The Importance of Unlikely Events*

Maya Eden
University of Zurich

Larry Samuelson
Yale University

Jakub Steiner
University of Zurich, CERGE-EI, and CTS

March 24, 2026

Abstract

We study an expected-utility maximizer who controls a stochastic growth process over a long horizon. Using large deviations theory, we show that optimal actions are generically driven by responses to arbitrarily unlikely contingencies. Unrealistic fears of ruin preclude extraordinary wealth, while unrealistic hopes for extraordinary wealth induce choices that almost surely disappoint. We show that a CRRA investor assigns zero value to perfect information at the exponential (growth-rate) scale: she will not sacrifice even an arbitrarily small fraction of long-run growth rate to learn the frequency of future economic shocks. This extends to broader utilities with hedging.

*We thank Pavel Kocourek for superb research assistance and Ian Martin, Meg Meyer, Antony Millner, Peter Bossaerts, and several seminar audiences for helpful comments.

1 Introduction

A long-horizon investor—perhaps a young person choosing a pension fund or a university endowment manager—faces an overwhelming number of possible future scenarios. The law of large numbers ensures that the realized frequencies of (for example) booms and busts will almost certainly be arbitrarily close to their expected frequency. However, an expected-utility maximizer recognizes that atypical scenarios, such as an unusually high frequency of busts (or booms), can have such a large impact on terminal wealth that they cannot be neglected, even though they are vanishingly rare.

This paper examines long-horizon growth problems in which histories arise naturally that can be critical for expected-utility maximization, despite occurring with vanishing probability. Our structure covers many environments in finance and beyond: the choice of a pension fund, a government choosing growth policy, or a demographic setting in which population growth is stochastic and "ruin" corresponds to extinction. The common elements of the applications are that decisions are made in advance, shocks accumulate over time, and long-run performance is governed by shock frequencies.

Using techniques from large deviations theory, we establish a *dominating frequency principle*. Out of the astronomical number of possible future shock frequencies, the decision-maker needs to track—and optimally respond to—only one or two. We thus share with the law-of-large-numbers argument the conclusion that only a tiny set of shock frequencies matter. However, for a generic utility function, the frequencies that matter are *atypical*. The probability of observing them vanishes as the horizon grows, yet they are pivotal because they generate large utility or disutility. Expected-utility maximization thus requires the manager of a long-horizon growth process to put aside the law of large numbers and focus on unlikely scenarios. We begin on relatively familiar ground by studying decision-makers with CRRA utility. We then move beyond CRRA to a broad class of utilities that we call two-tailed. These utilities behave like CRRA in the far left and far right tails, but with different curvatures, allowing distinct risk attitudes toward ruin and toward extraordinary wealth.

Our central results identify the frequencies that are relevant for long-run growth as the solution to a variational representation that trades off the rarity of a scenario against the utility level it induces. As a result, the complexity of controlling stochastic growth collapses asymptotically. For CRRA expected utility maximizers with moderate risk aversion, long-horizon expected utility is dominated by a single frequency that is more optimistic than the law-of-large-numbers benchmark. For higher degrees of risk aversion, long-horizon expected utility is dominated by a single, more pessimistic frequency. For two-tailed utilities, we characterize optimal choice via a *ruin-robust growth program*. Intuitively, it can be interpreted as a disciplined "hope–fear" tradeoff. The decision-maker chooses an action to benefit from a favorable rare shock frequency, while limiting exposure to adverse rare frequencies so that the fear of ruin does not dominate the hope for growth. This ruin-robust growth program pins down the two relevant large deviations and the optimal action.

Our analysis has stark economic implications. We establish a *long-run irrelevance of information* result for CRRA decision-makers: access to a perfect forecast of future shock frequencies does not raise expected utility at the exponential (growth-rate) scale. This follows from our characterization—once long-run expected-utility maximization collapses to a single dominating shock frequency, the decision-maker already "knows" this frequency, so foresight has no first-order effect. Similarly, access to hedging opportunities does not improve the CRRA objective at this scale. For our two-tailed class of utilities, the value of information remains zero at the exponential scale once the decision-maker can implement a simple hedging strategy.

2 Literature

Our study of long-run stochastic growth control recalls a debate between John Kelly and Paul Samuelson. Kelly's (1956) starting point in the development of his influential betting principle was that the future is essentially predictable. The law of large numbers ensures that the proportion of good and bad years over a long horizon hardly ever surprises; accordingly, Kelly advised, the long-run investor should optimize for the typical future evolution of shocks. Samuelson (1971, 1979) disagreed, demonstrating by examples that a generic expected-utility maximizer will not comply with Kelly's advice. Instead, very rare but highly payoff-consequential shock frequencies may be pivotal for optimal choice, rendering the law of large numbers an inappropriate guide.

We operationalize Samuelson's critique by solving for the expected-utility-maximizing portfolios. Kelly portfolios—though suboptimal for generic utilities—serve as basic building blocks for our characterization. A long-horizon CRRA investor chooses a Kelly portfolio optimized for a distorted (rather than typical) shock distribution, and an investor with two-tailed utility hedges across two Kelly portfolios, fine-tuned to optimistic and pessimistic distortions of the shock distribution.

Barro (2006, 2009) and Rietz (1988) (see Barro and Ursúa (2012) for a survey) emphasize that rare disasters can be quantitatively important for an expected-utility maximizer. Barro and Jin (2011) argue that the prospect of rare disasters can account for the equity premium puzzle. These models typically assume an exogenous disaster state that occurs with a fixed probability. For example, in Barro and Jin, shocks are drawn independently each period from a normal distribution capturing ordinary fluctuations, together with a second component that is usually zero but occasionally produces a catastrophe. In contrast, we do not single out any particular disaster state. Instead, we treat the realized empirical distribution of states as the relevant object, so that a continuum of atypical distributions can arise, and then characterize the atypical distribution that drives expected-utility maximization.

Weitzman (1998) and Gollier and Weitzman (2010) show that when shocks to the stochastic discount factor have a perfectly persistent component, the far-distant future should be discounted at the lowest feasible rate: the expected discount factor is dominated by the most patient realizations. Their insight is a special case of the Laplace principle, which states that a sum of exponentials is dominated by the term with the largest exponent. We too rely on the Laplace principle. In our setting, however, shocks do not persist forever, so the dominant contribution is not an extreme "most patient" regime, but rather an atypical empirical distribution of future shocks. Generically, the resulting empirical frequency is interior.

We build on the approach of Robson et al. (2023), Samuelson and Steiner (2025), and Millner (2025), who study stochastic growth using tools from large deviations theory. The first two papers study hedging in the log-utility setting under i.i.d. shocks. In a significant generalization, Millner (2025) allows for general serial correlation in the state-generating process and works with general rate functions. He characterizes present values of long-horizon investments, a setting which formally corresponds in our framework to CRRA utility with a particular coefficient of relative risk aversion. All three papers derive variational characterizations of optimal actions: Millner emphasizes the connection to ambiguity aversion, while Robson et al. derive a rational-inattention representation for hedging and log utility, and relate the risk-neutral case to the wishful-thinking representation of Caplin and Leahy (2019). The current paper allows for a general utility function, characterizes how utility curvature shapes the variational representation of optimal long-run investment, and derives implications for the value of information and hedging.

Using related tools, Stutzer (2003) employs large deviations theory to show that a non-expected-utility criterion—minimizing the probability of rare underperformance—is, in the long run, equivalent to expected-utility maximization with CRRA utility. Hansen and Scheinkman (2009) provide an operator-based eigenvalue characterization of long-run valuation that is conceptually analogous to our large-deviations characterization of the distortion that dominates long-horizon expected utility.

Weitzman (2009) drew attention to the importance of rare events for expected-utility maximization under structural uncertainty. In Weitzman, rare events become prominent because structural uncertainty generates fat tails. In our setting, rare events arise in the absence of structural uncertainty and matter without fat tails, due to the cumulative effects of multiplicative growth.

Millner (2013) examines responses to Weitzman that have appeared in the literature. Millner argues that the weakest point of Weitzman's argument is the reliance on CRRA utility. Indeed, it remains an open empirical question whether CRRA utility adequately represents people's attitudes toward risk, and there is also a normative question about the risk attitude policymakers should adopt (e.g., Fleurbaey (2010); Thoma (2023)). Nonetheless, given the dominance of CRRA specifications in macroeconomic analysis, it is useful to understand their implications in long-run optimization problems. Our analysis therefore begins with CRRA utility, clarifying its behavioral implications in the context of long-term investment. We then move beyond CRRA to utilities that are unbounded both above and below.

The robust-control approach of Hansen and Sargent (2001) and the variational preferences of Maccheroni et al. (2006) represent a decision-maker's concern about model misspecification. In our work, related variational representations arise for an expected-utility maximizer with no such concerns; instead, the representation captures the decision-maker's reasoning about rare events. Because we derive this representation from primitive properties of the utility function, we can link risk aversion to the decision-maker's taste for or aversion to rare events (an object that corresponds to ambiguity concerns in the standard literature).

Using an identity from large deviations theory, Strzalecki (2011) establishes an observational equivalence between expected-utility theory and the robust-control representation of Hansen and Sargent (2001). This equivalence corresponds to the special case of our setting with CRRA utility and i.i.d. shocks. By allowing for serial correlation and general utility functions, we generalize this connection and derive new applied results in the context of long-run growth control.

We do not take a normative stance on the expected-utility maximizer's focus on rare events. On the one hand, Chichilnisky (2009) argues that fear of catastrophes should be incorporated into the standard expected-utility framework. We show that fear and hope regarding rare events arise naturally within expected utility as the decision horizon grows. On the other hand, Russell (2024) views decision-makers whose choices are driven by negligible probabilities of enormous rewards as fanatics.

3 Model

A decision-maker, often abbreviated as DM, chooses an action that affects the stochastic growth of wealth (our running interpretation). The DM is born with one unit of initial wealth and, in period 0, chooses an action a from the compact and convex action space $A \subset \mathbb{R}^m$ once and for all. This once-and-for-all decision encompasses finite-memory Markov decision plans, accommodated by augmenting the payoff state with the relevant history and interpreting a as a stationary policy, such as continual portfolio rebalancing as a function of finitely many past returns.

The chosen action and the sequence of payoff states $\theta^T = (\theta_1, \ldots, \theta_T)$, where each θ_t takes values in a finite set Θ, determine the DM's wealth w_T at horizon T as follows. The DM's initial endowment, normalized to 1, is multiplied by the gross return $R(a, \theta_t)$ in each period t yielding terminal wealth:

$$w_T(a, \theta^T) = \prod_{t=1}^{T} R(a, \theta_t) = \exp\left[\sum_{t=1}^{T} r(a, \theta_t)\right], \tag{1}$$

where we introduce the log-return function $r(a, \theta) = \ln R(a, \theta)$ for the second equality.

We will make extensive use of the *empirical distribution* $q_T \in \Delta(\Theta)$, where

$$q_T(\theta) = \frac{1}{T} \sum_{t=1}^{T} \mathbf{1}_{\theta_t = \theta}$$

is the fraction of periods in which state $\theta \in \Theta$ occurs in the sequence θ^T, regardless of order. In our setting, terminal wealth depends on the state sequence θ^T only through its empirical distribution q_T. For example, the terminal wealth of an investor who reinvests her returns will depend on the shares of good and bad years during her investment period, but not on their order.[1] We then abuse notation and write the terminal wealth $w_T(a, q_T)$ as a function of action and state frequencies and rewrite (1) to obtain:

$$w_T(a, q_T) = \exp\left[\sum_{\theta \in \Theta} q_T(\theta) r(a, \theta) T\right] = \exp[r(a, q_T)T],$$

where the final equality simply extends the log-return function to $r(a, q) := \sum_\theta q(\theta) r(a, \theta)$ to make the notation more concise.

[1] This assumption that payoffs are determined by the *frequencies* alone rules out some settings, such as that of a gambler who follows a martingale betting system (e.g., initially bet one dollar, double the stake after each failure, and quit after the first success).

It follows from the definition that the log return $r(a, q)$ is linear in q. We assume that $r(a, q)$ is concave and continuous in a. Concavity in a captures familiar diminishing returns: diversifying by mixing actions weakly increases the average log return.

The DM has a prior belief about the stochastic state sequence θ^T, allowing for serial correlations. Since the empirical distribution q_T is a sufficient statistic for payoff evaluation, we only track the DM's period-0 belief over q_T, denoted by π_T.

The DM maximizes the expected utility of her wealth w_T and thus solves

$$\max_{a \in A} \mathrm{E}\, u\big(w_T(a, q_T)\big), \tag{2}$$

where the expectation is with respect to $q_T \sim \pi_T$, and u is a continuous and strictly increasing utility function. We seek to characterize the optimal action and the associated value for large but finite horizons T.[2]

Our focus on utility over terminal wealth isolates the role of long-horizon uncertainty.[3] It can be interpreted as a subproblem of a broader intertemporal allocation problem. Consider a DM who maximizes a discounted sum of flow utilities and, at date 0, allocates her initial endowment across future periods and contingencies. Our analysis concerns the investment of the resources earmarked for a distant date and contingency.

This setting accommodates a variety of interpretations. For a start, consider an example from **finance**. An investor chooses a *constant-weight rebalanced portfolio* of assets, where each asset j from a finite set J has log returns $r(j, \theta)$ in each state θ. Formally, a portfolio is an action $a = (a_j)_{j \in J} \in A \equiv \Delta(J)$. The investor chooses a at $t = 0$ and commits to automatic periodic rebalancing of the investments so that she invests the share a_j of her current wealth into asset j at each t.[4] A portfolio a has a log return $r(a, \theta) = \ln \sum_j a_j \exp r(j, \theta)$, which is concave in a as assumed.

The model can characterize the **valuation of long-term investments**, which is a focus of Martin (2012) and Millner (2025). These papers let $R(a, \theta_t) = M(a, \theta_t) R'(a, \theta_t)$, where M is the stochastic discount factor and R' is the gross return and they focus on $\mathrm{E} \prod_{t=1}^T R(a, \theta_t)$.

