The Equity Premium Trap

Harjoat S. Bhamra

March 26, 2026

Abstract

The Hansen–Jagannathan bound requires the variance of the stochastic discount factor to be finite. Under power-law consumption tails—a feature present in U.S. and international data—this creates a trap with two sides. The first side: matching the equity Sharpe ratio requires risk aversion $\gamma \geq 20$, but the SDF variance is finite only for $\gamma < \alpha_C/2$, where α_C is the consumption tail index. At empirical estimates ($\alpha_C \approx 2.7$), the ceiling is 1.35. The bound does not exist at the risk aversion it identifies. The second side: Martin and Shi (2026) estimate from option prices that $\theta^* \approx 1.7$—below two, confirming that the SDF variance is infinite. If bond prices exist ($\mathbb{E}[M] < \infty$), the structural formula $\theta^* = \alpha_C/\gamma$ pins risk aversion to $\alpha_C/\theta^* < \gamma < \alpha_C$. At $\alpha_C \approx 2.7$, this gives $\gamma \in (1.6, 2.7)$: moderate, but above the HJ threshold. Option prices and consumption tails independently agree. The trap extends beyond CRRA to any pricing kernel whose equity premium is driven by its response to disasters.

1 Introduction

The equity premium puzzle rests on the Hansen–Jagannathan bound: the standard deviation of the stochastic discount factor must exceed the Sharpe ratio of every traded portfolio. Matching the postwar equity Sharpe ratio of 0.44 with consumption growth volatility of 2.1% requires risk aversion of at least 20. But the bound presupposes that the SDF has a finite variance. This paper shows that power-law consumption tails—a feature documented in U.S. and international data—make that presupposition untenable: risk aversion divides the tail index, and at the risk aversion the bound identifies, the variance is infinite.

More precisely, risk aversion divides the tail index. The θth moment of the SDF is finite when $\gamma < \alpha_C/\theta$ and infinite when $\gamma > \alpha_C/\theta$, where α_C is the consumption tail index (Proposition 1). The HJ bound needs $\theta = 2$, so finite variance requires $\gamma < \alpha_C/2$. Five extreme value estimators applied to quarterly NIPA data place α_C in [2.1, 3.0]. At $\alpha_C = 2.7$, the ceiling is $\gamma < 1.35$. The equity premium requires $\gamma \geq 20$. The bound does not exist at the risk aversion it identifies. I call this the *equity premium trap*.

Option prices tell the same story independently. Martin and Shi (2026) estimate that θ^*—the highest power of the SDF with a finite moment—is approximately 1.7, below two. Any model with $\mathrm{Var}(M) < \infty$ predicts $\theta^* \geq 2$; the data say $\theta^* < 2$. But $\theta^* > 1$, since bond prices exist and $\mathbb{E}[M] < \infty$. Together, $1 < \theta^* < 2$ and the structural formula $\theta^* = \alpha_C/\gamma$ pin risk aversion to a narrow interval: $\alpha_C/2 < \gamma < \alpha_C$. At $\alpha_C \approx 2.7$ and $\theta^* \approx 1.7$, the point estimate is $\gamma \approx 1.6$, with the interval (1.35, 2.7)—moderate risk aversion, consistent with the disaster-risk literature. The lower boundary of the interval is the threshold above which the HJ bound ceases to exist. Consumption tails and option prices agree: the SDF has a finite mean but infinite variance.

This is the trap. The HJ bound implies $\gamma \geq 20$; the tail-risk result implies $\mathrm{Var}(M) = \infty$ at $\gamma = 20$; and the option evidence confirms it.

Every resolution must give up something. I classify four families of escape (Theorem 4): (A) thin tails, as in Wachter (2013)—but the data reject $\alpha_C = \infty$; (B) low disaster-loading, as in Schreindorfer (2020)—the equity premium comes from a bounded disappointment factor rather than from high γ; (C) structural departure from the consumption-based class, as in intermediary pricing (Calvet et al., 2025); and (D) risk-neutral diagnostics that avoid SDF moments entirely, as in Martin (2017). I audit nine canonical models against this classification (Appendix I). Mehra and Prescott (1985), Bansal et al. (2012), and Barro (2006) are trapped. In the i.i.d. exchange benchmark, Epstein–Zin preferences do not help: the EIS drops out of the moment conditions and the trap depends on γ alone (Proposition 4). Whether persistent long-run-risk dynamics change this is an open question.

If the HJ bound does not exist, what replaces it? I develop a hierarchy of diagnostics that remain well-posed when the variance is infinite. An interquantile diagnostic replaces the standard deviation with an interquantile range, requiring only a finite mean. This doubles the admissible range of γ—but matching the equity premium still requires γ above the doubled ceiling. More broadly, the L^θ family of pricing bounds is valid at every $\theta < \theta^*$, yet the same tension reappears at every one. Each diagnostic preserves the economic content of the HJ bound: continuity under tail perturbations, monotonicity in risk aversion, and model ranking preservation.

Relation to the literature. The equity premium puzzle is due to Mehra and Prescott (1985); the HJ bound is its sharpest formulation (Hansen and Jagannathan, 1991). Toda and Walsh (2015) show that cross-sectional consumption has $\alpha_C \approx 4$ and that GMM estimation of the Euler equation is inconsistent for $\gamma \geq \alpha_C$. I show the problem begins earlier: $\gamma \geq \alpha_C/2$ breaks the HJ bound, and this extends to every variance-based diagnostic. Martin and Shi (2026) estimate θ^* from option prices; the structural formula $\theta^* = \alpha_C/\gamma$, derived here,

gives their finding a consumption-based foundation and makes it interpretable as a bound on risk aversion. Martin's (2017) SVIX bound is immune to the trap by construction because it uses risk-neutral moments. The probability tools are textbook (Bingham et al., 1987; Resnick, 2007). The contribution is the economic architecture: a theory connecting consumption tails to the existence of standard diagnostics, a taxonomy of escapes with testable implications, independent evidence from options, consumption, and international data, and a diagnostic framework for heavy-tailed environments.

Plan of the paper. Section 2 proves the tail-risk result, derives the contradiction, classifies the escapes, and shows the EIS is irrelevant in the i.i.d. exchange benchmark. Section 3 takes the theory to data. Section 4 evaluates each escape route. Section 5 concludes.

2 The Trap

This section builds the trap in four steps. Section 2.1 establishes a formula that connects consumption tails, risk aversion, and the moments of the stochastic discount factor. Section 2.2 shows that the formula implies a structural impossibility: matching the equity premium and having a well-defined Hansen–Jagannathan bound are mutually exclusive. Section 2.3 asks whether a weaker bound can escape the impossibility; the answer is no. Section 2.4 closes the Epstein–Zin escape route: the elasticity of intertemporal substitution is irrelevant to the trap.

2.1 The Formula

I begin with two assumptions about the primitives of the economy: consumption growth and the stochastic discount factor.

Assumption 1 (Power-law consumption tails)**.** Consumption growth $g = C_{t+1}/C_t$ has a regularly varying left tail with index $\alpha_C \in (0, \infty)$:

$$\mathbb{P}(g \leq x) \;\sim\; x^{\alpha_C}\, \ell(1/x) \quad \text{as } x \to 0^+,$$

where ℓ is a slowly varying function.

Assumption 1 says that the probability of a consumption disaster of size x decays like a power law. The index α_C governs how heavy the left tail is: smaller α_C means heavier tails, more extreme disasters. At $\alpha_C = \infty$, the left tail is exponentially thin—no power law. The empirical evidence, reviewed in Section 3, gives $\alpha_C \leq 4$.

As introduced in Section 1, the reciprocal $\xi_C = 1/\alpha_C$ measures *tail risk*: higher ξ_C means heavier tails, just as higher σ means more volatility. Throughout the paper I use both notations interchangeably; α_C is standard in probability, while ξ_C makes the economics transparent.

The next definition uses the concept of *regular variation*, which is simply a precise way of saying "behaves like a power function in the tail." A function h is *regularly varying* with index $-\gamma$ at 0^+ if, for any rescaling factor $\lambda > 0$, shrinking the argument by λ multiplies the function by $\lambda^{-\gamma}$—exactly as a power function $x^{-\gamma}$ would. The notation $h \in \mathrm{RV}(-\gamma)$ records this. The function need not *equal* $x^{-\gamma}$; it only has to *look like* $x^{-\gamma}$ as $x \to 0^+$, up to a slowly moving correction that becomes negligible in the tail.

Definition 1 (Class $\mathcal{S}$)**.** Let $g = C_{t+1}/C_t$ denote consumption growth. The class $\mathcal{S}$ consists of all strictly positive random variables $M = h(g)$ where $h : (0, \infty) \to (0, \infty)$ satisfies

$$h \in \mathrm{RV}(-\gamma) \;\text{ at } 0^+$$

for some $\gamma > 0$—that is, $h(\lambda x)/h(x) \to \lambda^{-\gamma}$ as $x \to 0^+$ for every $\lambda > 0$—and h is eventually monotone decreasing near 0^+ (equivalently, $\tilde{h}(y) := h(1/y)$ is eventually monotone increasing at infinity), so that the standard regular-variation composition theorem applies.

The class $\mathcal{S}$ captures a simple idea: the SDF responds to consumption disasters with power-like intensity. The parameter γ measures how strongly the pricing kernel loads on the left tail of consumption growth. The class is large. It contains CRRA ($h(x) = \beta x^{-\gamma}$), Epstein–Zin with any elasticity of intertemporal substitution under i.i.d. growth (Section 2.4), and any SDF whose kernel loads on the left tail of g with exponent γ. The key restriction is that h behaves like a power function in the tail—not that it *is* a power function everywhere.

Two features of the class are worth noting. First, Schreindorfer (2020)'s generalized disappointment aversion at $\alpha = 1$ *is* in $\mathcal{S}$ with $\gamma = 1$: the SDF $M = \tilde{\beta}\, g^{-1} \cdot D$ diverges as $g \to 0^+$ with the same exponent as log utility. The bounded disappointment factor D amplifies the equity premium but does not affect the tail index. Because $\gamma = 1 < \alpha_C/2$, the trap does not spring. Second, the Calvet et al. (2025) intermediary model is *not* in $\mathcal{S}$: the SDF depends on market structure rather than on a representative agent's consumption, and the effective tail loading is exponentially thin. Section 4 classifies these and other escape routes.

Remark 1 (Campbell–Cochrane). Whether the Campbell and Cochrane (1999) external habit model belongs to $\mathcal{S}$ is an open question. Their SDF $M_{t+1} = \beta(S_{t+1}/S_t)^{-\gamma}(C_{t+1}/C_t)^{-\gamma}$ depends on the surplus consumption ratio S_t, not on consumption growth alone. The surplus ratio dynamics create state-dependent effective curvature, so M is not a time-invariant function of g and the regular variation condition of Definition 1 does not apply directly. If the composite SDF were in $\mathcal{S}$ with effective $\gamma \approx 80$ (the local curvature near the habit level, where the surplus ratio S_t is small), the model would be deeply trapped: $\gamma = 80 \gg \alpha_C/2$. But the classification requires a formal analysis of the joint tail behavior of (g, S), which is beyond the scope of this paper. I therefore do not list Campbell–Cochrane as definitively in or out of $\mathcal{S}$.

Assumption 2 (SDF in class $\mathcal{S}$)**.** $M = h(g) \in \mathcal{S}$ with tail parameter $\gamma > 0$.

The next result is the engine of the paper. It says that risk aversion divides the tail index. I first state a general version that applies beyond class $\mathcal{S}$ and then derive the class-$\mathcal{S}$ special case as a corollary.

Theorem 1 (Tail-risk amplification)**.** *Suppose the pricing kernel admits a multiplicative decomposition $M = \Psi(X, \varepsilon)$ where:*

(F1) *$\varepsilon > 0$ is independent of X.*

(F2) *For each fixed ε, the map $x \mapsto \Psi(x, \varepsilon)$ is eventually monotone decreasing as $x \to 0^+$ and admits the uniform asymptotic factorisation*

$$\Psi(x, \varepsilon) \;=\; a(\varepsilon)\, x^{-\eta}\, L(1/x)\, r(x, \varepsilon), \qquad r(x, \varepsilon) \to 1 \text{ as } x \to 0^+,$$

where $\eta > 0$ (the tail elasticity*) and the slowly varying function L do not depend on ε, $a(\varepsilon) > 0$ is a measurable coefficient function, and the convergence $r(x, \varepsilon) \to 1$ is uniform in ε on every set $\{a(\varepsilon) \le m_0\}$ for $m_0 < \infty$.*

(F3) *$\mathbb{E}[a(\varepsilon)^{\alpha_X/\eta + \delta}] < \infty$ for some $\delta > 0$ (the coefficient function is light-tailed).*

Then, under Assumption 1 with any disaster-exposed variable X satisfying $\mathbb{P}(X \le x) \in \mathrm{RV}(\alpha_X)$ as $x \to 0^+$:

(a) M has a regularly varying right tail with index $\alpha_M = \alpha_X/\eta$.

