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

Lorenzo Stanca†

August 15, 2025

Abstract

This paper investigates a novel behavioral feature of recursive preferences: aversion to risks that persist over time, or simply *correlation aversion*. Greater persistence provides information about future consumption but reduces opportunities to hedge consumption risk. I show that, for recursive preferences that exhibit a preference for early resolution of uncertainty, correlation aversion is equivalent to increasing relative risk aversion. To quantify correlation aversion, I develop the concept of the persistence premium, which measures how much an individual is willing to pay to eliminate persistence in consumption. I provide an approximation of the persistence premium in the spirit of Arrow–Pratt, which provides a quantitative representation of the trade-off between information and hedging. I show that correlation-averse preferences have a variational representation, linking correlation aversion to concerns about model misspecification. I present several applications. I first illustrate how correlation aversion shapes portfolio choices, and then show how the persistence premium can improve the calibration of macro-finance models. In an optimal taxation model, I show that recursive preferences—unlike standard preferences—lead to redistributive tax policies that increase social mobility.

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

JEL classification: C61, D81.

*This paper is a revised and extended version of Chapter 1 of my dissertation at Northwestern University; several results also appeared in my job market paper which shared the same title. I thank four anonymous referees for their constructive comments and suggestions. I am grateful to Nabil Al-Najjar, Peter Klibanoff, and Marciano Siniscalchi for their help and guidance. I thank Antoine Bommier, Francesco Fabbri, Veronica Merlone, Ignacio Monzón, Daniele Pennesi, and Giulio Principi for their feedback. Finally, I also thank several anonymous referees whose comments improved the quality of the paper.

†ESOMAS Department and Collegio Carlo Alberto, Università di Torino, 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, including models of consumption-based asset pricing (Epstein and Zin, 1989, 1991), precautionary savings (Weil, 1989; Hansen et al., 1999), business cycles (Tallarini, 2000), progressive taxation and inequality (Benabou, 2002), and risk-sharing (Epstein, 2001; Anderson, 2005). Recursive preferences have also been applied to climate change (Bansal et al., 2017; Cai and Lontzek, 2019), optimal fiscal policy (Karantounias, 2018), and repeated games (Kochov and Song, 2023).

A key feature of recursive preferences is their ability to distinguish between risk aversion and intertemporal substitution—two important preference parameters that, for both theoretical and empirical reasons, should be disentangled. In this paper, I show that recursive preferences also exhibit sensitivity to a third behavioral trait: *aversion to risks that persist over time*. This trait, which has received less attention in the past, plays a significant role in many economic applications.

I introduce a new axiom, which I call *correlation aversion*, that captures a preference for avoiding risks that are correlated over time. Greater persistence in risks reduces the ability to hedge, which a risk averse decision maker (DM) dislikes. However, increased correlation also means greater informativeness about future consumption. This fact creates a trade-off between *hedging* and *information*, which is central to this paper.

To illustrate, compare two gambles: in gamble A, a single coin flip at $t = 1$ determines all future consumption (all 1's or all 0's), whereas in gamble B a fair coin is tossed each period (yielding 1 or 0 each time). A hedging motive suggests a preference for B over A, but A resolves all risk at $t = 1$, providing more non-instrumental information about future consumption.[1] Kreps and Porteus (1978) show that recursive preferences imply a preference for early resolution of uncertainty, or equivalently, non-instrumental information. Hence, the comparison is not straightforward: A resolves risk early, while B is preferred for its hedging value. My notion of correlation aversion requires the hedging motive to outweigh the preference for non-instrumental information, leading a DM to prefer B over A.

Correlation aversion, therefore, should presumably be reflected in a DM's prefer-

[1] Here I refer to risk concerning consumption, not income; hence, early coin tosses offer no planning advantage—information is non-instrumental.

ences through a limited willingness to pay for *non-instrumental information* about future consumption. I characterize correlation aversion for recursive preferences that exhibit a preference for early resolution of uncertainty, showing that it is equivalent to increasing relative risk aversion (IRRA). Consistent with the intuition, IRRA imposes bounds on the demand for non-instrumental information, and a mild strengthening of IRRA guarantees that recursive preferences admit a representation reflecting robustness to model misspecification.

Further, I introduce the *persistence premium*, a measure of correlation aversion that reflects the willingness to pay to eliminate risk persistence. I derive an approximation of this premium in the spirit of Arrow-Pratt that links correlation aversion to preference parameters such as risk aversion and preference for non-instrumental information.

I then apply these results to asset pricing and income taxation. I illustrate how correlation aversion shapes portfolio choices and asset prices and how the persistence premium improves macro-finance model calibrations. Finally, I show how correlation-averse preferences shape progressive tax structures by increasing redistribution in a way that promotes social mobility.

Preview of results

I consider recursive preferences characterized by three components (ϕ, u, β): ϕ reflects risk attitudes, u determines elasticity of intertemporal substitution (EIS), and β captures time preference.

First, I re-frame preferences for early resolution of uncertainty in terms of the Blackwell order of informativeness (Blackwell, 1951). A preference for early resolution of uncertainty implies the DM prefers more informative lotteries (see the discussion after Definition 3). Propositions 1 and 2 connect preferences for early resolution of uncertainty with decreasing absolute risk aversion.

I then use this result to introduce a measure of attitudes toward early resolution of uncertainty, which I refer to as ER_ϕ (see equation 4 and Appendix A.2). I develop a foundation of this measure by providing an approximation of the early resolution premium, which quantifies the willingness to pay to have uncertainty resolve early. This approximation shows that the premium depends positively on ER_ϕ (see Corollary 2).

Next, I introduce a novel definition of correlation aversion. More correlation adds informativeness in the Blackwell sense (Proposition 3), creating conflicting incentives for a DM with recursive preferences, who prefers less persistent lotteries but values the information. This result formalizes the key trade-off explored in this paper between intertemporal hedging and non-instrumental information, as illustrated in the initial example: A is more informative than B, but B provides better hedging value.

The main result, Theorem 1, shows that for a DM who prefers early resolution for every possible value of β, correlation aversion is equivalent to ϕ satisfying IRRA. I further show that IRRA limits how much a DM values non-instrumental information (see equation 5). Notably, IRRA encompasses common recursive utility models like Epstein-Zin.

To measure correlation aversion, I introduce the *persistence premium*, which quantifies how much a DM is willing to pay to eliminate consumption persistence (see equation 6). Using an approximation à la Arrow-Pratt I find a formula that connects correlation aversion with risk aversion, EIS, persistence, and preference for information.

To illustrate, in the Epstein–Zin case—where $1-\alpha$ is the coefficient of relative risk aversion, $\frac{1}{1-\rho}$ is the elasticity of intertemporal substitution, and $\varepsilon \in [0,1]$ measures consumption persistence—the premium is approximately given by

$$\tilde{a} + \tilde{b}\,\varepsilon\Big(1 - \frac{\alpha}{\rho}\Big)\Big(\frac{1}{c_H} + \frac{1}{c_L}\Big) \;-\; \tilde{c}\,(\varepsilon^2 - 1)\,ER_\phi\;,$$

where $\tilde{a}, \tilde{b}, \tilde{c} > 0$ and consumption can be either high (c_H) or low (c_L). Hence, the premium rises with consumption persistence but at a decreasing rate. In particular, higher risk aversion makes the premium increase more rapidly as ε grows, while a higher ER_ϕ moderates this increase. Under IRRA, this result generalizes (see Corollary 1 and the related discussion), formalizing the quantitative trade-off between information and hedging.

To further clarify the role of risk attitudes, Theorem 2 shows that a mild strengthening of IRRA implies that recursive preferences admit a representation reflecting robustness to model misspecification—a concern that future consumption distributions may be wrong. This result extends earlier work (e.g., Hansen et al. (1999)) that links multiplier preferences to robustness, generalizing it to all correlation-averse preferences.

I then provide applications of these results.

Asset pricing. These results imply that correlation aversion significantly influences portfolio choices at both individual investor and macroeconomic levels. At the investor level, I discuss how risk preferences affect investment strategies. All other things being equal, investors characterized by high relative risk aversion and a low preference for information are more likely to favor bonds over stocks. Conversely, investors who exhibit a strong preference for information relative to their degree of risk aversion will find stocks more attractive due to the news they provide about long-run consumption growth.

At the macro-financial level, correlation aversion helps explain the equity premium puzzle—the high observed excess returns investors require for holding equities. The long-run risk model of Bansal and Yaron (2004), combined with Epstein-Zin preferences and persistent consumption growth, successfully matches the observed equity premium. Theorem 1 implies that the equity premium is higher under Epstein-Zin preferences because they exhibit IRRA and therefore correlation aversion, but *despite* their preference for non-instrumental information.

This analysis highlights the need to calibrate preference parameters to achieve a reasonable level of correlation aversion. Drawing inspiration from Epstein et al. (2014), I introduce an analogue of the persistence premium in this macro-finance setting which asks: "What fraction of your wealth would you sacrifice to eliminate all persistence in consumption growth?" Using existing experimental evidence, I show that—under standard parameters commonly used in the literature—the persistence premium proves unreasonably high (see Section 4.1).

This issue arises because Epstein-Zin preferences do not fully disentangle risk aversion from preferences for early resolution. To address this evidence, I explore a generalization of Epstein-Zin preferences to hyperbolic absolute risk aversion (HARA), which partially separates these preference parameters (see equation (11)). I show that these preferences can exhibit a level of correlation aversion comparable to the standard Epstein-Zin parametrization used in the literature, but with a lower level of risk aversion that is more consistent with empirical evidence. Hence, this new model suggests that one can explain investors' preference for bonds over stocks without assuming unrealistically high risk aversion.

Income taxation and social mobility. Progressive income taxation is often viewed as a key tool for addressing income inequality (Diamond and Saez, 2011). I show that correlation aversion introduces social mobility as an additional motive for progressive taxation. Under recursive utility with correlation aversion, I obtain a normative foundation for dynamic redistribution that takes the form of an “inheritance” tax: redistribution goes from historically high human capital households to historically low human capital ones. Correlation aversion would also favor other policies that target persistent inequalities rather than just smoothing short-term shocks—such as redistributive education financing.

I consider a simplified version of Benabou's (2002) stochastic model of human capital accumulation. While standard discounted expected utility implies that the optimal level of progressive taxation is largely unaffected by human capital persistence, recursive utility with correlation aversion implies that greater persistence significantly increases progressivity of the optimal income tax. Consequently, social mobility is higher under correlation aversion (see Section 4.2).

The intuition is that higher inheritability of human capital increases consumption persistence: individuals from historically high human-capital families tend to have higher incomes and consumption, and vice versa. Correlation aversion therefore favors reducing this persistence.

Related literature

The theoretical literature on dynamic choice has considered a notion of correlation aversion derived from the literature on risk aversion with multiple commodities started by Kihlstrom and Mirman (1974) (see also Richard 1975 or Epstein and Tanny 1980). In particular, Bommier (2007) considers a notion of correlation aversion based on the Kihlstrom and Mirman approach in a continuous time setting. 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. 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. I show that within the class of recursive preferences identified by (ϕ, u, β)—which I refer to as Kreps-Porteus (KP) preferences—intertemporal hedging is equivalent to ϕ being concave, i.e. risk aversion (see Section A.3).

I also consider the notion of strong correlation aversion, which strengthens IRRA and implies a robust representation of preferences. Epstein and Zin's (1989) preferences and Hansen and Sargent's (2001) multiplier preferences satisfy this condition.

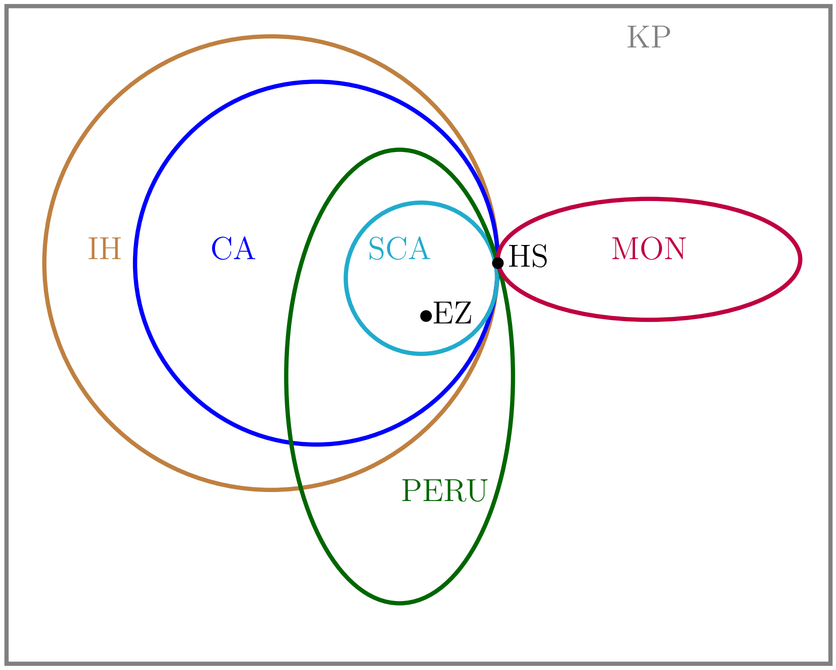

Figure 1: Relationship between correlation averse (CA) preferences and other recursive (KP) preferences: recursive preferences that satisfy intertemporal-hedging (IH), Epstein-Zin (EZ) preferences, multiplier-preferences (HS), monotone recursive preferences (MON), preferences that exhibit a preference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

Meyer-Gohde (2019) first provided a connection between Epstein-Zin preferences and model misspecification. Within the Kreps-Porteus setting, multiplier preferences are the only ones to jointly satisfy strong correlation aversion, preference for early resolution of uncertainty, 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. (2024) on preferences that satisfy stochastic impatience in more depth in Section 5.

Similarly to Andreasen and Jørgensen (2020), I introduce a generalization of Epstein-Zin preferences in order to disentangle risk aversion from attitudes toward non-instrumental information. Their generalization is able to resolve puzzles in the long-run risk model. The main difference with my approach is that they propose a more general form for the utility function u, while I propose a more general formulation of ϕ.

Grant et al. (1998) also provide a connection between preference for non-instrumental information and the Blackwell order in a setting in which preferences are defined over two-stage lotteries (see also Dillenberger 2010). In their framework, each information system induces a two-stage lottery. In contrast, in the present setting with temporal lotteries, consumption in one period serves as a signal for information in the subsequent period.

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) describes the setting for an infinite horizon, i.e., $T = \infty$. I assume that the consumption set C satisfies either $C = [0, \infty)$ or $C = (0, \infty)$, depending on the specific recursive representation under consideration. Given a Polish space X, let $\Delta_s(X), \Delta_b(X)$ denote the space of simple (i.e., finite support) and Borel probability measures with bounded support over X, respectively. Observe that $\Delta_s(X) \subseteq \Delta_b(X)$, and that both are convex spaces.

Given $\ell, m \in \Delta_b(X)$ such that ℓ is absolutely continuous with respect to m (denoted by $\ell \ll m$), I denote the Radon-Nikodym derivative by $\frac{d\ell}{dm}$. For $x \in X$, let $\delta_x \in \Delta_b(X)$ represent 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 convex combination of n probabilities $(m_i)_{i=1}^n$ in $\Delta_b(X)$ with a probability vector $(\pi_i)_{1 \leq i \leq n}$. Note that every two-stage lottery $m \in \Delta_s(\Delta_s(X))$ can be associated to a matrix-vector pair $(M[m], \mu[m])$ where $M[m]$ is a stochastic matrix whose rows describe each probability $M[m](\cdot|i) \in \operatorname{supp} m$ in the support of m for $i = 1, \dots, |\operatorname{supp} m|$, and $\mu[m]$ is a probability row vector which describes the probability of each i.

I consider temporal lotteries that are deterministic in the first period. 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 $D_{T,s} := C$ and recursively

$$D_{t,s} := C \times \Delta_s(D_{t+1,s}),$$

for every $t = 0, \dots, T-1$. Simple temporal lotteries can be intuitively represented

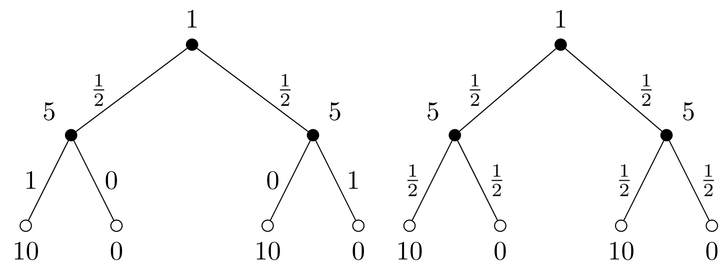

Figure 2: Probability tree representation of two temporal lotteries with $T = 2$

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 $\operatorname{marg}_X m$ the marginal probability over X, i.e., $\operatorname{marg}_X m(A) = m(A \times Y)$ for every measurable set $A \subseteq X$.

