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

Lorenzo Stanca†

July 5, 2023

Abstract

This paper investigates a novel behavioral feature exhibited by recursive preferences: aversion to risks that are persistent through time. I introduce a formal notion of correlation aversion to capture this phenomenon and provide a characterization based on risk attitudes. Furthermore, correlation averse preferences admit a specific variational representation, which connects correlation aversion to fear of model misspecification. These findings imply that correlation aversion is the main driver in many applications of recursive preferences 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 am grateful to Nabil Al-Najjar, Peter Klibanoff, and Marciano Siniscalchi for their help and guidance. I thank Antoine Bommier for providing useful comments. Additionally, I thank Francesco Fabbri, Ignacio Monzón, and Giulio Principi for their feedback.

†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 applications. 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)). A crucial behavioral property of recursive preferences in these applications is their ability to distinguish between risk aversion and intertemporal substitution.

The standard model of discounted expected utility in its recursive form can be written as

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

where β is the discount factor and p describes the distribution of future consumption. 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—to the demand for non-instrumental information that are necessary and sufficient for recursive preferences to satisfy correlation aversion.

To illustrate, consider two gambles: A and B. 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 following period. Otherwise, it is constant at the level 0 at every period. In gamble B, consumption is determined by tossing a fair coin at every period, giving a level

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

of consumption equal to 1 if heads and 0 otherwise. A hedging motive suggests that a decision maker should prefer B to A. But at the same time B resolves gradually while for A all uncertainty resolves at $t = 1$. As a result, gamble A provides more non-instrumental information than B regarding future consumption. The comparison between these two gambles is therefore non-obvious: A has the advantage of resolving all uncertainty at $t = 1$, while B is more desirable because of its hedging value.[2] My notion of correlation aversion requires the hedging motive to dominate the preference for non-instrumental information so a decision maker prefers B to A.

I consider a setting in which preferences are defined over temporal lotteries. I introduce a novel notion for comparing the correlation of temporal lotteries. A temporal lottery is more correlated than another if it displays a higher degree of (positive) correlation between consumption at two different time periods. Proposition 4 shows that increasing correlation increases (non-instrumental) informativeness in the Blackwell sense. This result formalizes the trade-off between intertemporal hedging and non-instrumental information.

The first main result, Theorem 1, shows a decision maker is always averse to increasing correlation (meaning that the hedging motive dominates) if and only if ϕ satisfies increasing relative risk aversion (IRRA). I further show that IRRA puts bounds to the demand for non-instrumental information. In other words, it sets a limit on how much a decision maker with these risk attitudes values non-instrumental information. Theorem 1 shows that this bound is strong enough so that the value of intertemporal hedging always dominates the value of non-instrumental information. Notably, IRRA includes the Epstein-Zin and Hansen-Sargent formulations of recursive preferences as special cases, which are common in applications.[3]

Under a mild strengthening of IRRA, Theorem 2 shows that recursive preferences admit a variational 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 $I^t_{\phi,u,\beta}(q||p)$ is a (generalized) statistical distance that measures the dissimilarity of any other distribution of future consumption q from p. A strand of the literature

[2]The work of Kreps and Porteus (1978) established that recursive preferences have a preference for non-instrumental information. Here we are referring to risk about consumption, and not about income. Therefore there is no planning advantage to tossing the coin early.

[3]Furthermore, IRRA is one of the most common assumptions on risk attitudes used in applications (e.g., see Arrow (1971), p. 96).

(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. The representation in (2) provides a connection between correlation aversion and robustness to model misspecification.[4] The 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, all other possible distributions q are considered plausible and evaluated depending on their dissimilarity from p as measured by $I^t_{\phi,u,\beta}(q\|p)$. The functions $(I^t_{\phi,u,\beta})_t$ can thus be equivalently interpreted as measures of correlation aversion or fear of model misspecification.[5]

When ϕ satisfies constant absolute risk aversion, $I^t_{\phi,u,\beta}(q\|p)$ is given by relative entropy (as shown by Strzalecki (2011)). Theorem 2 generalizes this connection with fear of model misspecification for a much broader class of preferences that satisfy correlation aversion. In the Epstein-Zin case $I^t_{\phi,u,\beta}(q\|p)$ 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 bounds imposed by IRRA to the demand for non-instrumental information are such that in certain domains of temporal lotteries there is no preference for non-instrumental information. In these domains, temporal lotteries can be ranked in terms of persistence, as is often the case in applications. Proposition 6 demonstrates that such domains can separate risk aversion from intertemporal substitution without a preference for non-instrumental information. However, I show that recursive preferences in (1) cannot disentangle risk aversion from correlation aversion, emphasizing the need for more general models of recursive preferences.

These results demonstrate that correlation aversion is the main driver in several applications of recursive preferences, such as asset pricing, climate policy, and fiscal policy.

The long-run risk model of Bansal and Yaron (2004) explains several puzzles in the asset pricing literature, including the equity premium puzzle. In this model, consumption growth contains a small, persistent component. This persistence amplifies the equity premium such that it matches the observed values in the data. Persistence

[4] To better appreciate the connection with fear of model misspecification, observe that gamble A is like flipping a biased coin with uncertainty about the bias, while for gamble B there is no such uncertainty. Preferring B to A reveals aversion to model misspecification.