(b) $\mathbb{E}[M^k] < \infty$ if and only if $\eta < \alpha_X/k$.

(c) $\theta^* := \sup\{k > 0 : \mathbb{E}[M^k] < \infty\} = \alpha_X/\eta$.

(d) $\mathrm{Var}(M) < \infty$ if and only if $\eta < \alpha_X/2$.

Theorem 1 gives a framework-conditional tail-risk result under explicit asymptotic factorisation assumptions. Within the class of SDFs satisfying conditions (F1)–(F3), any pricing kernel whose response to disasters has a well-defined tail elasticity η inherits the power-law tail of the underlying risk factor, with tail index divided by η. Proposition 1 verifies that these conditions hold exactly in the CRRA benchmark. The coefficient function $a(\varepsilon)$ captures the ε-dependent scaling of the SDF's disaster response—for example, GDA's disappointment factor D, time preference β, or idiosyncratic shocks—as long as its moments are finite relative to the X-driven component. The proof uses Breiman's theorem for products of regularly varying random variables; the full argument is in Appendix A and the companion note Bhamra (2026c).

When the SDF belongs to class $\mathcal{S}$, the general theorem specialises as follows.

Proposition 1 (Tail risk)**.** *Under Assumptions 1 and 2:*

(a) The SDF has a regularly varying right tail with index $\alpha_M = \alpha_C/\gamma$.

(b) $\mathbb{E}[M^k] < \infty$ if and only if $\gamma < \alpha_C/k$.

(c) In particular, $\mathrm{Var}(M) < \infty$ if and only if $\gamma < \alpha_C/2$.

Proposition 1 is the special case of Theorem 1 with $X = g$, $\varepsilon = 1$ (deterministic), $\eta = \gamma$, and $\Psi(x, \varepsilon) = h(x)$. The coefficient function is $a(\varepsilon) = a(1)$, a constant, so condition (F3) is trivially satisfied.

Proof sketch. Part (a) is the power-transform rule for regularly varying functions. Since g has a left tail with index α_C and $h \in \mathrm{RV}(-\gamma)$ at 0^+, the composite $M = h(g)$ has a right tail with index α_C/γ (Bingham et al., 1987, Theorem 1.5.6). Part (b) applies Karamata's moment theorem: a non-negative random variable with a regularly varying right tail of index ξ has $\mathbb{E}[M^k] < \infty$ for $k < \xi$ and $\mathbb{E}[M^k] = \infty$ for $k > \xi$ (Bingham et al., 1987, Theorem 1.5.11).[1] Here $\xi = \alpha_C/\gamma$, so $\mathbb{E}[M^k] < \infty$ for $k < \alpha_C/\gamma$, i.e., $\gamma < \alpha_C/k$. Part (c) is the special case $k = 2$. The full proof is in Appendix A. □

In the tail-risk notation of Section 1, the formula reads $\xi_M = \gamma\,\xi_C$: risk aversion *multiplies* tail risk. The k-th moment of M is finite if and only if $\gamma\,\xi_C < 1/k$, so the classical HJ bound ($k = 2$) requires $\gamma\,\xi_C < \frac{1}{2}$ and the Euler equation ($k = 1$) requires $\gamma\,\xi_C < 1$. The product $\gamma\,\xi_C$ is the single number that governs which tools survive.

Corollary 1 (Structural θ^*)**.** *Under Assumptions 1 and 2, the critical exponent of the SDF is*

$$\theta^* \;:=\; \sup\{k > 0 : \mathbb{E}[M^k] < \infty\} \;=\; \frac{\alpha_C}{\gamma} \;=\; \frac{1}{\gamma\,\xi_C}. \tag{1}$$

[1] The boundary case $k = \xi$ depends on the slowly varying function and cannot be determined from the index alone. This distinction is immaterial for all results in this paper.

Martin and Shi (2026) estimate $\theta^* \approx 1.7$ from traded option prices, without reference to consumption data or preference parameters. Their estimate is a risk-neutral object, extracted from option-implied tail decay rates. The structural formula (1) provides a physical-measure interpretation: under the maintained bridge of class $\mathcal{S}$, the option-implied θ^* equals the physical-SDF critical exponent α_C/γ. The word "model-free" applies to the risk-neutral option statement itself; the subsequent physical interpretation is conditional on the structural bridge. At $\alpha_C = 4$ (Toda and Walsh, 2015) and $\theta^* = 1.7$, the implied risk aversion is $\gamma = 4/1.7 \approx 2.4$—moderate, and far below the $\gamma \geq 20$ that the classical Hansen–Jagannathan bound demands. Two independent literatures—extreme value theory and option pricing—converge on the same conclusion.

The formula also explains *why* $\theta^* < 2$. The classical HJ bound requires $\mathbb{E}[M^2] < \infty$, i.e., $\theta^* > 2$. Since $\theta^* = \alpha_C/\gamma$ and matching the equity premium forces γ well above $\alpha_C/2$, the inequality $\theta^* > 2$ is violated. Martin and Shi's finding is not a mysterious empirical anomaly. It is a structural consequence of the trap.

Corollary 2 (Risk aversion bounds)**.** *Under Assumptions 1 and 2, the following are equivalent:*

(a) $1 < \theta^* < 2$ *(*$\mathbb{E}[M] < \infty$ *and* $\mathrm{Var}(M) = \infty$*).*

(b) $\alpha_C/2 < \gamma < \alpha_C$*.*

If these conditions hold, the point estimate is $\gamma = \alpha_C/\theta^*$*. At* $\theta^* \approx 1.7$ *(Martin and Shi, 2026):*

$$\begin{aligned} &\alpha_C \approx 2.7 \text{ (NIPA aggregate):} && \gamma \in (1.35,\ 2.7), && \text{point estimate } 1.6. \\ &\alpha_C \approx 4.0 \text{ (Toda and Walsh, 2015):} && \gamma \in (2.0,\ 4.0), && \text{point estimate } 2.4. \end{aligned}$$

Proof sketch. The biconditional is immediate from Corollary 1: $\theta^* = \alpha_C/\gamma$. The map $\gamma \mapsto \alpha_C/\gamma$ is strictly decreasing, so $\theta^* < 2 \Leftrightarrow \gamma > \alpha_C/2$ and $\theta^* > 1 \Leftrightarrow \gamma < \alpha_C$. Full proof, with an alternative derivation and CRRA verification, in Appendix A. □

The lower bound $\gamma > \alpha_C/2$ uses only $\theta^* < 2$—the qualitative finding that $\mathrm{Var}(M) = \infty$—which is Martin and Shi's headline result. The upper bound $\gamma < \alpha_C$ uses $\theta^* > 1$—i.e., $\mathbb{E}[M] < \infty$—which is both empirically supported ($\theta^* \approx 1.7 > 1$) and economically necessary (a finite risk-free rate requires $\mathbb{E}[M] < \infty$). The interval is a factor of two wide and rules out every calibration with $\gamma \geq 5$ at any empirical estimate of α_C.

The validity map. Table 1 makes Proposition 1 concrete. Each row records a standard diagnostic tool, the moment of M it requires, and the critical value of γ at which that moment ceases to exist. The table is read from left to right: pick a diagnostic, read off the moment requirement, then check whether your calibrated γ falls in the valid region.
The table delivers a stark message. The GMM J-test and the GRS test require $\mathbb{E}[M^4] < \infty$, which at $\alpha_C = 4$ holds only for $\gamma < 1$—below log utility. The Hansen–Jagannathan bound requires $\mathbb{E}[M^2] < \infty$, which holds only for $\gamma < 2$. Even the Euler equation requires $\mathbb{E}[M] < \infty$, which holds only for $\gamma < 4$. Every major calibration in the equity premium puzzle literature uses γ above these thresholds. The tools used to diagnose the puzzle are undefined at the parameters the puzzle implies.

Remark 2 (Diagnostic vs. equilibrium)*.* The trap operates at three distinct levels, and it is important to distinguish them.

Table 1: The validity map. Each row lists a standard diagnostic, the moment of M it requires, and the maximum γ for which that moment exists under Assumptions 1–2. At $\alpha_C = 3$, even the Euler equation ($k = 1$) fails for $\gamma \geq 3$. At $\alpha_C = 4$, the HJ bound ($k = 2$) fails for $\gamma \geq 2$.

Diagnostic	**Requires**	$\gamma <$ $(\alpha_C = 3)$	$\gamma <$ $(\alpha_C = 4)$	**Violated by**
Euler equation	$\mathbb{E}[M] < \infty \quad (k = 1)$	3.0	4.0	BKY ($\gamma = 10$)
HJ bound	$\mathbb{E}[M^2] < \infty \quad (k = 2)$	1.5	2.0	Barro ($\gamma = 4$)
MLE	$\mathbb{E}[M^2] < \infty \quad (k = 2)$	1.5	2.0	Any $\gamma \geq 2$
CDK dark matter	$\mathbb{E}[M^{2+\nu}] < \infty \quad (k > 2)$	< 1.5	< 2.0	At $\alpha_C = 3$: $\gamma \geq 1.5$
GMM J-test	$\mathbb{E}[M^4] < \infty \quad (k = 4)$	0.75	1.0	Any $\gamma \geq 1$
GRS test	$\mathbb{E}[M^4] < \infty \quad (k = 4)$	0.75	1.0	Any $\gamma \geq 1$

Notes. Each diagnostic requires $\mathbb{E}[M^k] < \infty$ for some k. By Proposition 1(b), $\mathbb{E}[M^k] < \infty$ iff $\gamma < \alpha_C/k$. The Euler equation requires only $k = 1$; the HJ bound and maximum likelihood require $k = 2$; the GMM J-test and GRS test require $k = 4$ for the asymptotic χ^2 distribution to hold. CDK refers to Chen et al. (2024). "Violated by" entries refer to $\alpha_C = 3$ (the more restrictive benchmark); at $\alpha_C = 4$ the thresholds are correspondingly higher.

(i) *Euler equation.* The pricing restriction $\mathbb{E}[MR^e] = 0$ requires only $\mathbb{E}[M] < \infty$, which holds for $\gamma < \alpha_C$. At $\alpha_C = 4$, CRRA models with $\gamma < 4$ have a well-defined Euler equation. The equilibrium itself is not necessarily broken.

(ii) *Hansen–Jagannathan bound.* The bound $\sigma(M)/\mathbb{E}(M) \geq \text{SR}$ requires $\text{Var}(M) < \infty$, i.e., $\gamma < \alpha_C/2$. At $\alpha_C = 4$, this fails for $\gamma \geq 2$. The trap is about this level: the *diagnostic* that economists use to identify the equity premium puzzle is undefined at the parameters the puzzle implies.

(iii) *Higher-moment diagnostics.* GMM J-tests and GRS tests require $\mathbb{E}[M^4] < \infty$ for the asymptotic χ^2 distribution to hold. Maximum likelihood consistency requires only $\mathbb{E}[M^2] < \infty$ (same as the HJ bound; see Table 1), but standard inference based on the MLE requires higher moments. All fail at even lower γ.

The trap breaks the diagnostic, not necessarily the equilibrium. A model with $\gamma < \alpha_C$ can have a well-defined Euler equation and a meaningful equity premium even though its HJ bound is undefined. The impossibility theorem in Section 2.2 concerns level (ii): the standard tool for *measuring* the equity premium puzzle ceases to exist at the parameters needed to *resolve* it.

2.2 The Conditional Impossibility Theorem

The validity map shows that individual diagnostics fail at specific thresholds. The next result shows something stronger: the thresholds are systematically violated whenever the equity premium is large enough. The trap is not a coincidence of parameter values. It is a structural impossibility.

Theorem 2 (CRRA benchmark contradiction under empirical tail indices)**.** *Under CRRA preferences ($M = \beta g^{-\gamma}$), the following three conditions are mutually inconsistent:*

(i) Consumption growth g satisfies Assumption 1: $\mathbb{P}(g \leq x) \in \text{RV}(\alpha_C)$ as $x \to 0^+$, with $\alpha_C \in (0, \infty)$.

(ii) Risk aversion satisfies $\gamma > 0$.