Alternatively, the model captures a **public economics** problem in which a policy a generates the economic growth rate $r(a, \theta)$ in state θ. It also captures a **population dynamics** interpretation in which w_T denotes population size at horizon T and $R(a, \theta_t)$ is the gross

[2] An optimal action exists since $r(a, \theta)$ is continuous in a, u is continuous, and the action set A is compact.

[3] Optimization of expected utility over terminal wealth is a common abstraction in theoretical finance; see e.g. Martin and Papadimitriou (2022), Walden (2019), and Chabakauri et al. (2022).

[4] We can interpret a as an instruction to the investor's broker (e.g., "Keep 70% of my wealth in S&P500 and 30% in US bonds until retirement."), in line with the standard advice to lay investors.

demographic growth factor.

4 Tools

Adaptations of two standard objects—probability distributions and certainty equivalents—prove useful for studying long-horizon stochastic growth.

4.1 Rate Functions

An expected-utility maximizer considers events even if these are astronomically unlikely, as long as their implications for her utility are large enough. To describe how such "black swans" arise in long-horizon environments, we turn to large deviations theory (cf. Dembo and Zeitouni (1998) and den Hollander (2000)). We assume that there exists a *rate function* $I : \Delta(\Theta) \to \mathbb{R}_+ \cup \{+\infty\}$ such that for any small neighborhood $N(q)$ of a distribution q, the probability

$$\pi_T\big(q_T \in N(q)\big) \approx \exp[-I(q)T], \tag{3}$$

for large T.[5] Thus, when $I(q) > 0$, the event that $q_T \approx q$ becomes increasingly rare at the exponential rate $I(q)$ as T grows. We refer to an empirical distribution q with $I(q) > 0$ as a *large deviation*, emphasizing its rarity, whereas any p with $I(p) = 0$ is a *typical distribution.*

The rate function depends only on the collection of distributions $\{\pi_T\}_{T=1}^{\infty}$ of q_T induced by the state-generating process. Ascertaining the rate function thus requires only a coarse description of the state-generating process. Rate functions are known for many stochastic processes, including i.i.d. draws, general Markov chains, and mixing/ergodic processes. An applied researcher working with, for example, a Markov model can then estimate the parameters of the transition matrix and compute the implied rate function $I(q)$.

Example 1 (Rate Functions: i.i.d. and Markov)**.** When the states are i.i.d. draws from a full-support distribution $p(\theta)$, Sanov's theorem (Dembo and Zeitouni, 1998, p. 16) indicates that $I(q) = \mathrm{KL}(q \parallel p)$, where $\mathrm{KL}(q \parallel p) = \mathrm{E}_q \ln \frac{q(\theta)}{p(\theta)}$ is the Kullback-Leibler divergence. The KL divergence is a (non-symmetric) pseudo-distance measuring how close the empirical distribution q is to the generating distribution p. In line with the law of large numbers, the typical empirical distribution equals the state-generating process p (i.e., $\mathrm{KL}(p \parallel p) = 0$),

[5] More precisely, we assume that the family of probability measures $(\pi_T)_{T=1}^{\infty}$ satisfies the large deviation principle with rate function I. This requires that, for any closed set $F \subseteq \Delta(\Theta)$ and any open set $G \subseteq \Delta(\Theta)$, $\limsup_{T\to\infty} \frac{1}{T} \ln \pi_T(F) \leq -\inf_{q\in F} I(q)$ and $\liminf_{T\to\infty} \frac{1}{T} \ln \pi_T(G) \geq -\inf_{q\in G} I(q)$.

and greater departures from the generating process, measured by this rate function, become increasingly rare.

To allow for serial correlations, consider a Markov chain with two states 1 and 2 and the transition matrix

$$\mathbf{M} := \begin{bmatrix} 1-\rho p_2 & \rho p_2 \\ \rho p_1 & 1-\rho p_1 \end{bmatrix} \tag{4}$$

with strictly positive p_1 and p_2, $p_1 + p_2 = 1$. In each period, with probability ρ, the state is perturbed and redrawn from p. When $\rho = 1$, states are drawn i.i.d., while the process becomes more persistent as ρ shrinks, approaching perfect persistence as ρ approaches zero. The stationary distribution of this Markov chain is p regardless of the value of the persistence parameter ρ. The rate function is given by (Dembo and Zeitouni, 1998, Theorem 3.1.6, p. 76)

$$I(q) = \max_{s \in \mathbb{R}^2_{++}} \sum_{j=1}^{2} q_j \ln \frac{s_j}{s_1 \mathbf{M}_{1j} + s_2 \mathbf{M}_{2j}}, \tag{5}$$

which reduces to $\mathrm{KL}(q \parallel p)$ for the independent-draws case $\rho = 1$. For all ρ, $I(p) = 0$, giving the natural result that the stationary distribution is typical. This rate function has natural comparative statics with respect to the persistence parameter: for $q \neq p$, $I(q)$ increases in ρ. The more persistent the state process is, the less rare large deviations are. Intuitively, high persistence is akin to taking fewer effectively independent draws over any finite horizon T, allowing π_T to exhibit greater dispersion. ▲

We impose the regularity condition that the rate function $I(q)$ be convex, and hence continuous. This convexity holds for i.i.d. processes, Markov chains, and a broad class of stationary processes with sufficiently fast decay of serial dependence; see Bryc and Dembo (1996). The intuition is as follows. An empirical distribution q over a sequence of length T can be obtained by two consecutive blocks of lengths λT and $(1-\lambda)T$, whose empirical distributions are q_1 and q_2, respectively, so that $q = \lambda q_1 + (1-\lambda) q_2$. Stationarity, together with approximate independence between the two blocks, then yields $I(\lambda q_1 + (1-\lambda) q_2) \leq \lambda I(q_1) + (1-\lambda) I(q_2)$.

Finally, we assume that a typical distribution, at which the rate function attains its minimum value 0, exists.[6]

[6] We allow for multiple typical distributions. This arises, for instance, under a hierarchical state-generating process with a distribution p drawn once and for all from a distribution on a convex set $P \subset \Delta(\Theta)$, and states subsequently generated i.i.d. from p. The induced rate function then satisfies $I(p) = 0$ for all $p \in P$.

4.2 Certainty Equivalents

To characterize the solution to the DM's problem for large horizons T, we adapt the standard definition of the certainty equivalent to long-term growth.[7]

The usual certainty equivalent of a lottery is the deterministic wealth level that leaves an expected-utility maximizer indifferent to the lottery. Analogously, for each action a, we define a (growth-rate) *certainty equivalent* $C_T(a) \in \mathbb{R}$ as the deterministic per-period growth rate of wealth that generates the same expected utility as the action a at horizon T. It solves[8]

$$\mathrm{E}\, u\left(w_T(a, q_T)\right) = u\left(\exp\left[C_T(a)T\right]\right),$$

where the expectation is with respect to $q_T \sim \pi_T$. The maximization of the expected utility in (2) is then equivalent to maximizing $C_T(a)$.

We approximate the optimal action for large horizons T by solving an asymptotic version of the DM's problem that maximizes the limiting certainty equivalent as $T \to \infty$. The advantage is that certainty equivalents typically admit well-behaved limits, unlike expected utility, which may diverge. The main text therefore focuses on the asymptotic regime. Appendix B briefly discusses finite-horizon corrections in several tractable special cases.

The *asymptotic certainty equivalent* of the action a is

$$C(a) = \lim_{T \to \infty} C_T(a),$$

whenever this limit exists. It is the per-period deterministic growth rate of wealth that matches the action's expected utility at the exponential scale, asymptotically. We prove existence and characterize the asymptotic certainty equivalents in a broad class of settings below.[9]

While the ranking of the finite-horizon certainty equivalents $C_T(a)$ completely represents the expected-utility preferences, the asymptotic certainty equivalent represents them only partially. If $C(a) > C(a')$, then the DM strictly prefers a to a' for all sufficiently large horizons T. However, $C(a) = C(a')$ does not imply $a \sim_T a'$ for finite T. For instance, two actions whose wealth paths differ by a multiplicative constant have the same asymptotic certainty equivalent, yet the DM strictly prefers one to the other for every T. Equality of asymptotic certainty equivalents instead yields a weaker implication: if $C(a) = C(a')$, then

[7] An alternative is to work with an overtaking criterion (e.g., Brock (1970), Rubinstein (1979)).

[8] The monotonicity and continuity of the utility function ensure that $C_T(a)$ exists and is unique.

[9] Related objects summarize long-horizon stochastic growth processes in Martin (2012), Robson et al. (2023), Millner (2025), and elsewhere.

the DM is not willing to sacrifice any positive amount of per-period growth to switch from a to a': for any $\varepsilon > 0$, the DM eventually prefers log returns $r(a,\theta)$ to $r(a',\theta)-\varepsilon$. Thus, ties in $C(\cdot)$ may still allow strict preference, but only at a subexponential scale. When $C(a) = C(a')$ we say that a and a' generate the same expected utility at the *exponential scale*.

5 Constant Relative Risk Aversion

We start by assuming that the DM has a constant relative risk aversion (CRRA) utility function. This utility family yields a particularly simple characterization, which we will build upon when studying value of information and hedging in Section 6 and more general utilities in Section 7.

We distinguish two classes of CRRA utilities,

$$u(w) = \frac{w^{1-\eta}}{1-\eta} \qquad (\eta \neq 1).$$

When the coefficient of relative risk aversion η lies in $[0,1)$, utility u is bounded below and unbounded above; we refer to such preferences as *growth-seeking*. When $\eta > 1$, u is bounded above and diverges to $-\infty$ as $w \downarrow 0$; we refer to such preferences as *ruin-averse*.

Lemma 1. *The asymptotic certainty equivalent exists for each action a. Specifically:*

1. *for the growth-seeking DM ($0 \leq \eta < 1$),*

$$C_{\mathsf{CRRA}}(a;\eta) = \max_{q\in\Delta(\Theta)} \left\{ r(a,q) - \frac{1}{1-\eta} I(q) \right\}, \tag{6}$$

2. *for the ruin-averse DM ($\eta > 1$),*

$$C_{\mathsf{CRRA}}(a;\eta) = \min_{q\in\Delta(\Theta)} \left\{ r(a,q) + \frac{1}{\eta-1} I(q) \right\}. \tag{7}$$

Thus, under CRRA preferences, the asymptotic certainty equivalent admits a variational representation. The DM evaluates each action's consequences under a single empirical distribution, albeit generically, not the typical one. The growth-seeking DM evaluates each action optimistically relative to the typical distribution: $C_{\mathsf{CRRA}}(a;\eta) > r(a,p)$ for $\eta \in [0,1)$ and a generic rate function. Similarly, the ruin-averse DM is pessimistic relative to the evaluation under the typical path: $C_{\mathsf{CRRA}}(a;\eta) < r(a,p)$ for $\eta > 1$.

Nature in the variational representation (6) and (7) appears benevolent for the growth-seeking DM and adversarial for the ruin-averse DM. The intuition is that, when $\eta \in [0, 1)$, expected utility is dominated by rare empirical distributions that deliver an exceptionally high growth rate of utility as wealth grows, whereas, when $\eta > 1$, it is dominated by rare empirical distributions that deliver a high growth rate of disutility (i.e., very negative utility) as wealth approaches ruin.[10]

We excluded the CRRA case with $\eta = 1$, which corresponds to logarithmic utility and is the preference underlying the Kelly criterion. The results of this section nonetheless extend to log utility as $\eta \to 1$: Nature's cost of distorting the empirical state distribution diverges in this limit, so the relevant distortion vanishes and the DM maximizes the growth rate under the typical state distribution, as in Kelly.

For a more detailed intuition, consider the risk-neutral DM ($\eta = 0$) whose expected utility for action a is

$$\begin{aligned} \int \exp\big[r(a,q)T\big]\,d\pi_T(q) \;&\approx\; \int \exp\big[\big(r(a,q) - I(q)\big)T\big]\,dq \\ &\approx\; \exp\big[\underbrace{\max_{q\in\Delta(\Theta)} \big\{r(a,q) - I(q)\big\}}_{C_{\mathsf{CRRA}}(a;0)}\,T\big]. \end{aligned}$$

The first approximation rewrites the expected utility on the left using the rate function and the second approximation applies the Laplace principle that approximates an integral over exponential functions by the dominating exponential term. The proof in Appendix A.1 applies Varadhan's lemma (Dembo and Zeitouni, 1998, p. 137) to conclude that the two approximations become precise as T grows large.

The next proposition presents our central result for the case of CRRA utilities. The solution to the DM's problem for a large finite horizon T is approximated by the solution to the asymptotic problem—a sequential game against Nature, in which the DM first selects an action, followed by Nature's choice of an empirical distribution, that isolate the atypical scenarios relevant for long-horizon expected utility maximization. We denote the DM's value achieved in the asymptotic problem as

$$\mathbb{C}_{\mathsf{CRRA}}(\eta) = \max_{a\in A} C_{\mathsf{CRRA}}(a;\eta), \tag{8}$$

[10] Certainty equivalents for particular values of the parameter η appear in Millner (2025) and, for i.i.d. shocks, in Robson et al. (2023). Here, we emphasize comparative statics with respect to risk aversion and then use the CRRA certainty equivalents as building blocks in the subsequent sections.

where we omit the argument η whenever it is fixed. To simplify the statement, we assume that this asymptotic problem has a unique maximizer a^*_{CRRA}.[11]

Proposition 1. *The solution to the asymptotic problem of the CRRA decision-maker approximates the solution to the finite-horizon problem:*

$$\lim_{T\to\infty} a^*_T = a^*_{\mathsf{CRRA}},$$

for any sequence of the finite-horizon optimizers $(a^*_T)_T$. Moreover,

$$\lim_{T\to\infty} \max_{a\in A} C_T(a) = \mathbb{C}_{\mathsf{CRRA}}.$$

Example 2 (Welfare Cost of Business Cycles)**.** Suppose the state evolves according to a Markov chain characterized by a persistence parameter ρ that affects the speed of transitions but not the stationary distribution. See Example 1 for a concrete parametrization. As ρ declines, and hence states become more persistent, the wealth generated by a typical path of states is unchanged, reflecting the invariance of the stationary distribution with respect to ρ. Greater persistence, however, reduces the effective number of independent draws over a horizon T, so atypical empirical distributions become more likely—the rate function $I(q)$ in (5) decreases. Persistence thus affects the welfare of a CRRA decision-maker even when the stationary distribution is unchanged.[12]

When, as is typically assumed, the DM's relative risk aversion exceeds 1, Nature is adversarial; the decrease in the rate function induced by persistence then lowers welfare, reflecting that ruins are less rare. In this sense, independence assumptions such as in Lucas (1987), while innocuous on typical paths, may understate the welfare cost of business cycles over long horizons through the effect of large deviations. ▲

6 Information and Hedging

In this section, we introduce two familiar instruments—information provision and hedging—and ask whether they improve long-run welfare. For starters, we analyze their impact under the CRRA preferences, and obtain stark results: neither instrument increases the asymptotic

[11] If multiple maximizing actions exist in problem (8), then any converging subsequence of $(a^*_T)_T$ has a limit in $\arg\max_{a\in A} C_{\mathsf{CRRA}}(a)$.