[5] When the horizon is infinite, the cost function is unique.

implies a trade-off between intertemporal hedging and non-instrumental information about consumption growth for an investor with Epstein-Zin preferences. By Theorem 1, aversion to persistence entails a restriction on preferences for non-instrumental information. Therefore, the equity premium in this model is higher relative to the discounted expected utility benchmark because Epstein-Zin preferences satisfy IRRA and therefore strict correlation aversion.[6] In light of this fact, I re-examine Epstein et al. (2014)'s result which suggests that timing premia for the long-run risk model seem implausibly high based on introspection. I ask a different question: "What fraction of your wealth would you give up to remove all persistence in consumption growth?" I show that an investor would be willing to give up a share of his wealth which is not consistent with the experimental evidence. This issue arises from the inability of Epstein-Zin preferences to disentangle risk aversion from correlation aversion. Since correlation aversion is entirely determined by the level of risk aversion, an excessive amount of correlation aversion is required to match the observed equity premium. I briefly discuss a potential extension of Epstein-Zin preferences that would address this problem.

Long-run risk models have also been used in climate models to estimate the social cost of carbon (SCC), which is the marginal economic loss caused by an extra emission of carbon. Cai and Lontzek (2019) show that combining factor productivity growth that displays long-run risk with Epstein-Zin preferences substantially increases the SCC. My results imply that such estimates rely on correlation aversion.

In the context of optimal fiscal policy, Karantounias (2018) shows that a planner with Epstein-Zin preferences adopts a policy of fiscal hedging. This policy mitigates income shocks by taxing less during unfavorable conditions and more during favorable shocks. I illustrate how fiscal hedging is a direct consequence of recursive preferences that exhibit correlation aversion, but is unrelated to preferences for non-instrumental information.

Related literature. The literature on dynamic choice has considered a notion of correlation aversion derived from the literature on risk aversion with multiple com-

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

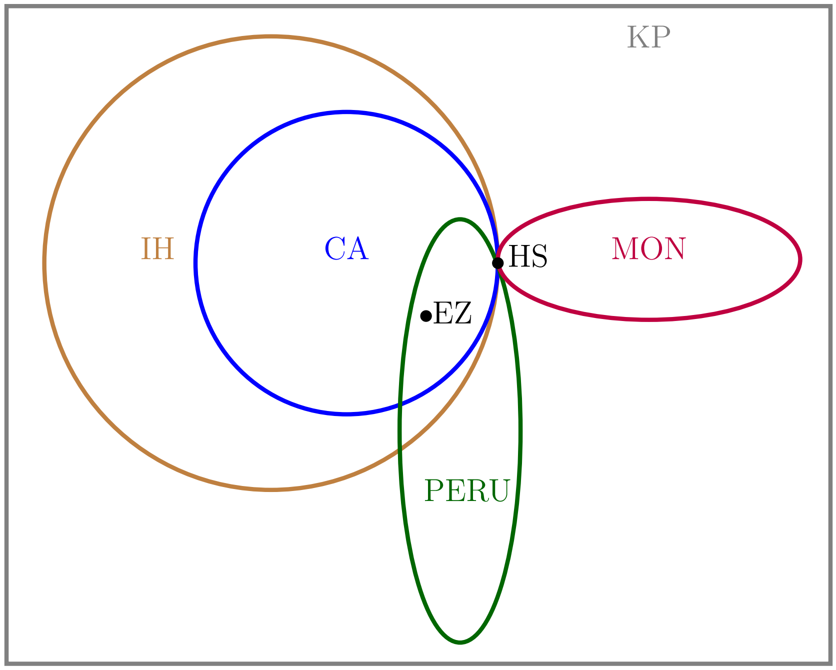

Figure 1: Relationship between correlation averse (CA) preferences and other recursive risk preferences: recursive preferences that satisfy intertemporal-hedging (IH), Epstein-Zin (EZ) preferences, multiplier-preferences (HS), monotone recursive preferences (MON), and preferences that exhibit a preference for early resolution of uncertainty (PERU). HS preferences are the only ones that exhibit all these features at the same time.

modities 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. Kochov (2015) and Bommier et al. (2019) study the extension to a purely subjective setting of this property, 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. An important consequence of the present paper is that within the Kreps-Porteus setting, multiplier preferences are the only 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 in more depth in Section 5.4.

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. This novel treatment is used in Section 3 to establish the main results related to correlation aversion. Section 4 examines the relationships among correlation aversion, risk aversion, and intertemporal substitution. The major implications of these results for the applied literature are examined in Section 5. Section 6 concludes the paper. The proofs are in the Appendix, which is completed by the Supplemental Appendix.

2 Preliminaries

Choice setting. I assume that time is discrete and varies over a finite horizon $2 \leq T < \infty$. The Supplemental Appendix (see Section S.1) discusses the case of an infinite horizon $T = \infty$. The consumption set C is assumed to be $C = [0, \infty)$.[7] 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 both are mixture spaces.

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

[7] 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.