Example 1. Assume $T = 2$. Let $d = (c_0, m) = \left(1, \frac{1}{2}(5, 10) \oplus \frac{1}{2}(5, 0)\right)$ and $d' = (c_0, m') = (1, 5, (\frac{1}{2}10 \oplus \frac{1}{2}0))$. Figure 2 provides a graphical representation of these two temporal lotteries. We have $|\operatorname{supp} m| = 2$, $|\operatorname{supp} m'| = 1$ and

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

Moreover, $\mu[\operatorname{marg}_{\Delta_s(D_{2,s})} m'] = [1]$ and $\mu[\operatorname{marg}_{\Delta_s(D_{2,s})} m] = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix}$. △

The preferences of a DM 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.

Definition 1 (Kreps-Porteus preferences)**.** *Preferences $\succeq$ admit a Kreps-Porteus (KP) recursive representation (ϕ, u, β) if 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*

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

where $\beta \in (0,1]$ is a discount factor, $u : C \to \mathbb{R}$ is continuous and strictly increasing, with image either $u(C) = [0,\infty)$ or $u(C) = (0,\infty)$, and $\phi : u(C) \to \mathbb{R}$ is a continuous and strictly increasing function.

This representation of preferences separates risk aversion (captured by the function ϕ) from the EIS (modeled by the utility function u). It is ordinally equivalent to more common formulations used in applications (see, e.g., Werner (2024)). The axiomatic foundations for this representation with bounded u are well known (e.g., Proposition 4 in Sarver 2018); the unbounded case can be covered by results from Bleichrodt et al. (2008). The assumption of unboundedness is necessary for the characterization provided later in Proposition 1. The parameter β is unique, whereas u is cardinally unique, and ϕ is cardinally unique given u.[2]

Two notable cases are the **Epstein–Zin (EZ) preferences**, defined by

$$u(c) = \frac{c^\rho}{\rho}, \quad \phi(x) = \frac{1}{\alpha}(\rho x)^{\frac{\alpha}{\rho}}, \quad c \in C,\ x \in u(C), \tag{1}$$

with parameters $0 \neq \alpha < 1$, $\rho \in (0,1)$, and $\alpha \leq \rho$; and the **Hansen–Sargent (HS) multiplier preferences**, given by

$$\phi(x) = -\exp\left(-\frac{x}{\theta}\right), \quad x \in u(C), \tag{2}$$

with parameter $0 < \theta < \infty$.[3]

I will typically consider KP representations that satisfy certain differentiability assumptions to employ standard tools from the theory of risk aversion. Write $\phi \in \mathcal{C}^r$ if ϕ 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 is 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

[2]In applied literature, the function ϕ is often referred to as *risk adjustment*; see, e.g., Hansen et al. (2007). Moreover, $\beta = 1$ is permitted due to the finite horizon, which would not be the case under an infinite horizon.

[3]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. In this case, when $C = (0,\infty)$ we have $u(x) = \log(x)$ and $\phi(x) = -\exp(\alpha x)$ where $\alpha = -\frac{1}{\theta}$.

non-increasing, it is increasing absolute risk averse (IARA) if its index A_ϕ is non-decreasing, and it is constant absolute risk averse (CARA) if it is both DARA and IARA. Increasing (IRRA), decreasing (DRRA), and constant (CRRA) relative risk averse functions are defined analogously by replacing the index A_ϕ with R_ϕ.

2.1 Preference for (non-instrumental) information

To formally model the trade-off between intertemporal hedging and 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 using a version of the Blackwell order, which allows them to be compared in terms of their (non-instrumental) informativeness. Beyond its theoretical appeal and generality, this approach permits building a formal link between correlation and information, which is central to the main results presented in the next section (see Proposition 3).

Similar to the Blackwell order, this ranking is based on the concept of *garbling*. Consider $m, m' \in \Delta_s(\Delta_s(X))$ such that $\cup_{\ell \in \text{supp} m} \text{supp} \ell = \cup_{\ell \in \text{supp} m'} \text{supp} \ell$, meaning they have the same support over terminal outcomes. Then m' is a garbling of m if they can be associated with $(\mu[m], M[m])$ and $(\mu[m'], M[m'])$, where the columns of $M[m]$ and $M[m']$ represent the same outcomes, and there exists a stochastic matrix G such that $M[m'] = GM[m]$ and $\mu[m']G = \mu[m]$.

Definition 2 (Temporal Blackwell)**.** *Consider $d, d' \in D_{0,s}$ such that $d = (c^{t+1}, m)\,, d' = (c^{t+1}, m')$ for some $t \leq T-2$, $c^{t+1} \in C^{t+1}$, and $m, m' \in \Delta_s(C \times \Delta_s(D_{t+2,s}))$. Say that d is more informative than d', denoted $d \geq_B d'$, if $\text{marg}_C\, m' = \text{marg}_C\, m$ and $\text{marg}_{\Delta_s(D_{t+2,s})}\, m'$ is a garbling of $\text{marg}_{\Delta_s(D_{t+2,s})}\, m$.*

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. A full characterization of this order is left for future research. A natural starting point would be to adopt the methods from Kihlstrom (1984). The following examples illustrate this notion of comparative information.

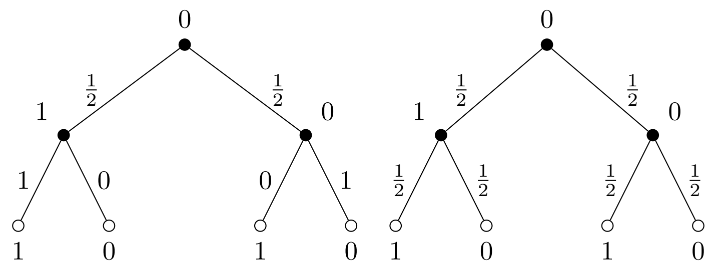

Figure 3: Probability tree representation of a temporal lottery

Example 2 (Example 1 continued)**.** Recall that here we have $d = \left(1, \frac{1}{2}(5,10) \oplus \frac{1}{2}(5,0)\right)$ and $d' = (1, 5, (\frac{1}{2}10 \oplus \frac{1}{2}0))$. If we let $G = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix}$

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

and

$$\mu[\operatorname{marg}_{\Delta_s(D_{2,s})} m']G = \begin{bmatrix} 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix} = \mu[\operatorname{marg}_{\Delta_s(D_{2,s})} m].$$

Furthermore, $\operatorname{marg}_C m' = \operatorname{marg}_C m = \delta_5$, 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'. △

Example 3. Again assume $T = 2$. Consider $d = (1, m), d' = (1, m')$ 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

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

which implies that m' is a garbling of m. Furthermore, $\operatorname{marg}_C m' = \operatorname{marg}_C m$, so that $d \geq_B d'$. In words, d' is an "iid" temporal lottery while d is perfectly correlated. △

I follow Bommier et al. (2017) to describe an agent who exhibits a preference for early resolution of uncertainty (see their Definition 2 and also Strzalecki 2013).

Definition 3 (PERU)**.** *Preferences $\succeq$ exhibit a preference for early resolution of uncertainty (PERU) if for every $n > 0$, $c_0, c_1 \in C$, $(c_{2i}, m_i)_{i=1}^n \in D_{2,s}$, and probability vector $(\pi_i)_{1 \leq i \leq n}$*

$$d_{KP}^{early} := \left(c_0, \bigoplus_{i=1}^{n} \pi_i (c_1, c_{2i}, m_i)\right) \succeq_0 \left(c_0, c_1, \bigoplus_{i=1}^{n} \pi_i (c_{2i}, m_i)\right) := d_{KP}^{late}$$

As I show in the Appendix (see Lemma 2), $d_{KP}^{early} \geq_B d_{KP}^{late}$, which explains why preferring d_{KP}^{early} to d_{KP}^{late} can equivalently be understood as a preference for non-instrumental information. The next result characterizes PERU.

Proposition 1. *Assume $\succeq$ admit a KP representation (ϕ, u, β) with $\phi \in \mathcal{C}^2$. Then preferences $\succeq$ exhibit PERU if and only if*

$$-\beta \frac{\phi''(\beta x + y)}{\phi'(\beta x + y)} \leq -\frac{\phi''(x)}{\phi'(x)}, \tag{3}$$

for every $x, y \in \operatorname{int} u(C)$.[4]

Proof. See the Appendix. □

Observe that the quantity defined by

$$ER_\phi(x, y) = -\frac{\phi''(x)}{\phi'(x)} + \beta \frac{\phi''(\beta x + y)}{\phi'(\beta x + y)} \quad \text{for every } x, y \in \operatorname{int} u(C), \tag{4}$$

can be considered as a local measure of strength of preference for non-instrumental information. In Appendix A.2, I introduce the notion of *early resolution premium*, and show that it is a function of the weighted average of different values of ER_ϕ. To illustrate, when $\beta = 1$ and $\phi(x) = -\exp\left(-\frac{x}{\theta}\right)$ we obtain that

$$ER_\phi(x, y) = \frac{1}{\theta} - \frac{1}{\theta} = 0,$$

which implies indifference to non-instrumental information. The same applies if ϕ is the identity.

Here I focus on risk attitudes that exhibit a preference for information regardless of the level of impatience or intertemporal substitution.

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

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 (3). □

[4]Condition (3) is due to Strzalecki (2013); see p. 1051.

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, I consider the case in which there are two risky periods, i.e., $T = 2$. The Supplemental Appendix (see Section S.1) extends the results to an infinite horizon $T = \infty$.

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. The advantage is that lotteries within this class can be ordered based on their correlation.

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).

The structure of these lotteries can be used to introduce the following notion of increasing correlation.

Definition 5 (IECIT)**.** *Consider $d = (c_0, m), d' = (c_0, m') \in D_{0,s}^*$. Say that d differs from d' by an* ***intertemporal elementary correlation increasing transformation (IECIT)*** *if and only if $m_1 = m_1'$ and there exist $\varepsilon \geq 0$ and a pair (c, c') such that $c \neq c', m_1(c), m_1(c') \neq 0$ and*

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

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

and $m_2 = m_2'$ *otherwise.*

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 4 (Example 3 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. $\triangle$

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 persistence.

Definition 6 (Correlation order)**.** *Given* $d, d' \in D^*_{0,s}$ *say that* d *is more correlated than* d'*, denoted* $d \geq_C d'$*, if* d *differs from* d' *by a finite amount of IECITs.*

Observe that $\geq_C$ is transitive and thus a partial order. 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 $\ell \in \Delta_s(C)$ by $d^{iid}(\ell) = (c, m)$ where $m_2(\cdot|c) = \ell(\cdot)$ for every $c \in C$. In words, the distribution of consumption at each period is described by ℓ.

Proposition 3. *Consider* $\ell \in \Delta_s(C)$ *and* $d, d' \in D^*_{0,s}$*. 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.

Definition 7 (Correlation aversion)**.** *Preferences $\succeq$ exhibit correlation aversion if 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 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 $\phi \in \mathcal{C}^3$ that satisfies UPI. Then every preference relation $\succeq$ with KP representation (ϕ, u, β) exhibit correlation aversion if and only if ϕ is concave and satisfies IRRA.*

Proof. See the Appendix. □

Thus, when ϕ satisfies UPI, every KP preference with representation (ϕ, u, β) is correlation averse exactly when it is risk averse (i.e., ϕ is concave) and exhibits increasing relative risk aversion (IRRA). Observe that IRRA is one of the most important classes of utility functions (e.g., see Arrow (1971), p. 96), and notably includes the Epstein-Zin and Hansen-Sargent preferences. Moreover, empirical findings support DARA and IRRA (Wakker (2010), p. 83). This result further implies that indifference to correlation occurs only under a linear adjustment factor, i.e., $\phi(x) = x$.[5]

A bound on preference for information. A central implication of Theorem 1 is that IRRA constrains preference for information. To illustrate, assume that ϕ exhibits HARA: suppose that for $x > 1$

$$\phi(x) = \frac{1-\gamma}{\gamma}\left(\frac{x}{1-\gamma} + b\right)^{\gamma},$$

[5] Note that the "boundary" case of indifference to correlation occurs when the relative risk aversion function R_ϕ is constant and equal to zero. Indeed, as a consequence of the proof of Theorem 1, whenever $R_\phi(x) > 0$ for some x, one can construct an iid lottery that is strictly better than a lottery that differs from it by an IECIT.

with $0 \neq \gamma < 1$ and $b \in \left[\frac{1}{\gamma-1}, \infty\right)$. Then, given $x, y > 1$, it is easy to check that when $\beta = 1$ the local measure of strength of preference for non-instrumental information

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

is decreasing in the parameter b. In words, the smaller the parameter b, the greater the preference for information. In this case, we have

$$R'_\phi(x) = \frac{(\gamma-1)^2 b}{(b(1-\gamma)+x)^2},$$

so that IRRA implies $b \geq 0$, thus excluding the case $b \in \left[\frac{1}{\gamma-1}, 0\right)$, where the DM values information more. Therefore, IRRA limits preference for information.

More generally, note 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. Therefore, when β is close to unity, IRRA effectively imposes an upper bound on $-\frac{\phi''(x)}{\phi'(x)} + \beta\frac{\phi''(\beta x+y)}{\phi'(\beta x+y)}$ since

$$-\frac{\phi''(x)}{\phi'(x)} + \beta\frac{\phi''(\beta x+y)}{\phi'(\beta x+y)} \approx -A'_\phi(x)y \leq \frac{A_\phi(x)y}{x}. \tag{5}$$

The persistence premium. To measure the level of correlation aversion, I introduce the notion of the persistence premium. The persistence premium quantifies how much a DM is willing to pay to eliminate all persistence from consumption. As a byproduct of Theorem 1, I derive an approximation of the persistence premium in the spirit of Pratt (1964) and Bommier (2007), which connects the premium with preference for information.

Formally, given $c_0 > 0, x > y > 0$, let

$$d^{corr}(\varepsilon) = \left(c_0, \tfrac{1}{2}\left(x, \left(\left(\tfrac{1}{2}+\tfrac{\varepsilon}{2}\right)x \oplus \left(\tfrac{1}{2}-\tfrac{\varepsilon}{2}\right)y\right)\right) \oplus \tfrac{1}{2}\left(y, \left(\left(\tfrac{1}{2}-\tfrac{\varepsilon}{2}\right)x \oplus \left(\tfrac{1}{2}+\tfrac{\varepsilon}{2}\right)y\right)\right)\right),$$

and

$$d^{iid}(\pi) = \left(c_0, \tfrac{1}{2}\left(x(1-\pi), \left(\tfrac{1}{2}x(1-\pi) \oplus \tfrac{1}{2}y(1-\pi)\right)\right) \oplus \tfrac{1}{2}\left(y(1-\pi), \left(\tfrac{1}{2}x(1-\pi) \oplus \tfrac{1}{2}y(1-\pi)\right)\right)\right).$$

The correlated lottery $d^{corr}(\varepsilon)$ is a generalization to arbitrary consumption levels of the perfectly correlated lottery in Example 3. In this case, the level of correlation depends on the parameter $\varepsilon \in [0,1]$, where at $\varepsilon = 0$ one has no correlation and perfect correlation at $\varepsilon = 1$. In contrast, the lottery $d^{iid}(\pi)$ is iid but the payoffs are discounted by a factor of $(1-\pi) \in [0,1]$.

Consider preferences $\succeq$ with KP representation (ϕ, u, β), where $\phi \in \mathcal{C}^3$ is concave, and satisfies both IRRA and UPI. For simplicity, assume that $u(x) = x$. By Theorem 1, $V_0\left(d^{iid}(0)\right) \geq V_0\left(d^{corr}(\varepsilon)\right)$ for every $\varepsilon \in [0,1]$, and as we increase π, $V_0\left(d^{iid}(\pi)\right)$ strictly decreases. In particular, $V_0\left(d^{iid}(1)\right) < V_0\left(d^{corr}(\varepsilon)\right)$ for every $\varepsilon \in [0,1]$. Therefore, we can denote with $\pi(\varepsilon) \in [0,1]$ the unique solution to the equation

$$V_0\left(d^{corr}(\varepsilon)\right) = V_0\left(d^{iid}(\pi(\varepsilon))\right). \tag{6}$$

The persistence premium $\pi(\varepsilon)$ therefore quantifies how much of consumption one is willing to relinquish to have all persistence removed from consumption. The next result provides an approximation of $\pi(\varepsilon)$ near $\varepsilon = 1$.