[12] Comparative statics with respect to shock persistence was previously examined in Millner (2025) in a related setting.

certainty equivalent. We show this by proving that, at the exponential scale, an uninformed DM without access to hedging can attain the same certainty equivalent as a DM with perfect foresight. Any gains from information or hedging under CRRA are therefore confined to subexponential orders.

These zero-value results are specific to CRRA utility. Section 7 builds on the CRRA benchmark to establish strictly positive values of information and hedging, building on the CRRA benchmark. At the same time, the zero-information-value insight is robust: once hedging is available, additional information has no marginal value; we formalize this in Corollary 2.

6.1 Perfect Foresight

Below, we consider general forms of information structures under which the DM observes a signal at $t = 0$—a forecast of the empirical distribution of future shocks q_T—and then makes a signal-contingent choice. We start with a special case of such an information structure: Under *perfect foresight*, the DM observes a signal that perfectly predicts q_T. Then, the DM chooses the growth-maximizing action against each q_T and therefore achieves per-period wealth growth rate

$$v(q_T) := \max_{a \in A} r(a, q_T)$$

for each realization of the empirical distribution.[13]

We define the finite-horizon perfect-foresight certainty equivalent $\mathbb{C}_{\mathsf{pf},T}$ as the deterministic growth rate of wealth that matches the expected utility of an optimizing DM with perfect foresight:

$$u\Big(\exp\big[\mathbb{C}_{\mathsf{pf},T} T\big]\Big) = \mathrm{E}\, u\Big(\exp[v(q_T)T]\Big),$$

where the expectation is with respect to $q_T \sim \pi_T$. The *asymptotic perfect-foresight certainty equivalent* is its limit,

$$\mathbb{C}_{\mathsf{pf}} = \lim_{T \to \infty} \mathbb{C}_{\mathsf{pf},T}.$$

Perfect foresight plays a special role because it provides an upper bound on the DM's performance under any other circumstances. Yet, perhaps counterintuitively, the next result shows that the benefits of perfect foresight are limited.

Proposition 2. *The uninformed CRRA decision-maker achieves the same asymptotic cer-*

[13] We do not introduce notation for the signal for the perfect-foresight information structure since this signal equals q_T with certainty.

tainty equivalent as under perfect foresight:

$$\mathbb{C}_{\mathsf{CRRA}} = \mathbb{C}_{\mathsf{pf}}.$$

Thus, the baseline and perfect-foresight DMs attain the same expected utility at the exponential scale. The perfect-foresight DM may achieve higher expected utility than her baseline counterpart for any finite horizon, but any such gains are necessarily limited. In particular, the baseline DM would reject, for all sufficiently large T, any offer of perfect foresight that requires paying a positive fraction of wealth each period, however small.

The proof in Appendix A.3 makes use of the Minimax theorem. The representation in Proposition 1 portrays the baseline DM as the first mover in a sequential game against Nature: the DM chooses an action first, and Nature then selects a large deviation. The perfect-foresight DM corresponds to the game with the opposite order of moves, in which she moves second after observing Nature's choice of large deviation, q^*. The Minimax theorem implies that this reversal of moves is inconsequential, and thus the two DMs achieve the same payoffs.[14]

The advantage of perfect foresight, which may be substantial at short horizons, is that it allows the DM to tailor her action to each realized empirical distribution q. This advantage disappears in the long run because, by the Laplace principle, expected utility is asymptotically dominated by a single large deviation,

$$q^* \in \begin{cases} \arg\max_{q \in \Delta(\Theta)} \left\{ v(q) - \dfrac{I(q)}{1-\eta} \right\}, & 0 \leq \eta < 1, \\ \arg\min_{q \in \Delta(\Theta)} \left\{ v(q) + \dfrac{I(q)}{\eta - 1} \right\}, & 1 < \eta, \end{cases} \tag{9}$$

while the contributions of all other empirical distributions become negligible. Consequently, for the uninformed DM to match the perfect-foresight payoff at the exponential scale, it is enough to choose the action optimal for q^*. Although this uninformed action underperforms when $q \neq q^*$, the resulting loss relative to perfect foresight is negligible at the exponential scale.

We now leverage the second-mover representation of the DM's problem. It implies that a CRRA investor's long-horizon portfolio can be viewed as the Kelly portfolio, preceded by a dose of motivated reasoning.

[14] The reversal of moves is immediate for growth-seeking CRRA utility functions. We use the Minimax theorem to establish the proposition in the ruin-averse case. The theorem applies because $r(a, q)$ is concave in a and $I(q)$ is convex.

Example 3 (Kelly Meets Samuelson)**.** We revisit the constant-weight rebalanced-portfolio setting described at the end of Section 3. An action is a portfolio $a \in \Delta(J)$ over assets $j \in J$ chosen at $t = 0$, yielding log return $r(a, \theta) = \ln \sum_{j \in J} a_j \exp r(j, \theta)$. We interpret the restriction to history-independent portfolios as a simplicity constraint. (This is without loss under i.i.d. states). Assume that the rate function has a unique typical distribution p.

The Kelly portfolio maximizes the typical growth rate,

$$a^*_{\mathsf{Kelly}}(p) \in \arg\max_{a \in \Delta(J)} r(a, p),$$

thereby achieving typical growth rate $v(p)$.[15]

We now turn to the portfolio choice of a CRRA investor. Despite Samuelson's critique of the Kelly approach, the CRRA and Kelly portfolios are closely related: any CRRA optimal portfolio is a Kelly portfolio for some state distribution. In particular, the CRRA decision-maker chooses the Kelly portfolio

$$a^*_{\mathsf{CRRA}}(\eta) = a^*_{\mathsf{Kelly}}(q^*(\eta)),$$

for a distorted belief $q^*(\eta)$ given by (9). ▲

6.2 Value of Information and Hedging

We now consider a general form of information and hedging.

Information. We allow the DM to condition on imperfect forecasts of future state frequencies. Following the standard modeling of information, the DM first observes a signal x taking values in a set X, drawn according to a Blackwell experiment $\mu_T(\cdot \mid q_T)$, which specifies the distribution of x conditional on the realization of q_T. For a horizon T, an experiment μ_T, and a measurable choice rule $a : X \to A$, we define the certainty equivalent $C_{\mathsf{info},T}\big(a(\cdot)\big)$ as the deterministic per-period wealth growth rate that matches the induced expected utility:

$$u\Big(\exp\big[C_{\mathsf{info},T}\big(a(\cdot)\big)T\big]\Big) = \mathrm{E}\, u\Big(\exp[r(a(x), q_T)T]\Big),$$

[15]Kelly portfolios are well understood; see Section 16.2 of Cover and Thomas (2006) for an implicit characterization. Robson et al. (2023) establish an equivalence between Kelly portfolio choice and a rational inattention problem with entropic information costs. In particular, $v(p)$ is the value function of the rational inattention problem with payoff function $r(a, \theta)$, and can therefore be characterized via concavification as in Caplin and Dean (2015).

where the expectation is with respect to x and q_T drawn from $\pi_T(q_T)\mu_T(x \mid q_T)$. The *value of information* is

$$\mathsf{VoI} = \lim_{T\to\infty} \left(\sup_{a(\cdot)} C_{\mathsf{info},T}\big(a(\cdot)\big) - \max_{a\in A} C_T(a) \right),$$

that is, the asymptotic increase in the certainty equivalent due to information provision.

Hedging. Instead of observing signals, we now allow the DM to split her initial wealth across subfunds $k = 1, \ldots, K$, each with initial wealth share λ_k and subfund-specific action $a_k \in A$, where $\lambda_k \geq 0$ and $\sum_{k=1}^{K} \lambda_k = 1$. The form of simple hedging studied here involves no rebalancing between the subfunds after $t = 0$. For a horizon T, a *hedge* $\mathbf{a} = (\lambda_k, a_k)_{k=1}^{K}$ yields terminal wealth

$$w_T(\mathbf{a}, q_T) = \sum_{k=1}^{K} \lambda_k \exp[r(a_k, q_T)\,T]\,,$$

and attains certainty equivalent $C_{\mathsf{hedge},T}(\mathbf{a})$ defined by

$$u(\exp[C_{\mathsf{hedge},T}(\mathbf{a})\,T]) = \mathrm{E}\; u(w_T(\mathbf{a}, q_T))\,,$$

with the expectation with respect to $q_T \sim \pi_T$. The *value of hedging* is

$$\mathsf{VoH} = \lim_{T\to\infty} \left(\sup_{\mathbf{a}} C_{\mathsf{hedge},T}(\mathbf{a}) - \max_{a\in A} C_T(a) \right),$$

that is, the asymptotic increase in the certainty equivalent due to hedging.

The value of information and the value of hedging are both nonnegative since the DM can always replicate the baseline choice. Moreover, both values are bounded above by the gain from perfect foresight. Proposition 2 shows that the perfect-foresight gain is zero for CRRA decision-makers (on the exponential scale). It follows that the values of information and hedging are zero:

Corollary 1. *Information and hedging generate no value for a CRRA decision-maker at the exponential scale:*

$$\mathsf{VoI} = \mathsf{VoH} = 0.$$

Information and hedging may still be beneficial, but any such gains are subexponential as the time horizon grows. For sufficiently large T, a CRRA investor is not willing to incur even a small reduction in per-period returns in exchange for information or the ability to hedge.

For an idea of when the exponential scale may be relevant for ascertaining the value of information, consider a government considering climate change policy. The government may be willing to sacrifice current consumption in order to collect information or hedge mitigation actions, but Corollary 1 indicates that it would be unwilling to do so if it required even a tiny reduction in future growth rate. Similarly, a long-run investor would be unwilling to purchase information or hedging services at the cost of a perpetual fractional reduction of her returns, as is commonly required by financial funds.

7 Ruin Aversion Meets Growth Seeking

CRRA utilities do not allow for interactions between growth-seeking and ruin aversion: being unbounded on only one side, they cannot generate both forces simultaneously. To study the interaction between the two forces, we turn to two-tailed utility functions that are unbounded in both directions. These utilities admit a sharp large-deviation characterization—fear of ruin and hope for growth interact in a simple asymptotic formula—and this structure implies that, once simple hedging is allowed, the marginal value of information vanishes at the exponential scale.

To explore the interaction between ruin aversion and growth seeking, this section assumes utilities unbounded in both tails. Though we omit the details, the same methods employed here also characterize long-horizon optimal actions for bounded utility functions with two distinct CRRA tails. As in the setting studied here, actions optimized for bounded utilities respond to vanishingly rare empirical shock frequencies.[16]

7.1 Two-Tailed Utilities

We consider *two-tailed* utility functions $u : \mathbb{R}_{++} \to \mathbb{R}$ that are increasing and continuous, $\lim_{w \downarrow 0} u(w) = -\infty$, $\lim_{w\to\infty} u(w) = \infty$, and their tails satisfy

$$\lim_{w\downarrow 0} \frac{\ln(-u(w))}{\ln w} = 1 - \eta^-, \qquad \lim_{w\to\infty} \frac{\ln u(w)}{\ln w} = 1 - \eta^+. \tag{10}$$

[16] A simple example with a bounded utility can be constructed by starting with a CRRA utility with $\eta > 1$ and a log-return function $r(a, \theta)$ that is positive for all (a, θ). Since wealth is guaranteed to grow, the lower tail of the utility function can be modified to achieve boundedness without affecting the optimal action. Proposition 1 then implies that the asymptotically optimal action best responds to an adversarial large deviation.

for $\eta^- > 1$, $\eta^+ \in [0,1)$. Thus, the lower and upper tails of u behave like a CRRA utility with coefficients η^- and η^+, respectively. Two-tailed utility therefore makes the DM both growth-seeking (utility is unbounded above) and ruin-averse (ruin yields unbounded disutility). For our asymptotic results, only these tail behaviors matter; the shape of u on intermediate wealth levels is immaterial.[17]

7.2 Action Choice under Fear of Ruin

We characterize the asymptotic certainty equivalent under two-tailed utilities by combining the CRRA certainty equivalents for each tail as building blocks. To this end, recall from Lemma 1 that $C_{\mathsf{CRRA}}(a;\eta^+)$ and $C_{\mathsf{CRRA}}(a;\eta^-)$ are the asymptotic certainty equivalents under the growth-seeking and ruin-averse CRRA utilities, respectively; see (6) and (7). These two certainty equivalents are ordered because

$$C_{\mathsf{CRRA}}(a;\eta^+) \geq r(a,p) \geq C_{\mathsf{CRRA}}(a;\eta^-),$$

for any typical distribution p.

We define the *rate difference*

$$D(a) := (1-\eta^+)C_{\mathsf{CRRA}}(a;\eta^+) + (\eta^- - 1)C_{\mathsf{CRRA}}(a;\eta^-),$$

and we refer to the case $D(a) = 0$ as a *tie*.

Appendix A.4 proves:

Lemma 2. *Fix a two-tailed utility and an action a such that a tie does not occur. Then the asymptotic certainty equivalent for a exists and is given by*

$$C(a) = \begin{cases} C_{\mathsf{CRRA}}(a;\eta^+) & \text{if } D(a) > 0, \\ C_{\mathsf{CRRA}}(a;\eta^-) & \text{if } D(a) < 0. \end{cases} \tag{11}$$

[17] The fact that our characterization of the long-horizon optimal choice depends only on the asymptotic properties of the utility function is reminiscent of the portfolio turnpike literature; see, e.g., Dybvig et al. (1999). In that literature, a growing-market or discounting condition asymptotically pushes the optimization problem into the high-wealth region, making the upper tail of utility decisive for long-horizon choice. In our setting, by contrast, long-horizon payoffs are shaped by rare events that may generate either ruin or diverging wealth, so both tails of the utility function matter for optimization.