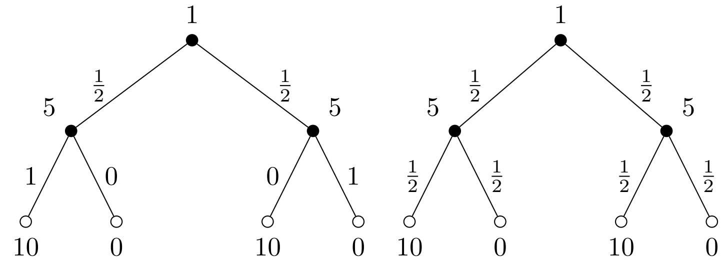

Figure 2: Probability tree representation of two temporal lotteries with $T = 2$

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 $\delta_x(A) = 0$ when $x \notin A$. 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 $i = 1, \dots, |\text{supp} m|$.

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 $t = 0, \dots, T-1$. Likewise, simple temporal lotteries are defined by $D_{T,s} := C$, $D_{t,s} := C \times \Delta_s(D_{t+1,s})$ for every $t = 0, \dots, T-1$. 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 $c^t = (c_0, \dots, c_{t-1}) \in C^t$ and $m \in \Delta_b(D_t)$, the temporal lottery $(c_0, (c_1, (c_2, (\dots, (c_{t-1}, m))))) \in D_0$ 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 set $A \subseteq X$. Finally, a function $I : X \times X \to [0, \infty]$ is a generalized (statistical) 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).

Definition 1. *Preferences $\succeq$ admit a Kreps-Porteus (KP) recursive representation (ϕ, u, β) if for $t = 0, \ldots, T$, each $\succeq_t$ is represented by $V_t : D_t \to \mathbb{R}$ such that $V_T(c) = u(c)$ for every $c \in C$ and recursively*

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

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. Because u is unbounded above, one can set without loss of generality $u(C) = [0, \infty)$.

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 $x \in u(C)$.[8]

I will typically consider KP representations with ϕ that is concave and satisfies certain differentiability assumptions to employ standard tools from the theory of risk aversion. Write $\phi \in \mathcal{C}^r$ if ϕ is continuous and has r continuous derivatives. Given $\phi \in \mathcal{C}^2$, the Arrow-Pratt index $A_\phi : \operatorname{int} u(C) \to \mathbb{R}$ is given by

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

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.

[8]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.

Decreasing (DRRA), increasing (IRRA), and constant (CRRA) relative risk averse functions are defined analogously by replacing the index A_ϕ with R_ϕ.

2.1 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 permits 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 there exists $t \leq T-2$, $c^t \in C^t$ and $m, m' \in \Delta_s(C \times \Delta_s(D_{t+1,s}))$ such that $d = (c^t, m)$, $d' = (c^t, m')$, $\operatorname{marg}_C m' = \operatorname{marg}_C m$ 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 (i.e., each row of G forms 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 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'.

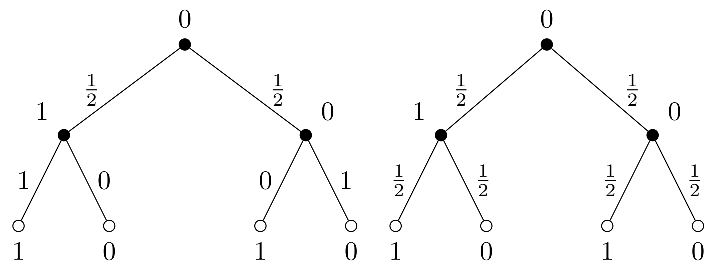

Figure 3: Probability tree representation of a temporal lottery

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

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

so that $d \geq_B d'$. In words, 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 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 for early resolution of uncertainty) over a domain of temporal lotteries is defined as monotonicity with respect to the order $\geq_B$.

Axiom 1. $\succeq$ exhibits a preference for information over $B \subseteq D_0$ 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 $t = 1$ is deterministic.[9] In this way consumption at $t = 1$ is informative but not correlated with consumption at $t = 2$. Formally, in this case the set B is given by

$$B := \left\{\left(c^t, m\right) \in D_{0,s} : \text{ there exists } \bar{c} \in C \text{ such that } \operatorname{marg}_C m = \delta(\bar{c})\right\}. \quad (4)$$

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

Observe that for these temporal lotteries, consumption at time $t+1$ is deterministic at level $\bar{c}$, which implies that it is uncorrelated with consumption in the following periods. The next result characterizes preferences valuing non-instrumental information over B.

Proposition 1. *Assume $\succeq$ admits a KP representation (ϕ, u, β) with $\phi \in \mathcal{C}^2$. 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 $x, y \in \operatorname{int} u(C)$.[10]

Proof. See the Appendix. □

The quantity

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

can be considered as a local measure of strength of preferences for non-instrumental information. Here I focus on risk attitudes that exhibit a preference for information regardless of the level of impatience or intertemporal substitution.

Definition 4. *Say that ϕ satisfies a uniform preference for information (UPI) if every $\succeq$ with KP representation (ϕ, u, β) exhibits a preference for information.*

The next simple result provides a connection between classical risk attitudes and preference for information.

Proposition 2. *If $\phi \in \mathcal{C}^2$ satisfies UPI then it also satisfies DARA.*

Proof. Immediate from (5). □

3 Main results: correlation aversion

I introduce a general notion of an increase in positive correlation between consumption at two distinct periods. I then characterize recursive preferences that are averse to correlation. 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.

[10]Condition (5) is due to Strzalecki (2013) (see p. 1051).