Corollary 1. *There exist constants $k_1, k_2, k_3 > 0$ such that for every $\varepsilon \in [0,1]$*

$$\begin{aligned}
\pi(\varepsilon) = {} & k_1\left(V_0(d^{iid}(0)) - V_0(d^{corr}(1))\right) \\
& + k_2(\varepsilon - 1)\int_y^x \frac{\phi'(z(1+\beta))}{\phi'(z)} \frac{\{R_\phi(z(1+\beta)) - R_\phi(z)\}}{z}\, dz \\
& - k_3(\varepsilon-1)^2\left\{\frac{\phi'(x(1+\beta))}{\phi'(x)} ER_\phi(x,x) + \frac{\phi'(y(1+\beta))}{\phi'(y)} ER_\phi(y,y)\right\} \\
& + o((\varepsilon-1)^3).
\end{aligned} \tag{7}$$

Proof. See the Appendix. □

The first term of this approximation shows that the persistence premium positively depends on the difference between the value of the iid and the perfectly correlated lottery, independent of the level of correlation ε. In parametric formulations, such as EZ or HS preferences, the term $V_0(d^{iid}(0)) - V_0(d^{corr}(1))$ will increase as the parameter of risk aversion increases.

The second and third terms refer, respectively, to the speed and acceleration of the persistence premium as the level of correlation increases. Since ϕ satisfies IRRA,

$R_\phi(z(1+\beta)) - R_\phi(z)$ is non-negative, and since additionally $\phi' > 0$ it follows that the premium is increasing in the level of persistence. In particular, the greater the derivative of R_ϕ, the faster the premium grows with persistence. However, the third term depends on an average of the measure of preference for information, $ER_\phi(x,x)$ and $ER_\phi(y,y)$. The greater the degree of preference for information, the more the premium grows at a decreasing pace. Hence, this approximation reflects the trade-off between hedging and information.

To illustrate this approximation in a practical case, quick calculations reveal that when $\phi(x) = \frac{1}{\alpha}(\rho x)^{\frac{\alpha}{\rho}}$ we obtain that $ER_\phi(x,y) = \frac{y\left(1-\frac{\alpha}{\rho}\right)}{x(\beta x+y)}$ and for some $\tilde{a}, \tilde{b}, \tilde{c} > 0$

$$\pi(\varepsilon) = \tilde{a} + \tilde{b}\,\varepsilon\Big(1 - \frac{\alpha}{\rho}\Big)\Big(\frac{1}{x} + \frac{1}{y}\Big) \;-\; \tilde{c}\,(\varepsilon^2 - 1)\,ER_\phi(x,y) + o\Big((\varepsilon-1)^3\Big)\,.$$

This formula provides the approximation introduced earlier. It illustrates how the trade-off depends on the preference parameters: higher risk aversion causes the premium to rise more rapidly as ε increases, while higher EIS slows it when $\alpha < 0$ but speeds it up when $\alpha > 0$. At the same time, a higher ER_ϕ moderates the increase. This moderating effect itself depends on both risk aversion and EIS—since ER_ϕ depends on both α and ρ—reflecting the fact that correlation aversion and PERU are not fully distinguishable in the EZ model. I examine the implications of this result for asset pricing in Section 4.1.

Correlation aversion and model misspecification. A further examination of IRRA reveals a tight connection between correlation aversion and fear of model misspecification. Consider the following condition which strengthens IRRA by requiring that the index of relative risk aversion increases sufficiently rapidly.

Definition 8 (SCA)**.** *Say that $\phi \in \mathcal{C}^4$ satisfies strong correlation aversion (SCA) if it is concave, 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 since in both cases $R''_\phi = 0$. Proposition S.2 in the Supplemental Appendix provides an axiomatic foundation for SCA.

To talk about model misspecification, one needs a notion of distance between probabilistic models. A function $I : X \times X \to [0,\infty]$ is called a *statistical distance* (Liese and Vajda, 1987) if $I(\cdot\|x)$ is convex and lower semicontinuous, and $I(x\|x) = 0$ for all $x \in X$.

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

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

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

Proof. See the Appendix. □

This result formalizes the connection between robustness to model misspecification and correlation aversion. By introducing a mild strengthening of risk attitudes related to correlation aversion, recursive preferences naturally reflect a fear of model misspecification. The extent to which SCA is necessary for this representation is discussed in the proof.

The intuition behind this representation is that the decision-maker is concerned about the potential misspecification of the distribution of future consumption. As a result, alternative distributions are evaluated based on their distance from m, as measured by the statistical distance $I^t_{\phi,u,\beta}$ (the general formulation of these statistical distances is discussed in the Appendix). Therefore, $I^t_{\phi,u,\beta}(\ell \| m)$ quantifies the "cost" of considering alternative distribution ℓ at time t. The dependence on t arises from the finite horizon, while the Supplemental Appendix establishes uniqueness when $T = \infty$.

As with multiplier preferences (Strzalecki, 2011), these statistical distances represent the fear of model misspecification—lower values imply lower costs for alternative distributions. However, unlike the framework in Strzalecki (2011), here the cost evaluates discrepancies between temporal lotteries rather than subjective beliefs.

For HS preferences, where $\phi(x) = -e^{-\frac{x}{\theta}}$, the cost function corresponds to Relative Entropy (Strzalecki, 2011):

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

Unlike HS preferences, the cost function for EZ preferences depends additionally on continuation utility. Moreover, Meyer-Gohde (2019) shows the cost function in this case can be expressed via the Tsallis entropy (see also equation 19 for an equivalent characterization in terms of Rényi divergence). Theorem 2 extends this model misspecification interpretation to all recursive preferences satisfying SCA.

4 Applications

4.1 Asset pricing

A key finding in macro-finance is that stocks are procyclical—they generate high returns during economic expansions but tend to crash in recessions (see, for example, Cochrane 2008). As a result, stocks are poor hedging instruments compared to bonds, which offer a relatively stable yield regardless of the business cycle. However, this limited hedging capability alone does not justify the substantial excess returns that investors require for holding equities. This inconsistency is known as the equity premium puzzle.

When economic growth is persistent, correlation aversion makes equity even worse for a hedging instrument: stock prices rise with favorable long-term economic prospects and fall sharply when those prospects deteriorate. Consequently, Theorem 1 and Corollary 1 imply that investors who are correlation averse will demand an even higher premium for bearing equity risk relative to bonds. In general, the relative attractiveness of bonds versus stocks depends on the strength of the preference for early resolution compared to risk aversion. Investors who are risk averse but place little value on early information will tend to favor bonds, whereas those who value early resolution strongly relative to their risk aversion will prefer stocks, since stocks convey news about long-run consumption growth.

This observation carries important macro-financial implications. In particular, the long-run risk model of Bansal and Yaron (2004) introduces a persistent component in consumption growth. If the representative investor has Epstein and Zin's preferences, this model is able to match the observed equity premium. In particular, this model relies on a consumption process (case I) that satisfies for $t = 0, \ldots$

$$\begin{aligned} \log\left(\frac{c_{t+1}}{c_t}\right) &= m + x_{t+1} + \sigma\epsilon_{c,t+1}, \\ x_{t+1} &= ax_t + \varphi\sigma\epsilon_{x,t+1}, \\ \epsilon_{c,t+1}, \epsilon_{x,t+1} &\sim \text{ iid } N(0,1), \end{aligned} \tag{8}$$

where $d_t := \log\left(\frac{c_{t+1}}{c_t}\right)$ denotes consumption growth, $m, \sigma, \phi > 0$, and $a \in [0,1)$ is the persistence parameter. The representative investor has EZ preferences with discount factor $\beta \in (0,1)$ so that the utilities U_t satisfy the recursion

$$U_t = u^{-1}\left((1-\beta)u(c_t) + \beta u\left(\left[\mathbb{E}_t\left(U_{t+1}^{\alpha}\right)\right]^{1/\alpha}\right)\right),$$

where $u(x) = x^\rho$ for $\rho \in [0, 1)$, under the convention that for $\rho = 0$, $u(x) = \log(x)$. This formulation of EZ preferences is ordinally equivalent to the one in (1) (e.g., see Werner 2024).

Such a model faces the trade-off discussed previously. An investor with recursive preferences values both the early resolution of uncertainty and intertemporal hedging. Theorem 1 and Corollary 1 suggest that an investor with EZ preferences is worse off as the persistence a increases, but due to preference for information, at a decreasing rate. It follows that the persistent component of consumption inflates the equity premium because of correlation aversion, *despite* preference for non-instrumental information. Preferences for early resolution of uncertainty still play a role in other contexts such as the macroeconomic announcement premium; see for example Ai and Bansal (2018).[6]

The persistence premium under 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 preferences for early resolution of uncertainty, which, when $\rho = 0$, is given by:

$$1 - \exp\left\{\frac{\alpha}{2}\frac{\beta^2\sigma^2}{1-\beta^2}\left(1 + \frac{\varphi\beta^2}{(1-\beta a)^2}\right)\right\}.$$

Under the standard parameters of the model from the literature, they note that the resulting timing premium seems excessively high compared to introspective assessments.

In light of my analysis of correlation aversion, I ask a different question: "What fraction of your wealth would you give up to remove all persistence in consumption?" Formally, here the persistence premium is defined as

$$\pi := 1 - \frac{U_0(d^{corr})}{U_0(d^{iid})},$$

where d^{iid} and d^{corr} are the stochastic processes of consumption in (8) with $a = 0$ (no persistence) and $a = 0.9790$, respectively. Given that EZ utility is positively homogeneous, the persistence premium, as in the previous section, quantifies the proportion of consumption an investor would be willing to forgo to eliminate all persistence from consumption.

[6] Note that in the long-run risk model, there are other features of preferences at play, unrelated to those discussed here. For instance, in this model it is crucial that the EIS 1, as this ensures that the substitution effect outweighs the income effect.

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)

When $\rho = 0$ the persistence premium is given by (see Appendix A.5)

$$\pi = 1 - \exp\left\{\frac{\beta}{1-\beta a}x_0 - \beta x_0 + \frac{\alpha}{2}\frac{\beta\sigma^2}{1-\beta}\left(\frac{\varphi^2\beta^2}{(1-\beta a)^2} - \varphi^2\beta^2\right)\right\},$$

Thus, π increases with the degree of persistence a, and grows more rapidly with higher risk aversion $1-\alpha$, as well as with the volatility parameters σ and φ. Moreover, the timing premium also depends on these parameters in a similar fashion.

This pattern mirrors that of the approximate persistence premium discussed in the preceding section, based on the approximation in equation (7). In this case, when $\phi(x) = -e^{\alpha x}$, for some constants $\tilde{a}, \tilde{b} > 0$, and for values of β close to unity, we have that for $x > y$:

$$\pi(\varepsilon) \approx \tilde{a} - \tilde{b}\varepsilon\left(e^{\alpha x} - e^{\alpha y}\right).$$

Hence, this premium also increases with the persistence parameter $\varepsilon \in [0,1]$. Moreover, for sufficiently small α, the rate of increase is amplified by higher risk aversion $1-\alpha$ and by greater consumption volatility, which is driven by the difference between x and y.

Table 1 summarizes the parameters of the model. In particular, with a risk aversion level of $1-\alpha = 7.5$, I obtain $\pi \approx 30\%$, while $\pi \approx 40\%$ when $1-\alpha = 10$ (see again Appendix A.5). In other words, an investor with such preferences would be willing to give up either 30% or 40% of their wealth to remove persistence of consumption. Since in the long-run risk model the persistent component is small, it seems unreasonable at the level of introspection to have such a high time premium. To better understand whether this intuition is correct, I will now examine what reasonable levels of the persistence premium are supported by the experimental evidence on correlation aversion.

Measuring risk aversion via correlation aversion. Rohde and Yu (2024) propose a model-free method to measure correlation aversion. Their approach quantifies the degree of positive correlation aversion (Rohde and Yu, 2024, p. 3496) by computing the difference in present certainty equivalents between an iid temporal lottery d^{iid} and a perfectly positively correlated temporal lottery d^{corr} relative to the present certainty equivalent of the iid lottery, that is:

$$\Delta_{\mathrm{POS}}^{\%} := \frac{u^{-1}(V_0(d^{iid})) - u^{-1}(V_0(d^{corr}))}{u^{-1}(V_0(d^{iid}))} = 1 - \frac{u^{-1}(V_0(d^{corr}))}{u^{-1}(V_0(d^{iid}))}. \tag{9}$$

Assuming that $\beta = 0.998$ and $\rho = \frac{1}{3}$—standard parameter specifications in the macro-finance literature—I find that one can match the observed level of $\Delta_{\mathrm{POS}}^{\%}$ with a level of risk aversion such that $1 - \alpha \approx 1.2$. When $\rho = 0$, the observed level of $\Delta_{\mathrm{POS}}^{\%}$ can instead be matched with a level of risk aversion satisfying $1-\alpha \approx 1$ (see Section A.6). These results are consistent with existing estimates, which find $1 - \alpha$ to be between 1 and 2 (see, for example, the discussion on p. 154 in Mehra and Prescott 1985).

So why is the persistence premium π under long-run risk so high? Rohde and Yu (2024) find that correlation aversion does not depend on preference for information. However, as shown previously, Epstein-Zin preferences cannot distinguish between risk aversion and preference for information. Indeed, when relative risk aversion $1-\alpha$ increases, the measure of preference for information increases as well since

$$ER_{\phi}(x, y) = \frac{y(1 - \frac{\alpha}{\rho})}{x(\beta x + y)}. \tag{10}$$

Therefore, based on the approximation in equation (7), one can see that in the EZ case, a higher relative risk aversion parameter, $1 - \alpha$, leads to greater values of $ER_{\phi}(x, x)$ and $ER_{\phi}(y, y)$. This effect is amplified given that $\alpha < 0$ and that the EIS is large in the long-run risk model. This, in turn, causes the persistence premium to grow at a slower rate as the level of persistence increases. This result explains why one may need very high levels of risk aversion and persistence to achieve the degree of correlation aversion necessary to match the equity premium.

An extension of Epstein-Zin preferences. I introduce a new model where the parameter of risk aversion does not necessarily increase preference for information. With this model, one can match the elicited level of correlation aversion in the experiment of Rohde and Yu (2024) with much lower levels of preference for information. This point is connected to the one raised by Meyer-Gohde (2019) who showed that a generalization of EZ preferences can produce Sharpe ratios comparable to the empirical values using more realistic parameters.

I consider the following HARA risk adjustment factor $\phi_{\gamma,b}$ (Merton, 1971, p. 389) given by

$$\phi_{\gamma,b}(x) = \frac{1-\gamma}{\gamma}\left(\frac{x}{1-\gamma}+b\right)^{\gamma} \text{ for every } x \in C, \tag{11}$$

where $0 \neq \gamma < 1$ and $b \geq 0$. Preferences $\succeq$ admit a HARA recursive representation if they admit a KP recursive representation $(\phi_{\gamma,b}, u, \beta)$.

The risk adjustment factor of EZ preferences corresponds to the case $b = 0$ and $\gamma = \frac{\alpha}{\rho}$. Notice that this risk adjustment factor permits a partial separation between risk aversion and preference for information, meaning that high levels of risk aversion can coexist with either a high or low degree of preferences for early resolution of uncertainty. See Appendix A.7 for a formal discussion of these facts. This reasoning motivates the following recursive representation.

In Appendix A.7, I show that $\phi_{\gamma,b}$ satisfies UPI. Consequently, by Theorem 1, correlation aversion corresponds to IRRA, which is equivalent to $b \geq 0$. Moreover, consistent with the intuition discussed earlier, I show that HARA recursive preferences can replicate the experimentally observed correlation aversion reported by Rohde and Yu (2024), using *significantly lower—and thus more realistic—levels of risk aversion* compared to standard EZ parametrizations commonly employed in asset pricing. In short, with a more reasonable utility specification, bonds can be attractive without assuming implausibly high risk aversion.

4.2 Application: income taxation and social mobility

In a setting of intergenerational mobility, where multiple dynasties care about both today's consumption and future generations, the trade-off between hedging and information becomes a trade-off between social mobility and the predictability of income status—that is, the extent to which future positions are dictated by one's current place in the distribution (Shorrocks, 1978).