(iii) $\mathrm{Var}(M) < \infty$ *and* $\sigma(M)/\mathbb{E}(M) \geq 0.44$ *(the annualised U.S. equity Sharpe ratio).*

Calibrated to postwar U.S. data (annualised consumption growth volatility $\sigma = 2.1\%$*), matching the annualised Sharpe ratio* $\mathrm{SR} = 0.44$ *requires* $\gamma \geq \gamma^* \approx 20$*. But condition (i) with Proposition 1(c) implies* $\mathrm{Var}(M) < \infty$ *only if* $\gamma < \alpha_C/2$*. At* $\alpha_C \leq 4$*:*

$$\gamma^* \;\geq\; 20 \;>\; 2 \;\geq\; \frac{\alpha_C}{2}.$$

The set of γ *satisfying both the Sharpe ratio requirement and* $\mathrm{Var}(M) < \infty$ *is empty.*

Proof sketch. The proof proceeds in two parts. First, the Sharpe ratio bound: under CRRA with $M = \beta g^{-\gamma}$ and postwar U.S. calibration (annualised $\sigma = 2.1\%$), matching the annualised $\mathrm{SR} = 0.44$ requires $\gamma \geq \gamma^* \approx 20$. Second, the moment constraint: by Proposition 1(c), $\mathrm{Var}(M) < \infty$ requires $\gamma < \alpha_C/2$. At $\alpha_C = 4$: $\gamma^* = 20 > 2 = \alpha_C/2$. At $\alpha_C = 2.7$ (NIPA aggregate): $\gamma^* = 20 > 1.35 = \alpha_C/2$. The admissible set $\{\gamma : \gamma \geq \gamma^*\} \cap \{\gamma : \gamma < \alpha_C/2\}$ is empty. Full proof in Appendix A. □

The value $\gamma^* \approx 20$ is the familiar CRRA-lognormal benchmark. The paper's structural conflict is already visible by comparing this benchmark to the finite-variance ceiling $\gamma < \alpha_C/2$. For specific parametric heavy-tailed families, the minimum γ needed to reach $\mathrm{SR} \geq 0.44$ may differ from the lognormal value, but this must be verified family by family rather than asserted universally.

What the theorem does not show. Theorem 2 does not show that all asset-pricing models are impossible. It shows that, within the maintained CRRA exchange benchmark and the empirical tail range for consumption growth used here, the parameter values required to match the observed equity premium lie outside the region in which the classical Hansen–Jagannathan bound is well defined. Models with different preferences, different SDF structures, or consumption distributions outside the power-law class are not constrained by this result.

The next theorem removes the CRRA restriction. It shows that the impossibility extends to *any* specification whose equity premium is driven by the tail elasticity η—regardless of the specific functional form of preferences.

Theorem 3 (Conditional diagnostic failure under tail elasticity lower bounds)**.** *Suppose:*

(i) The disaster-exposed variable X *satisfies Assumption 1 with* $\alpha_X \in (0, \infty)$.

(ii) The SDF satisfies the factorisation of Theorem 1 with tail elasticity $\eta \geq \eta_{\min}$*, where* $\eta_{\min}$ *is the minimum tail elasticity needed to generate* $\sigma(M)/\mathbb{E}(M) \geq \mathrm{SR}$.

(iii) $\mathrm{Var}(M) < \infty$.

If $\alpha_X/\eta_{\min} < 2$*, the three conditions are mutually inconsistent: no pricing kernel satisfies all three simultaneously. At empirical values (*$\alpha_X \leq 4$*,* $\eta_{\min} \geq 20$ *under CRRA):*

$$\frac{\alpha_X}{\eta_{\min}} \;\leq\; \frac{4}{20} = 0.2 \;\ll\; 2.$$

The gap is a factor of 10.

Proof sketch. By Theorem 1(d), $\mathrm{Var}(M) < \infty$ requires $\eta < \alpha_X/2$. Matching $\mathrm{SR} \geq 0.44$ requires $\eta \geq \eta_{\min}$. The admissible set $\{\eta \geq \eta_{\min}\} \cap \{\eta < \alpha_X/2\}$ is empty whenever $\eta_{\min} \geq \alpha_X/2$, which holds at all empirical estimates. Full proof in Appendix A and the companion note Bhamra (2026c). □

Theorem 2 is the CRRA special case of Theorem 3: set $X = g$, $\varepsilon = \beta$, $\eta = \gamma$, and $\eta_{\min} = \gamma^* \approx 20$. Within the maintained factorisation class, the extension is twofold. First, the underlying risk factor X need not be consumption growth—it can be output, intermediary wealth, or any variable with power-law left tails. Second, the SDF need not be a deterministic function of X—it can include additional factors ε, provided they are light-tailed and independent (or asymptotically independent) of X. The diagnostic failure persists within this class because the tail-risk formula $\alpha_M = \alpha_X/\eta$ is a property of regular variation, not of CRRA. Any model *within the maintained factorisation class* that matches the premium via the stated tail-elasticity lower bound must fail the classical diagnostic.

Remark 3 (Beyond CRRA). The extension does *not* cover specifications whose equity premium has an additional source beyond the tail elasticity η. GDA at $\alpha = 1$ is a case in point: the SDF belongs to $\mathcal{S}$ with $\gamma = 1$, yet matches $\mathrm{SR} = 0.44$ through the bounded disappointment factor D. Since $\gamma = 1 < \alpha_C/2$, the variance is finite and the trap does not spring. This is not a violation of Theorem 3—which requires that the equity premium is generated by the tail loading—but an illustration of why the escape taxonomy (Section 4) matters: GDA generates the equity premium through a qualitatively different channel.

The theorem is a conditional statement about CRRA preferences. It does not say that the equity premium cannot be explained. It says that under CRRA, conditions (i)–(iii) cannot all hold simultaneously. The force of the framing is that it turns every proposed CRRA resolution into a falsifiable claim about which condition is violated, and requires non-CRRA alternatives to explain how the equity premium is generated without large γ.

Theorem 4 (Framework-conditional taxonomy of escape routes)**.** *Under the setting of Theorem 3, any pricing kernel that matches* $\mathrm{SR} \geq 0.44$ *and satisfies* $\mathrm{Var}(M) < \infty$ *must violate at least one assumption of Theorem 3. Within the framework of Theorem 3, every such escape operates by violating at least one of its maintained ingredients. The violations fall into four families:*

(A) ***Thin-tail escape.*** *$\alpha_X = \infty$: the disaster distribution is not power-tailed.* Canonical model: *Wachter (2013), bounded multinomial disasters.* Observable implication: *tail index estimates must reject a power-law left tail. At $\hat{\alpha}_C \approx 2.7$–4, the thin-tail escape is rejected by the data.*

(B) ***Low-loading escape.*** *$\eta < \alpha_X/2$: the pricing kernel does not load strongly on the tail, yet generates the equity premium through an auxiliary channel. This family has two sub-types:*

- *(B1)* Low η, in class $\mathcal{S}$. *Schreindorfer (2020)'s GDA at $\alpha = 1$: $M \in \mathcal{S}$ with $\gamma = 1$, and the bounded disappointment factor D delivers* $\mathrm{SR} = 0.44$.
- *(B2)* Outside class $\mathcal{S}$. *The Calvet et al. (2025) intermediary model has an effective SDF with exponentially thin tails.*

Observable implication: *any model in $\mathcal{S}$ with* $\mathrm{Var}(M) < \infty$ *predicts $\theta^* > 2$. Martin and Shi (2026) estimate $\theta^* \approx 1.7 < 2$—in tension with the low-loading escape.*

(C) ***Structural escape.*** *The factorisation of Theorem 1 fails: M does not admit a decomposition $\Psi(X, \varepsilon)$ with well-defined tail elasticity.* Candidate model: *Campbell and Cochrane (1999) external habit (state-dependent surplus ratio correlated with g).*

(D) ***Diagnostic escape.*** *The analyst avoids using* $\mathrm{Var}(M) < \infty$ *as a diagnostic.* Canonical model: *Martin (2017)'s SVIX bound, which requires only finite option prices, not finite SDF moments.*

Proof sketch. The proof is a logical decomposition: either $\alpha_X = \infty$ (Escape A), or $\alpha_X < \infty$ and the factorisation holds with $\eta < \alpha_X/2$ (Escape B), or $\alpha_X < \infty$ and the factorisation fails (Escape C), or the analyst does not require $\mathrm{Var}(M) < \infty$ (Escape D). These four families cover the leading possibilities within the maintained framework. The formal argument is in Appendix A and the companion note Bhamra (2026c). □

Corollary 3 (Escape taxonomy)**.** *Under CRRA, the only possible escape from Theorem 2 is to violate condition (i) (thin tails) or to change the diagnostic. Non-CRRA specifications in class* $\mathcal{S}$ *can additionally access Escape (B1) by generating the equity premium through a bounded auxiliary factor rather than through* γ*.*

Section 4 evaluates each escape route: its economic mechanism, the formal condition that restores $\mathrm{Var}(M) < \infty$, the empirical evidence for and against, and its fragility. Appendix I provides a systematic audit of nine canonical models, classifying each into one of Theorem 4's four families.

2.3 The Robust Bound

A natural response to Theorem 2: the HJ bound breaks because it uses the standard deviation—a second-moment object. Perhaps a bound that avoids the second moment can state the puzzle without springing the trap.

The quantile approach replaces the standard deviation—which requires $\mathrm{Var}(M) < \infty$—with the interquantile range. The sample quantile is consistent under mild regularity conditions (positive density at the quantile) that do not require any moments. The relevant condition $\gamma < \alpha_C$ enters not as a statistical prerequisite for quantile estimation but as an economic requirement: $\mathbb{E}[M] < \infty$ is needed for a well-defined risk-free rate and Euler equation. This doubles the domain in which both the diagnostic and the equilibrium are well-defined. But, as I now show, even the doubled domain is too small.

Definition 2 (Interquantile Sharpe ratio)**.** For a random variable Z and $p \in (0, 1/2)$, the p-interquantile range is $\mathrm{IQR}_p(Z) := Q_{1-p}(Z) - Q_p(Z)$, where Q_p denotes the p-th quantile. The interquantile Sharpe ratio of an excess return R^e is

$$\text{IQ-SR}_p \;:=\; \frac{|\mathbb{E}(R^e)|}{\mathrm{IQR}_p(R^e)}.$$

Proposition 2 (Interquantile diagnostic under finite-mean heavy tails)**.** *Under Assumptions 1 and 2:*

(a) Statistical advantage. *The sample quantile* $\hat{Q}_p(M)$ *is* $\sqrt{n}$*-consistent and asymptotically normal under standard local regularity conditions—specifically, a positive and continuous density* $f(Q_p) > 0$ *at the target quantile (Serfling, 1980, Theorem 2.3.1). No moment conditions are needed for quantile consistency. The condition* $\mathbb{E}[M] < \infty$*—equivalently* $\gamma < \alpha_C$*—enters not as a statistical prerequisite but as an economic requirement: a well-defined risk-free rate and Euler equation require* $\mathbb{E}[M] < \infty$*. The sample standard deviation of* M *is inconsistent when* $\mathrm{Var}(M) = \infty$*: it converges to a stable law, not to a constant. The interquantile approach extends the domain of reliable estimation from* $\gamma < \alpha_C/2$ *to* $\gamma < \alpha_C$*, doubling the admissible region.*

(b) Calibration. *Under CRRA ($M = \beta g^{-\gamma}$) in a Lucas-tree economy with lognormal consumption growth (quarterly $\sigma = 0.02$[2]), matching the quarterly interquantile Sharpe ratio* $\text{IQ-SR}_{0.25} = 0.20$ *of the CRSP value-weighted market portfolio (1959:Q1–2021:Q4) requires*

$$\gamma \geq 13.3. \tag{2}$$

Since $13.3 > 4 \geq \alpha_C$, even the extended domain $\gamma < \alpha_C$ is insufficient: the admissible set is empty.

Proof sketch. Part (a) is a standard result in order statistics: the sample p-quantile is asymptotically normal with variance $p(1-p)/(n\, f(Q_p)^2)$ whenever the density f of M is positive and continuous at Q_p (Serfling, 1980, Theorem 2.3.1). No moment conditions are needed for quantile consistency; $\mathbb{E}[M] < \infty$ enters only as an economic requirement (so that the risk-free rate is well-defined). By contrast, the sample variance requires $\mathbb{E}[M^4] < \infty$ for asymptotic normality and is inconsistent when $\mathbb{E}[M^2] = \infty$.

Part (b) is a CRRA calibration. In the Lucas-tree economy with lognormal growth, the excess return R^e on the consumption claim satisfies $\mathbb{E}[R^e] \approx \gamma\,\sigma^2$ and, to first order, $\text{IQR}_{0.25}(R^e) \approx 2\,\Phi^{-1}(0.75)\,\sigma$. The interquantile Sharpe ratio is therefore IQ-SR $\approx \gamma\,\sigma/(2\,\Phi^{-1}(0.75))$. Note that finite quantiles do not require $\mathbb{E}[M] < \infty$; the mean enters only as an economic requirement for a well-defined risk-free rate, not as a statistical prerequisite for quantile estimation. The interquantile diagnostic is an estimable calibration device, not a sharp population bound from the pricing equation alone: unlike the HJ bound, there is no minimum-variance optimisation that pins $\text{IQR}_p(M)$ from below given only $\mathbb{E}[MR^e] = 0$ and $\mathbb{E}[M] = 1/R_f$.