3.1 The case $T = 2$

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 previous 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 $m_1(c_1 \leq \cdot)$, $m_2(c_2 \leq \cdot | c_1 \leq \cdot)$.

These lotteries can be ordered based on their correlation using the following class of transformations.

Definition 5. *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*

$$\begin{aligned} m_2(c|c) &= m_2'(c|c) + \frac{\varepsilon}{m_1'(c)}, \\ m_2(c'|c) &= m_2'(c'|c) - \frac{\varepsilon}{m_1'(c)}, \\ m_2(c'|c') &= m_2'(c'|c') + \frac{\varepsilon}{m_1'(c')}, \\ m_2(c|c') &= m_2'(c|c') - \frac{\varepsilon}{m_1'(c')}, \end{aligned}$$

and $m_2 = m_2'$ otherwise.

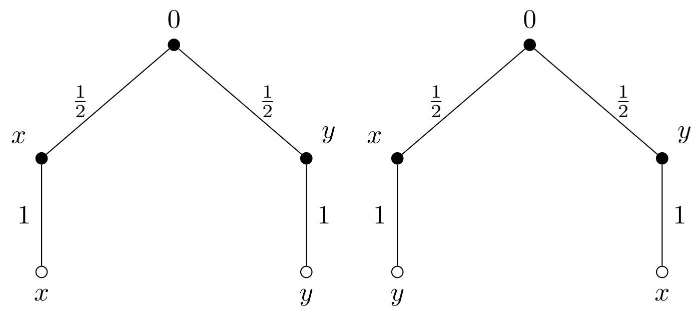

Figure 4: Negative vs positive correlation

In simpler terms, these transformations increase the probability that, if consumption at $t = 1$ is either c or c', it will remain the same at $t = 2$, and, concurrently, decrease the probability that consumption will shift to a different level. 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 d can be obtained from the "iid" lottery d' by means of an IECIT. In this case, an IECIT also 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 for some $x, y \in C$ we have $m_1(x)' = m_1(x) = \frac{1}{2}$, $m_2(x|x) = m_2(y|y) = 1$, and $m_2'(y|x) = m_2'(x|y) = 1$. 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 their being equally informative only hedging considerations affect the evaluations of these two lotteries. Proposition 1 in the Supplemental Appendix demonstrates that intertemporal hedging is equivalent to the concavity of ϕ (i.e., risk aversion).

The concept of an IECIT 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 6. *Given $d, d' \in D^*_{0,s}$ say that d is more informative than d', denoted $d \geq_C d'$, 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.

Proposition 3. *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 3 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 7. *Given $d = (c, m), d' = (c, m') \in D^*_0$, say that d is more correlated than d', written $d \geq_C d'$, 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 $d_n \geq_C d'_n$ 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 4. *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. We can define correlation aversion as aversion towards increasing correlation to an "iid" temporal lottery.

Axiom 2. $\succeq$ exhibits correlation aversion 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 there is a trade-off between intertemporal hedging and non-instrumental information.

Theorem 1. *Consider $\phi \in \mathcal{C}^2$ that is concave 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. □

In words, if ϕ satisfies IRRA, then any KP preference (ϕ, u, β) will exhibit correlation aversion; conversely, if ϕ is such that every KP preference (ϕ, u, β) exhibits correlation aversion, then ϕ must satisfy IRRA.

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 limits preferences for information. Indeed, observe that IRRA means that R_ϕ is non-decreasing. Therefore, when R_ϕ is differentiable we have:

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

for every $x \neq 0$. Under DARA it holds $A'_\phi \leq 0$, so that we obtain

$$A'_\phi(x) \in \left[-\frac{A_\phi(x)}{x}, 0\right].$$

This means that IRRA limits the reduction of absolute risk aversion for a given increase in utility. When β is close to unity, IRRA effectively imposes a lower bound on $\frac{\phi''(x)}{\phi'(x)} - \beta\frac{\phi''(\beta x+y)}{\phi'(\beta x+y)}$. Hence, IRRA introduces a constraint on preferences for information. Finally, observe that as a consequence of Theorem 1, indifference to correlation occurs only under risk neutrality, i.e. $\phi(x) = x$.

The following condition strengthens IRRA by requiring that the index of relative risk aversion increases sufficiently rapidly.

Definition 8. *Say that $\phi \in \mathcal{C}^2$ satisfies strong correlation aversion (SCA) if it satisfies IRRA and $R''_\phi(x) \geq 0$ for every $x \in (0, \infty)$.*

Thus, SCA requires not only that the index of relative risk aversion R_ϕ is increasing, but also that it increases at a sufficiently fast pace. Observe that both EZ and HS preferences satisfy this condition. The next key result formalizes the connection between robustness to model misspecification and correlation aversion.

Theorem 2. *Assume that $\succeq$ admits a KP representation (ϕ, u, β) with $\phi \in \mathcal{C}^4$ that is concave and satisfies UPI. If ϕ satisfies SCA, then $\succeq$ admits the recursive representation $(V_t)_{t=0}^2$ given by $V_2(c) = u(c)$ and*

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

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

Proof. See the Appendix. □

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 $I^t_{\phi,u,\beta}$.

