u, β) satisfies intertemporal hedging if and only if ϕ is concave.*

Proof. Observe that intertemporal hedging is equivalent to

$$\frac{1}{2}\phi(x+\beta x)+\frac{1}{2}\phi(y+\beta y)\leq\frac{1}{2}\phi(y+\beta x)+\frac{1}{2}\phi(x+\beta x),$$

for every $x, y \in u(X)$. Therefore the statement follows by a straightforward application Theorem 4(a) in Epstein and Tanny (1980). □