In this setting, correlation aversion has a significant impact on redistribution policies. Indeed, I show that redistribution policies—which resemble an "inheritance" tax based on historical family income—weaken persistent links between parental and child outcomes, increasing social mobility compared to standard preferences. The same result can also be achieved through alternative policies that reduce long-term consumption inequality, such as the redistribution of education expenditures.

More specifically, I consider a modified version of Benabou's (2002) dynamic model of optimal income taxation with the additional assumption that the innate ability shock can be persistent. In this model, progressive income taxation can serve as a welfare-enhancing tool due to imperfections in credit and insurance markets. However, redistribution introduces distortions in agents' effort or savings decisions, reflecting the classic trade-off between equity and efficiency. One might expect that, as the persistence of innate ability increases, the optimal progressive tax rate (i.e., the one maximizing steady-state aggregate welfare) would become more progressive in order to mitigate the heightened risks arising from imperfections in insurance markets.

Contrary to this intuition, the tax rate that maximizes steady state welfare remains around 33–35%, regardless of the level of persistence of innate ability. However, with recursive utility, correlation aversion amplifies the impact of greater persistence, leading to a significant increase in the optimal progressive tax rate, rising from 45% to 51%. As a consequence, correlation aversion leads to higher social mobility (see equation 12 and the related discussion).

The model. Consider Park's (2009) modified version of Benabou's model. There is a continuum of infinitely-lived agents or dynasties, indexed by $i \in [0,1]$. In each period $t = 0, 1, 2, \ldots$, agent i chooses consumption c_t^i and labor supply l_t^i to maximize intertemporal utility U_t^i, defined recursively by:

$$U_t^i = \max_{c_t^i, l_t^i} \exp\left\{(1-\beta)\left(\ln c_t^i - (l_t^i)^\eta\right) + \beta \ln\left[\mathbb{E}_t\left((U_{t+1}^i)^\gamma\right)\right]^{1/\gamma}\right\},$$

subject to the constraints:

$$y_t^i = (h_t^i)^\lambda (l_t^i)^\mu, \tag{i}$$

$$\hat{y}_t^i = c_t^i + e_t^i, \tag{ii}$$

$$h_{t+1}^i = k\,\xi_{t+1}(h_t^i)^\alpha (e_t^i)^\rho. \tag{iii}$$

Therefore, each dynasty has EZ preferences that satisfy correlation aversion. In this model, $\frac{1}{\eta-1}$ is the elasticity of labor supply and $1-\gamma$ is relative risk aversion. The parameter k scales human-capital formation; α is the elasticity with respect to parental human capital; and λ and μ are the output shares of human capital and labor. Income (y_t^i) and disposable income ($\hat{y}_t^i$) depend on labor supply (l_t^i) and human capital (h_t^i). Human capital (h_{t+1}^i) is determined by the innate ability shock (ξ_{t+1}), parental human capital (h_t^i, where $\ln h_0^i \sim N(m_0, \Delta_0)$), and investment in education (e_t^i). The expectation operator $\mathbb{E}_t$ is conditional on the realized human capital h_t^i.

The constraints require that (i) the income of each generation is produced by combining inherited human capital with labor supply, (ii) that entire income must be allocated between consumption and educational investment in the next generation, and (iii) the child's human capital next period arises from a technology that mixes parental human capital, the education investment just made, and an uninsurable innate ability shock.

The innate ability shock ξ_t can be interpreted as reflecting among other things cognitive ability, and evolves according to the relationship

$$\log(\xi_t) = \phi \log(\xi_{t-1}) + \varepsilon_t \quad \text{where } \varepsilon_t \sim N\left(\mu_\varepsilon, \sigma_\varepsilon^2\right),$$

where ϕ denotes the level of *persistence* of innate ability. Durlauf (1996) provides a theoretical foundation for persistence in innate ability.

The break-even level income $\tilde{y}_t$ is defined implicitly by the balanced-budget constraint:

$$\int_0^1 \left(y_t^i\right)^{1-\tau} (\tilde{y}_t)^\tau \, di = \int_0^1 \left(h_t^i\right)^\lambda \left(l_t^i\right)^\mu di,$$

and the disposable income $\hat{y}_t^i$ is a loglinear function of market income,

$$\hat{y}_t^i = \left(y_t^i\right)^{1-\tau} (\tilde{y}_t)^\tau,$$

where the elasticity τ measures the rate of progressivity of fiscal policy.

The taxation mechanism operates as follows: after determining the market income for all agents, the government calculates the break-even income level, $\tilde{y}_t$. Agents then report their market income to the tax agency. If an agent's income, y_t^i, exceeds $\tilde{y}_t$, they pay a positive tax; otherwise, they receive a subsidy. The elasticity, τ, reflects the progressivity of fiscal policy, with both average and marginal tax rates increasing when $\tau > 0$.

The planner's objective is to optimize the steady-state aggregate welfare given by $\lim_{t\to\infty} W_t$, where $W_t = \int_0^1 \ln U_t^i di$. Park (2009) shows that under standard discounted expected utility (i.e., $1-\gamma = 1$, which corresponds to logarithmic utility) the optimal progressive tax rate is 33% with no persistence of innate ability ($\phi = 0$) and equal to 35% when persistence increases to $\phi = 0.6$. This result suggests that greater persistence has minimal impact on redistribution through progressive taxation. However, I show that with correlation averse preferences satisfying $1 - \gamma = 10$

(see Appendix A.8) one has significantly higher variability: the optimal progressive tax rate is approximately $\tau^* = 45.25\%$ for $\phi = 0$ and $\tau^* = 51.72\%$ for $\phi = 0.6$.

Notably, correlation aversion increases social mobility. To see this formally, observe that $\log(h_t)^i$ follows an AR(2) process with autoregressive coefficients $(\alpha + \lambda\beta(1-\tau)) + \phi$ and $-(\alpha + \lambda\beta(1-\tau))\phi$. As an inverse measure of social mobility, we can therefore take the sum of these two coefficients:

$$(\alpha + \lambda\beta(1-\tau))\,(1-\phi) + \phi. \tag{12}$$

Because correlation averse preferences lower this inverse measure relative to the standard case, they imply higher social mobility (see Appendix A.8 for details.)

This result relates to Gottschalk and Spolaore's (2002) argument that employing recursive utility generates a preference for social mobility. Nevertheless, in this model society aggregates utility additively, so the effect depends not on the welfare criterion but rather on individual preferences, which exhibit correlation aversion and favor mobility over the predictability of income status.

5 Concluding remarks and discussion

This paper has explored the relationship between non-instrumental information and intertemporal hedging within the framework of recursive preferences. I have shown that under reasonable constraints on risk attitudes, a decision maker will value intertemporal hedging more than early resolution of uncertainty. I highlighted the importance of this trade-off in applications such as asset pricing and intergenerational mobility.

A limitation of existing models is that both correlation aversion and preference for information are entirely determined by risk aversion. To address this issue, the paper proposed a generalization of Epstein-Zin preferences to partially separate risk aversion and preference for non-instrumental information.

Further research is needed to understand the implications of more general recursive utility models. Below, I connect my findings to related literature in economics as potential avenues for further research.

Stochastic impatience. DeJarnette et al. (2020) and Dillenberger et al. (2024) study stochastic impatience, an axiom that extends impatience to risky environments.

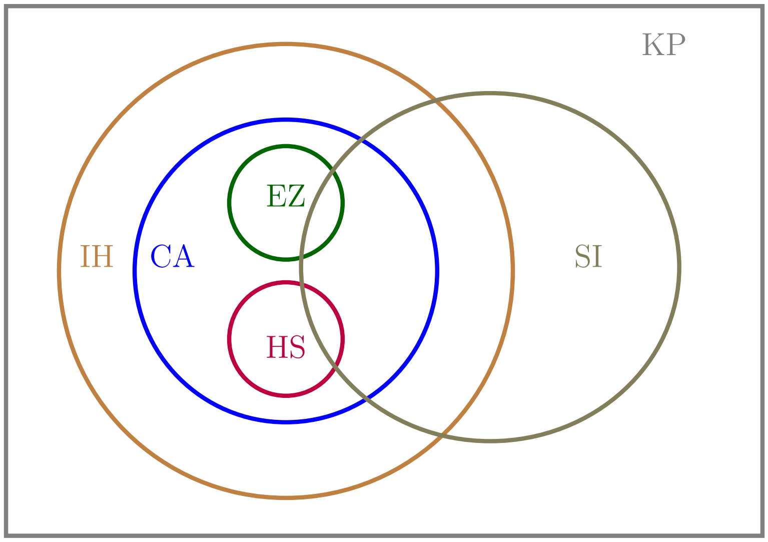

Figure 4: Relationship between correlation averse (CA) preferences and recursive preferences that satisfy intertemporal-hedging (IH), Epstein-Zin (EZ) preferences, and stochastic impatience (SI)

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 of the EIS parameter. The relationship between correlation aversion and stochastic impatience is represented in Figure 4. 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.

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 (8). 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). The persistence premium developed in this model can be used to quantify the cost of long-run climate risks.

Utility smoothing and fiscal hedging. Karantounias (2018, 2022) shows that

Epstein-Zin recursive preferences significantly alter standard Ramsey tax-smoothing policies. The planner adopts fiscal hedging, taxing less during downturns and more during upturns to mitigate income shocks, driven by aversion to volatility in future utilities (Karantounias (2018), p. 2284).

Such a feature of preferences emerges *in spite* of the fact that recursive preferences value early resolution of uncertainty. Instead, this feature emerges from correlation aversion. As shown by Theorems 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.

A Appendix

A.1 Acronyms, Notation, and Technical Definitions

Polish Spaces. A *Polish space* is a topological space that is:

1. Completely metrizable (i.e., there exists a metric that induces the topology and makes the space complete), and

2. Separably metrizable (i.e., there exists a countable, dense subset).

Simple Probability Measures ($\Delta_s(X)$): Let X be a Polish space. The space of *simple probability measures* on X, denoted $\Delta_s(X)$, is the set of probability measures on X with *finite support*. That is:

$$\Delta_s(X) = \Big\{\mu \in \Delta(X) : \text{there exists a finite subset } \{x_1, \dots, x_n\} \subseteq X \\ \text{such that } \mu(\{x_1, \dots, x_n\}) = 1\Big\}.$$

Borel Probability Measures with Bounded Support ($\Delta_b(X)$): Let X be a Polish space. The space of *Borel probability measures with bounded support* on X, denoted $\Delta_b(X)$, is the set of all Borel probability measures μ on X such that the support of μ, $\mathrm{supp}(\mu)$, is compact:

$$\Delta_b(X) = \{\mu \in \Delta(X) : \mathrm{supp}(\mu) \text{ is compact}\}.$$

Acronym/Symbol	Description
V_t	Recursive utility representation at time t
u	Utility function capturing intertemporal substitution
ϕ	Risk adjustment factor
β	Discount factor
EIS	Elasticity of Intertemporal Substitution
IRRA	Increasing Relative Risk Aversion
DARA	Decreasing Absolute Risk Aversion
CRRA	Constant Relative Risk Aversion
HARA	Hyperbolic Absolute Risk Aversion
PERU	Preference for information/early resolution of uncertainty
UPI	Uniform preference for information
SCA	Strong correlation aversion
KP	Kreps-Porteus preferences
EZ	Epstein-Zin preferences
HS	Hansen and Sargent's multiplier preferences
$\pi(\varepsilon)$	Persistence premium given a level of persistence $\varepsilon \in [0,1]$
$\Delta^{\%}_{\text{POS}}$	Measure of correlation aversion from Rohde and Yu (2024)
C	Consumption set: $[0,\infty)$ or $(0,\infty)$
$\Delta_b(X), \Delta_s(X)$	Borel probability measures with bounded support on X
$D_{t,s}$	Simple (finite support) temporal lotteries at time t
$D^*_{0,s}$	Set of lotteries with defined correlation structures
T	Finite time horizon, with $T=2$ in Section 3

Table 2: List of Acronyms and Symbols

Absolute Continuity ($\ell \ll m$). Let $\ell, m \in \Delta_b(X)$. The measure ℓ is *absolutely continuous* with respect to m (denoted $\ell \ll m$) if for every Borel set $A \subseteq X$,

$$m(A) = 0 \implies \ell(A) = 0.$$

Radon-Nikodym Derivative ($\frac{d\ell}{dm}$). If $\ell \ll m$, the *Radon-Nikodym derivative* $\frac{d\ell}{dm}$ is a Borel-measurable function $f : X \to \mathbb{R}_+$ such that:

$$\ell(A) = \int_A f(x)\, dm(x), \quad \text{for all Borel sets } A \subseteq X.$$

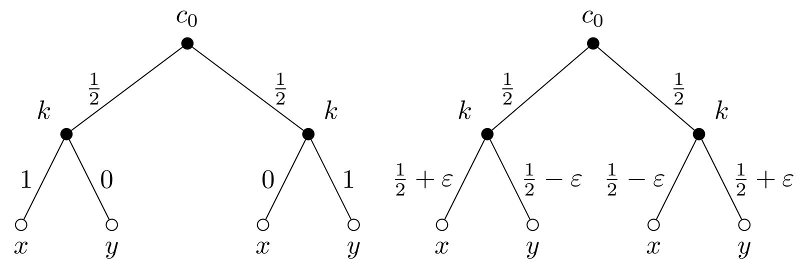

Figure 5: Probability tree representation of two temporal lotteries with $T = 2$

Weak* Topology. The *weak* topology* on $\Delta_b(X)$ is the coarsest topology such that for all continuous and bounded functions $f : X \to \mathbb{R}$, the map:

$$\mu \mapsto \int_X f(x)\, d\mu(x)$$

is continuous. With this topology, μ_n converges to μ in the weak* topology if and only if:

$$\int_X f(x)\, d\mu_n(x) \to \int_X f(x)\, d\mu(x) \quad \text{for all such } f.$$

A.2 Measuring preference for information

To measure attitudes toward early resolution, I introduce the following notion of *early resolution premium*. This notion quantifies how much a DM is willing to pay to have risk resolve at $t = 1$ rather than gradually. Assume for simplicity that $T = 2$, and consider the temporal lotteries given $c_0, k > 0, x > y > 0$

$$d^{gradual}(\varepsilon) = \left(c_0, \frac{1}{2}\left(k, \left(\left(\frac{1}{2}+\frac{\varepsilon}{2}\right)x \oplus \left(\frac{1}{2}-\frac{\varepsilon}{2}\right)y\right)\right)\right.$$
$$\left.\oplus \frac{1}{2}\left(k, \left(\left(\frac{1}{2}+\frac{\varepsilon}{2}\right)x \oplus \left(\frac{1}{2}-\frac{\varepsilon}{2}\right)y\right)\right)\right).$$

and

$$d^{early}(\pi) = \left(c_0, \frac{1}{2}\left(k(1-\pi), x(1-\pi)\right) \oplus \frac{1}{2}\left(k(1-\pi), y(1-\pi)\right)\right).$$

In words, the gradual lottery resolves late when $\varepsilon = 0$ and early when $\varepsilon = 1$. In contrast, the early lottery resolves always early but the payoffs are discounted by a factor of $(1 - \pi)$. See Figure 5 for a graphical representation of these lotteries.

Consider preferences $\succeq$ with KP representation (ϕ, u, β), where $\phi \in \mathcal{C}^2$ is concave and satisfies UPI. Moreover, I assume $u(x) = x$ to simplify calculations. Because ϕ

satisfies UPI, by Proposition 1 we have that $V_0\left(d^{early}(0)\right) \geq V_0\left(d^{gradual}(\varepsilon)\right)$, and as we increase π, $V_0(d^{early}(\pi))$ decreases, so that we can denote with $\pi(\varepsilon)$ the unique solution to the equation

$$V_0(d^{gradual}(\varepsilon)) = V_0(d^{early}(\pi)).$$

The timing premium $\pi(\varepsilon)$ therefore quantifies how much one is willing to pay to have risk resolve at $t = 1$ rather than more gradually as measured by the parameter ε. The next corollary of Proposition 1 provides an approximation of $\pi(\varepsilon)$ near $\varepsilon = 1$.

Corollary 2. *There exists a constant $k_1 > 0$ such that for every $\varepsilon \in [0, 1]$.*

$$\pi(\varepsilon) = k_1 \int_y^x \frac{\phi'(k + \beta z)}{\phi'(z)} ER_\phi(z, k) dz \, (1 - \varepsilon) + \mathrm{o}\,(\varepsilon - 1)\,.$$

Proof. See the Appendix. □

Therefore, the premium is approximated by an average of the values of $ER_\phi(z, k)$ over the interval $[y, x]$. As the measure increases, the premium increases.