The quantity $I^t_{\phi,u,\beta}(\ell \| m)$ measures the cost paid when considering the alternative distribution ℓ. When $\phi(x) = -e^{-\frac{x}{\theta}}$, $I^t_{\phi,u,\beta}$ is given by Relative Entropy (see Strzalecki (2011)), that is

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

when $\ell \ll m$ and $I^t_{\phi,u,\beta}(\ell \| m) = \infty$ otherwise.[11] Theorem 2 implies that this interpretation in 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 interpreted as a 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.[12]

[11] See for example Strzalecki (2011).

[12] The general formulation of each $I^t_{\phi,u,\beta}$ is discussed in the Appendix. The multiplicity of cost functions arises because the setting is not fully stationary when there is a finite horizon. In the Supplemental Appendix, I show that there is a unique cost function when the horizon is infinite.

To illustrate, consider the common parametrization of Epstein-Zin used in asset pricing with an 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

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

and $I^t_{\phi,u,\beta}(\ell\|m) = \infty$ otherwise, 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 depends also on the continuation utility V_{t+1}. In particular, temporal lotteries with higher expected continuation utility will be more costly to consider.

3.2 The case $T < \infty$

The previous results can be 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 S.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 $d \sim_t (c, 0, \ldots, 0)$. Now observe that every $m \in \Delta_b(C \times \Delta_b(D_{t+1}))$ and $\succeq$ with KP representation (ϕ, u, β) induces 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 $\ell_\succeq(A) = \ell(\{d \in D_{t+1} : PE_{\succeq_t}(d) \in A\})$.[13] 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.

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

Definition 9. *Consider $d = (c, m), d' = (c, m') \in D_0$. Say that d is more correlated than d', written $d \geq_C d'$, 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 $m_{\succeq}(A \times B) = \ell(A)$ whenever $\ell \in B$ and $m_{\succeq}(A \times B) = 0$ otherwise.

Axiom 3. $\succeq$ exhibits correlation aversion if 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 1 and 2 to this setting with an arbitrary finite horizon.

Theorem 3. *Consider $\phi \in \mathcal{C}^2$ that is concave and satisfies UPI. Then every $\succeq$ with KP representation (ϕ, u, β) exhibits correlation aversion if and only if ϕ satisfies IRRA. Further, if $\succeq$ that admits a KP representation (ϕ, u, β) with $\phi \in \mathcal{C}^4$ that satisfies SCA, then $\succeq$ admits the recursive representation $(V_t)_{t=0}^T$ given by $V_T(c) = u(c)$ and*

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

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

Proof. The proof follows the same steps as the proof of Theorems 1 and 2 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 $c_0 \in C$, $\ell \in \Delta_s(C)$ and $\varepsilon \in [0, 1]$, define $d_{\varepsilon, c_0}(\ell)$ recursively as follows: $d_{T-1,c,\varepsilon}(\ell) = (c, m_{T-1,c,\varepsilon}(\ell))$ where $m_{T-1,c,\varepsilon} \in \Delta_s(C)$ satisfies

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

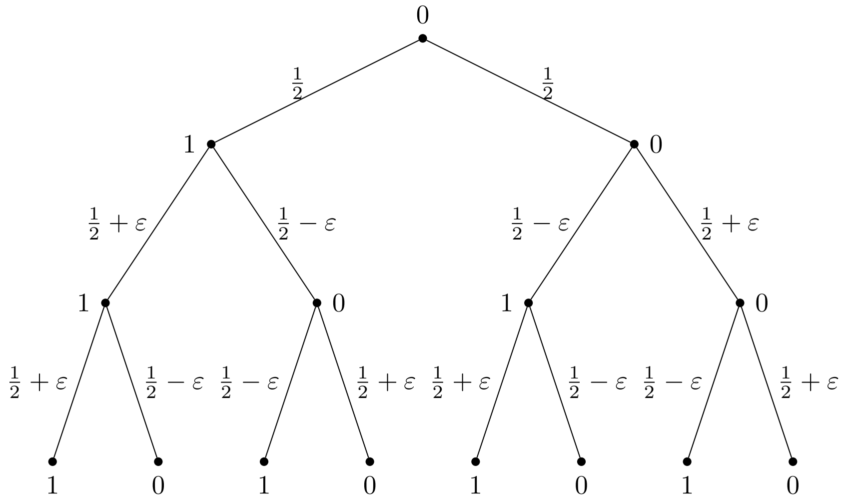

Figure 5: Example of $d_{\varepsilon,0}(\ell)$ with $T=3$ and $\ell(1)=\ell(0)=\frac{1}{2}$

and recursively for $2 \leq t \leq T-2$ define $d_{t-1,c,\varepsilon}(\ell) = (c, m_{t-1,c,\varepsilon}(\ell))$ by

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

and finally set $d_{\varepsilon,c_0}(\ell) = (c_0, m_{1,\varepsilon})(\ell)$ 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 5. *Consider $\phi \in \mathcal{C}^2$ that is concave and satisfies both IRRA and UPI. Consider $\succeq$ with KP representation (ϕ, u, β). Then for every $\ell \in \Delta_s(C)$ and $c_0 \in C, \varepsilon, \varepsilon' \in [0,1]$*

$$\varepsilon \geq \varepsilon' \implies d_{\varepsilon',c_0}(\ell) \succeq_0 d_{\varepsilon,c_0}(\ell).$$