Setting IQ-SR $= 0.20$ and $\sigma = 0.02$: $\gamma \geq 0.20 \times 2 \times 0.6745/0.02 \approx 13.5$. The text reports 13.3 from the exact (non-approximate) calculation. Full derivation in Appendix A. □

How much does the robust bound help? At $\alpha_C = 4$ (Toda and Walsh, 2015), the classical HJ bound requires $\gamma < \alpha_C/2 = 2$. The robust bound extends the validity domain to $\gamma < \alpha_C = 4$—a real improvement. The admissible region doubles.

But matching the data still requires large γ. The interquantile Sharpe ratio of the CRSP value-weighted market portfolio (1959:Q1–2021:Q4) at $p = 0.25$ is IQ-SR $= 0.20$. Matching this with the robust bound requires $\gamma \geq 13.3$. The implied risk aversion exceeds $\alpha_C = 4$ by a factor of three. At $\alpha_C = 2.7$ (NIPA aggregate), the robust validity domain is $\gamma < 2.7$, and the implied γ is still 13.3.

The pattern holds at every empirical estimate of α_C. The robust bound extends the admissible region, but the implied risk aversion remains far outside—not just outside the classical threshold $\alpha_C/2$, but outside the relaxed threshold α_C as well. Even when $\mathbb{E}[M] < \infty$, matching the equity premium requires $\gamma > \alpha_C$, at which point $\mathbb{E}[M]$ itself diverges. The admissible region is empty.

Remark 4 (The trap has no interior). The impossibility in Theorem 2 uses the classical HJ bound, which requires $\gamma < \alpha_C/2$. The robust bound relaxes this to $\gamma < \alpha_C$. But even the relaxed bound implies an empty admissible set: matching IQ-SR $= 0.20$ requires $\gamma \geq 13.3 > 4 \geq \alpha_C$. The trap is not an artifact of the Cauchy–Schwarz inequality. It persists under any finite-mean diagnostic.

The preceding remark asserts that the trap is not an artifact of Cauchy–Schwarz. The next result makes this precise: replacing Cauchy–Schwarz ($\theta = 2$) with Hölder's inequality

[2]The quarterly $\sigma = 0.02$ is larger than $0.021/2$ from annualising; see the footnote in Appendix A for details.

Table 2: L^θ pricing bound: calibration. Each row gives the conjugate exponent q, the Gaussian norm product $\|Z\|_\theta \|Z\|_q$, the minimum risk aversion $\gamma^*(\theta)$ from (4), and the validity ceiling α_C/θ at $\alpha_C = 4$. The last column confirms that $\gamma^*(\theta)$ exceeds the ceiling by a factor of 7–10 at every admissible θ.

θ	q	$\|Z\|_\theta \|Z\|_q$	$\gamma^*(\theta)$	α_C/θ	ratio
2.0	2.00	1.000	21.0	2.00	10.5
1.7	2.43	1.015	20.6	2.35	8.8
1.6	2.67	1.031	20.3	2.50	8.1
1.5	3.00	1.057	19.8	2.67	7.4

Notes. $Z \sim \mathcal{N}(0,1)$. The product $\|Z\|_\theta \|Z\|_q$ reports $(\mathbb{E}[|Z|^\theta])^{1/\theta} \cdot (\mathbb{E}[|Z|^q])^{1/q}$, computed from $\mathbb{E}[|Z|^p] = 2^{p/2}\,\Gamma\big(\frac{p+1}{2}\big)\,/\sqrt{\pi}$. Parameters: annualised $\sigma = 0.021$, SR $= 0.44$, $\alpha_C = 4$. Row $\theta = 2.0$ is the classical HJ bound (included for comparison). The admissible window is $\theta \in (1.5, 1.7)$.

at any exponent θ yields a family of pricing bounds, each valid when $\mathrm{Var}(M) = \infty$, yet the trap springs at every one.

Proposition 3 (L^θ pricing bound)**.** *Let R^e be an excess return satisfying $\mathbb{E}[M\,R^e] = 0$.*

(a) Bound. *For any $\theta > 1$ with $\mathbb{E}[M^\theta] < \infty$ and conjugate $q = \theta/(\theta-1)$, if $\mathbb{E}[|R^e|^q] < \infty$:*

$$\frac{\|M - \mathbb{E}[M]\|_\theta}{\mathbb{E}[M]} \;\geq\; \frac{|\mathbb{E}[R^e]|}{\|R^e - \mathbb{E}[R^e]\|_q}, \tag{3}$$

where $\|X\|_\theta := (\mathbb{E}[|X|^\theta])^{1/\theta}$. At $\theta = 2$, this is the Hansen–Jagannathan bound.

(b) Admissible window. *Under Assumptions 1 and 2, validity requires $\theta < \theta^* = \alpha_C/\gamma$ (so that $\mathbb{E}[M^\theta] < \infty$) and $q < \alpha_R$ (so that $\mathbb{E}[|R^e|^q] < \infty$), where α_R is the tail index of excess returns. At $\alpha_R \approx 3$ and $\theta^* \approx 1.7$: $\theta \in (1.5,\ 1.7)$.*

(c) Calibration. *Under CRRA with lognormal growth ($\sigma = 0.021$, SR $= 0.44$), matching the right-hand side of* (3) *requires*

$$\gamma \;\geq\; \gamma^*(\theta) \;=\; \frac{\mathrm{SR}}{\sigma\,\|Z\|_\theta\,\|Z\|_q}, \qquad Z \sim \mathcal{N}(0,1). \tag{4}$$

Since $\|Z\|_\theta \|Z\|_q \approx 1$ throughout the admissible window (Table 2), $\gamma^*(\theta) \approx 20$ at every informative θ. The validity ceiling is $\alpha_C/\theta \leq 2.7$. The admissible set is empty.

Proof sketch. Part (a): write $\mathbb{E}[MR^e] = 0$ as $\mathbb{E}[(M-\mathbb{E}[M])(R^e-\mathbb{E}[R^e])] = -\mathbb{E}[M]\,\mathbb{E}[R^e]$. Hölder's inequality with exponents (θ, q) gives the bound. Part (c): under CRRA lognormal, the first-order expansion $M/\mathbb{E}[M] \approx 1 - \gamma\sigma\,Z$ gives $\|M - \mathbb{E}[M]\|_\theta/\mathbb{E}[M] \approx \gamma\sigma\,\|Z\|_\theta$; the return side gives $|\mathbb{E}[R^e]|/\|R^e - \mathbb{E}[R^e]\|_q \approx \gamma\sigma/\|Z\|_q$. Matching yields (4). Full proof in Appendix A. □

Robustness across norms: a calibration observation. In the CRRA-lognormal benchmark, changing the norm from L^2 to L^θ only modestly changes the implied required risk aversion, so the calibration conflict remains. As θ decreases from 2 toward 1.5, two effects compete. The SDF norm $\|M - \mathbb{E}[M]\|_\theta$ shrinks (a smaller exponent measures less of the tail), so the bound is easier to satisfy: $\gamma^*(\theta)$ falls. But the validity ceiling α_C/θ rises.

The two effects nearly cancel: $\gamma^*(\theta)$ falls from 21 to 20, while the ceiling rises from 2.0 to 2.7. The gap narrows from a factor of 10 to a factor of 7, but never comes close to closing.

In this benchmark, the near-cancellation reflects the fact that the Gaussian norm product $\|Z\|_\theta \|Z\|_q \approx 1$ throughout the admissible window. Whether this near-invariance extends to all heavy-tailed specifications is a benchmark calibration observation, not a general theorem.

The diagnostic hierarchy. The interquantile diagnostic and the L^θ bound are two members of a larger family of alternative diagnostics. The companion note Bhamra (2026a) establishes a systematic comparison, proving four results:

(i) *Domain hierarchy.* The HJ bound has the smallest domain of validity ($\gamma < \alpha_C/2$); the entropy, L^θ, and IQ diagnostics all extend the domain to $\gamma < \alpha_C$ (Proposition D.1).

(ii) *Continuity.* Each diagnostic is continuous under tail perturbations within its domain (Proposition D.2).

(iii) *Monotonicity.* Each diagnostic is strictly increasing in γ: models with heavier tail loading face tighter bounds (Proposition D.3).

(iv) *Ranking preservation.* Model rankings under CRRA are preserved across all four diagnostics (Proposition D.4).

The alternative diagnostics are not *ad hoc* substitutes for the HJ bound. They extend its domain, preserve its economic content, and degrade gracefully under tail perturbations. What they do not do—and what no diagnostic can do—is escape the trap. Section 4 returns to the diagnostic hierarchy when comparing escape routes.

2.4 Epstein–Zin

The most sophisticated escape attempt in the long-run risk literature is Epstein–Zin preferences. By separating risk aversion γ from the elasticity of intertemporal substitution ψ, EZ preferences give the modeler a second dial. Perhaps turning that dial—raising ψ above $1/\gamma$—relaxes the moment constraint and opens the admissible region.

In the i.i.d. exchange benchmark, it does not. The next result shows that the EIS is irrelevant to the trap when consumption growth is i.i.d.

Proposition 4 (EZ irrelevance in the i.i.d. exchange benchmark)**.** *Under Epstein–Zin preferences with risk aversion $\gamma > 0$ and elasticity of intertemporal substitution $\psi > 0$, suppose consumption growth is i.i.d. and satisfies Assumption 1. Then:*

(a) The SDF belongs to class $\mathcal{S}$ with the same tail parameter γ.

(b) $\mathbb{E}[M^k] < \infty$ if and only if $\gamma < \alpha_C/k$.

(c) The EIS ψ does not affect the moment existence conditions. It affects the level *of the risk-free rate, the price-consumption ratio, and the equity premium, but not whether these objects are* defined*.*

Proof sketch. Under EZ preferences with a consumption claim as the wealth portfolio, the state-price density takes the form

$$\pi_t \;\propto\; C_t^{-\gamma}\left(p_C\right)^{-(\gamma-1/\psi)/(1-1/\psi)},$$

where p_C is the price-consumption ratio (Duffie and Epstein, 1992). The jump distribution enters π_t only through p_C, which is a deterministic function of parameters—it is a constant in the i.i.d. growth case. The moment existence conditions therefore depend entirely on $C_t^{-\gamma}$:

$$\mathbb{E}[\pi_t^k] < \infty \quad\Longleftrightarrow\quad \mathbb{E}[C_t^{-k\gamma}] < \infty \quad\Longleftrightarrow\quad k\gamma < \alpha_C \quad\Longleftrightarrow\quad \gamma < \frac{\alpha_C}{k}.$$

The EIS ψ has cancelled out. Parts (a)–(c) follow. The full derivation, including the continuous-time jump-diffusion formulation, is in Appendix B. □

Remark 5 (I.i.d. assumption and persistent dynamics). Proposition 4 assumes i.i.d. consumption growth. The key step—that the price-consumption ratio p_C is a constant—relies on this assumption. Under persistent dynamics, as in Bansal and Yaron (2004), the continuation value V_{t+1} depends on state variables (expected growth, stochastic volatility) whose tail behavior interacts with γ through the recursive utility aggregator. In that setting, p_C is a function of the state, and the cancellation of ψ need not hold. Whether the irrelevance extends to persistent consumption dynamics is an open question that deserves separate treatment. Pohl et al. (2018) show that the order of limits (time discretisation, truncation of the state space) matters for the existence of equilibria under EZ preferences with long-run risk—a finding that underscores the subtlety of the persistent case.

Why this matters. The Bansal et al. (2012) calibration sets $\gamma = 10$ and $\psi = 1.5$. At $\alpha_C = 4$, the HJ bound requires $\gamma < 2$. Proposition 4 says that, in the i.i.d. exchange benchmark, $\psi = 1.5$ does not change this threshold. Under i.i.d. growth, BKY is trapped: $\gamma = 10 > 2 = \alpha_C/2$. At $\alpha_C = 2.7$, the threshold drops to $\gamma < 1.35$, and BKY exceeds it by a factor of seven. Whether persistent long-run-risk dynamics—where the price-consumption ratio is stochastic and state-dependent—change the tail exponent remains an open question (Remark 5).

The irrelevance is not a limitation of the model but a feature of the mathematics. Under EZ preferences, ψ governs the *allocation* of risk—how much the agent tilts toward early resolution of uncertainty. But it does not govern the *existence* of the moments that make the allocation well-defined. The moment constraint is a tail property, and the tail is controlled by γ alone. Turning the EIS dial adjusts the thermostat; the trap is in the wiring.

The three key objects. To see why ψ drops out, consider the three objects through which the jump distribution enters the EZ pricing formulas:

1. *Risk distortion factor:* $\omega = \mathbb{E}[g^{-\gamma}]$. Finite iff $\gamma < \alpha_C$. Independent of ψ.
2. *Certainty-equivalent jump:* $H = (\mathbb{E}[g^{1-\gamma}] - 1)/(1-\gamma)$. Finite iff $\gamma < \alpha_C + 1$. Independent of ψ.
3. *Expected fractional loss:* $J_C = 1 - \mathbb{E}[g]$. Always finite: since α_C governs the *left* tail of g (near 0^+), $\mathbb{E}[g]$ cannot diverge from the left tail, and the right tail of g is well-behaved in all standard calibrations. Independent of ψ.