A.3 Intertemporal hedging

Here I discuss the difference between my notion of correlation aversion with Kochov's (2015) notion of intertemporal hedging. 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 6 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 with their being equally informative only hedging considerations affect the evaluations of these two lotteries. The next result demonstrates that intertemporal hedging is equivalent to the concavity of ϕ (i.e., risk aversion).

Proposition 4. *Preferences $\succeq$ with KP representation (ϕ, u, β) exhibit intertemporal hedging if and only if ϕ is concave.*

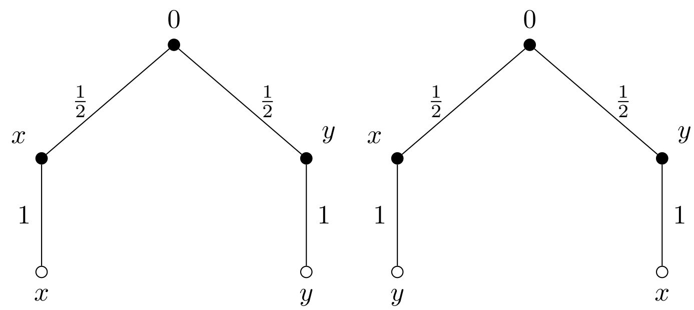

Figure 6: Negative vs positive correlation

Proof. Observe that intertemporal hedging is equivalent to

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

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

Observe that under the assumptions of Theorem 1, correlation aversion implies that ϕ is concave. Hence, from Proposition 4 we can infer that under these assumptions correlation aversion implies intertemporal hedging.

A.4 Extension to lotteries with infinite support

Note that here we operate under the convention that if $Y \subseteq X$, then $\Delta_b(Y) \subseteq \Delta_b(X)$ by identifying each probability measure in $\Delta_b(Y)$ with the equivalent probability measure in $\Delta_b(X)$ that assigns probability 1 to Y. Endow $\Delta_b(X)$ with the weak* topology.

The extension uses the following notion a uniformly bounded sequence of temporal lotteries.

Definition 9 (Uniformly bounded temporal lotteries)**.** *Say that a sequence $(c, m_n)_{n=0}^{\infty}$ in D_0 is uniformly bounded if there exists a compact set $K \subseteq \mathbb{R}$ such that for some K*

$$m_n \in \Delta_b(K \times \Delta_b(K)) \text{ eventually.}$$

The idea is that $d \in D_0$ is considered more correlated than $d' \in D_0$ if both can be approximated by sequences of uniformly bounded simple lotteries $(c, m_n)_{n=0}^{\infty}$ and

$(c, m'_n)_{n=0}^{\infty}$, respectively, where each (c, m_n) is more correlated than (c, m'_n). Correlation aversion then implies that d' is preferred to d.

Proposition 5. *Assume that the preferences $\succeq$ exhibit correlation aversion. Consider $d, d' \in D_0$ such that there exist uniformly bounded sequences $(c, m_n)_{n=0}^{\infty}$ and $(c, m'_n)_{n=0}^{\infty}$ in $D^*_{0,s}$, and a sequence $(\ell_n)_{n=0}^{\infty}$ in $\Delta_s(C)$, satisfying:*

$$\lim_{n\to\infty} (c, m_n) = d, \quad \lim_{n\to\infty} (c, m'_n) = d',$$

and for every $n \geq 0$,

$$(c, m_n) \geq_C (c, m'_n) \geq_C d^{iid}(\ell_n).$$

Then $d' \succeq_0 d$.

A.5 The persistence premium and long-run risk

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

$$\log U_0(d^{corr}) = \log c_0 + \frac{\beta}{1-\beta a} x_0 + \frac{\beta}{1-\beta} m + \frac{\alpha}{2} \frac{\beta\sigma^2}{1-\beta} \left(1 + \frac{\varphi^2\beta^2}{(1-\beta a)^2}\right),$$

and

$$\log U_0(d^{iid}) = \log c_0 + \beta x_0 + \frac{\beta}{1-\beta} m + \frac{\alpha}{2} \frac{\beta\sigma^2}{1-\beta} \left(1 + \varphi^2\beta^2\right).$$

Therefore, we obtain

$$\pi = 1 - \frac{U_0(d^{corr})}{U_0(d^{iid})} = 1 - e^{\frac{\beta}{1-\beta a} x_0 - \beta x_0 + \frac{\alpha}{2}\frac{\beta\sigma^2}{1-\beta}\left(\frac{\varphi^2\beta^2}{(1-\beta a)^2} - \varphi^2\beta^2\right)}.$$

$$\begin{aligned}
\pi &= 1 - \exp\left(-6.5 \times 0.998 \times \tfrac{0.0078^2}{2(1-0.998)} \left(0.044^2 \times \tfrac{0.998^2}{(1-0.998\times 0.979)^2} - 0.044^2 \times 0.998^2\right)\right) \\
&\approx 0.302, \\
\pi &= 1 - \exp\left(-9 \times 0.998 \times \tfrac{0.0078^2}{2(1-0.998)} \left(0.044^2 \times \tfrac{0.998^2}{(1-0.998\times 0.979)^2} - 0.044^2 \times 0.998^2\right)\right) \\
&\approx 0.393.
\end{aligned}$$

Therefore, we have that $\pi \approx 30\%$ with $1-\alpha = 7.5$ and $\pi \approx 40\%$ with $1-\alpha = 10$.

Note that these results do not change significantly if one increases the long-run volatility of the iid process to match the volatility of the persistent process. Consider for example the case $1-\alpha = 7.5$. Observe that

$$\lim_{t\to\infty} \mathrm{Var}\left(\log \frac{c_{t+1}}{c_t}\right) = \sigma^2 + \frac{\varphi^2\sigma^2}{1-a^2},$$

so that by setting $\sigma \approx 0.0079719$ we obtain that the two processes have the same long-run volatility:

$$0.0078^2 + \frac{0.044^2 \times 0.0078^2}{1 - 0.979^2} = \sigma^2 + 0.044^2\sigma^2.$$

With this level of persistence in the i.i.d. process, when $1 - \alpha = 7.5$, we obtain the persistence premium:

$$1 - \exp\Bigg(- \frac{6.5}{2} \times \frac{0.998 \times 0.0078^2}{1 - 0.998} \left(1 + 0.044^2 \times \frac{0.998^2}{(1 - 0.979 \times 0.998)^2}\right) \\ + \frac{6.5}{2} \times \frac{0.998 \times 0.0079719^2}{1 - 0.998} \left(1 + 0.044^2 \times 0.998^2\right) \Bigg) \approx 0.299790.$$

A.6 Measuring risk aversion

Recall that from equation (9) we have that the measure of correlation aversion is given by

$$\Delta^{\%}_{POS} = 1 - \frac{u^{-1}(V_0(d^{corr}))}{u^{-1}(V_0(d^{iid}))}.$$

The lotteries considered in Rohde and Yu (2024) feature risk at $t = 0$. However, since EZ preferences are stationary, one can equivalently consider the pair of temporal lotteries given by

$$d^{corr} = \left(0, \tfrac{1}{2}\,(10, 10) \oplus \tfrac{1}{2}\,(5, 5)\right).$$

and

$$d^{iid} = \left(0, \tfrac{1}{2}\left(10, \left(\tfrac{1}{2}10 \oplus \tfrac{1}{2}5\right)\right) \oplus \tfrac{1}{2}\left(5, \left(\tfrac{1}{2}10 \oplus \tfrac{1}{2}5\right)\right)\right).$$

Here, I assume that $t = 1$ equals 4 weeks, as in the first time frame considered by Rohde and Yu (2024). Since the time unit is 4 weeks, we can apply the monthly discount factor $\beta = 0.998$ used in Bansal and Yaron (2004). In this case, we have

$$u^{-1}\left(V_0\left(d^{corr}\right)\right) = u^{-1}\left\{0.998\phi^{-1}\left(\phi\left(u(5) + 0.998u(5)\right) + \phi\left(u(10) + 0.998u(10)\right)\right)\right\},$$

and

$$u^{-1}\left(V_0\left(d^{iid}\right)\right) = u^{-1}\Bigg\{0.998\phi^{-1}\Bigg(\phi\left(u(5) + 0.998\phi^{-1}\left(\frac{1}{2}\phi(u(5)) + \frac{1}{2}\phi(u(10))\right)\right) \\ + \phi\left(u(10) + 0.998\phi^{-1}\left(\frac{1}{2}\phi(u(5)) + \frac{1}{2}\phi(u(10))\right)\right)\Bigg)\Bigg\}$$

I consider the mean value

$$\Delta^{\%}_{POS} = 0.008$$

found in their experiment (Rohde and Yu, 2022, p. 55). This value corresponds to the time-risk framing of their experiment, which encouraged subjects to consider correlation over time.

When $\frac{1}{1-\rho} = 1.5$—a common specification in the long-run risk literature (Bansal and Yaron, 2004) we can match this level of correlation aversion by setting $\alpha = -\frac{0.61}{3} \approx -0.2$ since in this case we obtain

$$1 - \frac{u^{-1}(V_0(d^{corr}))}{u^{-1}(V_0(d^{iid}))} \approx 0.008.$$

When $\rho \approx 0$, by setting $\alpha \approx -0.0345$ we obtain

$$1 - \frac{u^{-1}(V_0(d^{corr}))}{u^{-1}(V_0(d^{iid}))} \approx 1 - \frac{\left(\frac{5^{1.998\alpha}+10^{1.998\alpha}}{2}\right)^{\frac{0.998}{\alpha}}}{\left(\frac{1}{2}\left(5^{\alpha}+10^{\alpha}\right)\right)^{\frac{1.996}{\alpha}}} \approx 0.008.$$

Note that when $\alpha \approx -0.0345$, we have that the persistence premium π is close to zero since

$$\pi = 1 - \exp\left((-0.711) \times \frac{0.998 \times 0.0078^2}{2 \times (1 - 0.998)}\left(\frac{0.044^2 \times 0.998^2}{(1 - 0.998 \times 0.979)^2} - 0.044^2 \times 0.998^2\right)\right) \approx 0.0019,$$

a more reasonable value in terms of introspection and consistent with the evidence in Meissner and Pfeiffer (2022), which shows that 40% of subjects have a zero timing premium.

A.7 The persistence premium and HARA recursive preferences

First observe that for every $\beta \in (0, 1]$ and $x, y \geq 0$

$$-\frac{\phi''_{\gamma,b}(x)}{\phi'_{\gamma,b}(x)} + \beta\frac{\phi''_{\gamma,b}(\beta x + y)}{\phi'_{\gamma,b}(\beta x + y)} = \left(\frac{1}{\frac{x}{1-\gamma} + b} - \beta\frac{1}{\frac{\beta x + y}{1-\gamma} + b}\right)$$
$$= \left(\frac{1}{\frac{x}{1-\gamma} + b} - \frac{1}{\frac{x + \frac{y}{\beta}}{1-\gamma} + \frac{b}{\beta}}\right) \geq 0,$$

which implies that UPI is satisfied. Further, we have that for $x > 0$

$$R_{\phi_{\gamma,b}}(x) = \frac{x}{\frac{x}{1-\gamma} + b} = \frac{1}{\frac{1}{1-\gamma} + \frac{b}{x}}.$$

so that IRRA is satisfied whenever $b \geq 0$. Now observe that in the EZ case $b = 0$ so that:

$$-\frac{\phi''_{\gamma,0}(x)}{\phi'_{\gamma,0}(x)} + \beta\frac{\phi''_{\gamma,0}(\beta x + y)}{\phi'_{\gamma,0}(\beta x + y)} = \frac{(1-\gamma)y}{x(\beta x + y)} \geq 0,$$

which implies that if risk aversion goes to infinity, i.e. if $\gamma \to -\infty$ then

$$-\frac{\phi''_{\gamma,0}(x)}{\phi'_{\gamma,0}(x)} + \beta\frac{\phi''_{\gamma,0}(\beta x + y)}{\phi'_{\gamma,0}(\beta x + y)} \to +\infty. \tag{13}$$

Finally, we have that when β is close to unity

$$\lim_{\gamma\to-\infty} -\frac{\phi''_{\gamma,0}(x)}{\phi'_{\gamma,0}(x)} + \beta\frac{\phi''_{\gamma,0}(\beta x + y)}{\phi'_{\gamma,0}(\beta x + y)} = \frac{1-\beta}{b} \approx 0.$$

Hence, high levels of risk aversion are compatible with a small demand for non-instrumental information if we assume "large" values of b, while by (13) for $b \approx 0$ one can have high levels of risk aversion compatible with high demand of non-instrumental information.

To illustrate, assume that $u(x) = 3x^{1/3}$ and $\phi_{\gamma,b}$ with $\gamma = -2$, and $b = 0.72$. Consider the temporal lotteries from Rohde and Yu (2024)

$$d^{corr} = \left(0, \tfrac{1}{2}(10, 10) \oplus \tfrac{1}{2}(5, 5)\right).$$

and

$$d^{iid} = \left(0, \tfrac{1}{2}\left(10, \left(\tfrac{1}{2}10 \oplus \tfrac{1}{2}5\right)\right) \oplus \tfrac{1}{2}\left(5, \left(\tfrac{1}{2}10 \oplus \tfrac{1}{2}5\right)\right)\right).$$

We obtain that

$$\Delta^{\%}_{POS} = 1 - \frac{u^{-1}(V_0(d^{corr}))}{u^{-1}(V_0(d^{iid}))} \approx 0.0341.$$

If we assume that $\gamma = -27$ and $b = 0$, that is a standard EZ parametrization in which $\gamma = \frac{\alpha}{\rho}$, $1 - \alpha = 10$, $\frac{1}{1-\rho} = 1.5$. In this case I obtain

$$\Delta^{\%}_{POS} = 1 - \frac{u^{-1}(V_0(d^{corr}))}{u^{-1}(V_0(d^{iid}))} \approx 0.0341.$$

Hence, these different formulation of recursive HARA preferences and EZ preferences attain the same level of correlation aversion. However, in the former case, we obtain an average value of the measure of preference for information $ER_\phi(x,x)$ equal to

$$\int_{u(5)}^{u(10)} ER_\phi(x,x)dx = \int_{u(5)}^{u(10)} \left(\frac{1}{\frac{x}{3} + 0.722} - \frac{1}{\frac{1}{3}\left(x + \frac{x}{0.998}\right) + \frac{0.722}{0.998}} \right) dx = 0.212242,$$

as opposed to the EZ preferences:

$$\int_{u(5)}^{u(10)} ER_\phi(x,x)dx = \int_{u(5)}^{u(10)} \frac{(1+27)x}{x(0.998x+x)} dx = 3.23792,$$

and a level of relative risk aversion

$$\int_{u(5)}^{u(10)} R_\phi(x)dx = \int_{u(5)}^{u(10)} \frac{1}{\left(\frac{1}{3} + \frac{0.722}{x}\right)5} dx = 0.581891,$$

as opposed to relative risk aversion of $1-(-27) = 28$ under EZ preferences. Hence, recursive HARA preferences achieve a comparable level of correlation aversion as EZ preferences under the standard parametrization, but they exhibit a much more limited level of relative risk aversion that is consistent with empirical evidence and a small preference for information.

A.8 Income taxation and social mobility

The steady level of aggregate welfare is derived in Appendix B in Park (2009) under the difference that ρ and β are switched in the present notation. Here we assume that $\sigma_\varepsilon^2 = \omega^2$ and $\mu_\varepsilon = -\frac{\omega^2}{2}$. The values of the parameters are those in Table 3.

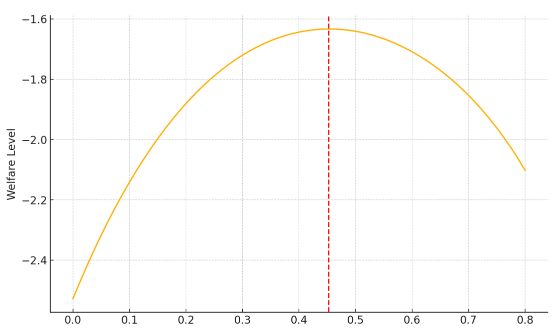

Figure 7: Optimal progressive tax rate as a function of $\tau(\gamma = -9, \phi = 0)$.