Proof. See the Appendix. □

Increasing the value of ε involves again a trade-off between non-instrumental information aversion and intertemporal hedging.[14] Under the interpretation that $\varepsilon \in [0,1]$

[14] Again, it holds that $d_{\varepsilon,c_0}(\ell) \geq_B d_{\varepsilon',c_0}(\ell)$ just like in Proposition 4. The proof is omitted for brevity.

models the persistent component in consumption, the above result establishes a connection between IRRA (and therefore correlation aversion) and aversion to 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 disentangling these two is that it leads to preferences for non-instrumental information in certain domains such as B 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

$$\mathcal{R} = \{d_{\varepsilon,c_0}(\ell) : c_0 \in C, \ell \in \Delta_s(C), \varepsilon \in [0,1]\} \cup C^T.$$

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 in the entire class of temporal lotteries D_0.

Definition 10. *Say that $\succeq^1$ is more risk averse than $\succeq^2$ 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

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

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

Proposition 6. *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 $A_{\phi_1} \geq A_{\phi_2}$.*

Proof. See the Appendix. $\square$

Now consider preferences $\succeq$ with KP representation (ϕ, u, β), where ϕ satisfies IRRA. These preferences do not exhibit a preference for information over $\mathcal{R}$, since by Proposition 5 it holds that $d_{\varepsilon',c_0}(\ell) \succeq_0 d_{\varepsilon,c_0}(\ell)$ when $\varepsilon \geq \varepsilon'$. Moreover, the preference $d_{\varepsilon',c_0}(\ell) \succeq_0 d_{\varepsilon,c_0}(\ell)$ when $\varepsilon \geq \varepsilon'$ is the same exhibited by preferences that are correlation averse but indifferent to non-instrumental information. By Proposition 1, these are characterized by $\beta = 1$ and $\phi(x) = -\exp\left(-\frac{x}{\theta}\right)$. Thus, under the empirically relevant restriction on risk attitudes of IRRA, the consumption domain $\mathcal{R}$ is able to distinguish between risk aversion and intertemporal substitution without a preference for non-instrumental information.

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 2, 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 fully disentangled. For example, for EZ and HS preferences, correlation aversion and risk aversion are both determined by the parameters α and θ, respectively.

This last point does not hold for other classes of recursive preferences, such as Epstein-Uzawa. These preferences admit the recursive representation $V_T(c) = u(c)$ and

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

for some continuous functions $u : X \to \mathbb{R}$ and $b : X \to (0, 1)$. Epstein-Uzawa 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

This section discusses the implications of the previous results, emphasizing the importance of correlation aversion in applications of recursive preferences such as asset pricing, climate policy, and fiscal policy.

5.1 Asset pricing and long-run risk

The long-run risk model of Bansal and Yaron (2004) is a cornerstone in the consumption-based asset pricing literature for its ability to account for a wide range of asset pricing puzzles. This model relies on Epstein and Zin's preferences and a consumption process (case I) that satisfies for $t = 0, \ldots$

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

where $d_t := \log\left(\frac{c_{t+1}}{c_t}\right)$ denotes consumption growth.

Such a model faces the trade-off discussed in previous sections.[15] An investor with recursive preferences values both early resolution of uncertainty and intertemporal hedging, with intertemporal hedging being more valuable 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. However, some limitations of the model stem from

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

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

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)

the fact that the key feature of preferences–correlation aversion–is based solely on risk attitudes, which I will discuss next.

How much would you pay to remove long-run risk? Epstein et al. (2014) suggest 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 (6) with $a = 0$ (no persistence) and $a = 0.9790$, respectively.

Table 1 summarizes the parameters of the model.[16] 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 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

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

finding implies that either the level of persistence is unrealistic or the degree of risk aversion is too elevated. A potential solution 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.

5.2 Climate policy

Cai and Lontzek (2019) develop a dynamic stochastic general equilibrium model to estimate the effect of economic and climate risks on the social cost of carbon (SCC). They consider productivity shocks that exhibit persistence, leading to consumption growth rates that display long-run risk as in (6). Combined with Epstein-Zin preferences, the inclusion of persistent productivity shocks results in substantially higher social cost of carbon compared to scenarios without productivity shocks (see pp. 2705-2706 in Cai and Lontzek (2019)). My analysis implies that these estimates rely on the fact that Epstein-Zin preferences exhibit correlation aversion, which is the key behavioral feature that amplifies the SCC. Importantly, these estimates are not driven by preferences for non-instrumental or irrelevant information. This distinction improves the credibility of climate models.

5.3 Utility smoothing and fiscal hedging

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).

Correlation aversion is a behavioral formulation of this property. Consider again gambles A and B described in the introduction. In gamble B, at $t = 0$ future utility is constant, while for gamble A future utility is volatile. Thus, preferring B to A indicates aversion to volatility in future utility.