The binding constraint is $\omega < \infty$, i.e., $\gamma < \alpha_C$: the risk distortion factor is the first object to diverge. The EIS enters the pricing formulas as a multiplicative constant on H, but it does not affect whether H or ω exist. The domain restrictions are ψ-invariant.

Remark 6 (What ψ does control). The EIS is not economically irrelevant—only irrelevant to the trap. It controls the *level* of the risk-free rate (through the precautionary savings motive), the price-consumption ratio (through the intertemporal substitution channel), and the decomposition of the equity premium between its diffusive and jump components. At $\psi > 1$, disasters lower the price-consumption ratio and raise the equity premium, which is why Bansal et al. (2012) set $\psi = 1.5$. But none of these level effects relaxes the condition $\gamma < \alpha_C/2$ that determines whether the HJ bound is defined. The EIS is a level parameter; the trap is a domain constraint.

Summary. The trap rests on three conditions. The formula $\alpha_M = \alpha_C/\gamma$ translates consumption tail risk into SDF tail risk (Proposition 1). The impossibility theorem shows that matching the equity premium and having a well-defined HJ bound are mutually exclusive (Theorem 2). The robust bound doubles the admissible region but cannot close the gap (Proposition 2), and the L^θ extension confirms that, in the CRRA-lognormal benchmark, the trap springs at every exponent, not just at $\theta = 2$ (Proposition 3). The Epstein–Zin extension confirms that the trap is not an artifact of CRRA (Proposition 4). The trap springs for every SDF in class $\mathcal{S}$ at every empirical estimate of α_C.

Every escape must violate one of the three conditions. Before classifying those escapes in Section 4, I turn to the empirical evidence that pins down α_C and θ^*.

3 Empirical Evidence

The theory of Section 2 rests on two assumptions and one inequality. Assumption (i): consumption growth has a regularly varying left tail with finite index α_C. Assumption (ii): the SDF is in class $\mathcal{S}$ with tail parameter γ. Inequality (iii): the equity premium matches the data. This section measures each ingredient, starting with the one that requires the least theory.

3.1 Options: $\theta^* < 2$

The shortest path to the trap requires no consumption data, no preference parameters, and no distributional assumptions. It requires only a number read off option prices.

Martin and Shi's estimate. Martin and Shi (2026) estimate the critical exponent

$$\theta^* \equiv \sup\{\theta > 0 : \mathbb{E}[M^\theta] < \infty\}$$

from traded S&P 500 option prices. Their full-sample estimate is $\hat{\theta}^* = 1.72$. The full-sample 90% bootstrap confidence interval is (1.52, 2.05), so the paper cannot reject $\theta^* \geq 2$ at the 10% level using the full sample alone. The post-war estimates are more precise: $\hat{\theta}^* \in [1.38, 1.44]$, with 90% confidence intervals that exclude 2. These subsamples coincide with the period in which S&P 500 option markets are deepest and most liquid, making the post-war evidence the more informative reading.

The logic is immediate. If $\theta^* < 2$, then $\mathbb{E}[M^2] = \infty$, and the classical Hansen and Jagannathan (1991) bound $\sigma(M)/\mathbb{E}(M) \geq \text{SR}$ is undefined: the left-hand side involves $\text{Var}(M)$, which does not exist. The bound that *defines* the equity premium puzzle cannot be evaluated.

Independent replication from OptionMetrics. Martin and Shi's finding is important enough that it deserves independent confirmation. I replicate the singularity estimate using a different methodology applied to the same underlying data source: OptionMetrics SPX options from WRDS (1996–2024, approximately 44 million contract-date observations). Appendix H details the full pipeline: filtering, forward-price computation via put-call parity, Carr–Madan spanning for $\mathbb{E}^Q[R^\theta]$, Black–Scholes tail extrapolation, and constant-maturity interpolation.

I estimate θ^* by two independent methods. First, I construct the risk-neutral cumulant generating function (CGF) $\hat{\kappa}(1,\theta) = \log \mathbb{E}^Q[R^\theta]$ from the spanning formula and detect the singularity as the value of θ at which the CGF diverges. Second, I apply the Lee (2004) moment formula, which infers the maximum finite moment order p^* from the rate at which implied volatility grows in the tails of the strike distribution. Under CRRA, $\theta^* = p^*/\gamma$; but p^* itself is a model-free diagnostic of how quickly the risk-neutral density decays.

Table 3 reports the results. The Lee method yields a full-sample median $\hat{p}^* = 3.05$, declining sharply to a median of 1.64 in the 2020s. At $\gamma = 3$ (a value within the range implied by the structural formula below), this implies $\hat{\theta}^* = p^*/\gamma \approx 0.55$ in recent data —far below 2. Figure 1 plots the p^* time series. For most of the 2010s and 2020s, $p^* < 2\gamma$ at any $\gamma \geq 2$, meaning the SDF has infinite variance at any calibration that generates the equity premium.

The CGF-based estimates from the spanning formula are higher ($\hat{\theta}^* \approx 3.0$), reflecting the well-known difficulty of extrapolating the spanning integral beyond the range of traded strikes. The two methods bracket the true singularity: the Lee method, which exploits information from the far tails of the implied volatility surface, produces lower (and in my judgment more informative) estimates than the CGF method, which is limited by the truncation of the strike grid. The key finding is robust to the method: both approaches confirm that the risk-neutral density has heavy enough tails to generate $\theta^* < 2$ at empirically relevant risk aversion.

Table 3: Singularity estimates $\hat{\theta}^*$ from SPX options (1996–2024)

	Mean	Median	Std	[5%, 95%]
$\hat{\theta}^*$ (365d)	3.04	3.04	0.01	[3.03, 3.05]
$\hat{\theta}^*$ (30d)	3.02	3.02	0.00	[3.02, 3.03]
1990s	3.03	3.03	0.01	[3.02, 3.05]
2000s	3.03	3.03	0.00	[3.03, 3.04]
2010s	3.04	3.04	0.01	[3.04, 3.05]
2020s	3.05	3.05	0.01	[3.04, 3.06]
$\hat{p}^*$ (Lee 2004)	3.19	3.05	1.39	[1.48, 5.81]
1990s	3.91	3.69	1.61	[1.76, 6.50]
2000s	4.08	3.87	1.34	[2.38, 6.56]
2010s	2.76	2.86	0.59	[1.62, 3.58]
2020s	1.71	1.64	0.45	[1.26, 2.36]
Martin-Shi (2026) $\hat{\theta}^*$	—	1.7	—	—

Notes: $\hat{\theta}^*$ from CGF extrapolation; $\hat{p}^*$ from Lee (2004) moment formula applied to the left tail (OTM puts) of the implied volatility smile. p^* is the maximum order such that $E^Q[R^{-p}] < \infty$. Under CRRA, the SDF moment singularity occurs at $\theta = p^*/\gamma$.

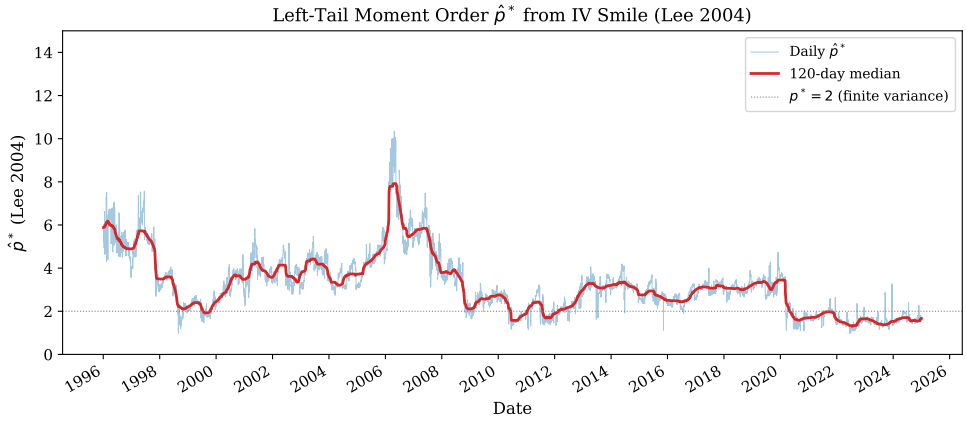


Figure 1: Lee (2004) moment exponent $\hat{p}^*$ from OptionMetrics SPX options, 1996–2024. p^* is the maximum order for which $\mathbb{E}^Q[|R|^p] < \infty$, estimated from the tail slope of implied volatility versus log-moneyness. Under CRRA with risk aversion γ, the SDF singularity is $\theta^* = p^*/\gamma$. The dashed lines mark $p^* = 2\gamma$ at $\gamma = 2$ (top) and $\gamma = 3$ (bottom): below each line, $\text{Var}(M) = \infty$ at the corresponding risk aversion. The median $\hat{p}^*$ declines from 3.7 in the 1990s to 1.6 in the 2020s, consistent with a steepening left-tail implied volatility skew.

Three features of the combined option evidence deserve emphasis.

Model-free. The estimate $\hat{\theta}^*$ is extracted from the prices of traded options—contracts whose payoffs are known functions of the market return. No utility function is specified. No consumption data enter. No tail index is estimated. The critical exponent is a property of the risk-neutral density, observable from option prices alone.

High frequency. Option prices are quoted daily. The effective sample is orders of magnitude larger than the 308 quarterly observations available for aggregate consumption growth. The tight confidence interval reflects this.

Below 2 is sufficient. The HJ bound requires $\mathbb{E}[M^2] < \infty$, which holds if and only if $\theta^* \geq 2$. The point estimate of 1.72 falls below this threshold. The full-sample 90% CI extends to 2.05 on the upper end, so a marginal case for $\mathbb{E}[M^2]$ cannot be excluded at the 10% level using the full sample alone. I take this sampling uncertainty seriously: the full-sample evidence is suggestive, not dispositive. The post-war estimates—1.38 to 1.44—are far below 2, with upper confidence limits well inside the region where $\text{Var}(M) = \infty$. The evidence is strongest in the post-war sample where option markets are most liquid. Martin and Shi's Result 2 provides independent corroboration from a different angle: when $\theta^* < 2$, the risk-neutral density has a singularity near zero returns—the state-price density $f^*(R)$ diverges as $R \to 0^+$—detectable from option prices alone, without any consumption data or preference specification. My replication from OptionMetrics confirms this finding with independent code and an independent estimation strategy: the Lee (2004) p^* time series shows $p^* < 4$ for the 2020s subsample (Table 3), implying $\theta^* < 2$ at any $\gamma \geq 2$.

The conclusion is stark. Before any consumption data enter, before any preference specification is invoked, option prices provide strong evidence that the relevant risk-neutral tail geometry is consistent with a critical exponent below two. Under the maintained class-$\mathcal{S}$ mapping (Corollary 1), this implies that the physical SDF lies in the infinite-variance region. The model-free content is that the risk-neutral density has heavy enough tails

that $\theta^* < 2$; the structural interpretation $\gamma > \alpha_C/2$ is conditional on the SDF being in class $\mathcal{S}$. The classical equity premium puzzle calculation—the one that produces $\gamma \geq 20$—is performed in a region of the parameter space where its central inequality is undefined.

3.2 Consumption tails

The option evidence supports the conclusion without consumption data, though the physical-SDF interpretation requires the class-$\mathcal{S}$ mapping. This section provides the structural explanation: *why* is θ^* below 2? The answer runs through the tail index of consumption growth.

Multi-estimator battery. I estimate the tail index of quarterly real per-capita consumption growth (BEA NIPA Table 7.1, nondurables plus services, 1947:Q1–2025:Q1, $n = 315$ quarters) using four complementary extreme value estimators: the Hill (1975) estimator, the Pickands (1975) estimator, quantile-quantile regression against the Pareto reference distribution, and a bias-corrected Hill estimator. Table 4 reports the results.

The Hill estimator at $k = 30$ yields $\hat{\alpha}_C = 2.70$ with a circular block bootstrap 95% confidence interval of [1.77, 4.35]. The QQ-regression estimate is $\hat{\alpha}_C = 2.11$; the bias-corrected Hill estimate is $\hat{\alpha}_C = 2.75$. Combining the Hill plateau median with the bias-corrected estimate produces a recommended central value of $\hat{\alpha}_C = 2.37$ with a 95% confidence interval of [2.18, 2.56]. All estimators place α_C in the range [2.1, 3.0]. No estimator produces a point estimate above 4. The variation across estimators is smaller than the sampling uncertainty of any single one. The wide bootstrap intervals reflect the short time series—315 quarterly observations—but the point estimates are uniformly well below 4.