The proof shows that expected utility approximates

$$e^{(1-\eta^+)C_{\mathsf{CRRA}}(a;\eta^+)T} - e^{-(\eta^- -1)C_{\mathsf{CRRA}}(a;\eta^-)T}, \tag{12}$$

where the first term collects the contribution of large deviations under which wealth grows, while the second term collects the contribution of those q under which wealth shrinks. If $D(a) > 0$, then the first term dominates—that is,

$$(1-\eta^+)C_{\mathsf{CRRA}}(a;\eta^+) > -(\eta^- - 1)C_{\mathsf{CRRA}}(a;\eta^-),$$

so large deviations leading to ruin are negligible at the exponential scale. The DM is then effectively growth-seeking and $C(a) = C_{\mathsf{CRRA}}(a;\eta^+)$. Under the opposite inequality, the second term dominates, the DM is effectively ruin-averse, and $C(a) = C_{\mathsf{CRRA}}(a;\eta^-)$. Accordingly, we say that the *upper tail is selected* when $D(a) > 0$, and the *lower tail is selected* when $D(a) < 0$.

Lemma 2 does not characterize the asymptotic certainty equivalent for actions that generate a tie because then the two terms in (12) have the same exponential rates. To determine the DM's certainty equivalent without subexponential analysis, we impose a regularity condition throughout the section.

Regularity condition 1 For any action a that generates a tie, there exists a sequence of actions $(a_k)_k$ converging to a such that the upper tail is selected for each a_k.

The regularity condition and continuity of $C_{\mathsf{CRRA}}(a;\eta^+)$ in a ensure that the certainty equivalent of any action a generating a tie can be approximated by certainty equivalents of actions for which the upper tail is selected.[18]

To characterize the solution to the DM's problem, we define the *ruin-robust growth program* as

$$\begin{aligned} \max_{a\in A} \quad & C_{\mathsf{CRRA}}(a;\eta^+) \\ \text{s.t.} \quad & D(a) \geq 0, \end{aligned} \tag{13}$$

[18]Since $r(a,q) - \frac{1}{1-\eta^+}I(q)$ is continuous and $\Delta(\Theta)$ is compact, the Maximum theorem implies that $C_{\mathsf{CRRA}}(a;\eta^+)$ is continuous in a.

with the convention that if no action in A satisfies the constraint (13), then the DM solves

$$\max_{a \in A} C_{\mathsf{CRRA}}(a; \eta^-). \tag{14}$$

Let $\mathbb{C}_{\mathsf{tt}}$ be the value of the *ruin-robust growth program* and a^*_{tt} an optimizer. We can then extend our central result to two-tailed utilities. To simplify the statement of the next result, assume that a^*_{tt} is unique.

Proposition 3. *The solution to the ruin-robust growth program approximates the solution to the finite-horizon problem for the DM with a two-tailed utility:*

$$\lim_{T \to \infty} a^*_T = a^*_{\mathsf{tt}},$$

for any sequence of the finite-horizon optimizers $(a^*_T)_T$, and

$$\lim_{T \to \infty} \max_{a \in A} C_T(a) = \mathbb{C}_{\mathsf{tt}}.$$

In words, for long horizons we can represent the DM as choosing an action a that maximizes the growth promise $C_{\mathsf{CRRA}}(a; \eta^+)$, subject to not being overwhelmed by fear of ruin, as captured by the constraint (13). If no action satisfies this constraint, then fear of ruin dominates for all a, and the DM instead chooses the action that makes ruin least severe in the exponential sense, i.e., the maximizer in (14).

Example 4 (Stock Market Participation Puzzle)**.** We revisit the constant-weight rebalanced portfolio choice to illustrate how fear of ruin constrains the investor's portfolio choice.

For the sake of concreteness, we consider two states, $\theta \in \{H, L\}$, and two assets, $j \in \{s, c\}$, with returns $R(j, \theta)$ as specified in the table:

	H	L
s	1.10	0.20
c	0.97	0.97

The state is high, $\theta = H$, with probability $p = 0.95$ independently across periods. Cash ($j = c$) is risk-free but delivers a negative real return, whereas the stock ($j = s$) usually pays a modest return but carries a small probability of a disastrous outcome.

The DM has a two-tailed utility $u(w) = w - 1/w$ and thus is approximately risk-neutral ($\eta^+ = 0$) for large wealth levels, but also fears bankruptcy ($\eta^- = 2$), as captured by the diverging disutility for $w \downarrow 0$. She chooses a constant-weight rebalanced portfolio, investing a share $a \in [0,1]$ of wealth in the stock each period and the remaining share in cash. The resulting log return is $r(a,\theta) = \ln\big(aR(s,\theta) + (1-a)R(c,\theta)\big)$.

The long-run investor solves the ruin-robust growth program: she maximizes $C_{\mathsf{CRRA}}(a;0)$ subject to $D(a) = C_{\mathsf{CRRA}}(a;0) + C_{\mathsf{CRRA}}(a;2) \geq 0$. The constraint binds, and the constrained-optimal portfolio assigns the weight $a^*_{\mathsf{tt}} \approx 0.87$ to the stock, generating certainty equivalent $\mathbb{C}_{\mathsf{tt}} = C_{\mathsf{CRRA}}(a^*_{\mathsf{tt}};0) \approx 0.043$; see Figure 1.

In contrast, an analogous investor without ruin-aversion—that is, with the CRRA utility with $\eta = \eta^+ = 0$—would fully invest in stock ($a^*_{\mathsf{CRRA}}(\eta^+) = 1$) and achieve a higher certainty equivalent $C_{\mathsf{CRRA}}(1;\eta^+) \approx 0.054$. Since the stock generates a positive growth rate on the typical path, $r(1,p) > 0$, this full exposure to stock almost never results in ruin for large horizons. Yet, our two-tailed DM finds the full exposure to stock too fearful: $D(a) = C_{\mathsf{CRRA}}(1;0) + C_{\mathsf{CRRA}}(1;2) < 0$, indicating that the possibility of ruin, despite its vanishing probability, dominates her expected-utility evaluation of the full-exposure portfolio. Indeed, $a^*_{\mathsf{tt}} \approx 0.87$ is the highest exposure to stock at which our two-tailed investor is not overcome by the fear of ruin.

Suppose an analyst models this DM as a CRRA utility maximizer with η calibrated from the observed state sequences. The analyst will typically observe state sequences that generate diverging wealth, and so will approximate the DM's utility by the CRRA utility with $\eta = 0$ at the observed wealth range. The analyst will then conclude that the apparently risk-neutral DM should invest all of her wealth in stock and that the DM's cautious attitude to stock is puzzling. In this setting, the equity premium puzzle is driven by contingencies that are effectively never observed. The analyst observes no evidence of the lower tail of the utility function, and so proceeds as if this tail does not exist, even though it explains why the DM limits her exposure to the risky asset. ▲

7.3 Valuable Information

Information has no value for the CRRA decision-maker at the exponential scale because she already attains the perfect-foresight payoff with only prior information. We now show that under two-tailed utilities information may be valuable: perfect foresight relaxes the fear-of-ruin constraint and can therefore deliver a strictly higher payoff than the uninformed benchmark. Since perfect foresight is a specific information structure, this gap implies that the value of information can be strictly positive for two-tailed utilities.

Recall that $\mathbb{C}_{\mathsf{CRRA}}(\eta^+)$ and $\mathbb{C}_{\mathsf{CRRA}}(\eta^-)$ defined in (8) denote the optimized asymptotic cer-

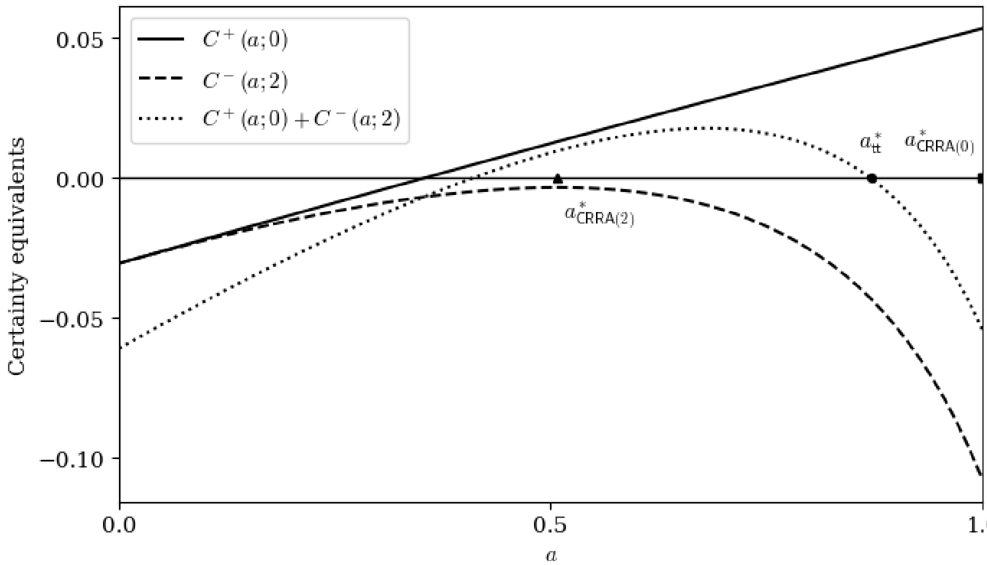

Figure 1: The optimal action a^*_{tt} from Example 4 maximizes $C_{\mathsf{CRRA}}(a;0)$ subject to the fear-of-ruin constraint $C_{\mathsf{CRRA}}(a;0) + C_{\mathsf{CRRA}}(a;2) \geq 0$. When hedging is allowed, as in Example 6, then a pair of subfunds with actions $a^*_{\mathsf{CRRA}}(0)$ and $a^*_{\mathsf{CRRA}}(2)$ achieves the perfect-foresight certainty equivalent.

tainty equivalents of the CRRA decision-maker with the relative risk aversion coefficients η^+ and η^-, respectively. Analogously to the rate difference $D(a)$ defined for each action, we let

$$\mathbb{D} = (1-\eta^+)\mathbb{C}_{\mathsf{CRRA}}(\eta^+) + (\eta^- - 1)\mathbb{C}_{\mathsf{CRRA}}(\eta^-),$$

stand for the optimized rate difference.

To streamline the exposition, we impose a regularity condition throughout the remainder of the section:[19]

Regularity condition 2 Assume $\mathbb{D} \neq 0$.

Appendix A.4 proves:

Lemma 3. *The asymptotic perfect-foresight certainty equivalent for the DM with a two-tailed*

[19]In the nongeneric case of a tie, $\mathbb{D} = 0$, the asymptotic behavior of $C_{\mathsf{pf},T}$ is determined by subexponential effects.

utility exists and is given by

$$\mathbb{C}_{\mathsf{pf}} = \begin{cases} \mathbb{C}_{\mathsf{CRRA}}(\eta^+) & \textit{if } \mathbb{D} > 0, \\ \mathbb{C}_{\mathsf{CRRA}}(\eta^-) & \textit{if } \mathbb{D} < 0. \end{cases}$$

When $\mathbb{D} > 0$, then the perfect-foresight DM achieves the same certainty equivalent as the growth-seeking CRRA decision-maker with risk-parameter $\eta = \eta^+$. Thus, in this case, perfect foresight eliminates the fear-of-ruin constraint (13) from the ruin-robust growth program solved by the uninformed DM. This constraint relaxation may generate a strictly positive increase in the certainty equivalent, as the next example illustrates.

Example 5 (Positive Value of Information)**.** To demonstrate that information can be valuable, we now endow the DM from Example 4 with perfect foresight. Since $\mathbb{D} = \mathbb{C}_{\mathsf{CRRA}}(0) + \mathbb{C}_{\mathsf{CRRA}}(2) > 0$, she achieves the same certainty equivalent $\mathbb{C}_{\mathsf{pf}} = \mathbb{C}_{\mathsf{CRRA}}(0) \approx 0.054$ as the growth-seeking CRRA decision-maker with the risk parameter $\eta = \eta^+ = 0$ corresponding to risk neutrality. Recall that the uninformed two-tailed DM from Example 4 achieved a lower certainty equivalent $C_{\mathsf{CRRA}}(a^*_{\mathsf{tt}}; 0) \approx 0.043$. Hence, the information provided allows the DM to obtain a strictly higher certainty equivalent. ▲

Assume $\mathbb{D} > 0$. To understand why the perfect-foresight DM achieves the same payoff as the CRRA decision-maker with the upper-tail risk-aversion coefficient, let (a^*_+, q^*_+) and (a^*_-, q^*_-) be the optimizing pairs of actions and large deviations for the CRRA decision-maker with $\eta = \eta^+$ and $\eta = \eta^-$, respectively. Since the perfect-foresight DM chooses $a = a^*_+$ at $q = q^*_+$ and $a = a^*_-$ at $q = q^*_-$, and the expected-utility contributions of all other large deviations are negligible, her expected utility approximates

$$e^{(1-\eta^+)\mathbb{C}_{\mathsf{CRRA}}(\eta^+)T} - e^{-(\eta^- - 1)\mathbb{C}_{\mathsf{CRRA}}(\eta^-)T}.$$

The inequality $\mathbb{D} > 0$ implies that the positive term dominates and thus, at the exponential scale, the perfect-foresight DM achieves the same payoff as the CRRA decision-maker with $\eta = \eta^+$.

7.4 Hedging Matches Perfect Foresight

Hedging is not valuable for the CRRA decision-maker because she attains the perfect-foresight payoff without it. By contrast, the two-tailed DM may fall short of this upper bound. We show that the two-tailed DM benefits from hedging and, in fact, attains the perfect-foresight payoff by hedging between just two actions.

We let

$$\mathbb{C}_{\mathsf{hedge}} = \lim_{T\to\infty} \sup_{\mathbf{a}} C_{\mathsf{hedge},T}(\mathbf{a})$$

denote the asymptotic certainty equivalent under optimal hedging.