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 Theo-

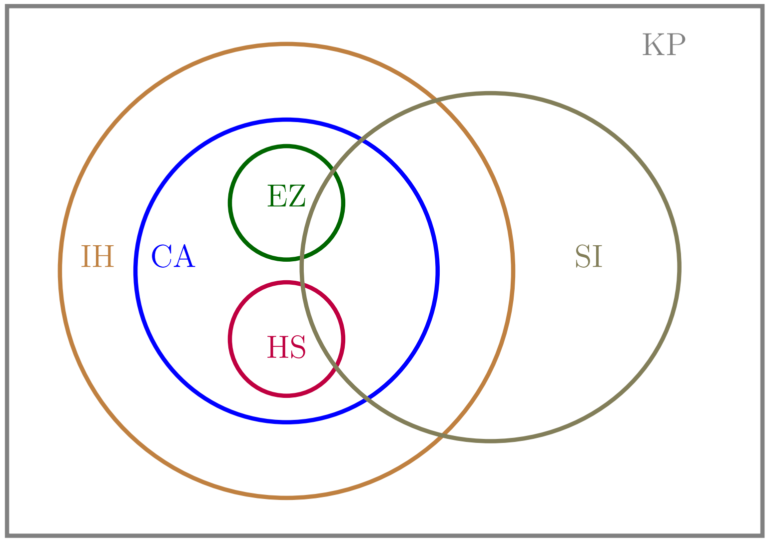

Figure 6: Relationship between correlation averse (CA) preferences and recursive preferences that satisfy intertemporal-hedging (IH), Epstein-Zin (EZ) preferences, and stochastic impatience (SI)

rems 1 and 2, 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.

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 general 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, preferences value intertemporal hedging 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: model misspecification, asset pricing, climate policy, and optimal taxation. 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 Proposition 1

Outline of the proof. The key idea behind the proof is to demonstrate, using established results from information economics (e.g., 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

$$U_t(m) = \phi\left(u(\bar{c}) + \beta\phi^{-1}\left(\mathbb{E}_m\phi(V_{t+1})\right)\right) \quad \text{for every } m \in \Delta_s(D_{t+1,s}), \tag{7}$$

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

Lemma 1. *Each U_t defined in (7) is convex if and only if (5) holds.*

Proof. First I claim that each U_t defined in (7) is convex if and only if the function $\Phi : \phi(u(C)) \to \mathbb{R}$ defined by $x \mapsto \phi(\bar{c} + \beta\phi^{-1}(x))$ is convex. To see this point, observe that for every $\bar{c} \in C$ we have that

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

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). $\square$

Proof of Proposition 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

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

and

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

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 3

Proof. Take $d, d' \in D^*_{0,s}$. Without loss of generality we can assume m_1, m'_1 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 $g_{ij} = m_2(c_i|c_j)m_1(c_i), f_{ij} = m'_2(c_i|c_j)m'_1(c_i)$ 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

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

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

7.1.3 Proof of Proposition 4

Proof. Denote with $\{c, c', \ldots, c_N\}$ 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' differs from $d^{iid}(\ell)$ by an IECIT. Then for some $\varepsilon \geq 0$ and (c, c') it holds that

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

for some $x \in [0,1]$, so that $d' \geq_B d^{iid}(\ell)$. Using the same reasoning it is immediately shown that $d \geq_B d'$. $\square$

7.1.4 Proof of Theorem 1

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

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

The statement can be extended to arbitrary elements in D_0 by means of continuity of preferences.[17] I provide 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 weakly 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 Section S.2 in the Supplemental Appendix. □

It is now possible to prove Theorem 1. 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 1. 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 $U(0) = V_0(d^{iid}(\ell))$, $U(q_1) = V_0(d')$, $U(q_2) = V_0(d)$, $U(1) = V_0(d^{corr}(\ell))$, $\lim_{\varepsilon \to 0} U'(\varepsilon) \leq 0$, and $U'(\varepsilon) \geq 0$ for every $\varepsilon \in (0,1)$ (where derivatives are intended in the weak sense, see Section 8.2 in Brezis (2010)).[18] I claim that it also holds that

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

[17] 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 $\lim_n d_n = d$, $\lim_n d'_n = d'$, $\lim_n d^{iid}(\ell_n) = d^{iid}(\ell)$ 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 of preferences in the weak* topology.

[18] 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 find $U : [0,1] \to \mathbb{R}$ that is continuous and weakly differentiable by constructing $(U_i)_{i=1}^{N}$ using Lemma 2 and setting $U(x) = U_i\left(\frac{Nx}{i}\right)$ for $x \in \left[\frac{i-1}{N}, \frac{i}{N}\right)$, $i = 1, \ldots, N-1$, and $U(x) = U_N(x)$ for $x \in \left[\frac{N-1}{N}, 1\right]$.

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

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

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

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

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

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

and

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

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

Conversely, assume that ϕ does not satisfy IRRA. Then there exists $\underline{z} < \bar{z}$ in $\operatorname{int} u(C)$ 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

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

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

7.1.5 Proof of Theorem 2

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.[19] This result allows us to utilize the Fenchel-Moreau duality theorem, revealing that the certainty equivalent can be represented dually as $\phi^{-1}\left(\mathbb{E}_m \phi\left(V_{t+1}\right)\right) = \min_\ell \mathbb{E}_\ell V_{t+1} + I^t_{\phi,u,\beta}(\ell \| m)$.

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 Dunford and Schwartz (1988), 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 $\nu \mapsto \frac{d\nu}{dp}$).

Turning to the proof of Theorem 1, 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 concave conjugate $M^*_{\phi,p}$ of the quasi-arithmetic mean $M_{\phi,p}$.

[19] 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.