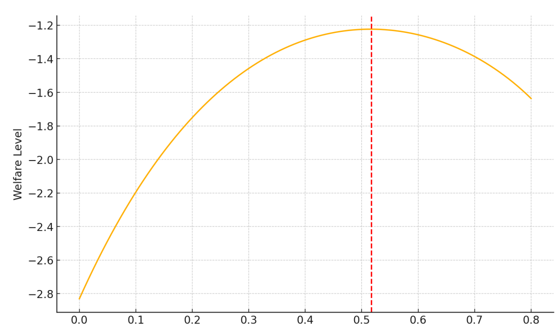

Figure 8: Optimal progressive tax rate as a function of $\tau(\gamma = -9, \phi = 0.6)$.

β	γ	$\frac{1}{\eta-1}$	ϕ	λ	μ	ρ	α	k	ω
0.2939	-9	0.2	0.6	0.625	0.375	$\frac{0.25}{0.625}$	0.35	1	1

Table 3: Parameters in the model

The steady state aggregate welfare as a function of $\tau \in [0,1]$ is given by

$$\frac{(1-\beta)\lambda\rho}{(1-\beta(\alpha+\rho\lambda))(1-\alpha-\rho\lambda)}\left(\ln(1-\tau)+\ln(\rho\beta\lambda)-\ln(1-\beta\alpha)+\frac{\mu}{\eta}\ln\left(\frac{\mu}{\eta}\right)+\right.$$
$$\left.\frac{\mu}{\eta}\ln(1-\beta\alpha)-\frac{\mu}{\eta}\ln(1-\beta(\alpha+\rho\lambda(1-\tau)))+\frac{\tau(2-\tau)\lambda^2\omega^2}{2(1-\phi^2)(1-\alpha-\rho\lambda+\rho\lambda\tau)^2}\right)$$
$$+(1-\beta\alpha)(1-\beta(\alpha+\rho\lambda))^{-1}\left(\frac{\mu}{\eta}\ln\left(\frac{\mu}{\eta}\right)+\frac{\mu}{\eta}\ln(1-\beta\alpha)-\frac{\mu}{\eta}\ln(1-\beta(\alpha+\rho\lambda(1-\tau)))\right)$$
$$-\frac{(\mu/\eta)(1-\beta\alpha)}{1-\beta(\alpha+\rho\lambda(1-\tau))}$$
$$+\lambda\beta(1-\beta(\alpha+\rho\lambda))^{-1}(\ln k+\rho\ln(1-\tau)+\rho\ln(\rho\beta\lambda)-\rho\ln(1-\beta\alpha))$$
$$+\ln\left(1-\frac{(1-\tau)\rho\beta\lambda}{1-\beta\alpha}\right)$$
$$+\gamma\beta\frac{\lambda^2(1-\beta)(1-\tau)^2\omega^2}{2(1-\phi^2)(1-\beta\alpha-\beta\rho\lambda+\beta\rho\lambda\tau)^2}$$
$$+(1-\beta\alpha)(1-\beta(\alpha+\rho\lambda))^{-1}\tau(2-\tau)\frac{\lambda^2\omega^2}{2(1-\phi^2)(1-\alpha-\rho\lambda+\rho\lambda\tau)^2}$$

The optimal $\tau^* \in [0,1]$ maximizes the previous expression. When $\phi = 0$, I obtain that $\tau \approx 0.4525$. When $\phi = 0.6$, I obtain that $\tau \approx 0.5172$ (see Figures 7 and 8).

Finally, observe that social mobility is higher under correlation averse preferences, with $\tau \approx 0.5172$, rather than under correlation neutrality which implies $\tau \approx 0.35$ (Park (2009)). Indeed, when $\alpha = 0.35, \lambda = 0.625, \beta = 0.2939, \phi = 0.6$ we obtain

$$(\alpha+\lambda\beta(1-0.5172))(1-\phi)+\phi = 0.775 < 0.7878 = (\alpha+\lambda\beta(1-0.35))(1-\phi)+\phi.$$

Hence, social mobility is higher under correlation averse preferences.

A.9 Proofs

Proof of Proposition 1. The proof uses the functions $U_t \colon \Delta_s(D_{t+1,s}) \to \mathbb{R}$, defined for $\bar{c} \in C$ and $t = 1, \dots, T-2$ by

$$U_t(m) = \phi\left(u(\bar{c}) + \beta\phi^{-1}\left(\mathbb{E}_m\phi(V_{t+1})\right)\right) \quad \text{for every } m \in \Delta_s(D_{t+1,s}). \tag{14}$$

Lemma 1. *Each U_t defined in (14) is convex if and only if (3) holds.*

Proof. First I claim that each U_t defined in (14) 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 must 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). □

Lemma 2. *It holds that*

$$d_{KP}^{early} \geq_B d_{KP}^{late}.$$

Proof. First observe that since

$$d_{KP}^{early} = \left(c_0, \bigoplus_{i=1}^{n} \pi_i\left(c_1, c_{2i}, m_i\right)\right) \text{ and } \left(c_0, c_1, \bigoplus_{i=1}^{n} \pi_i\left(c_{2i}, m_i\right)\right) = d_{KP}^{late},$$

we have that $\operatorname{marg}_C \bigoplus_{i=1}^n \pi_i\left(c_1, c_{2i}, m_i\right) = \delta_{c_1} = \operatorname{marg}_C\left(c_1, \bigoplus_{i=1}^n \pi_i\left(c_{2i}, m_i\right)\right)$. Furthermore, the lotteries

$$\operatorname{marg}_{\Delta_s(D_{2,s})}\left(c_1, \bigoplus_{i=1}^{n} \pi_i\left(c_{2i}, m_i\right)\right) = \delta_{\bigoplus_{i=1}^n \pi_i(c_{2i}, m_i)} \in \Delta_s\left(\Delta_s\left(D_{2,s}\right)\right),$$

and

$$\operatorname{marg}_{\Delta_s(D_{2,s})} \bigoplus_{i=1}^{n} \pi_i\left(c_1, c_{2i}, m_i\right) = \bigoplus_{i=1}^{n} \pi_i \delta_{(c_{2i}, m_i)} \in \Delta_s\left(\Delta_s\left(D_{2,s}\right)\right),$$

can be associated with the matrix-vector pairs

$$M\left[\operatorname{marg}_{\Delta_s(D_{2,s})}\delta_{\bigoplus_{i=1}^n \pi_i (c_{2i},m_i)}\right] = \begin{bmatrix}\pi_1 & \dots & \pi_n\end{bmatrix},$$

$M\left[\operatorname{marg}_{\Delta_s(D_{2,s})}\bigoplus_{i=1}^n \pi_i \delta_{(c_{2i},m_i)}\right] = I$, where I denotes the identity matrix,

$$\mu\left[\operatorname{marg}_{\Delta_s(D_{2,s})}\delta_{\bigoplus_{i=1}^n \pi_i (c_{2i},m_i)}\right] = [1] \text{ and } \mu\left[\bigoplus_{i=1}^n \pi_i \delta_{(c_{2i},m_i)}\right] = [\pi_1 \dots \pi_n].$$

Now by setting $G := \begin{bmatrix}\pi_1 & \dots & \pi_n\end{bmatrix}$ we obtain that

$$M\left[\operatorname{marg}_{\Delta_s(D_{2,s})}\delta_{\bigoplus_{i=1}^n \pi_i (c_{2i},m_i)}\right] = \begin{bmatrix}\pi_1 & \dots & \pi_n\end{bmatrix} I = GM\left[\operatorname{marg}_{\Delta_s(D_{2,s})}\bigoplus_{i=1}^n \pi_i \delta_{(c_{2i},m_i)}\right],$$

and

$$\mu\left[\operatorname{marg}_{\Delta_s(D_{2,s})}\delta_{\bigoplus_{i=1}^n \pi_i (c_{2i},m_i)}\right] G = [1]\begin{bmatrix}\pi_1 & \dots & \pi_n\end{bmatrix} = [\pi_1 \dots \pi_n] = \mu\left[\bigoplus_{i=1}^n \pi_i \delta_{(c_{2i},m_i)}\right].$$

Hence, we can conclude that $d_{KP}^{early} \geq_B d_{KP}^{late}$ as desired. □

Proof of Proposition 1. Note that if ϕ satisfies (3), then each U_t is convex by Lemma 1. By Lemma 2, $\operatorname{marg}_{\Delta_s(D_{2,s})}\left(c_1, \bigoplus_{i=1}^n \pi_i (c_{2i}, m_i)\right) = \delta_{\bigoplus_{i=1}^n \pi_i (c_{2i},m_i)}$ is a garbling of $\operatorname{marg}_{\Delta_s(D_{2,s})}\bigoplus_{i=1}^n \pi_i (c_1, c_{2i}, m_i) = \bigoplus_{i=1}^n \pi_i \delta_{(c_{2i},m_i)}$. By Theorem 4 in Kihlstrom (1984), it follows that

$$\sum_{i=1}^n \pi_i W((c_{2i}, m_i)) \geq W\left(\bigoplus_{i=1}^n \pi_i (c_{2i}, m_i)\right),$$

for every convex function $W : \Delta_s(D_{2,s}) \to \mathbb{R}$.

Observe that

$$V_0\left(d_{KP}^{early}\right) \geq V_0\left(d_{KP}^{late}\right) \iff$$
$$\sum_{i=1}^n \pi_i \phi\left(u(c_1) + \beta V_2(c_{2i}, m_i)\right) \geq \phi\left(u(c_1) + \beta\phi^{-1}\left(\sum_{i=1}^n \pi_i \phi\left(V_2(c_{2i}, m_i)\right)\right)\right).$$

Since

$$\sum_{i=1}^n \pi_i U_1((c_{2i}, m_i)) = \sum_{i=1}^n \pi_i \phi\left(u(c_1) + \beta V_2((c_{2i}, m_i))\right)$$

and

$$U_1\left(\bigoplus_{i=1}^n \pi_i (c_{2i}, m_i)\right) = \phi\left(u(c_1) + \beta\phi^{-1}\left(\sum_{i=1}^n \pi_i \phi\left(V_2((c_{2i}, m_i))\right)\right)\right),$$

which by convexity of U_1 we obtain $\sum_{i=1}^{n} \pi_i U_1\left((c_{2i}, m_i)\right) \geq U_1\left(\bigoplus_{i=1}^{n} \pi_i (c_{2i}, m_i)\right)$, and therefore that $d_{KP}^{early} \succeq_0 d_{KP}^{late}$.

Conversely, consider $d, d' \in D_{0,s}$ 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. □

Proof of Proposition 3.

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 and that d differs from d' by an IECIT. First observe that $d = (c_0, m), d' = (c_0, m')$ where $\operatorname{marg}_C m' = \operatorname{marg}_C m$ since by definition of an IECIT $m_1 = m_1'$. Now consider the stochastic matrices of conditional distributions for d and d':

$$A := M\left[\operatorname{marg}_{\Delta_s(C)} m\right] = \begin{bmatrix} m_2(c|c) & m_2(c'|c) & \ldots & m(c_N|c) \\ m_2(c|c') & m_2(c'|c') & \ldots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ m_2(c|c_N) & m_2(c'|c_N) & \ldots & m_2(c_N|c_N) \end{bmatrix},$$

and

$$B := M\left[\operatorname{marg}_{\Delta_s(C)} m'\right] = \begin{bmatrix} m_2'(c|c) & m_2'(c'|c) & \ldots & m(c_N|c) \\ m_2'(c|c') & m_2'(c'|c') & \ldots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ m_2'(c|c_N) & m_2'(c'|c_N) & \ldots & m_2'(c_N|c_N) \end{bmatrix}.$$

Then for some $\varepsilon, \varepsilon' \geq 0$

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

and

$$\begin{bmatrix} m_2'(c|c) & m_2'(c'|c) \\ m_2'(c|c') & m_2'(c'|c') \end{bmatrix} = \begin{bmatrix} \ell(c) + \frac{\varepsilon'}{\ell(c)} & \ell(c') - \frac{\varepsilon'}{\ell(c)} \\ \ell(c) - \frac{\varepsilon'}{\ell(c')} & \ell(c') + \frac{\varepsilon'}{\ell(c')} \end{bmatrix}.$$

Hence, if we choose $x_1, x_2 \in [0,1]$ satisfying $x_1(\frac{\varepsilon+\varepsilon'}{\ell(c)}) - (1-x_1)(\frac{\varepsilon+\varepsilon'}{\ell(c')}) = \frac{\varepsilon'}{\ell(c)}$ and $x_2(\frac{\varepsilon+\varepsilon'}{\ell(c)}) - (1-x_2)(\frac{\varepsilon+\varepsilon'}{\ell(c')}) = \frac{\varepsilon'}{\ell(c')}$, then letting

$$G := \begin{bmatrix} x_1 & 1-x_1 & 0 & \dots & 0 \\ x_2 & 1-x_2 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \\ 0 & \vdots & \vdots & 0 & 1 \end{bmatrix}$$

it holds that

$$B = GA,$$

which implies that $d \geq_B d'$. If instead we choose $x_1 = x_2 = \frac{\frac{\ell(c)}{\ell(c')}}{1+\frac{\ell(c)}{\ell(c')}}$ then we obtain that

$$\begin{bmatrix} \ell(c) & \ell(c') & \dots & \ell(c_N) \\ \ell(c) & \ell(c') & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \ell(c) & \ell(c') & \dots & \ell(c_N) \end{bmatrix} = GB,$$

which implies $d' \geq_B d^{iid}(\ell)$ as desired. □

Proof of Theorem 1. Let $d, d' \in D^*_{0,s}$ and $\ell \in \Delta_s(C)$. I provide first the following preliminary result.

Lemma 3. *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 twice continuously differentiable function $U : [0,1] \to \mathbb{R}$ such that*

1. *$U(0) = V_0(d)$ and $U(1) = V_0(d')$;*
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.3 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. Assume that ϕ satisfies IRRA and that $d \geq_C d' \geq_C d^{iid}(\ell)$. I claim that

$$d^{iid}(\ell) \succeq_0 d' \succeq_0 d,$$

for every $\succeq$ with KP representation (ϕ, u, β). By Lemma 3, 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).[7] I claim that it also holds that

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

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

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

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

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

[7] By applying Lemma 3, if there is a sequence $(d_i)_{i=0}^N$ such that each d_i differs from d_{i-1} by an IECIT, then one can construct a continuous function $U : [0, 1] \to \mathbb{R}$ that admits weak derivatives U', U''. Indeed, one can first apply Lemma 3 to obtain functions $(U_i)_{i=1}^N$ 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 $x \in \left[\frac{N-1}{N}, 1\right]$. Note that each U_i is continuously differentiable, which implies that U is piecewise Lipschitz continuous. By Lemma 2 in Leobacher and Steinicke (2022), it follows that U is Lipschitz continuous and therefore weakly differentiable. The same reasoning also applies to U'.

Applying the fundamental theorem of calculus for weak derivatives (see Theorem 8.2 in Brezis 2010), it follows that $\lim_{\varepsilon\to 1} U'(\varepsilon) - U'(\tilde{\varepsilon}) = \int_{\tilde{\varepsilon}}^{1} U''(t)dt \geq 0$ for every $\tilde{\varepsilon} \in (0,1)$, which further implies that

$$V_0(d') - V_0(d^{iid}(\ell)) = \int_0^{q_1} U'(\tilde{\varepsilon})d\tilde{\varepsilon} \leq 0,$$

and

$$V_0(d') - V_0(d) = \int_{q_1}^{q_2} U'(\tilde{\varepsilon})d\tilde{\varepsilon} \leq 0.$$

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

As for the converse, assume that ϕ is not concave. Then there exist $x > y > 0$ such that for every $z \in (y, x)$,

$$\phi''\left(z(1+\beta)\right) > 0. \tag{15}$$

Consider the lottery $d^{iid}(\ell)$ with $\ell(x) = \ell(y) = \frac{1}{2}$. For each $\varepsilon \in [0,1]$, define $d^{\varepsilon}(\ell) = (c_0, m)$ by setting $m_2(x|x) = \ell(x) + \frac{1}{2}\varepsilon$ and $m_2(y|y) = \ell(y) + \frac{1}{2}\varepsilon$. Then, $d^{\varepsilon}(\ell) \geq_C d^{iid}(\ell)$. Define the function $U : [0,1] \to \mathbb{R}$ by $U(\varepsilon) = V_0(d^{\varepsilon}(\ell))$. Using equation (15) and the reasoning analogous to Lemma 3, we obtain

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

which implies there exists $\bar{\varepsilon} \in (0,1)$ such that $U'(\tilde{\varepsilon}) > 0$ for all $\tilde{\varepsilon} \in (0, \bar{\varepsilon})$. Consequently,

$$V_0(d^{\bar{\varepsilon}}(\ell)) - V_0(d^{iid}(\ell)) = \int_0^{\bar{\varepsilon}} U'(\varepsilon)\,d\varepsilon > 0,$$

and thus $d^{\bar{\varepsilon}}(\ell) \geq_C d^{iid}(\ell)$ but $d^{\bar{\varepsilon}}(\ell) \succ_0 d^{iid}(\ell)$. Hence, ϕ must be concave.