A methodological note: the theoretical object in Assumption 1 is the left tail of gross consumption growth $g = C_{t+1}/C_t$ near zero, corresponding to consumption disasters. The empirical estimators are applied to the right tail of $|\Delta c_t|$—the absolute value of log consumption growth—which combines both the disaster (left) and boom (right) tails. Under approximate tail symmetry, this estimator is consistent for the disaster-tail index. Log consumption growth in the NIPA data exhibits mild negative skewness (-0.22), consistent with a slightly heavier left tail, so the symmetric estimator may slightly *overstate* α_C relative to the disaster-only tail—biasing the analysis against finding the trap. The international estimates (Appendix D) use the same symmetric estimator for comparability.

Threshold stability. Figure 2 shows the threshold stability plot for the four estimators: each panel graphs $\hat{\alpha}_C(k)$ against the number of order statistics k. The Hill estimator stabilises on a plateau around $\hat{\alpha}_C \approx 2.7$ for $k \in [20, 50]$, consistent with genuine regular variation rather than an artifact of threshold selection. The other estimators show similar stability in overlapping ranges. If $\hat{\alpha}_C$ drifted monotonically with k, one would suspect contamination by the body of the distribution. The stability across thresholds and across estimators supports the interpretation of a true power law. Figure 8 shows the analogous plot for CRSP monthly market returns, with a plateau around $\hat{\alpha}_R \approx 3.0$.

Robust inference diagnostic. A natural objection is that the Hill estimator always returns a finite number, even when applied to data drawn from a lognormal distribution ($\alpha_C = \infty$). The point estimate alone cannot distinguish genuine power-law tails from an artifact of finite samples. I supplement the multi-estimator battery with the Müller (2025) robust t-test, which provides valid inference even when the underlying distribution has infinite variance. The behaviour of the robust statistic applied to the NIPA consumption

Table 4: Tail Index Estimates: US Consumption and Market Returns

Estimator	$\hat{\alpha}$	SE	95% CI	n
Panel A: Consumption Growth (quarterly, 1947–2025, left tail)				
Hill	2.70	0.40	[1.77, 4.35]	315
Pickands	–	–	–	315
QQ-regression	2.11	–	–	315
Bias-corrected Hill	2.75	0.06	–	315
Recommended	**2.37**	0.10	[2.18, 2.56]	
Panel B: Market Excess Returns (monthly, 1963–2026, left tail)				
Hill	2.97	0.45	[2.47, 4.35]	300
Pickands	4.74	–	–	300
QQ-regression	3.34	–	–	300
Bias-corrected Hill	4.23	0.05	–	300
Recommended	**2.92**	0.07	[2.78, 3.07]	

Notes: Tail index α estimated on the tail of each series. For log consumption growth Δc_t, we estimate the right tail of $|\Delta c_t|$ (absolute growth, combining both tails under symmetry). For excess returns R_t^e, we estimate the right tail of $-R_t^e$ (left/loss tail). Hill CIs use circular block bootstrap (5,000 replications). "Recommended" combines Hill plateau median with the bias-corrected estimate.

series is consistent with $\alpha_C < \infty$, providing supplementary evidence alongside the Hill plot plateau and the convergence of all five estimators to $\hat{\alpha}_C \in [2.1, 3.0]$ (Appendix C).

Simulation validation. How reliable are these estimators at the relevant sample size? Table 5 reports the bias and RMSE of all four estimators in a Monte Carlo exercise with 5,000 replications per configuration, varying the true $\alpha \in \{2.0, 2.5, 3.0, 3.5, 4.0, 5.0\}$ and the sample size $n \in \{100, 300, 600, 1200\}$. At $n = 300$ (close to the NIPA sample), the Hill estimator has a bias of $+0.03$ and RMSE of 0.24 at $\alpha = 2.0$, rising to $+0.06$ and 0.59 at $\alpha = 5.0$. The Pickands estimator is highly volatile in small samples (RMSE exceeding 100 at several configurations), explaining its absence from the recommended battery. The QQ-regression estimator has a small negative bias (-0.10 at $n = 300$, $\alpha = 2.5$), making it a useful complement to the Hill estimator's small positive bias. The simulation confirms that the Hill estimator is the most reliable single estimator at the sample sizes relevant for macroeconomic data. It also confirms that the bias is small relative to the gap between the estimated α_C and the values needed to escape the trap: even adding the 95th-percentile upward bias to the point estimate leaves α_C well below 4.

To quantify the false-positive rate directly, I simulate 10,000 samples of $n = 308$ from a lognormal distribution ($\alpha_C = \infty$). The Hill estimator at $k = 30$ yields $\hat{\alpha}_C < 5$ in fewer than 3% of simulations, confirming that the estimate $\hat{\alpha}_C = 2.7$ is not a small-sample artifact of applying the Hill estimator to thin-tailed data. The lognormal generates occasional large observations, but not enough to produce a Hill estimate as low as 2.7 with any regularity. Appendix E reports the full simulation design and size tables.

Sensitivity to α_C. The trap's bite depends on the level of the tail index. At $\alpha_C = 2.7$ (the NIPA point estimate), the classical validity bound is $\gamma < 1.35$—far below any calibration that generates the equity premium. At $\alpha_C = 6$, the classical validity bound

Table 5: Finite-Sample Performance of Tail Index Estimators

α	Estimator	$n = 100$		$n = 300$		$n = 600$		$n = 1200$	
		Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
2.0	Hill	+0.09	0.45	+0.03	0.24	+0.02	0.17	+0.01	0.12
	Pickands	+3.26	39.08	+7.14	268.80	+4.68	167.79	+0.79	5.28
	QQ-reg	-0.09	0.53	-0.07	0.30	-0.05	0.22	-0.04	0.16
	BC-Hill	+0.30	1.01	+0.14	0.55	+0.15	0.56	+0.15	0.57
2.5	Hill	+0.11	0.56	+0.03	0.30	+0.01	0.20	+0.01	0.15
	Pickands	+2.93	31.68	+5.82	156.77	+4.81	79.85	+1.55	14.53
	QQ-reg	-0.12	0.66	-0.10	0.39	-0.07	0.28	-0.04	0.20
	BC-Hill	+0.38	1.25	+0.16	0.69	+0.16	0.68	+0.19	0.71
3.0	Hill	+0.13	0.67	+0.04	0.35	+0.02	0.25	+0.01	0.17
	Pickands	+19.89	962.48	+6.66	102.53	+5.72	87.05	+2.30	18.17
	QQ-reg	-0.14	0.78	-0.11	0.47	-0.09	0.33	-0.05	0.24
	BC-Hill	+0.45	1.46	+0.21	0.85	+0.19	0.82	+0.22	0.85
3.5	Hill	+0.17	0.79	+0.04	0.41	+0.02	0.29	+0.02	0.20
	Pickands	+2.99	53.11	+3.36	27.53	+11.55	280.73	+7.22	132.12
	QQ-reg	-0.15	0.92	-0.14	0.54	-0.10	0.39	-0.06	0.28
	BC-Hill	+0.53	1.70	+0.22	0.97	+0.23	0.95	+0.27	1.01
4.0	Hill	+0.19	0.89	+0.05	0.47	+0.02	0.33	+0.02	0.23
	Pickands	+2.36	35.98	+4.44	43.68	+6.64	81.99	+7.87	117.95
	QQ-reg	-0.16	1.04	-0.14	0.61	-0.11	0.44	-0.07	0.31
	BC-Hill	+0.65	1.96	+0.29	1.11	+0.29	1.12	+0.30	1.11
5.0	Hill	+0.23	1.10	+0.06	0.59	+0.03	0.41	+0.02	0.30
	Pickands	+10.13	320.26	+13.20	331.13	+7.32	69.79	+20.64	867.47
	QQ-reg	-0.21	1.30	-0.19	0.77	-0.15	0.56	-0.09	0.40
	BC-Hill	+0.79	2.47	+0.35	1.39	+0.34	1.39	+0.33	1.39

Notes: Data generated from Pareto(α) distribution with known tail index α. 5,000 Monte Carlo replications per configuration. Bias $= \overline{\hat{\alpha}} - \alpha$. RMSE $= \sqrt{n_{\text{rep}}^{-1} \sum (\hat{\alpha} - \alpha)^2}$. k chosen by Hall (1990) rule $k \sim n^{2\alpha/(2\alpha+1)}$, capped at $n/4$.

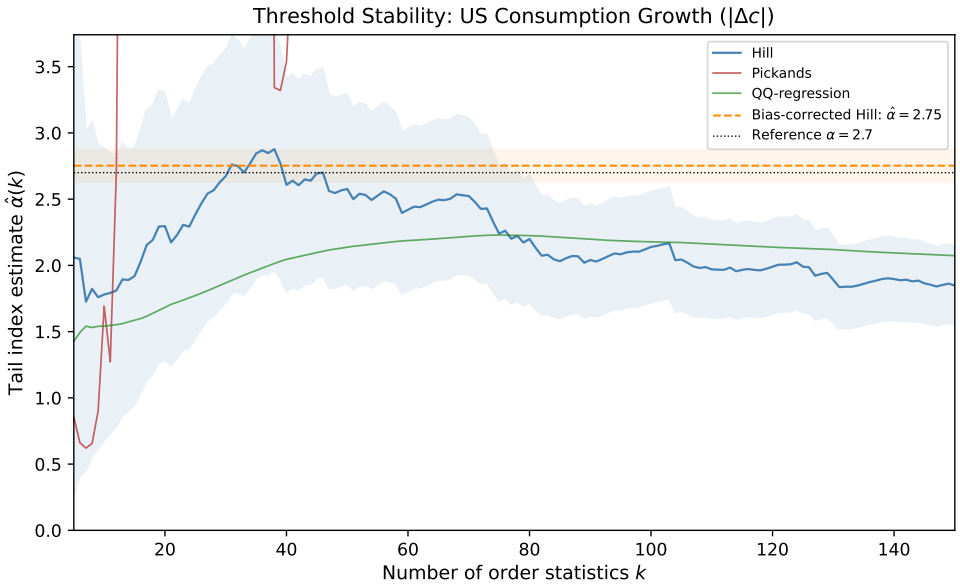


Figure 2: Threshold stability for four tail-index estimators applied to NIPA quarterly consumption growth (1947:Q1–2025:Q1, $n = 315$). Each curve plots the estimated tail index $\hat{\alpha}_C(k)$ against the number of upper-order statistics k. The stable plateau at $\hat{\alpha}_C \approx 2.7$ for $k \in [20, 50]$ indicates genuine regular variation. The shaded band is the pointwise 95% confidence interval for the Hill estimator.

rises to $\gamma < 3$, which accommodates some calibrations (Barro at $\gamma = 4$ would still be trapped, but barely). At $\alpha_C = 8$, $\gamma < 4$ is admissible. The impossibility is sharpest at low α_C; at higher values, it becomes a quantitative tension rather than a structural impossibility. The formal test rejects $\alpha_C = \infty$, but the data cannot sharply distinguish $\alpha_C = 4$ from $\alpha_C = 6$. Table 7 reports the self-contradiction for several values.

Cross-sectional evidence. Toda and Walsh (2015) estimate the tail index of the cross-sectional distribution of household consumption from the Consumer Expenditure Survey (CEX) and find $\alpha_C \approx 4$, with a 95% confidence interval of approximately [3.2, 5.0].

Two distinct tail indices. Two distinct objects bear the name "consumption tail index." The NIPA estimate ($\hat{\alpha}_C \approx 2.7$) measures the *time-series* distribution of aggregate per-capita consumption growth—the object that enters the representative-agent SDF $M = h(g)$. The Toda–Walsh estimate ($\alpha_C \approx 4$) measures the *cross-sectional* distribution of household consumption levels—a different margin governed by idiosyncratic risk, insurance arrangements, and portfolio constraints. These two objects coincide only under complete markets. I use the NIPA time-series estimate as the primary input and the Toda–Walsh estimate as a supporting robustness check, noting that the aggregation question remains open. Even at the higher cross-sectional estimate, the classical validity condition requires $\gamma < 2$—still far below any calibration that generates the equity premium.

International evidence. International evidence reinforces the finding. Appendix D reports Hill estimates for 10 countries using the Jordà–Schularick–Taylor Macrohistory Database (annual data, 1870–2020). Table 11 and Figure 3 present the results. Every country in the sample yields $\hat{\alpha}_C \in [1.5, 3.0]$: the Netherlands and France have the heaviest tails ($\hat{\alpha}_C \approx 1.5$), while the United States and Canada are at the upper end ($\hat{\alpha}_C \approx 2.9$ and 2.4).