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 concave 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 concave 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,

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

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 $L^\infty_+(p) := \{\xi \in L^\infty(p) : \xi \geq 0\}$ be the non-negative orthant of $L^\infty(p)$.

Theorem 4 (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 $\frac{1}{A_{\phi|(0,\infty)}}$ is concave.*

Proof. If A_ϕ is convex, 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 Theorem 1 and Theorem 5 in Chudziak et al. (2019) to show that $M_{\phi,p}|L^\infty_{s,+}(p)$ is concave. Concavity 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 4, 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 $\phi \in \mathcal{C}^4$ is concave and satisfies UPI over $(0, \infty)$. Then if R_ϕ is convex $M_{\phi,p}|L^\infty_+(p)$ is concave.*

Proof. First observe that if ϕ satisfies UPI, then by DARA we have $A'_\phi \leq 0$. Further, it is immediate to see that $\frac{1}{A_\phi}$ is concave whenever

$$A''_\phi(x)A_\phi(x) \geq 2(A'_\phi(x))^2,$$

for every $x \in (0, \infty)$. This condition is equivalent to

$$xA''_\phi(x) \geq 2x\frac{(A'_\phi(x))^2}{A_\phi(x)}, \tag{8}$$

for every $x \in (0, \infty)$. Since $R'_\phi \geq 0$, we obtain that for every $x \in (0, \infty)$ it holds that

$$A(x) \geq -xA'_\phi(x).$$

From this last condition we obtain

$$-2A'_\phi(x) \geq 2x\frac{(A'_\phi(x))^2}{A_\phi(x)}. \tag{9}$$

Therefore since

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

if $R''_\phi \geq 0$ it follows that $xA''_\phi(x) \geq -2A'_\phi(x)$ which by (9) implies that (8) is satisfied. Hence we conclude that if ϕ satisfies SCA then $\frac{1}{A_\phi}$ is concave. The result therefore follows by Theorem 4. $\square$

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

It is important to observe that both EZ and HS preferences satisfy SCA.

Corollary 2. *Assume that ϕ is given by $\phi(x) = \frac{x^\lambda}{\lambda}$ for $\lambda < 1$ or $\phi(x) = -e^{-\frac{x}{\theta}}$ with $\theta > 0$ for every $x \in \mathbb{R}_+$. Then $M_{\hat{\phi},p}$ is concave.*

Proof. Immediate from Theorem 4. □

It is now possible to deliver a proof of Theorem 2.

Proof of Theorem 2. Given $(V_t)_{t=0}^T$ 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 $V_t : D_t \to \mathbb{R}$, $t = 0, \dots, T$ 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 $t = 0, \dots, T-1$. 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

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

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

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

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 $t = 0, \dots, T-1$ we set

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

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 $t = 0, \ldots, T-1$. Finally, observe that each $I^t_{\phi,u,\beta}$ is a premetric or 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

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

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 $q = \frac{\alpha}{\alpha-\rho}$,

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

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

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

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 5

The proof is a straightforward consequence of Theorem 1. Define $U : [0,1] \to \mathbb{R}$ by

$$U(\tilde{\varepsilon}) = \sum_{c\in C}\ell(c)\phi(c + \beta\phi^{-1}(\mathbb{E}_{m_{2,c,\varepsilon}}\phi(V_2)),$$

for every $\tilde{\varepsilon} \in [0,1]$ and observe $d_{\varepsilon',c_0}(\ell) \succeq_0 d_{\varepsilon,c_0}(\ell)$ if and only if $U(\varepsilon') \geq U(\varepsilon)$. Notice that by UPI the function U satisfies $U''(\tilde{\varepsilon}) \geq 0$ for every $\tilde{\varepsilon} \in [0,1]$. Further observe that by concavity of ϕ we obtain $\lim_{\varepsilon\to 0} U'(\varepsilon) \leq 0$ and by IRRA we obtain

$\lim_{\varepsilon \to 1} U'(\varepsilon) \le 0$.[20] Hence for every $\varepsilon \ge \varepsilon'$ we obtain that

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

which implies $d_{\varepsilon', c_0}(\ell) \succeq_0 d_{\varepsilon, c_0}(\ell)$ as desired.

7.1.7 Proof of Proposition 6

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 $\bigoplus_{i=1}^n \pi_i (c_i, \dots, c_i)$ for every probability vector $(\pi_i)_{i=1}^n$ and constant consumption streams $(c_i, \dots, c_i) \in C^T$ to obtain that $A_{\phi_1} \ge A_{\phi_2}$. The converse follows by Jensen's inequality.

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 $\ell = V_{\#} \bigoplus_{i=1}^n \pi_i \bar{c}_i$.[21] Since $\succeq_1$ is more risk averse than $\succeq_2$, it follows that since $\left(c_0, \bigoplus_{i=1}^n \pi_i(\bar{c}_i, \dots, \bar{c}_i)\right), (c_0, \mathbf{0}) \in \mathcal{R}$

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

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

[20] The explicit calculations for these results are not presented here, as they mirror those provided in the proof of Theorem 1.

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

7.2 The persistence premium

7.2.1 Long-run risk

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

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

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 in the literature that quantifies the persistence premium. 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

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

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

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

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

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.

Dunford, N. and Schwartz, J. T. (1988). *Linear operators, part 1: general theory*, volume 7. John Wiley & Sons.

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.