Finally, suppose that ϕ does not satisfy IRRA. Since $\phi \in \mathcal{C}^3$, the function R_ϕ is continuously differentiable. Thus, there exist $\underline{z} < \bar{z}$ in $\operatorname{int} u(C)$ such that R_ϕ is non-increasing on the interval $[\underline{z}, \bar{z}]$ and $R_\phi(\bar{z}) < R_\phi(\underline{z})$. Choose $\beta \in (0,1]$ so that $\frac{\bar{z}}{1+\beta} > \underline{z}$, and set $x = \frac{\bar{z}}{1+\beta}$ and $y = \underline{z}$. Consider again the lottery $d^{iid}(\ell)$ with $\ell(x) = \ell(y) = \frac{1}{2}$, and define the lottery $d^{\varepsilon}(\ell) = (c_0, m)$ with $m_2(x|x) = \ell(x) + \frac{1}{2}\varepsilon$ and $m_2(y|y) = \ell(y) + \frac{1}{2}\varepsilon$. For $\varepsilon \geq \varepsilon'$, it holds that $d^{\varepsilon}(\ell) \geq_C d^{\varepsilon'}(\ell) \geq_C d^{iid}(\ell)$.

As before, define $U : [0,1] \to \mathbb{R}$ by $U(\varepsilon) = V_0(d^{\varepsilon}(\ell))$. Applying analogous reasoning to Lemma 3, we have

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

which implies the existence of some $\bar{\varepsilon} < 1$ such that $U'(\tilde{\varepsilon}) > 0$ for all $\tilde{\varepsilon} \in [\bar{\varepsilon}, 1)$. Hence,

$$V_0(d^1(\ell)) - V_0(d^{\bar{\varepsilon}}(\ell)) = \int_{\bar{\varepsilon}}^1 U'(\varepsilon)\, d\varepsilon > 0,$$

thus establishing $d^1(\ell) \geq_C d^{\bar{\varepsilon}}(\ell) \geq_C d^{iid}(\ell)$ but $d^1(\ell) \succ_0 d^{\bar{\varepsilon}}(\ell)$. Therefore, ϕ must satisfy IRRA. □

Proof of Corollary 1.

Proof. Observe that if we set

$$\begin{aligned} f(\varepsilon) = &\frac{1}{2}\phi\left(x + \beta\phi^{-1}\left(\phi(x)\left(\frac{1}{2} + \frac{\varepsilon}{2}\right) + \phi(y)\left(\frac{1}{2} - \frac{\varepsilon}{2}\right)\right)\right) \\ &+ \frac{1}{2}\phi\left(y + \beta\phi^{-1}\left(\phi(x)\left(\frac{1}{2} - \frac{\varepsilon}{2}\right) + \phi(y)\left(\frac{1}{2} + \frac{\varepsilon}{2}\right)\right)\right), \end{aligned}$$

and

$$\begin{aligned} g(\pi) = &\frac{1}{2}\phi\left(x(1-\pi) + \beta\phi^{-1}\left(\phi(x(1-\pi))\frac{1}{2} + \phi(y(1-\pi))\frac{1}{2}\right)\right) \\ &+ \frac{1}{2}\phi\left(y(1-\pi) + \beta\phi^{-1}\left(\phi(x(1-\pi))\frac{1}{2} + \phi(y(1-\pi))\frac{1}{2}\right)\right), \end{aligned}$$

then the equation $V_0(d^{corr}(\varepsilon)) = V_0(d^{iid}(\pi))$ is equivalent to $f(\varepsilon) = g(\pi)$.

Since $\phi \in \mathcal{C}^3$, we can take the second and first order approximations of f and g, respectively. We obtain that $f(\varepsilon) = f(1) + f'(1)(\varepsilon - 1) + f''(1)(\varepsilon-1)^2 + \mathrm{o}\left((\varepsilon-1)^2\right)$ and $g(\pi) = g(0) + g'(0)\pi + \mathrm{o}\left(\pi\right)$. Setting these expressions equal, we obtain

$$\pi(\varepsilon) = \frac{f(1) - g(0)}{g'(0)} + \frac{f'(1)}{g'(0)}(\varepsilon - 1) + \frac{f''(1)}{g'(0)}(\varepsilon-1)^2 + \mathrm{o}\left((\varepsilon-1)^2\right).$$

Furthermore, note that $f(1) = V_0(d^{corr}(1))$, $g(0) = V_0(d^{iid}(0))$, and $g'(0) < 0$, since the value of the lottery obviously is decreasing in π. The same calculations as in the proof of Theorem 1 reveal that

$$f'(1) = -\beta\frac{(\phi(x) - \phi(y))}{4}\int_y^x \frac{\phi'(z(1+\beta))}{\phi'(z)}\frac{\{R_\phi(z(1+\beta)) - R_\phi(z)\}}{z}dz,$$

and

$$f''(1) = \beta^2\frac{(\phi(x) - \phi(y))^2}{8}\left\{\frac{\phi'(x(1+\beta))}{\phi'(x)}ER_\phi(x,x) + \frac{\phi'(y(1+\beta))}{\phi'(y)}ER_\phi(y,y)\right\}.$$

which by setting $k_1 = -\frac{1}{g'(0)}$, $k_2 = -\beta\frac{(\phi(x)-\phi(y))}{4g'(0)}$, and $k_3 = \beta^2\frac{(\phi(x)-\phi(y))^2}{8g'(0)}$ implies the desired result. □

Proof of Theorem 2. *Outline of the proof.* Using a general result from 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.[8] 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 (see Dunford and Schwartz (1958), Theorem IV.8.16) to the dual of $L^\infty(p) := L^\infty(S, \Sigma, p)$ and $ca(\Sigma, p) = ca(\Sigma) \cap ba\left(\Sigma, p\right)$ is (isometrically isomorphic to) $L^1(p)$ (via the Radon-Nikodym derivative $\nu \mapsto \frac{d\nu}{dp}$).

Turning to the proof of Theorem 2, I first introduce important notions related to quasi-arithmetic certainty equivalent functionals: given $p \in \Delta(\Sigma)$, let $M_{\phi,p} : L^\infty(p) \to \mathbb{R} \cup \{-\infty\}$ be defined by

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

assuming that $\phi : \mathbb{R} \to \mathbb{R} \cup \{-\infty\}$ is non-decreasing and upper semicontinuous. I provide an important result concerning the concave conjugate $M^*_{\phi,p}$ of the quasi-arithmetic mean $M_{\phi,p}$. Recall 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 $M^*_{\phi,p} : ba(\Sigma, p) \to \mathbb{R} \cup \{\infty\}$ defined by

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

Lemma 4. *It holds that $M^*_{\phi,p} \leq 0$ and $M^*_{\phi,p}(p) = 0$ for every $p \in \Delta(\Sigma)$. Moreover, $M^*_{\phi,p}$ is upper semicontinuous and concave when $M_{\phi,p}$ is concave. Moreover, if $\phi(x) = -\infty$ for any $x < 0$, $\phi'(x) > 0$, $\phi''(x) < 0$ for every $x > 0$, and $p \in \Delta(\Sigma)$ has finite support, then the concave conjugate satisfies $M^*_{\phi,p}(q) = -\infty$ whenever $q \notin \Delta(\Sigma) \cap ca(\Sigma, p)$.*

[8] Cerreia-Vioglio et al. (2011) provide a similar representation under the assumption that ϕ is strictly increasing and concave (see their Theorem 24). 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.

Proof. Omitted. □

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

Theorem 3 (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 3, 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 3. *Assume that $\phi \in \mathcal{C}^4$ is concave and satisfies UPI over $(0, \infty)$. Then $R''_\phi \geq 0$ implies that $M_{\phi,p}|L^\infty_+(p)$ is concave. Conversely, if there exist $x \in (0, \infty)$ such that $R'_\phi(x) < 0$ and $R''_\phi(x) < 0$ then $M_{\phi,p}|L^\infty_+(p)$ is not concave.*

Proof. First observe that if ϕ satisfies UPI, then by DARA we have $A'_\phi \leq 0$. Further, it is immediately evident 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

$$xA''_\phi(x) \geq 2x\frac{(A'_\phi(x))^2}{A_\phi(x)}, \tag{16}$$

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

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

From this last condition we obtain that for every $x \in (0, \infty)$

$$-2A'_\phi(x) \geq 2x\frac{(A'_\phi(x))^2}{A_\phi(x)}. \tag{17}$$

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 (17) implies that (16) is satisfied. Hence we conclude that if ϕ satisfies SCA then $\frac{1}{A_\phi}$ is concave. The result therefore follows by Theorem 3.

Conversely, if there exist $x \in (0, \infty)$ such that $R'_\phi(x), R''_\phi(x) < 0$ we obtain that $-2A'_\phi(x) < 2x\frac{(A'_\phi(x))^2}{A_\phi(x)}$ and $xA''_\phi(x) < -2A'_\phi(x)$ which implies that (16) is violated, and so by Theorem 3 we can conclude that $M_{\phi,p}|L^\infty_+(p)$ is not concave. □

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

Proof. The proof is a simple consequence of Corollary 3 and therefore is omitted. □

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

Corollary 4. *Assume that ϕ is given by $\phi(x) = \frac{x^\lambda}{\lambda}$ for $0 \neq \lambda < 1$ 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 3. □

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

Proof of Theorem 2. Consider the utility functions $(V_t)_{t=0}^T$ from the KP representation (ϕ, u, β), 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 $V_t : D_t \to \mathbb{R}$, $t = 0, \ldots, T$ is upper semicontinuous in the weak* topology, 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 5 the function $M_{\hat{\phi},m_t}$ is concave for each $t = 0, \ldots, T-1$. By applying the Fenchel-Moreau theorem (see Phelps (2009), p. 42) and Lemma 5 it follows that

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

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

$$I^t_{\phi,u,\beta}(\ell || m_t) := \begin{cases} -M^*_{\hat{\phi},m_t}(\ell), & \\ +\infty & \text{otherwise,} \end{cases}$$

then one obtains that for every $(c, m_t) \in D_{t,s}$

$$V_t(c, m_t) = u(c) + \beta \min_{\ell \ll m_t} \left\{ \mathbb{E}_\ell V_{t+1} + I^t_{\phi,u,\beta}(\ell \| m_t) \right\}, \tag{18}$$

where the infimum is attained because $\{\ell \in \Delta_b(D_{t+1}) : \ell \ll m_t\}$ is a compact subset of $\Delta_b(D_{t+1})$. Further, observe that by Lemma 4, each is $I^t_{\phi,u,\beta}(\cdot \| m_t)$ is a convex statistical distance in the sense of Liese and Vajda (1987).

Now consider the common parametrization of Epstein-Zin preferences used in asset pricing with $\frac{1}{1-\rho} > 1$ and $\alpha < 0$ (see Bansal and Yaron 2004). In this case, one obtains (see Section 5.2 in Frittelli and Bellini 1997) that by setting $q = \frac{\alpha}{\alpha - \rho}$,

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

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

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

we obtain that for $t = 0, \ldots, T-1$

$$I_{\phi,u,\beta}(\ell \| m_t) = \mathbb{E}_\ell V_{t+1} \left[e^{\frac{1-q}{q} R_q(\ell \| m_t)} - 1 \right]. \tag{19}$$

The Rényi divergence has applications in a variety of fields, including information theory, statistics, and machine learning (see Sason 2022 for a review). This result was already observed by Meyer-Gohde, who showed that in the EZ case, the cost function can be expressed using Tsallis entropy (Rényi and Tsallis entropies are monotonic functions of each other; see, for example, Wong and Zhang 2022).

Finally, notice that if there exists $x \in (0, \infty)$ such that $R'_\phi(x) < 0$ and $R''_\phi(x) < 0$ then by Lemma 3 $M_{\phi,p} | L^\infty_+(p)$ is not concave, and so equation (18) cannot hold for every $(c, m_t) \in D_{t,s}$. $\square$

Proof of Proposition 5. It suffices to prove the following result.

Lemma 6. *If* $(c, m_n)_{n=0}^{\infty}$ *is a uniformly bounded sequence that converges to some* $d \in D_0$*, then* $\lim_n V_0(c, m_n) = V_0(d)$

Proof. Observe that

$$K \times \Delta_b(K), \tag{20}$$

is a compact set whenever K is compact (e.g., see Theorem 15.11 in Aliprantis and Border (2006)). Further observe that $V_2(c) = u(c)$ when restricted to a compact set K is continuous and therefore bounded. It follows that $V_1(c, m) = u(c) + \beta\phi^{-1}\mathbb{E}_m(\phi(V_2))$ is continuous and therefore also bounded when restricted to the set (20). Finally, since the set in (20) is compact, it follows that the set

$$\Delta_b(K \times \Delta_b(K)), \tag{21}$$

is also compact. Therefore, we can conclude that when restricted to the set in (21) the function $V_0(c, m) = u(c) + \beta\phi^{-1}\mathbb{E}_m(\phi(V_1))$ is continuous. □

Proof of Corollary 2.

Proof. Observe that the equation $V_0(d^{gradual}(\varepsilon)) = V_0(d^{early}(\pi))$ is equivalent to $f(\varepsilon) = g(\pi)$ if we set

$$\begin{aligned} f(\varepsilon) = &\frac{1}{2}\phi\left(k + \beta\,\phi^{-1}\left(\phi(x)\left(\frac{1}{2} + \frac{\varepsilon}{2}\right) + \phi(y)\left(\frac{1}{2} - \frac{\varepsilon}{2}\right)\right)\right) \\ &+ \frac{1}{2}\phi\left(k + \beta\,\phi^{-1}\left(\phi(x)\left(\frac{1}{2} - \frac{\varepsilon}{2}\right) + \phi(y)\left(\frac{1}{2} + \frac{\varepsilon}{2}\right)\right)\right), \end{aligned}$$

and

$$\begin{aligned} g(\pi) = &\frac{1}{2}\phi\left(k(1-\pi) + \beta\,\phi^{-1}\left(\phi(x(1-\pi))\frac{1}{2} + \phi(y(1-\pi))\frac{1}{2}\right)\right) \\ &+ \frac{1}{2}\phi\left(k(1-\pi) + \beta\,\phi^{-1}\left(\phi(x(1-\pi))\frac{1}{2} + \phi(y(1-\pi))\frac{1}{2}\right)\right). \end{aligned}$$

Since $\phi \in \mathcal{C}^2$, one can take the first order approximations of f and g so that $f(\varepsilon) = f(1) + f'(1)(\varepsilon - 1) + \mathrm{o}((\varepsilon - 1))$ and $g(\pi) = g(0) + g'(0)\pi + \mathrm{o}(\pi)$. Setting these expressions equal we obtain that

$$\pi(\varepsilon) = \frac{f(1) - g(0)}{g'(0)} + \frac{f'(1)}{g'(0)}(\varepsilon - 1) + \mathrm{o}((\varepsilon - 1)).$$

Furthermore, note that since $V_0(d^{gradual}(1)) = V_0(d^{early}(0))$, it follows that $f(1) = g(0)$. Moreover $g'(0) < 0$, since $V_0(d^{early}(\pi))$ obviously decreases with π. The same calculations as in the proof of Proposition 1 reveal that

$$f'(1) = \beta \frac{\phi(x) - \phi(y)}{4} \int_y^x \frac{\phi'(k + \beta z)}{\phi'(z)} ER(z, k) dz,$$

which setting $k_1 = -\beta \frac{\phi(x)-\phi(y)}{4g'(0)}$ implies the desired result. $\square$

References

Ai, H. and Bansal, R. (2018). Risk preferences and the macroeconomic announcement premium. *Econometrica*, 86(4):1383–1430.

Aliprantis, C. and Border, K. (2006). *Infinite dimensional analysis*. Springer.

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.

Andreasen, M. M. and Jørgensen, K. (2020). The importance of timing attitudes in consumption-based asset pricing models. *Journal of Monetary Economics*, 111:95–117.