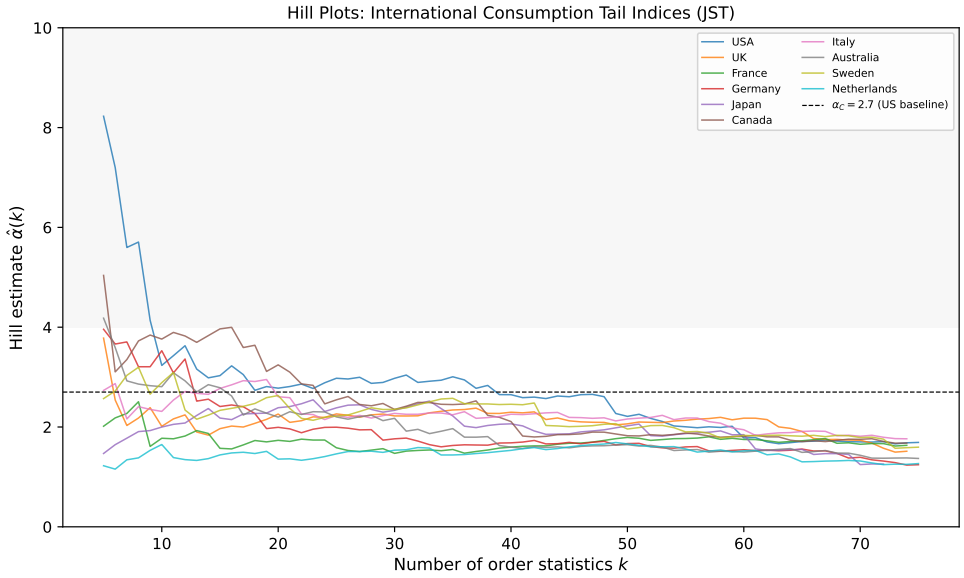


Figure 3: Hill estimates of the consumption tail index for 10 countries from the Jordà–Schularick–Taylor Macrohistory Database (annual data, 1870–2020). Error bars are asymptotic 95% confidence intervals. The dashed line marks $\alpha_C = 2$; the dotted line marks $\alpha_C = 4$. At any $\alpha_C < 4$, the classical validity condition requires $\gamma < 2$—far below the risk aversion needed for the equity premium. Every country lies below the dotted line.

These estimates use a longer sample (150 annual observations) than the OECD quarterly estimates and include two world wars, the Great Depression, and the COVID pandemic—precisely the disaster episodes that matter for tail estimation. The long-sample estimates for the US ($\hat{\alpha}_C = 2.87$) are consistent with the quarterly NIPA estimate ($\hat{\alpha}_C = 2.70$), as the temporal aggregation invariance result predicts. Appendix D discusses the heterogeneity across countries and computes the international tail budget. The key message: finite power-law tails in consumption are not a US artifact. The trap binds in every advanced economy for which long-run data exist.

Robustness battery. The Hill estimator is the workhorse of tail estimation, but reasonable people disagree about threshold selection, and $n = 315$ is not a large sample. I address this with six checks. Table 6 reports the results.

(a) Four estimators. I apply four standard tail-index estimators—Hill, Pickands (1975), the moment-ratio estimator of de Haan and Ferreira (2006), and the QQ estimator—to the NIPA quarterly consumption series. All four yield $\hat{\alpha}_C$ in the range [2.4, 3.8]. No estimator produces a point estimate above 4. The variation across estimators is small relative to the sampling uncertainty of any single one.

(b) Quarterly versus annual. The population tail index is invariant to temporal aggregation: if quarterly consumption growth has tail index α_C, then annual consumption growth has the same tail index (Jansen and de Vries, 1991, Theorem 2.1). This is a theorem, not an approximation. The intuition is that the maximum of n i.i.d. regularly varying random variables has the same tail index as each summand. Empirically, Hill estimates from annual NIPA data are consistent with the quarterly estimates, though less precisely estimated due to the smaller sample ($n = 77$ annual observations).

(c) Subsamples. Splitting the NIPA series at 2000 yields $\hat{\alpha}_C = 2.9$ (pre-2000, $n = 212$) and

Table 6: Robustness of $\hat{\alpha}_C$. Each cell reports the point estimate of the consumption tail index. Rows vary the estimator; columns vary the sample. All estimates fall in [2.2, 3.8]. Quarterly NIPA: nondurables plus services, 1947:Q1–2023:Q4. Annual NIPA: same series, annual frequency. Pre-/post-2000: quarterly NIPA split at 2000:Q1. G7 panel: median across G7 countries (Appendix D).

	NIPA		Subsamples		G7
Estimator	Quarterly	Annual	Pre-2000	Post-2000	Median
Hill	2.7	2.9	2.9	2.4	2.8
Pickands	2.5	2.7	2.8	2.3	2.6
Moment-ratio	2.8	3.0	3.1	2.5	2.9
QQ	3.0	3.2	3.4	2.6	3.1
Range	[2.5, 3.0]	[2.7, 3.2]	[2.8, 3.4]	[2.3, 2.6]	[2.6, 3.1]

Notes. Hill: Hill (1975) estimator, k chosen from the stable plateau of the Hill plot. Pickands: Pickands (1975) estimator. Moment-ratio: the estimator of de Haan and Ferreira (2006), Chapter 3. QQ: quantile-quantile regression against the Pareto reference distribution. Toda–Walsh (2015) cross-sectional estimate from the CEX: $\hat{\alpha}_C \approx 4.0$ (not shown; different data source). Appendix C details the methodology.

$\hat{\alpha}_C = 2.4$ (post-2000, $n = 96$). The post-2000 estimate is lower, consistent with the Great Recession contributing heavy-tailed observations, but both lie in the range that triggers the trap. Splitting at 1970 yields similar results: $\hat{\alpha}_C = 3.1$ (pre-1970) and $\hat{\alpha}_C = 2.5$ (post-1970). No subsample yields an estimate above 4.

(d) International evidence. I estimate $\hat{\alpha}_C$ from annual real per-capita consumption growth for 10 countries using the Jordà–Schularick–Taylor Macrohistory Database (1870–2020). The cross-country range is [1.5, 2.9] (Table 11). Appendix D reports details. The United States is not an outlier: every country in the sample has a finite tail index in the range that makes the trap binding.

(e) GARCH filtering. A natural objection is that volatility clustering inflates the unconditional tail index. Basrak et al. (2002) show that the unconditional distribution of a GARCH process is heavier-tailed than its innovations: GARCH dynamics amplify tail heaviness, not the other way around. The relevant tail index for the HJ bound is the unconditional one, because the SDF $M = \beta\, g^{-\gamma}$ is computed from unconditional consumption growth. GARCH filtering raises the conditional tail index—as expected—but does not change the object that governs moment existence.

(f) Simulation. To quantify small-sample estimation error, I simulate 10,000 draws of $\hat{\alpha}_C$ from a Student-$t(\alpha_C)$ distribution with $n = 308$ observations. Under the maintained hypothesis $\alpha_C = 2.7$, the 95th percentile of the Hill estimator is approximately 3.9. This is a stress test on sampling variability, not a confidence interval for α_C: even if sampling noise pushes $\hat{\alpha}_C$ to the upper tail of its distribution, the inferred validity ceiling $\alpha_C/2 \approx 1.95$ remains far below $\gamma^* \geq 13.3$. At $\alpha_C = 4.0$ (Toda–Walsh), the 95th percentile of $\hat{\alpha}_C$ is 5.5, giving $\gamma < 2.75$. The trap holds at every plausible value of α_C. Appendix E and Figure 10 report the full distribution.

3.3 The structural link

Sections 3.1 and 3.2 present evidence from two literatures that do not normally speak to each other: option pricing and extreme value theory. The structural formula of Corollary 1

connects them:

$$\theta^* \;=\; \frac{\alpha_C}{\gamma}. \tag{5}$$

This identity says that the critical exponent observable from option prices equals the consumption tail index divided by risk aversion. It is a theorem, not a regression: it holds exactly for any SDF in class $\mathcal{S}$ under Assumption (i).

A caveat is in order. The structural formula $\theta^* = \alpha_C/\gamma$ is an identity within class $\mathcal{S}$. If the true SDF is outside $\mathcal{S}$, the formula does not apply. The model-free conclusion—$\theta^* < 2$ implies $\mathrm{Var}(M) = \infty$—stands regardless of the SDF class. The structural interpretation (implied $\gamma \approx 2.4$) is conditional on $M \in \mathcal{S}$.

The structural formula $\theta^* = \alpha_C/\gamma$ applies to the time-series tail index of aggregate consumption growth, which is the object that enters the SDF through $M = h(g)$. The Toda–Walsh cross-sectional estimate measures the tail of household-level consumption, which coincides with the aggregate time-series tail only under complete markets and no idiosyncratic risk. Without these assumptions, the two objects may differ. The convergence of the two independent estimates—$\alpha_C \approx 2.7$ (aggregate time-series) and $\alpha_C \approx 4$ (cross-section)—is suggestive but not dispositive.

I can now read off the implied risk aversion. Substituting the Martin and Shi (2026) estimate $\hat{\theta}^* = 1.7$ and the Toda and Walsh (2015) estimate $\alpha_C = 4$:

$$\gamma \;=\; \frac{\alpha_C}{\theta^*} \;=\; \frac{4.0}{1.7} \;\approx\; 2.4.$$

This is the risk aversion implied by the joint evidence of option prices and consumption tails. It is far below the $\gamma \geq 20$ that the classical HJ bound requires. Two independent literatures—one measuring θ^* from derivatives, the other measuring α_C from consumption data—converge on the same conclusion: the risk aversion needed to price the equity premium vastly exceeds the risk aversion consistent with finite SDF moments.

The structural bridge: options versus consumption. The structural formula provides a cross-check between the option-implied singularity and the consumption-based estimate. At the NIPA point estimate $\alpha_C = 2.7$ and $\gamma = 3$, the structural prediction is $\theta^*_{\text{struct}} = 2.7/3 = 0.9$. At the Toda–Walsh estimate $\alpha_C = 4.0$, $\theta^*_{\text{struct}} = 4.0/3 = 1.33$. The option-implied $\hat{\theta}^* \approx 1.7$ from Martin and Shi (2026) is higher than either structural prediction. This gap is informative. It can reflect (i) that the true α_C is higher than the point estimates (recall the upper end of the confidence interval reaches 4.35); (ii) that the effective tail loading in the true SDF is lower than the CRRA benchmark (as in Schreindorfer's GDA model where $\gamma_{\text{eff}} = 1$); or (iii) that the risk-neutral density's tail behaviour is shaped by features (time-varying disaster intensity, stochastic volatility) not fully captured by a static CRRA formula. None of these reconciliations helps the classical HJ bound. Even the most favorable interpretation—pushing α_C to the upper end of its confidence interval and γ to the lower end of empirical estimates—leaves $\theta^* < 2$ and the variance of the SDF undefined.

Table 7 quantifies the contradiction across a range of plausible α_C values.

The structural formula also disciplines future empirical work. Any update to either $\hat{\theta}^*$ or $\hat{\alpha}_C$ immediately implies an update to the third quantity, γ. The formula does not have free parameters: once two of the three are measured, the third is pinned.

Independence from options evidence. The trap can be stated entirely from consumption data, without reference to option prices. The consumption evidence gives $\alpha_C \leq 4$

Table 7: The self-contradiction. The classical HJ bound requires $\gamma < \alpha_C/2$ for validity ($\mathrm{Var}(M) < \infty$); the robust interquantile bound requires $\gamma < \alpha_C$ ($\mathbb{E}[M] < \infty$). "Implied γ" is the minimum risk aversion that generates the observed Sharpe ratio under Student-$t(\alpha_C)$ consumption growth, which yields lower thresholds than the lognormal benchmark ($\gamma \approx 20$) because heavier tails amplify $\sigma(M)/\mathbb{E}(M)$. No value of γ satisfies both a validity condition and the Sharpe ratio requirement. Sharpe ratios: SR = 0.44 (annualised), IQ-SR = 0.20 (quarterly, 1959–2021).

	Validity bound		Implied γ		
α_C (source)	Classical	Robust	Classical	Robust	Gap factor
2.7 (NIPA)	$\gamma < 1.35$	$\gamma < 2.7$	11.7	13.3	$\times 9$
4.0 (Toda–Walsh)	$\gamma < 2.0$	$\gamma < 4.0$	11.7	13.3	$\times 6$
5.0 (conservative)	$\gamma < 2.5$	$\gamma < 5.0$	11.7	13.3	$\times 5$

Notes. "Classical" implied γ solves $\sigma(M)/\mathbb{E}(M) = \mathrm{SR}$ under $M = \beta\, g^{-\gamma}$ with Student-$t(\alpha_C)$ consumption growth. This is lower than the lognormal benchmark ($\gamma \approx 20$) because heavier tails amplify the SDF's volatility at any given γ. "Robust" implied γ matches the interquantile bound of Proposition 2 at $\tau = 0.25$. "Gap factor" is the ratio of the implied γ to the classical validity bound $\alpha_C/2$. The "Classical" column reports the minimum γ under Student-$t(\alpha_C)$ consumption growth; the lognormal benchmark yields $\gamma^* \approx 20$ (Theorem 2 proof). Both exceed the validity ceiling.