Proposition 4. *The DM with two-tailed utility who can hedge achieves the same payoff at the exponential scale as under perfect foresight:*

$$\mathbb{C}_{\mathsf{hedge}} = \mathbb{C}_{\mathsf{pf}}.$$

To understand how hedging generates value, consider first a DM with a two-tailed utility who cannot hedge. She must choose a single action that fares well against both Nature's benevolent and malevolent large deviations, which may require a compromise. We prove the proposition by considering a hedging strategy that splits the initial wealth into two subfunds and assigns actions $a^*_{\mathsf{CRRA}}(\eta^+)$ and $a^*_{\mathsf{CRRA}}(\eta^-)$ to them, respectively. Since these two actions are optimized against the two dominating large deviations, this hedge achieves the perfect foresight payoff at the exponential scale.

Example 6 (Kelly Meets Samuelson Again)**.** We return to the constant-weight rebalanced-portfolio setting. The tight link between Kelly portfolios and the expected-utility-optimal strategy reappears for an investor with a two-tailed utility, provided she can split initial wealth across two subfunds. She achieves the upper bound $\mathbb{C}_{\mathsf{pf}}$ by investing the two subfunds in two Kelly portfolios. Specifically,

$$a^*_{\mathsf{CRRA}}(\eta^+) = a^*_{\mathsf{Kelly}}(q^*_+) \quad \text{and} \quad a^*_{\mathsf{CRRA}}(\eta^-) = a^*_{\mathsf{Kelly}}(q^*_-),$$

where q^*_+ and q^*_- solve

$$q^*_+ \in \arg\max_{q\in\Delta(\Theta)} \left\{ v(q) - \frac{1}{1-\eta^+} I(q) \right\},$$

$$q^*_- \in \arg\min_{q\in\Delta(\Theta)} \left\{ v(q) + \frac{1}{\eta^- - 1} I(q) \right\}.$$

Comparison of Examples 4 and 5 implies a strictly positive value of hedging for a two-tailed DM. By Proposition 4, hedging attains the same asymptotic certainty equivalent as perfect foresight. Example 5 shows that perfect foresight yields a strictly higher certainty equivalent than the uninformed benchmark in Example 4. Hence, $\mathsf{VoH} > 0$ in this investment problem. ▲

7.5 Hedging Versus Information

We conclude the section by noting that, if the two-tailed DM can hedge, information provides no additional value at the exponential scale. For this, we consider a DM who observes at $t = 0$ a signal x generated by a Blackwell experiment $\mu_T(x \mid q_T)$, and then chooses once and for all a hedge $\mathbf{a}(x)$ contingent on the realized signal. We let $C_{\mathsf{hedge,info},T}\big(\mathbf{a}(\cdot)\big)$ be the certainty equivalent for the hedging rule $\mathbf{a}(x)$; it is defined by

$$u\Big(\exp\big[C_{\mathsf{hedge,info},T}\big(\mathbf{a}(\cdot)\big)T\big]\Big) = \mathrm{E}\, u\big(w_T(\mathbf{a}(x), q_T)\big),$$

where the expectation is with respect to x and q_T drawn from $\pi_T(q_T)\mu_T(x \mid q_T)$, and $w_T(\mathbf{a}, q_T)$ is the terminal wealth for the hedge $\mathbf{a}$ and shock frequencies q_T. We define the *marginal value of information (given hedging)*,

$$\mathsf{mVoI} = \lim_{T\to\infty}\left(\sup_{\mathbf{a}(\cdot)} C_{\mathsf{hedge,info},T}\big(\mathbf{a}(\cdot)\big) - \sup_{\mathbf{a}} C_{\mathsf{hedge},T}(\mathbf{a})\right),$$

as the asymptotic increase in the certainty equivalent due to information relative to the benchmark setting in which the DM can hedge but has no information.

Corollary 2 (Zero Marginal Value of Information)**.** *The marginal value of information for a DM with a two-tailed utility who can hedge is zero:*

$$\mathsf{mVoI} = 0.$$

This extends Corollary 1 beyond CRRA: even when information can be valuable for a two-tailed DM, its marginal value disappears once hedging is allowed. The corollary is implied by Proposition 4. Since the perfect-foresight payoff is an upper bound on any payoff, and the hedging DM without information achieves it, information has no additional value.

8 Discussion

A long-run expected-utility maximizer faces a bewildering array of possible future scenarios. In the long run, however, she needs to identify and respond to just one or two; the vast array of alternatives is irrelevant. These relevant scenarios are pinned down not by the law of large numbers, but as solutions to a variational problem.

The rate function summarizes all relevant details of the stochastic state-generating process. By construction, it characterizes the probabilities of events that are essentially never observed. This raises interesting questions about the empirical content of expected-utility theory in long-horizon problems. Because rare events cannot be observed directly, both the decision-maker and the analyst can learn the rate function only indirectly, through a structural model of the state-generating process. Moreover, when rationalizing the decision-maker's choices, the analyst must assess preferences over atypical wealth ranges that are effectively unobservable in the data, yet loom large in the decision-maker's calculations.

Our framework points to a normative tension between growth and equity. As is well known and demonstrated formally in Robson et al. (2023), stochastic growth processes can generate extraordinary inequality. Just as the expected-utility criterion in our single-decision setting is dominated by wealth generated along a shrinking set of atypical sequences, long-run aggregate wealth in their population setting is dominated by a shrinking set of atypically lucky individuals. A utilitarian objective offers a social planner little basis for objecting to such an outcome, even if one is troubled by the inevitability of extreme inequality. This may motivate objectives that incorporate ex post equity considerations, as in Fleurbaey (2010) or the prioritarian literature Arneson (2022).

The decision-maker's focus on events that are overwhelmingly unlikely to occur may seem counterintuitive. We remain open to the interpretation that, while expected-utility theory is pervasive in economic models, its application to very long horizons may be stretched beyond plausibility in this setting. Our results thus prompt the question of whether it is "rational" to be an expected-utility maximizer in long-horizon problems. We view this as an important normative question for future work.

A Proofs

A.1 Proof of Lemma 1

Proof of Lemma 1. The definition of the certainty equivalent $C_T(a)$ for action a over a finite horizon T,

$$\mathrm{E}_{q_T \sim \pi_T}\, u\big(\exp[r(a, q_T)T]\big) \;=\; u\left(\exp\left[C_T(a)T\right]\right),$$

applied to the CRRA utility yields

$$C_T(a) = \frac{1}{T(1-\eta)} \ln \mathrm{E}_{q_T \sim \pi_T} \exp\left[(1-\eta) r(a, q_T) T\right].$$

We apply Varadhan's lemma (Dembo and Zeitouni, 1998, p. 137), noting that the moment condition for the lemma is satisfied due to the boundedness of $r(a,q)$. This yields

$$C(a) = \lim_{T\to\infty} C_T(a) = \frac{1}{1-\eta} \max_{q\in\Delta(\Theta)} \left\{(1-\eta) r(a,q) - I(q)\right\},$$

where the maximum exists due to the assumed continuity of the rate function and compactness of $\Delta(\Theta)$. Thus,

$$C_{\mathsf{CRRA}}(a;\eta) = \begin{cases} \max_{q\in\Delta(\Theta)} \left\{ r(a,q) - \frac{1}{1-\eta} I(q) \right\} & \text{if } 0 \le \eta < 1, \\ \min_{q\in\Delta(\Theta)} \left\{ r(a,q) + \frac{1}{\eta-1} I(q) \right\} & \text{if } \eta > 1, \end{cases}$$

as needed. □

A.2 Proof of Proposition 1

We fix η and omit it from the arguments of the certainty equivalents.

Lemma 4. *For a CRRA utility function, $C_T(a) \to C_{\mathsf{CRRA}}(a)$, uniformly on A.*

Proof. Lemma 1 established pointwise convergence $C_T(a) \to C_{\mathsf{CRRA}}(a)$. To establish uniform convergence, we now prove equicontinuity. Fix $\varepsilon > 0$. By continuity of r on the compact set $A \times \Delta(\Theta)$, there exists $\delta > 0$ such that

$$d(a,a') < \delta \quad \Longrightarrow \quad \sup_{q\in\Delta(\Theta)} |r(a,q) - r(a',q)| \le \varepsilon,$$

where $d(a,a')$ is the Euclidean metric. For

$$C_T(a) = \frac{1}{T(1-\eta)} \ln \mathrm{E}_{q_T\sim\pi_T}\left[\exp\bigl((1-\eta) r(a,q_T) T\bigr)\right],$$

$d(a,a') < \delta$ implies $|C_T(a) - C_T(a')| \le \varepsilon$ for all T.[20] Thus, $\{C_T\}_T$ is equicontinuous on A.

[20]From $|r(a,q) - r(a',q)| \le \varepsilon$ one has $e^{(1-\eta)r(a,q)T} \le e^{|1-\eta|\varepsilon T} e^{(1-\eta)r(a',q)T}$ and the reverse inequality with a, a' swapped; taking expectations and logs yields the bound.

By Berge's Maximum Theorem, $C_{\mathsf{CRRA}}(\cdot)$ is continuous on A. Since A is compact, equicontinuity of $\{C_T\}_T$ and pointwise convergence $C_T(a) \to C_{\mathsf{CRRA}}(a)$ imply uniform convergence $C_T \to C_{\mathsf{CRRA}}(\cdot)$ on A (by a standard Arzelà–Ascoli argument). □

Proof of Proposition 1. Uniform convergence implies

$$\left| \max_{a \in A} C_T(a) - \mathbb{C}_{\mathsf{CRRA}} \right| \leq \sup_{a \in A} |C_T(a) - C_{\mathsf{CRRA}}(a)| \to 0.$$

For the convergence of the optimizer, take any convergent subsequence $a^*_{T_k} \to \bar{a}$. Uniform convergence and optimality yield

$$C_{\mathsf{CRRA}}(\bar{a}) = \lim_k C_{T_k}(a^*_{T_k}) = \lim_k \max_{a \in A} C_{T_k}(a) = \mathbb{C}_{\mathsf{CRRA}},$$

hence $\bar{a} = a^*_{\mathsf{CRRA}}$ by the assumed uniqueness of the asymptotic optimizer. □

A.3 Proof of Proposition 2

Proof of Proposition 2. For each finite horizon T,

$$\begin{aligned} \mathbb{C}_{\mathsf{pf},T} &= \frac{1}{(1-\eta)T} \ln \mathrm{E}_{q_T \sim \pi_T} \exp[(1-\eta)v(q_T)T] \\ &\to \frac{1}{1-\eta} \sup_{q \in \Delta(\Theta)} \{(1-\eta)v(q) - I(q)\} \\ &= \mathbb{C}_{\mathsf{pf}}, \end{aligned}$$

where the convergence is implied by Varadhan's lemma applied to the continuous and bounded function $v(q)$. We proceed separately for the growth-seeking and ruin-averse DMs.

The growth-seeking DM, $(0 \leq \eta < 1)$**:**

$$\begin{aligned} \mathbb{C}_{\mathsf{pf}} &= \sup_{q \in \Delta(\Theta)} \left\{ v(q) - \frac{1}{1-\eta} I(q) \right\} \\ &= \sup_{q \in \Delta(\Theta), a \in A} \left\{ r(a,q) - \frac{1}{1-\eta} I(q) \right\} \\ &= \sup_{a \in A, q \in \Delta(\Theta)} \left\{ r(a,q) - \frac{1}{1-\eta} I(q) \right\} \\ &= \sup_{a \in A} C_{\mathsf{CRRA}}(a;\eta). \end{aligned}$$

The ruin-averse DM, ($\eta > 1$):

$$\begin{aligned}
\mathbb{C}_{\mathsf{pf}} &= \inf_{q\in\Delta(\Theta)} \left\{ v(q) + \frac{1}{\eta-1} I(q) \right\} \\
&= \inf_{q\in\Delta(\Theta)} \sup_{a\in A} \left\{ r(a,q) + \frac{1}{\eta-1} I(q) \right\} \\
&= \sup_{a\in A} \inf_{q\in\Delta(\Theta)} \left\{ r(a,q) + \frac{1}{\eta-1} I(q) \right\} \\
&= \sup_{a\in A} C_{\mathsf{CRRA}}(a;\eta),
\end{aligned}$$

where we have used the Minimax theorem for the third equality. The Minimax theorem applies because the objective is concave in a and convex in q, as implied by concavity of $r(a,q)$ with respect to a, linearity of $r(a,q)$ in q and convexity of $I(q)$, and the domains of optimization, A and $\Delta(\Theta)$, are compact and convex. □

A.4 Proofs of Lemmas 2 and 3

We explain below that Lemma 2 is implied by the following result. Let $f(q)$ be a continuous convex function and

$$C^+ = \sup_{q\in\Delta(\Theta)} \left\{ f(q) - \tfrac{1}{1-\eta^+} I(q) \right\}, \quad C^- = \inf_{q\in\Delta(\Theta)} \left\{ f(q) + \tfrac{1}{\eta^- -1} I(q) \right\}.$$

Lemma 5. *The asymptotic certainty equivalent of $w_T = \exp[f(q_T)T]$ under a two-tailed utility function is:*

$$C = \begin{cases} C^+ & \textit{if } (1-\eta^+)C^+ + (\eta^- - 1)C^- > 0, \\ C^- & \textit{if } (1-\eta^+)C^+ + (\eta^- - 1)C^- < 0. \end{cases} \tag{15}$$

To prove Lemma 5, we partition the simplex $\Delta(\Theta)$ into

$$S_+ = \{q : f(q) > 0\}, \quad S_- = \{q : f(q) < 0\}, \quad S_0 = \{q : f(q) = 0\},$$

and define the *constrained* asymptotic certainty equivalents for these sets as:[21]

$$\tilde{C}^+ = \sup_{q\in S_+} \left\{ f(q) - \tfrac{1}{1-\eta^+} I(q) \right\}, \quad \tilde{C}^- = \inf_{q\in S_-} \left\{ f(q) + \tfrac{1}{\eta^- -1} I(q) \right\}, \quad \tilde{C}^0 = \sup_{q\in S_0} \left\{ -I(q) \right\}.$$

[21] We use standard convention $\sup\emptyset = -\infty$ and $\inf\emptyset = +\infty$.