Arrow, K. J. (1971). *Essays in the theory of risk-bearing*. Markham Publishing.

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.

Benabou, R. (2002). Tax and education policy in a heterogeneous-agent economy: What levels of redistribution maximize growth and efficiency? *Econometrica*, 70(2):481–517.

Blackwell, D. (1951). Comparison of experiments. In *Proceedings of the second Berkeley symposium on mathematical statistics and probability*, volume 1, page 26.

Bleichrodt, H., Rohde, K. I., and Wakker, P. P. (2008). Koopmans' constant discounting for intertemporal choice: A simplification and a generalization. *Journal of Mathematical Psychology*, 52(6):341–347.

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.

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

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. (2008). Financial markets and the real economy, in 'handbook of the equity risk premium'. *Red. R. Mehra. Elsevier, Amsterdam, Oxford*.

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.

Diamond, P. and Saez, E. (2011). The case for a progressive tax: From basic research to policy recommendation. *Journal of Economic Perspectives*, 25(4):165–190.

Dillenberger, D. (2010). Preferences for one-shot resolution of uncertainty and allais-type behavior. *Econometrica*, 78(6):1973–2004.

Dillenberger, D., Gottlieb, D., and Ortoleva, P. (2024). Stochastic impatience and the separation of time and risk preferences. *Theoretical Economics*. Forthcoming.

Dunford, N. and Schwartz, J. T. (1958). *Linear operators, part 1: general theory*, volume 7. John Wiley & Sons.

Durlauf, S. N. (1996). A theory of persistent income inequality. *Journal of Economic growth*, 1:75–93.

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.

Gottschalk, P. and Spolaore, E. (2002). On the evaluation of economic mobility. *The Review of Economic Studies*, 69(1):191–208.

Grant, S., Kajii, A., and Polak, B. (1998). Intrinsic preference for information. *Journal of Economic Theory*, 83(2):233–259.

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. (2023). Intertemporal hedging and trade in repeated games with recursive utility. *Econometrica*, 91(6):2333–2369.

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

Leobacher, G. and Steinicke, A. (2022). Exception sets of intrinsic and piecewise lipschitz functions. *The Journal of Geometric Analysis*, 32(4):118.

Liese, F. and Vajda, I. (1987). *Convex statistical distances.* Teubner, Leipzig.

Mehra, R. and Prescott, E. C. (1985). The equity premium: A puzzle. *Journal of monetary Economics*, 15(2):145–161.

Meissner, T. and Pfeiffer, P. (2022). Measuring preferences over the temporal resolution of consumption uncertainty. *Journal of Economic Theory*, 200:105379.

Merton, R. C. (1971). Optimum consumption and portfolio rules in a continuous-time model. *Journal of Economic Theory*, 3(4):373–413.

Meyer-Gohde, A. (2019). Generalized entropy and model uncertainty. *Journal of Economic Theory*, 183:312–343.

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.

Park, K. Y. (2009). Genes and social mobility: A case for progressive income tax. *Global Economic Review*, 38(4):431–459.

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

Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica*, 32(1-2):122–136.

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

Rohde, K. I. and Yu, X. (2022). Intertemporal correlation aversion-a model-free measurement. *Available at SSRN*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4079812.

Rohde, K. I. and Yu, X. (2024). Intertemporal correlation aversion—a model-free measurement. *Management Science*, 70(6):3493–3509.

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.

Shorrocks, A. F. (1978). The measurement of mobility. *Econometrica*, pages 1013–1024.

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.

Werner, J. (2024). Ordinal representations and properties of recursive utilities. Working Paper, University of Minnesota.

Wong, T.-K. L. and Zhang, J. (2022). Tsallis and rényi deformations linked via a new λ-duality. *IEEE Transactions on Information Theory*, 68(8):5353–5373.

Supplemental Appendix

This supplemental material contains two parts. Section S.1 extends the main results to the case of an infinite horizon. Section S.2 provides a behavioral characterization of strong correlation aversion. Section S.3 provides a proof of Lemma 3.

S.1 The case $T = \infty$

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. Here I consider the case of an infinite time horizon. As the consumption set $C = [0, \infty) = \mathbb{R}_+$ is identical to that of Epstein and Zin (1989), I follow their approach in introducing the set of temporal lotteries for the case of an infinite horizon, with specific reference to their discussion on pages 940-944. The only deviation in my approach is the use of $\Delta(X)$ to denote the set of Borel probabilities defined on a metric space X (equipped with the weak* topology). For every $b \geq 1$ and $l > 0$, the sets of temporal lotteries $D(b; l)$ and $D(b)$ are defined in equations 2.3 and 2.5. Moreover, equations 2.2-2.11 define all the relevant objects. I also make use of their characterization of temporal lotteries in $D(b)$.

Theorem S.1 (Theorem 2.2 in Epstein and Zin (1989))**.** *For every $b \geq 1$ we have that*

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

where

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

Because of this result, each $d \in D(b)$ can be identified with $(c, m) \in C \times \hat{\Delta}(D(b))$. Further, each $m \in \hat{\Delta}(D(b))$ can be equivalently identified with an element of

$$\hat{\Delta}(C \times \hat{\Delta}(D(b))).$$

Preferences are given by a weak order $\succeq$ over $D(b)$. The utility function $V : D(b) \to \mathbb{R}$ is called recursive if it satisfies the following equation for every $(c, m) \in C \times \hat{\Delta}(D(b))$,

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

where $\phi : [0, \infty) \to \mathbb{R}$. The next result shows that (1) always has a solution, thus making recursive utility well defined in this context.

Theorem S.2. *Suppose that ϕ is concave, $\rho > 0$ and $\beta b^\rho < 1$. Then there exists a $V : D(b) \to \mathbb{R}$ that satisfies (1).*

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

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

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

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

Let V^* be a continuous function such that

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

which exists uniquely by Theorem 3.1 in Epstein and Zin (1989) since $\rho > 0$ and $\beta b^\rho < 1$.

Let $T^0(V^*) = T(V^*)$ and $T^n(V^*) = T(T^{n-1}(V^*))$. By Jensen inequality

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

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

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

I now claim that V solves (1). Since

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

the statement follows by the fact that

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

□

Consider now a weak order $\succeq$ over $D(b)$. Say that $\succeq$ admits a KP representation (ϕ, ρ, β) if there exists $V : D(b) \to \mathbb{R}$ that satisfies (1) and such that represents $\succeq$. For every $d \in D(b)$ one can define the present equivalent $PE_{\succeq}(d)$ as the unique single period consumption level $c \in C$ such that $d \sim (c, \mathbf{0})$, where $\mathbf{0} \in D(b)$ is the temporal lottery that pays the constant zero level of consumption at every time period. Note that $PE_{\succeq}(d)$ is well defined since $V(c, \mathbf{0}) = c$.

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

Definition S.1 (Correlation order with $T = \infty$)**.** *Fix a weak order $\succeq$ over $D(b)$. Consider $d = (c, m), d' = (c, m') \in D(b)$. Say that d is more correlated than d', written $d \geq_C d'$, if*

$$m_{\succeq}, m'_{\succeq} \in \Delta_s(C \times \Delta_s(C)) \text{ and } (c, m_{\succeq}) \geq_C (c, m_{\succeq}).$$

Correlation aversion can then be defined as in the main text, where now $d^{iid}(\ell) = (c, m)$ denotes a temporal lottery such that $(c, m_{\succeq}) = d^{iid}(\ell)$ for some $\ell \in \Delta_s(C)$.

Definition S.2 (Correlation aversion with $T = \infty$)**.** *Say that $\succeq$ exhibits correlation aversion if and only if for every $l > 0$ and $d, d' \in D(b)$*

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

The main results of the paper carry over in the same way. Notice that due to the stationary setting, here there is a unique cost function $I_{(\phi,u,\beta)}(\cdot, \cdot)$.

Theorem S.3. *Consider $\phi \in \mathcal{C}^3$ that is concave and satisfies UPI. Then every $\succeq$ with KP representation (ϕ, ρ, β) exhibit correlation aversion if and only if ϕ satisfies*

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

IRRA. Further, if $\succeq$ admits a KP representation (ϕ,ρ,β) with $\phi \in \mathcal{C}^4$ that additionally satisfies SCA, then $\succeq$ admits the representation for every $(c,m) \in C \times \left(\hat{\Delta}(D(b)) \cap \Delta_s(D(b))\right)$

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

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

Proof. The proof follows the same steps as the proof of Theorems 1 and 2. □

To better understand the previous result, it is helpful to examine its implications. The following result demonstrates that if preferences satisfy this notion of correlation aversion, they will always prefer an iid lottery over a perfectly correlated one, where an iid lottery and a perfectly correlated lottery are straightforward extensions of those considered in the main text.

Formally, given $\ell \in \Delta_s(C)$, consider the perfectly correlated lottery $(c_0, m^{corr}(\ell))$, where $m^{corr} \in \Delta(C^\infty)$ satisfies $m^{corr}(c,c,\dots) = \ell(c)$ for every $c \in C$, and the iid lottery $(c_0, m^{iid}(\ell))$, where

$$m^{iid}(\ell)(c, m^{iid}(\ell)) = \ell(c) \text{ for every } c \in C.$$

These lotteries generalize the notions of iid lotteries and perfectly correlated lotteries from the case $T = 2$ to the case $T = \infty$.

Proposition S.1. *Consider preferences $\succeq$ with a KP representation (ϕ,ρ,β) such that ϕ satisfies IRRA, UPI and $\phi \in \mathcal{C}^3$. Then*

$$(c_0, m^{iid}(\ell)) \succeq (c_0, m^{corr}(\ell)),$$

for every $c_0 \in C$ and $\ell \in \Delta_s(C)$.

Proof. Observe that

$$\left(V\left(c_0, m^{corr}(\ell)\right)\right)^\rho = c_0^\rho + \lim_{T\to\infty} \beta\phi^{-1}\left(\sum_{c\in\operatorname{supp}\ell} \ell(c)\phi\left(\sum_{t=0}^{T-1}\beta^t c^\rho\right)\right)$$

and

$$\left(V\left(c_0, m^{iid}(\ell)\right)\right)^\rho = \lim_{T\to\infty} V_T\left(c_0, \ell\right),$$

where $V_0(c,\ell) = c^\rho$ and, recursively,

$$V_t(c,\ell) = c^\rho + \beta\phi^{-1}\left(\sum_{c'\in\operatorname{supp}\ell} \ell(c')\phi\Big(V_{t-1}(c',\ell)\Big)\right) \quad \text{for } t = 1,\ldots,T.$$

Since the preferences $\succeq$ satisfy correlation aversion, Theorem S.3 implies that ϕ satisfies IRRA. Additionally, by assumption, ϕ satisfies UPI and therefore DARA (Proposition 2). Hence, by Proposition 6 and Theorem 12 in Marinacci and Montrucchio (2010), the functional $(x_i)_{i=1}^n \mapsto \phi^{-1}\left(\sum_{i=1}^n \phi(x_i)\, q_i\right)$ is constant superadditive and subhomogeneous. Therefore, by repeatedly applying these results, we have that for every $T \geq 2$ it holds

$$V_T(c_0,\ell) - c_0^\rho \geq \sum_{t=0}^{T-1} \beta^t \beta\phi^{-1}\left(\sum_{c\in\operatorname{supp}\ell} \ell(c)\phi(c^\rho)\right) \geq \beta\phi^{-1}\left(\sum_{c\in\operatorname{supp}\ell} \ell(c)\phi\left(\sum_{t=0}^{T-1}\beta^t c^\rho\right)\right).$$

Consequently, since

$$\left(V\left(c_0, m^{iid}(\ell)\right)\right)^\rho = \lim_{T\to\infty} V_T(c_0,\ell) \geq \lim_{T\to\infty}\left(c_0^\rho + \beta\phi^{-1}\left(\sum_{c\in\operatorname{supp}\ell} \ell(c)\phi\left(\sum_{t=0}^{T-1}\beta^t c^\rho\right)\right)\right),$$

and

$$\lim_{T\to\infty}\left(c_0^\rho + \beta\phi^{-1}\left(\sum_{c\in\operatorname{supp}\ell} \ell(c)\phi\left(\sum_{t=0}^{T-1}\beta^t c^\rho\right)\right)\right) = \left(V\left(c_0, m^{corr}(\ell)\right)\right)^\rho,$$

it follows that

$$(c_0, m^{iid}(\ell)) \succeq (c_0, m^{corr}(\ell)),$$

as desired. □

S.2 Axiomatic foundation of SCA

Consider again the setting of Section 3. In order to provide an axiomatic foundation of SCA, I introduce the notion of *Correlation Aversion Attenuation* (CAA) transformation. This integral operator modifies a given function ϕ to produce a new function, denoted by $CAA(\phi)$, which attenuates correlation aversion. Formally, the CAA transformation is a nonlinear integral operator

$$CAA : \mathcal{C}^4 \to \mathcal{C}^3,$$

defined for every $\phi \in \mathcal{C}^4$ by:

$$CAA(\phi)(x) = \int_1^x \exp\left(-\int_1^t \frac{R'_\phi(s)}{s}\, ds\right) dt.$$

This integral transform attenuates correlation aversion in that it "flattens" the index of relative risk aversion R_ϕ, as I illustrate in the next example.

Example S.1. Given $\theta \in (0,1) \cup (1,\infty)$, let

$$\phi(x) = -\exp\left(-\frac{x}{\theta}\right).$$

Relative risk aversion is $R_\phi(x) = \frac{x}{\theta}$. Applying the CAA operator, we obtain

$$CAA(\phi)(x) = \frac{x^{1-\frac{1}{\theta}} - 1}{1 - \frac{1}{\theta}}.$$

In this case, relative risk aversion is flat at the level $R_{CAA(\phi)}(x) = \frac{1}{\theta}$. Applying the CAA operator once again yields:

$$CAA^2(\phi)(x) = x - 1,$$

which satisfies $R_{CAA^2(\phi)}(x) = 0$. Hence, repeated applications of the CAA operator progressively attenuate correlation aversion by flattening the index of relative risk aversion. △

Consider preferences $\succeq$ that admit a KP representation (ϕ, u, β), where $\phi \in \mathcal{C}^4$. Let $\succeq_{CAA}$ denote preferences with the KP representation $(CAA(\phi), u, \beta)$.

Definition S.3 (Strong correlation aversion)**.** *Preferences $\succeq$ exhibit strong correlation aversion if both $\succeq$ and $\succeq_{CAA}$ exhibit correlation aversion.*

Therefore, this notion of strong correlation aversion requires that preferences exhibit correlation aversion even after risk attitudes are adjusted to attenuate correlation aversion.

Proposition S.2. *Assume that $CAA(\phi)$ satisfies UPI. Every preference relation $\succeq$ with KP representation (ϕ, u, β) exhibit strong correlation aversion if and only if ϕ satisfies SCA.*

Proof. Straightforward calculations show that

$$R_{CAA(\phi)}(x) = R'_\phi(x).$$

Therefore since $CAA(\phi) \in \mathcal{C}^3$ and $CAA(\phi)$ satisfies UPI, the result follows by Theorem 1. □

This result shows that this behavioral notion of strong correlation aversion is effectively the behavioral counterpart of SCA.

S.3 Proof of Lemma 3

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

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

where for some $\underline{i}, \underline{j}$ it holds that $p^{\varepsilon}_{\underline{j}\underline{i}} = p_{\underline{j}\underline{i}} - p_{\underline{i}\underline{j}}\varepsilon$, $p^{\varepsilon}_{\underline{i}\underline{i}} = p_{\underline{i}\underline{i}} + p_{\underline{i}\underline{j}}\varepsilon$, $p^{\varepsilon}_{\underline{i}\underline{j}} = p_{\underline{i}\underline{j}} - p_{\underline{j}\underline{i}}\varepsilon$, $p^{\varepsilon}_{\underline{j}\underline{j}} = p_{\underline{j}\underline{j}} + p_{\underline{j}\underline{i}}\varepsilon$, and otherwise $p_{ji} = m_2(c_j|c_i)$ for every other j, i. Clearly, the function U defined in this manner is twice continuously differentiable and satisfies condition (1) of the statement.

To prove point (2), observe that in this case we have that for some $p, q \in (0,1)$, $k \in \phi(u(C))$ and $x, y \in u(C)$ with $x > y$

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

where the last inequality follows by the fact that ϕ is strictly increasing and concave.

Finally, to prove point (3), observe that the functions

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

and

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

are convex by Lemma 1 in the main text. Then we obtain

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

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

Bibliography

References

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.

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