Note that $\tilde{C}^+$ and $\tilde{C}^-$ constrain the optimization to S_+ and S_-, whereas C^+ and C^- optimize over all of $\Delta(\Theta)$. The next lemma characterizes the growth rate of expected utility in terms of $\tilde{C}^\pm$; we then use it to prove Lemma 5.

Lemma 6. *Assume that* $\arg\max_{q\in\Delta(\Theta)}\{f(q) - \frac{1}{1-\eta^+}I(q)\} \cap \operatorname{int}(S_0) = \emptyset$. *Then,*

$$\mathrm{E}\,u(w_T) = \begin{cases} \exp[(1-\eta^+)\tilde{C}^+T + o(T)] & \text{if } (1-\eta^+)\tilde{C}^+ + (\eta^- - 1)\tilde{C}^- > 0, \\ -\exp[-(\eta^- - 1)\tilde{C}^- T + o(T)] & \text{if } (1-\eta^+)\tilde{C}^+ + (\eta^- - 1)\tilde{C}^- < 0. \end{cases}$$

Proof. We decompose the expected utility $\mathrm{E}[u(w_T)]$ as follows:

$$\mathrm{E}[u(w_T)] = \underbrace{\int_{S_+} u(e^{f(q)T})d\pi_T}_{E_T^+} + \underbrace{\int_{S_-} u(e^{f(q)T})d\pi_T}_{E_T^-} + \underbrace{\int_{S_0} u(1)d\pi_T}_{E_T^0}.$$

Using the asymptotic properties of $u(w)$ (where $u(w) \sim w^{1-\eta^+}$ for large w and $u(w) \sim -w^{1-\eta^-}$ for small w), we apply Varadhan's lemma to each term:

Set $\phi_+(q) := (1-\eta^+)f(q)$, which is continuous and bounded on the compact set $\Delta(\Theta)$. Since $S_+ \subset \Delta(\Theta)$ is open, the Large Deviation Property for $(\pi_T)_T$ implies the Laplace–Varadhan bounds on sets:

$$\begin{aligned} \liminf_{T\to\infty} \frac{1}{T} \ln \int_{S_+} \exp\big(T\phi_+(q)\big)\, d\pi_T &\geq \sup_{q\in S_+}\{\phi_+(q) - I(q)\}, \\ \limsup_{T\to\infty} \frac{1}{T} \ln \int_{S_+} \exp\big(T\phi_+(q)\big)\, d\pi_T &\leq \limsup_{T\to\infty} \frac{1}{T} \ln \int_{\overline{S}_+} \exp\big(T\phi_+(q)\big)\, d\pi_T \\ &\leq \sup_{q\in\overline{S}_+}\{\phi_+(q) - I(q)\}. \end{aligned}$$

Because $\phi_+ - I$ is continuous, $\sup_{S_+}(\phi_+ - I) = \sup_{\overline{S}_+}(\phi_+ - I)$, hence the limit exists and

$$\lim_{T\to\infty} \frac{1}{T} \ln \int_{S_+} \exp\big(T\phi_+(q)\big)\, d\pi_T = \sup_{q\in S_+}\{(1-\eta^+)f(q) - I(q)\}.$$

Set $\phi_-(q) := (1-\eta^-)f(q)$, which is continuous and bounded. Applying the same open-set Laplace–Varadhan bounds to S_- yields

$$\lim_{T\to\infty} \frac{1}{T} \ln \int_{S_-} \exp\big(T\phi_-(q)\big)\, d\pi_T = \sup_{q\in S_-}\{(1-\eta^-)f(q) - I(q)\}.$$

Finally,[22]

$$
\begin{aligned}
\limsup_{T\to\infty} \frac{1}{T} \ln |E_T^0| &= \limsup_{T\to\infty} \frac{1}{T} \ln \left| \int_{S_0} u(1) d\pi_T \right| \\
&= \limsup_{T\to\infty} \frac{1}{T} \ln \left(|u(1)| \cdot \pi_T(S_0) \right) \\
&= \limsup_{T\to\infty} \frac{1}{T} \ln \pi_T(S_0) \\
&\leq \sup_{q\in S_0} \{-I(q)\} \\
&= \tilde{C}^0.
\end{aligned}
$$

The asymptotic behavior of the total expected utility is determined by the term with the strictly largest exponent (we show below that the strictly largest exponent exists).

If $\text{int}(S_0) \neq \emptyset$, then the convex function $f(q)$ is nonnegative and hence $S_- = \emptyset$. Then $\tilde{C}^- = \inf \emptyset = +\infty$, and thus $(1-\eta^+)\tilde{C}^+ + (\eta^- - 1)\tilde{C}^- > 0$. By assumption in the lemma, $\arg\max_{q\in\Delta(\Theta)}\{f(q) - \frac{1}{1-\eta^+} I(q)\} \subset \text{cl}(S_+)$. Thus, $\tilde{C}^0 \leq (1-\eta^+)\tilde{C}^+$, the term E_T^+ dominates, and the asymptotic growth rate of expected utility is $(1-\eta^+)\tilde{C}^+$, as needed.

Next, we assume $\text{int}(S_0) = \emptyset$. Because $f(q)$ and $I(q)$ are continuous, the suprema of the objective functions over the open sets S_+ or S_- are bounded below by the supremum over ∂S_0. Thus, $\tilde{C}^0 = \sup_{q\in S_0}\{-I(q)\} = \sup_{q\in\partial S_0}\{-I(q)\}$. Therefore, we have the inequalities:

$$
\tilde{C}^0 \leq (1-\eta^+)\tilde{C}^+ \quad \text{and} \quad \tilde{C}^0 \leq -(\eta^- - 1)\tilde{C}^-. \tag{16}
$$

We now analyze the two cases:

1. Case $(1-\eta^+)\tilde{C}^+ + (\eta^- - 1)\tilde{C}^- > 0$: Combining $(1-\eta^+)\tilde{C}^+ > -(\eta^- - 1)\tilde{C}^-$ with the inequality in (16) yields strict dominance: $(1-\eta^+)\tilde{C}^+ > -(\eta^- - 1)\tilde{C}^- \geq \tilde{C}^0$. Thus, E_T^+ dominates, and the asymptotic growth rate of expected utility is $(1-\eta^+)\tilde{C}^+$.

2. Case $(1-\eta^+)\tilde{C}^+ + (\eta^- - 1)\tilde{C}^- < 0$: Combining $-(\eta^- - 1)\tilde{C}^- > (1-\eta^+)\tilde{C}^+$ with (16) yields: $-(\eta^- - 1)\tilde{C}^- > (1-\eta^+)\tilde{C}^+ \geq \tilde{C}^0$. Thus, E_T^- dominates (in magnitude), and the asymptotic growth rate of expected utility is $-(\eta^- - 1)\tilde{C}^-$.

□

Proof of Lemma 5. We will use that if $(1-\eta^+)C^+ + (\eta^- - 1)C^- > 0$, then $C^+ > 0$. Similarly, if $(1-\eta^+)C^+ + (\eta^- - 1)C^- < 0$, then $C^- < 0$. For this, observe that $C^+ \geq f(p) \geq C^-$. To

[22] If $u(1) = 0$, then $E_T^0 = 0$ and its rate is $-\infty$.

prove the first implication, assume $(1-\eta^+)C^+ + (\eta^- - 1)C^- > 0$. For contradiction, assume $C^+ \leq 0$. Since $C^- \leq C^+$, $(1-\eta^+)C^+ + (\eta^- - 1)C^- \leq 0$, yielding a contradiction. The proof of the other implication is analogous.

First, consider the case where $\arg\max_{q\in\Delta(\Theta)} \left\{ f(q) - \frac{1}{1-\eta^+} I(q) \right\} \cap \operatorname{int}(S_0) \neq \emptyset$ and pick

$$q^* \in \arg\max_{q\in\Delta(\Theta)} \left\{ f(q) - \tfrac{1}{1-\eta^+} I(q) \right\} \cap \operatorname{int}(S_0).$$

Since f is convex and vanishes on the nonempty open set $\operatorname{int}(S_0)$, it follows that $f(q) \geq 0$ for all $q \in \Delta(\Theta)$. By optimality of q^*, for every $q \in \Delta(\Theta)$,

$$f(q) - \tfrac{1}{1-\eta^+} I(q) \leq f(q^*) - \tfrac{1}{1-\eta^+} I(q^*) = -\tfrac{1}{1-\eta^+} I(q^*).$$

Using $f(q) \geq 0$ we obtain that q^* is a global minimizer of I on $\Delta(\Theta)$, and therefore $I(q^*) = I(p) = 0$. Consequently, $C^+ = 0$. Additionally, since $f(q) \geq 0$, $C^- = f(q^*) + \frac{1}{\eta^- - 1} I(q^*) = 0$, the tie obtains, and the lemma is vacuous in this case.

For the remainder of the proof, assume that $\arg\max_{q\in\Delta(\Theta)}\{f(q) - \frac{1}{1-\eta^+} I(q)\} \cap \operatorname{int}(S_0) = \emptyset$, so that Lemma 6 applies. It remains to relate the constrained rates $\tilde{C}^\pm$ to the unconstrained rates $C^\pm$ and compute the certainty equivalents.

Since $S_+, S_- \subseteq \Delta(\Theta)$, we have

$$\tilde{C}^+ \leq C^+ \quad \text{and} \quad \tilde{C}^- \geq C^-.$$

We now consider two cases based on the sign of the unconstrained sum.

1. Case $(1-\eta^+)C^+ + (\eta^- - 1)C^- > 0$: Pick any $q_+^* \in \arg\max_{q\in\Delta(\Theta)}\{f(q) - \frac{1}{1-\eta^+} I(q)\}$ so that $(1-\eta^+)C^+ = (1-\eta^+)f(q_+^*) - I(q_+^*)$. By the definition of C^- as an infimum, $(\eta^- - 1)C^- \leq (\eta^- - 1)f(q_+^*) + I(q_+^*)$. Summing over these inequalities yields:

$$0 < (1-\eta^+)C^+ + (\eta^- - 1)C^- \leq (\eta^- - \eta^+) f(q_+^*).$$

Thus, since $(\eta^- - \eta^+)$ is positive, the unconstrained optimizer q_+^* satisfies the constraint $q \in S_+$. Consequently, the constrained supremum achieves the unconstrained value: $\tilde{C}^+ = C^+$, and the condition for Lemma 6 is satisfied:

$$(1-\eta^+)\tilde{C}^+ + (\eta^- - 1)\tilde{C}^- \geq (1-\eta^+)C^+ + (\eta^- - 1)C^- > 0.$$

Thus, $\mathrm{E}\, u(w_T) \sim \exp[(1-\eta^+)C^+ T]$. Additionally, $C^+ > 0$. Since the expected utility diverges to $+\infty$, its relevant asymptotic approximation is the CRRA with $\eta = \eta^+$, and thus the certainty equivalent is C^+.

2. Case $(1-\eta^+)C^+ + (\eta^- - 1)C^- < 0$: Pick any $q_-^* \in \arg\min_{q\in\Delta(\Theta)}\{f(q) + \frac{1}{\eta^- -1} I(q)\}$ so that $(\eta^- - 1)C^- = (\eta^- - 1)f(q_-^*) + I(q_-^*)$. By the definition of C^+ as a supremum, $(1-\eta^+)C^+ \geq (1-\eta^+)f(q_-^*) - I(q_-^*)$. Summing over these inequalities yields:

$$0 > (1-\eta^+)C^+ + (\eta^- - 1)C^- \geq (\eta^- - \eta^+) f(q_-^*).$$

Thus, the unconstrained optimizer q_-^* satisfies the constraint $q \in S_-$. Consequently, the constrained infimum achieves the unconstrained value: $\tilde{C}^- = C^-$, and the condition for Lemma 6 is satisfied:

$$(1-\eta^+)\tilde{C}^+ + (\eta^- - 1)\tilde{C}^- \leq (1-\eta^+)C^+ + (\eta^- - 1)C^- < 0.$$

Thus, $\mathrm{E}\, u(w_T) \sim -\exp[-(\eta^- - 1)C^- T]$. Additionally, $C^- < 0$. Since the expected utility diverges to $-\infty$, its relevant asymptotic approximation is the CRRA with $\eta = \eta^-$, and thus the certainty equivalent is C^-.

□

Proof of Lemma 2. Lemma 2 is implied by Lemma 5 with $f(q)$ specified as $f(q) = r(a, q)$. □

Proof of Lemma 3. The perfect-foresight DM achieves terminal wealth $w_T = \exp[v(q_T)T]$, where $v(q) = \max_{a\in A} r(a, q)$. Thus, Lemma 3 is implied by Lemma 5 with $f(q)$ specified as $f(q) = v(q)$. □

A.5 Proof of Proposition 3

Define

$$A^+ := \{a \in A :\ D(a) > 0\}, \qquad A^- := \{a \in A :\ D(a) < 0\}.$$

By Regularity condition 1, we have $\mathrm{cl}(A^+) = \{a \in A :\ D(a) \geq 0\}$. Recall that by Lemma 2,

$$C(a) = \begin{cases} C_{\mathsf{CRRA}}(a; \eta^+) & \text{if } a \in \mathrm{cl}(A^+), \\ C_{\mathsf{CRRA}}(a; \eta^-) & \text{if } a \in A^-, \end{cases}$$

where we extended $C(a) = C_{\mathsf{CRRA}}(a; \eta^+)$ on the event of a tie ($D(a) = 0$).

Recall $C_{\mathsf{CRRA}}(a;\eta^-) \le C_{\mathsf{CRRA}}(a;\eta^+)$ for all $a \in A$, and that $D(a) \ge 0$ implies $C_{\mathsf{CRRA}}(a;\eta^+) \ge 0$, while $D(a) < 0$ implies $C_{\mathsf{CRRA}}(a;\eta^-) < 0$. (See the beginning of Lemma 5 for the proof of the implications.)

We will write $\alpha^+ := 1-\eta^+ > 0$ and $\alpha^- := 1-\eta^- < 0$ for brevity. Recall that $d(a,\hat{a})$ is the Euclidean norm on A.

Lemma 7 (Local uniform upper bound near $\hat{a}$)**.** *Fix $\hat{a} \in A$ with $D(\hat{a}) \ge 0$. Then, for every $\hat{\varepsilon} > 0$ there exist $\hat{\delta} > 0$ and $\hat{T}$ such that for all $T \ge \hat{T}$ and all $a \in A$ with $d(a,\hat{a}) < \hat{\delta}$,*

$$C_T(a) \le C_{\mathsf{CRRA}}(\hat{a};\eta^+) + \hat{\varepsilon}.$$

Proof. Since $D(\hat{a}) \ge 0$, we have $C_{\mathsf{CRRA}}(\hat{a};\eta^+) \ge 0$. Write $U_T(a) := \int u(e^{r(a,q)T})\,\pi_T(dq)$ so that $u(e^{C_T(a)T}) = U_T(a)$. Fix $\kappa \in (0,\alpha^+)$. By (10), there exists $W_\kappa > 1$ such that $w^{\alpha^+-\kappa} \le u(w) \le w^{\alpha^++\kappa}$ for all $w \ge W_\kappa$, and $u(w) \le u(W_\kappa)$ for $w \le W_\kappa$. Hence,

$$U_T(a) \le u(W_\kappa) + \int \exp\big((\alpha^+ + \kappa)\, r(a,q)\, T\big)\, \pi_T(dq).$$

Fix $\varepsilon_1 > 0$. By continuity of $a \mapsto r(a,\theta)$ for all $\theta \in \Theta$, there exists $\hat{\delta} > 0$ such that $d(a,\hat{a}) < \hat{\delta}$ implies $r(a,q) \le r(\hat{a},q) + \varepsilon_1$ for all $q \in \Delta(\Theta)$. Let $R := \max_{a\in A,\theta\in\Theta} |r(a,\theta)|$. For $d(a,\hat{a}) < \hat{\delta}$ we have

$$U_T(a) \le u(W_\kappa) + e^{(\alpha^++\kappa)\varepsilon_1 T} \int \exp\big((\alpha^+ + \kappa)\, r(\hat{a},q)\, T\big)\, \pi_T(dq).$$

Therefore, when $U_T(a) > 0$ for large T, Varadhan's lemma yields the uniform bound[23]

$$\limsup_{T\to\infty} \sup_{d(a,\hat{a})<\hat{\delta}} \frac{1}{T} \ln U_T(a) \le \alpha^+ C_{\mathsf{CRRA}}(\hat{a};\eta^+) + \kappa R + (\alpha^+ + \kappa)\varepsilon_1.$$

Set $\hat{c} := C_{\mathsf{CRRA}}(\hat{a};\eta^+) + \hat{\varepsilon}$. If $C_T(a) \le \hat{c}$ we are done. Otherwise $C_T(a) > \hat{c} > 0$, so for all large T we may apply the lower bound $u(w) \ge w^{\alpha^+-\kappa}$ at $w = e^{C_T(a)T}$ to get

$$\frac{1}{T} \ln U_T(a) = \frac{1}{T} \ln u(e^{C_T(a)T}) \ge (\alpha^+ - \kappa)\, C_T(a).$$

Combining the last two displays and choosing $\kappa, \varepsilon_1 > 0$ sufficiently small (and then T large)

[23] Use $\ln(x+y) \le \max\{\ln x, \ln y\} + \ln 2$ and that $\frac{1}{T} \ln u(W_\kappa) \to 0$. Varadhan gives $\frac{1}{T} \ln \int e^{\beta r(\hat{a},q)T}\, d\pi_T \to \sup_q\{\beta r(\hat{a},q) - I(q)\}$. Finally, since $|r(\hat{a},q)| \le R$, $\sup_q\{(\alpha^++\kappa) r(\hat{a},q) - I(q)\} \le \sup_q\{\alpha^+ r(\hat{a},q) - I(q)\} + \kappa R = \alpha^+ C_{\mathsf{CRRA}}(\hat{a};\eta^+) + \kappa R$.

yields $C_T(a) \le C_{\mathsf{CRRA}}(\hat a;\eta^+) + \hat\varepsilon$ uniformly over $d(a,\hat a) < \hat\delta$. □

Lemma 8 (Uniform convergence on compact lower-tail sets)**.** *Let $K \subseteq A$ be compact and assume $D(a) < 0$ for all $a \in K$. Then, $\sup_{a\in K} \left|C_T(a) - C_{\mathsf{CRRA}}(a;\eta^-)\right| \to 0$.*

Proof. Since $D < 0$ on the compact set K, there exists $\gamma > 0$ such that $D(a) \le -\gamma$ for all $a \in K$. For $\beta \in \mathbb{R}$ define

$$\Lambda_T^\beta(a) := \frac{1}{T}\ln\int \exp\big(\beta\, r(a,q)\, T\big)\, \pi_T(dq), \qquad \Lambda^\beta(a) := \sup_{q\in\Delta(\Theta)} \{\beta\, r(a,q) - I(q)\}.$$

Then, for each fixed β, $\Lambda_T^\beta \to \Lambda^\beta$ uniformly on K.[24] Moreover $\Lambda^{\alpha^+}(a) = \alpha^+ C_{\mathsf{CRRA}}(a;\eta^+)$ and $\Lambda^{\alpha^-}(a) = \alpha^- C_{\mathsf{CRRA}}(a;\eta^-)$, hence $D(a) = \Lambda^{\alpha^+}(a) - \Lambda^{\alpha^-}(a)$ and therefore,

$$\Lambda_T^{\alpha^-}(a) \ge \Lambda_T^{\alpha^+}(a) + \gamma/2 \qquad (a \in K)$$

for all large T.

Fix $\kappa \in (0, |\alpha^-|)$ and choose $W_\kappa > 1$ such that

$$u(w) \le w^{\alpha^+ + \kappa} \quad (w \ge W_\kappa), \qquad w^{\alpha^- + \kappa} \le -u(w) \le w^{\alpha^- - \kappa} \quad (0 < w \le W_\kappa^{-1}).$$

Let $M_\kappa := \sup_{w\in[W_\kappa^{-1}, W_\kappa]} |u(w)| < \infty$ and define $U_T(a) := \int u(e^{r(a,q)T})\, \pi_T(dq)$ so that $u(e^{C_T(a)T}) = U_T(a)$. Decompose $U_T(a) = U_T^+(a) + U_T^0(a) + U_T^-(a)$ according to whether $e^{r(a,q)T} \ge W_\kappa$, $W_\kappa^{-1} < e^{r(a,q)T} < W_\kappa$, or $e^{r(a,q)T} \le W_\kappa^{-1}$. Then $|U_T^0(a)| \le M_\kappa$ and, for all a,

$$0 \le U_T^+(a) \le \int \exp\big((\alpha^+ + \kappa)\, r(a,q)\, T\big)\, \pi_T(dq),$$

while

$$\int_{\{e^{r(a,q)T} \le W_\kappa^{-1}\}} \exp\big((\alpha^- {+} \kappa)\, r(a,q)\, T\big)\, \pi_T(dq) \le -U_T^-(a) \le \int_{\{e^{r(a,q)T} \le W_\kappa^{-1}\}} \exp\big((\alpha^- {-} \kappa)\, r(a,q)\, T\big)\, \pi_T(dq).$$

Using the gap $\Lambda_T^{\alpha^-} \ge \Lambda_T^{\alpha^+} + \gamma/2$ on K, the positive-tail and middle terms are exponentially dominated by the negative-tail term, hence $U_T(a) < 0$ for all large T and

$$\sup_{a\in K} \left| \frac{1}{T} \ln\big(- U_T(a)\big) - \Lambda_T^{\alpha^-}(a) \right| \to 0.$$

[24] As in the CRRA lemma, $\{\Lambda_T^\beta\}_T$ is equicontinuous in a and Varadhan's lemma gives pointwise convergence to the continuous limit Λ^β (Berge), hence convergence is uniform on compact K.

Since $-U_T(a)$ grows exponentially on K, we are in the lower tail of u and may invert the lower-tail bounds to obtain

$$\sup_{a\in K}\left|C_T(a)-\frac{1}{\alpha^-}\frac{1}{T}\ln\big(-U_T(a)\big)\right|\to 0.$$

Finally, $\Lambda_T^{\alpha^-}\to\Lambda^{\alpha^-}=\alpha^- C_{\mathsf{CRRA}}(\cdot;\eta^-)$ uniformly on K, so $\sup_{a\in K}|C_T(a)-C_{\mathsf{CRRA}}(a;\eta^-)|\to 0$. □

Proof of Proposition 3. It suffices to show $\bar{a}=a^*_{\mathsf{tt}}$ for any converging subsequence $a^*_T\to\bar{a}$.

Case $D(a^*_{\mathsf{tt}})\geq 0$: We first show $D(\bar{a})\geq 0$. Suppose for contradiction that $D(\bar{a})<0$. Choose $a^+\in A^+$. Then, $C_{\mathsf{CRRA}}(a^+;\eta^+)>0$ and, by pointwise convergence $C_T(a)\to C_{\mathsf{CRRA}}(a;\eta^+)$ on A^+, we have $C_T(a^+)>0$ for all large T. On the other hand, since $D(\bar{a})<0$ and D is continuous, there exists $\delta>0$ such that $D<0$ on the closed ball $\overline{N}:=\{a\in A:\ d(a,\bar{a})\leq\delta\}$. Then, $C_{\mathsf{CRRA}}(a;\eta^-)<0$ on $\overline{N}$ and hence $c_-:=\max_{a\in\overline{N}}C_{\mathsf{CRRA}}(a;\eta^-)<0$. Applying Lemma 8 to $K=\overline{N}$, for all large T, $C_T(a)\leq c_-/2<0$ on $\overline{N}$. Since $a^*_T\to\bar{a}$,[25] for all large T we also have $a^*_T\in\overline{N}$, so $C_T(a^*_T)<0$, contradicting maximality because $C_T(a^*_T)\geq C_T(a^+)>0$ for all large T. Therefore, $D(\bar{a})\geq 0$.

Now, $D(\bar{a})\geq 0$ implies $C(\bar{a})=C_{\mathsf{CRRA}}(\bar{a};\eta^+)$. If $\bar{a}\neq a^*_{\mathsf{tt}}$, then by the assumed uniqueness of the asymptotic optimizer, $C(a^*_{\mathsf{tt}})>C(\bar{a})$. Fix $\varepsilon>0$ such that $C(\bar{a})+3\varepsilon<C(a^*_{\mathsf{tt}})$. Choose $a_\varepsilon\in A^+$ sufficiently close to a^*_{tt} so that, by continuity of $C_{\mathsf{CRRA}}(\cdot;\eta^+)$,

$$C(a_\varepsilon)=C_{\mathsf{CRRA}}(a_\varepsilon;\eta^+)\geq C_{\mathsf{CRRA}}(a^*_{\mathsf{tt}};\eta^+)-\varepsilon=C(a^*_{\mathsf{tt}})-\varepsilon.$$

By pointwise convergence on A^+, for all large T we have $C_T(a_\varepsilon)\geq C(a^*_{\mathsf{tt}})-2\varepsilon$, hence by optimality $C_T(a^*_T)\geq C_T(a_\varepsilon)\geq C(a^*_{\mathsf{tt}})-2\varepsilon$. By Lemma 7 applied to $\hat{a}=\bar{a}$, and since $a^*_T\to\bar{a}$, we have $C_T(a^*_T)\leq C(\bar{a})+\varepsilon$ for all large T. Thus, for all large T,

$$C(a^*_{\mathsf{tt}})-2\varepsilon\leq C_T(a^*_T)\leq C(\bar{a})+\varepsilon, \tag{17}$$

contradicting $C(\bar{a})+3\varepsilon<C(a^*_{\mathsf{tt}})$. Hence $\bar{a}=a^*_{\mathsf{tt}}$.

Since this holds for every convergent subsequence of $\{a^*_T\}$, we have $a^*_T\to a^*_{\mathsf{tt}}$. Inequality (17) (with $\bar{a}=a^*_{\mathsf{tt}}$) yields $\max_{a\in A}C_T(a)\to C(a^*_{\mathsf{tt}})=\mathbb{C}_{\mathsf{tt}}$.

Case $D(a^*_{\mathsf{tt}})<0$: Then, $C(a^*_{\mathsf{tt}})=C_{\mathsf{CRRA}}(a^*_{\mathsf{tt}};\eta^-)<0$. If there existed $a\in A$ with $D(a)\geq 0$, then $C(a)=C_{\mathsf{CRRA}}(a;\eta^+)\geq 0$, contradicting maximality of a^*_{tt}. Hence $D<0$ on A, so

[25] The convergence is along a convergent subsequence T_k of T, but we omit relabeling for brevity.

$C(\cdot) = C_{\mathsf{CRRA}}(\cdot;\eta^-)$ on A. Lemma 8 yields uniform convergence $C_T \to C$ on A. The standard uniform-convergence/uniqueness argument implies $a^*_T \to a^*_{\mathsf{tt}}$ and $\max_{a\in A} C_T(a) \to \mathbb{C}_{\mathsf{tt}}$. □

A.6 Proof of Proposition 4

Proof of Proposition 4. Consider a hedging strategy $\tilde{\mathbf{a}}$ with two subfunds of equal initial sizes and actions $a^*_+ = a^*_{\mathsf{CRRA}}(\eta^+)$ and $a^*_- = a^*_{\mathsf{CRRA}}(\eta^-)$ assigned to these subfunds, respectively. We have,

$$\mathbb{C}_{\mathsf{pf},T} \geq \sup_{\mathbf{a}} C_{\mathsf{hedge},T}(\mathbf{a}) \geq C_{\mathsf{hedge},T}(\tilde{\mathbf{a}}).$$

Since the first two quantities converge to $\mathbb{C}_{\mathsf{pf}}$ and $\mathbb{C}_{\mathsf{hedge}}$, respectively, it suffices to prove that

$$\liminf_{T\to\infty} C_{\mathsf{hedge},T}(\tilde{\mathbf{a}}) \geq \mathbb{C}_{\mathsf{pf}}.$$

The expected utility generated by $\tilde{\mathbf{a}}$ exceeds

$$\mathrm{E}\, u\left(\frac{1}{2}\exp\left[\tilde{v}(q_T)T\right]\right), \tag{18}$$

where

$$\tilde{v}(q) = \max\left\{r(a^*_+, q), r(a^*_-, q)\right\},$$

is the growth rate for the more highly performing subfund for the large deviation q.

Noting that the term 1/2 in (18) has only a subexponential impact on $C_{\mathsf{hedge},T}(\tilde{\mathbf{a}})$, we get that[26]

$$\liminf_{T\to\infty} C_{\mathsf{hedge},T}(\tilde{\mathbf{a}}) \geq \tilde{\mathbb{C}},$$

where $\tilde{\mathbb{C}}$ is the asymptotic certainty equivalent of $w_T = \exp\left[\tilde{v}(q_T)T\right]$.

We prove $\tilde{\mathbb{C}} = \mathbb{C}_{\mathsf{pf}}$. Noting from Lemma 3 that

$$\mathbb{C}_{\mathsf{pf}} = \begin{cases} \mathbb{C}_{\mathsf{CRRA}}(\eta^+) & \text{if } (1-\eta^+)\mathbb{C}_{\mathsf{CRRA}}(\eta^+) + (\eta^- - 1)\mathbb{C}_{\mathsf{CRRA}}(\eta^-) > 0, \\ \mathbb{C}_{\mathsf{CRRA}}(\eta^-) & \text{if } (1-\eta^+)\mathbb{C}_{\mathsf{CRRA}}(\eta^+) + (\eta^- - 1)\mathbb{C}_{\mathsf{CRRA}}(\eta^-) < 0, \end{cases}$$

[26] The factor 1/2 does not affect the exponential scale: by the two-tailed property, $u(w/2)$ is asymptotically proportional to $u(w)$ as $w \to \infty$ and as $w \downarrow 0$, so scaling wealth by half changes $\mathrm{E}\, u$ only by a constant that does not affect the growth rate of the expected utility.

and from Lemma 5 that

$$\tilde{\mathbb{C}} = \begin{cases} C^+ & \text{if } (1-\eta^+)C^+ + (\eta^- - 1)C^- > 0, \\ C^- & \text{if } (1-\eta^+)C^+ + (\eta^- - 1)C^- < 0, \end{cases}$$

with

$$C^+ = \sup_{q\in\Delta(\Theta)} \{\tilde{v}(q) - \tfrac{1}{1-\eta^+} I(q)\}, \quad C^- = \inf_{q\in\Delta(\Theta)} \{\tilde{v}(q) + \tfrac{1}{\eta^- - 1} I(q)\},$$

it suffices to prove $\mathbb{C}_{\mathsf{CRRA}}(\eta^+) = C^+$ and $\mathbb{C}_{\mathsf{CRRA}}(\eta^-) = C^-$.

Since,

$$\mathbb{C}_{\mathsf{CRRA}}(\eta^+) = \sup_{q\in\Delta(\Theta)} \{r(a_+^*, q) - \tfrac{1}{1-\eta^+} I(q)\},$$

and $\tilde{v}(q) \geq r(a_+^*, q)$, we have $\mathbb{C}_{\mathsf{CRRA}}(\eta^+) \leq C^+$. But by Proposition 2, $\mathbb{C}_{\mathsf{CRRA}}(\eta^+)$ is the perfect-foresight certainty equivalent for the CRRA(η^+) utility, and hence

$$\mathbb{C}_{\mathsf{CRRA}}(\eta^+) = \sup_{q\in\Delta(\Theta)} \{v(q) - \tfrac{1}{1-\eta^+} I(q)\}.$$

Since $v(q) \geq \tilde{v}(q)$, we have $\mathbb{C}_{\mathsf{CRRA}}(\eta^+) \geq C^+$. Hence $\mathbb{C}_{\mathsf{CRRA}}(\eta^+) = C^+$, as needed.

Similarly, since

$$\mathbb{C}_{\mathsf{CRRA}}(\eta^-) = \inf_{q\in\Delta(\Theta)} \{r(a_-^*, q) + \tfrac{1}{\eta^- - 1} I(q)\},$$

and $\tilde{v}(q) \geq r(a_-^*, q)$, we have $\mathbb{C}_{\mathsf{CRRA}}(\eta^-) \leq C^-$. But by Proposition 2, $\mathbb{C}_{\mathsf{CRRA}}(\eta^-)$ is the perfect-foresight certainty equivalent for the CRRA(η^-) utility. Thus, as in the previous paragraph, $\mathbb{C}_{\mathsf{CRRA}}(\eta^-) \geq C^-$. Hence also $\mathbb{C}_{\mathsf{CRRA}}(\eta^-) = C^-$, as needed. □

B Finite-Horizon Corrections

Unlike our asymptotic results, results for finite horizons must engage with the details of the state-generating process since these details determine the convergence rate at which the large-deviation approximation governed by $I(q)$ becomes a valid approximation. Additionally, finite-horizon versions of the analysis in Section 7 must engage with the quality of the CRRA tail approximation of the DM's utility. Below, we discuss three special cases in which finite-horizon corrections can be obtained by simple arguments.

CRRA and i.i.d. states: The asymptotic formula for the certainty equivalent is exact at every horizon since the objective factorizes into a product of expectations of independent variables,

$$\mathrm{E}\left[w_T^{1-\eta}\right] = \prod_{t=1}^{T} \mathrm{E}\left[R(a,\theta_t)^{1-\eta}\right].$$

Then, the certainty equivalents $C_T(a)$ are independent of the horizon T.

CRRA and Markov chains: The finite-horizon correction of the certainty equivalents is of order $\mathcal{O}(1/T)$. Let θ_t evolve according to a finite-state Markov chain with a transition matrix $\mathbf{P}$. Recall that

$$C_T(a) = \frac{1}{T(1-\eta)} \ln \mathrm{E}\left[\exp\left((1-\eta)\sum_{t=1}^{T} r(a,\theta_t)\right)\right].$$

The expectation can be written as a matrix product:

$$\mathrm{E}\left[\exp\left((1-\eta)\sum_{t=1}^{T} r(a,\theta_t)\right)\right] = \pi_0^\top \left[\mathbf{M}(a)\right]^T \mathbf{1},$$

where $\pi_0^\top$ is the transposed initial distribution vector and $\mathbf{M}$ is the tilted matrix with entries $M_{ij}(a) = P_{ij} \exp\left[(1-\eta) r(a,j)\right]$.

Spectral expansion yields

$$\pi_0^\top \left[\mathbf{M}(a)\right]^T \mathbf{1} = c(a)\left[\lambda_1(a)\right]^T + \mathcal{O}\left(|\lambda_2(a)|^T\right),$$

where $\lambda_1(a)$ is the strictly positive dominant eigenvalue and $\lambda_2(a)$ is the second-largest eigenvalue, and $c(a) > 0$.

Taking the logarithm and dividing by $T(1-\eta)$ yields the finite-horizon expansion:

$$C_T(a) = \underbrace{\frac{\ln \lambda_1(a)}{1-\eta}}_{C_{\mathsf{CRRA}}(a;\eta)} + \underbrace{\frac{\ln c(a)}{T(1-\eta)}}_{\mathcal{O}(1/T)\ \text{correction}} + \mathcal{O}\left(\frac{e^{-\Delta T}}{T}\right),$$

where $\Delta = \ln(\lambda_1(a)/|\lambda_2(a)|) > 0$.

Perfect foresight, CRRA, and i.i.d. states: The value of perfect foresight,

$$\mathbb{C}_{\mathsf{pf},T} - \mathbb{C}_T,$$

relative to the uninformed certainty equivalent is bounded by a term of the order $\mathcal{O}(\ln T/T)$. We sketch the argument using the method of types (assuming the risk-neutral DM, $\eta = 0$, to cut down on notation).

Let $\mathcal{Q}_T$ denote the set of empirical distributions (types) attainable at horizon T. Then,

$$\mathbb{C}_{\mathsf{pf},T} = \frac{1}{T} \ln \sum_{q \in \mathcal{Q}_T} \Pr(q_T = q) e^{Tv(q)}.$$

Under i.i.d. states, $\Pr(q_T = q) \le e^{-T\mathrm{KL}(q\|p)}$, and hence

$$\begin{aligned} \mathbb{C}_{\mathsf{pf},T} &\le \frac{1}{T} \ln \sum_{q \in \mathcal{Q}_T} e^{T(v(q) - \mathrm{KL}(q\|p))}. \\ &\le \max_{q \in \mathcal{Q}_T} \{v(q) - \mathrm{KL}(q \parallel p)\} + \frac{\ln |\mathcal{Q}_T|}{T}. \end{aligned}$$

Recall that the uninformed certainty equivalent $\mathbb{C}_T := \sup_a C_T(a)$ is independent of T for i.i.d. states and hence equal to the asymptotic certainty equivalent. We use $I(q) = \mathrm{KL}(q \parallel p)$, where p is the state-generating distribution, and Proposition 2 to get

$$\mathbb{C}_T = \max_{q \in \Delta(\Theta)} \{v(q) - \mathrm{KL}(q \parallel p)\}.$$

Thus, $\max_{q \in \mathcal{Q}_T} \{v(q) - \mathrm{KL}(q \parallel p)\} \le \mathbb{C}_T$, and

$$\mathbb{C}_{\mathsf{pf},T} \le \mathbb{C}_T + \frac{\ln |\mathcal{Q}_T|}{T}.$$

The result then follows from the fact that $|\mathcal{Q}_T|$ is polynomial in T.

References

Arneson, R. J. (2022). *Prioritarianism.* Cambridge University Press.

Barro, R. J. (2006). Rare disasters and asset markets in the twentieth century. *The Quarterly Journal of Economics 121*(3), 823–866.

Barro, R. J. (2009). Rare disasters, asset prices, and welfare costs. *American Economic Review 99*(1), 243–264.

Barro, R. J. and T. Jin (2011). On the size distribution of macroeconomic disasters. *Econometrica 79*(5), 1567–1589.

Barro, R. J. and J. F. Ursúa (2012). Rare macroeconomic disasters. *Annu. Rev. Econ. 4*(1), 83–109.

Brock, W. A. (1970). An axiomatic basis for the Ramsey-Weizsäcker overtaking criterion. *Econometrica: 38*(6), 927–929.

Bryc, W. and A. Dembo (1996). Large deviations and strong mixing. In *Annales de l'IHP Probabilités et statistiques*, Volume 32, pp. 549–569.

Caplin, A. and M. Dean (2015). Revealed preference, rational inattention, and costly information acquisition. *American Economic Review 105*(7), 2183–2203.

Caplin, A. and J. V. Leahy (2019). Wishful thinking. Technical report, National Bureau of Economic Research.

Chabakauri, G., K. Yuan, and K. E. Zachariadis (2022). Multi-asset noisy rational expectations equilibrium with contingent claims. *The Review of Economic Studies 89*(5), 2445–2490.

Chichilnisky, G. (2009). The topology of fear. *Journal of Mathematical Economics 45*(12), 807–816.

Cover, T. M. and J. A. Thomas (2006). *Elements of Information Theory* (2nd ed.). Hoboken, NJ: Wiley-Interscience.

Dembo, A. and O. Zeitouni (1998). *Large Deviations: Techniques and Applications* (Second ed.). New York: Springer.

den Hollander, F. (2000). *Large deviations.* Providence: American Mathematical Society.

Dybvig, P. H., L. Rogers, and K. Back (1999). Portfolio turnpikes. *The Review of Financial Studies 12*(1), 165–195.

Fleurbaey, M. (2010). Assessing risky social situations. *Journal of Political Economy 118*(4), 649–680.

Gollier, C. and M. L. Weitzman (2010). How should the distant future be discounted when discount rates are uncertain? *Economics Letters 107*(3), 350–353.

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

Hansen, L. P. and J. A. Scheinkman (2009). Long-term risk: An operator approach. *Econometrica 77*(1), 177–234.

Kelly, J. L. (1956). A new interpretation of information rate. *The Bell System Technical Journal 35*(4), 917–926.

Lucas, Jr., R. E. (1987). *Models of Business Cycles*. Yrjö Jahnsson Lectures. Oxford: Basil Blackwell.

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

Martin, I. (2012). On the valuation of long-dated assets. *Journal of Political Economy 120*(2), 346–358.

Martin, I. and D. Papadimitriou (2022). Sentiment and speculation in a market with heterogeneous beliefs. *American Economic Review 112*(8), 2465–2517.

Millner, A. (2013). On welfare frameworks and catastrophic climate risks. *Journal of Environmental Economics and Management 65*(2), 310–325.

Millner, A. (2025). Ambiguity and the language of long run risk. Technical report, National Bureau of Economic Research.

Rietz, T. A. (1988). The equity risk premium a solution. *Journal of Monetary Economics 22*(1), 117–131.

Robson, A., L. Samuelson, and J. Steiner (2023). Decision theory and stochastic growth. *American Economic Review: Insights 5*(3), 357–376.

Rubinstein, A. (1979). Equilibrium in supergames with the overtaking criterion. *Journal of Economic Theory 21*(1), 1–9.

Russell, J. S. (2024). On two arguments for fanaticism. *Noûs 58*(3), 565–595.

Samuelson, L. and J. Steiner (2025). Growth and redistribution: The hedging perspective. *American Economic Review: Insights 7*(2), 250–267.

Samuelson, P. A. (1971). The "fallacy" of maximizing the geometric mean in long sequences of investing or gambling. *Proceedings of the National Academy of Sciences 68*(10), 2493–2496.

Samuelson, P. A. (1979). Why we should not make mean log of wealth big though years to act are long. *Journal of Banking and Finance 3*, 305–307.

Strzalecki, T. (2011). Axiomatic foundations of multiplier preferences. *Econometrica 79*(1), 47–73.

Stutzer, M. (2003). Portfolio choice with endogenous utility: A large deviations approach. *Journal of Econometrics 116*(1-2), 365–386.

Thoma, J. (2023). Taking risks on behalf of another. *Philosophy Compass 18*(3), e12898.

Walden, J. (2019). Trading, profits, and volatility in a dynamic information network model. *The Review of Economic Studies 86*(5), 2248–2283.

Weitzman, M. L. (1998). Why the far-distant future should be discounted at its lowest possible rate. *Journal of Environmental Economics and Management 36*(3), 201–208.

Weitzman, M. L. (2009). On modeling and interpreting the economics of catastrophic climate change. *The Review of Economics and Statistics 91*(1), 1–19.