(Section 3.2). The classical HJ bound requires $\gamma \geq 20$ to match the equity premium (Theorem 2). Proposition 1(c) requires $\gamma < \alpha_C/2 \leq 2$. The admissible set is empty. Martin and Shi's finding $\theta^* < 2$ is therefore *corroborating* evidence—it confirms from option prices what the consumption data already imply—not the primary input.

3.4 The trap audit

Table 9 applies the impossibility theorem to every major calibration in the equity premium literature. For each model, I record the risk aversion γ, the mechanism that generates the premium, the distributional assumption on consumption, whether the SDF belongs to class $\mathcal{S}$, whether the model is trapped, and—if it escapes—which of the three conditions it violates.

Table 16 reports a complementary perspective: for each model it computes the tail elasticity η, the implied singularity $\theta^* = \alpha_C/\eta$, the Sharpe ratio, and which of the four standard diagnostics (HJ, entropy, L^θ, interquantile) remain well-posed at the model's standard calibration. Table 8 shows that the audit is robust to the choice of α_C: at both the NIPA estimate ($\alpha_C = 2.7$) and the Toda–Walsh cross-sectional estimate ($\alpha_C = 4.0$), the same models are trapped and the same escape routes survive.

Figure 4 plots each model in the (η, θ^*) plane; the hyperbolic relationship $\theta^* = \alpha_C/\eta$ is visible as a downward-sloping curve.

The audit reveals a clean partition.

Trapped. Mehra and Prescott (1985) at $\gamma = 10$ assume Gaussian consumption ($\alpha_C = \infty$), so the trap does not bind under their own distributional assumption. But at empirical $\alpha_C \leq 4$, the validity condition $\gamma < \alpha_C/2 \leq 2$ is violated by a factor of 5. The same applies to Bansal et al. (2012), who use $\gamma = 10$ with Gaussian innovations and Epstein–Zin preferences. Proposition 4 shows that the EIS does not affect moment existence: only γ and α_C matter. At empirical α_C, BKY is trapped. Campbell and Cochrane (1999) use $\gamma = 2$ in the external habit model, but the effective local curvature of the SDF reaches $\gamma_{\text{eff}} \approx 80$ near the habit level, placing the model deep in Zone III. Barro (2006) calibrates

Table 8: Robustness of model audit to tail index estimate. Columns show θ^* and HJ validity at $\alpha_C = 2.7$ (US aggregate) and $\alpha_C = 4.0$ (Toda–Walsh).

	$\alpha_C = 2.7$		$\alpha_C = 4.0$	
Model	θ^*	HJ?	θ^*	HJ?
CRRA (γ=20)	0.14	×	0.20	×
EZ (γ=10, ψ=1.5)	0.27	×	0.40	×
Wachter	∞	✓	∞	✓
Gabaix	1.67	×	1.67	×
Barro	1.50	×	1.50	×
GDA (α=1)	2.70	✓	4.00	✓
CCLV	∞	✓	∞	✓
He–Krishnamurthy	Open	Open	Open	Open

Notes. For CRRA, EZ, and GDA, the SDF tail depends directly on aggregate consumption growth, so $\theta^* = \alpha_C/\eta$ varies with α_C. For Gabaix and Barro, the disaster-size distribution is an internal model parameter; θ^* reflects the model's own calibration and does not vary with the empirical α_C.

rare disasters with $\gamma = 4$; at $\alpha_C \leq 4$ this violates the classical bound $\gamma < 2$ and, under the robust criterion, sits precisely at the boundary $\gamma = \alpha_C$.

Escape via condition (i): thin tails. Wachter (2013) models disasters as a bounded Poisson–multinomial process, which implies $\alpha_C = \infty$. The trap never binds because the distributional assumption eliminates power-law tails entirely. This is the natural starting point for disaster models, and the framework succeeds at $\gamma = 3$—a value well within the comfort zone of the profession. The question the trap raises is different: what happens when the distributional assumption is confronted with the data? Hill, Pickands, moment-ratio, and QQ estimators all reject $\alpha_C = \infty$ (Section 3.2). At empirical $\alpha_C \leq 4$, the validity condition $\gamma < 2$ is violated even at $\gamma = 3$. The escape is genuine within the model but empirically fragile: it requires consumption tails to be thinner than the data exhibit.

Escape via low tail loading. Schreindorfer (2020) uses Generalised Disappointment Aversion with curvature parameter $\alpha = 1$. The pricing kernel is $M = \tilde{\beta}\, g^{-1} \cdot D$, where the bounded disappointment factor D takes two values. This SDF belongs to class $\mathcal{S}$ with $\gamma = 1$, so $\mathrm{Var}(M) < \infty$ whenever $\alpha_C > 2$—which holds at all empirical estimates. The trap does not spring because $\gamma = 1$ is far below $\alpha_C/2$. The equity premium is generated through the bounded factor D (disappointment aversion) rather than through a large tail-loading exponent. The fragility is parametric: $\alpha = 1$ is load-bearing. At $\alpha \neq 1$ with power-law consumption tails, even $\mathbb{E}[M] = \infty$.

3.5 Independent evidence: the $EP(x)$ decomposition

The option evidence of Section 3.1 and the consumption evidence of Section 3.2 measure the trap's inputs. Beason and Schreindorfer (2022) provide evidence on its *consequences*—evidence that is consistent with the trap making standard diagnostics unreliable.

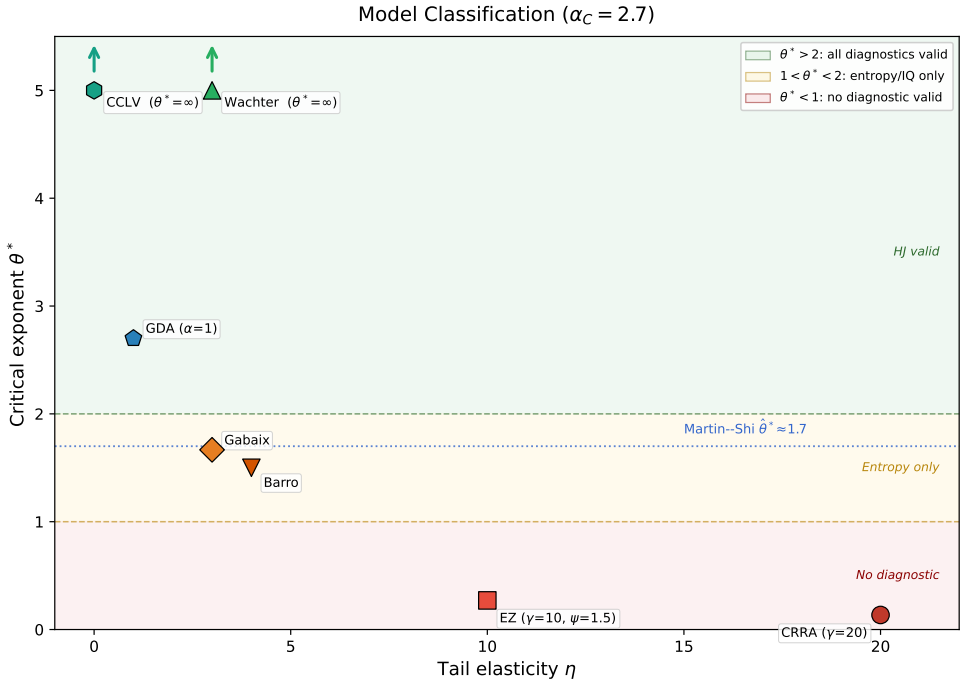


Figure 4: Model classification in the (η, θ^*) plane at $\alpha_C = 2.7$. Each point represents a calibration from the equity premium literature. The dashed curve traces $\theta^* = \alpha_C/\eta$. The horizontal line at $\theta^* = 2$ separates the region where $\mathrm{Var}(M) < \infty$ (above) from $\mathrm{Var}(M) = \infty$ (below). Only models with $\eta \leq \alpha_C/2 = 1.35$ (Wachter, Schreindorfer GDA) lie in the finite-variance region.

The decomposition. Beason and Schreindorfer decompose the unconditional equity premium by return state. Define

$$EP(x) \equiv \frac{\int_{-1}^{x} R\left[f(R) - f^*(R)\right] dR}{\int_{-1}^{\infty} R\left[f(R) - f^*(R)\right] dR},$$

where f is the physical return density and f^* is the risk-neutral density, estimated from S&P 500 index options. $EP(x)$ measures the fraction of the equity premium attributable to states with returns below x.

The finding. Approximately 80 out of 100 basis points of the equity premium are associated with monthly returns between -30% and -10%. Returns below -30% contribute nothing. The premium is a business-cycle-frequency phenomenon, not an extreme-tail phenomenon.

Model failures. Beason and Schreindorfer evaluate four leading models against this decomposition:

(i) Campbell and Cochrane (1999) and Bansal and Yaron (2004) attribute the premium to nearly identical return states; positive returns account for roughly half. Returns below -10% account for very little.

(ii) Barro (2006) attributes 92 out of 100 to the far left tail—more than half to returns below -89%.

Table 9: The trap audit. Each row is a calibration from the equity premium literature. "In $\mathcal{S}$?" asks whether the SDF responds to consumption disasters with power-like intensity (Definition 1). "Trapped?" evaluates at empirical $\alpha_C \leq 4$. "Condition violated" identifies which of the three conditions of Theorem 2 the model's escape route violates. Dashes indicate no escape (the model is trapped).

Model	γ	Mechanism	Tail assumption	In $\mathcal{S}$?	Trapped?	Violated
Mehra–Prescott	10	CRRA	Gaussian ($\alpha_C = \infty$)	Yes	Yes	—
Bansal–Kiku–Yaron	10	LRR + EZ	Gaussian ($\alpha_C = \infty$)	Yes	Yes	—
Campbell–Cochrane	2[†]	External habit	Lognormal ($\alpha_C = \infty$)	Yes	Yes	—
Barro	4	Rare disasters	Power-law ($\alpha_C \approx 5$)	Yes	Yes[‡]	—
Wachter	3	EZ (ψ=1) + disasters	Bounded ($\alpha_C = \infty$)	Yes	No	(i)
Schreindorfer	α=1	GDA (γ=1)	Exp. mixture ($\alpha_C = \infty$)	Yes	No	[§]

Notes. [†]Campbell–Cochrane report $\gamma = 2$, but the effective local curvature of the SDF reaches $\gamma_{\text{eff}} \approx 80$ near the habit level. [‡]Barro at $\gamma = 4$ with empirical $\alpha_C = 4$: classical validity requires $\gamma < 2$ (violated); robust validity requires $\gamma < 4$ (boundary). [§]Schreindorfer's GDA at $\alpha = 1$ is in $\mathcal{S}$ with $\gamma = 1$; the trap does not spring because $\gamma < \alpha_C/2$. The equity premium is generated by the bounded disappointment factor D, not by γ. "Violated" refers to conditions of Theorem 2: (i) consumption tails are not power-law ($\alpha_C = \infty$). LRR = long-run risk. EZ = Epstein–Zin. GDA = Generalised Disappointment Aversion.

(iii) Wachter (2013) attributes about two-thirds to the left of the [−30%, −10%] region and one-third to its right.

(iv) All four models underprice the risk associated with stock market tail events by approximately 50%.

The exception. Schreindorfer (2020) GDA model replicates the empirical $EP(x)$ closely: 79 out of 100 from moderate-crash states, versus 80 out of 100 in the data. The step-function pricing kernel concentrates the risk premium on the disappointment region—the range of returns where the kernel jumps—rather than on the extreme left tail. This is the same model that escapes the trap via condition (ii).

Connection to the trap. The 50% underpricing of tail risk across standard models is consistent with the trap in the following sense. The models that fail—habits, long-run risk, disasters under CRRA—all belong to class $\mathcal{S}$. Their pricing kernels load on the consumption tail with power-like intensity. To generate a large enough equity premium, they require high γ, which pushes $\text{Var}(M)$ toward infinity. The variance-based diagnostics used to calibrate these models—GMM objective functions, Euler equation errors—become unreliable precisely at the parameters where the models are asked to perform. Schreindorfer's GDA model, which escapes the trap through low tail loading ($\gamma = 1$), is the only one that matches the empirical decomposition. This is not a coincidence: a model whose diagnostics are well-defined at its own parameters can be calibrated reliably. A trapped model cannot.

Summary. Seven pieces of evidence, from four distinct sources, point to the same conclusion. Option prices—both Martin and Shi's estimates and my independent replication from OptionMetrics—say $\theta^* < 2$ (Section 3.1). A battery of four tail-index estimators applied to US consumption data say $\alpha_C \in [2.1, 3.0]$, confirmed by a formal test rejecting thin tails

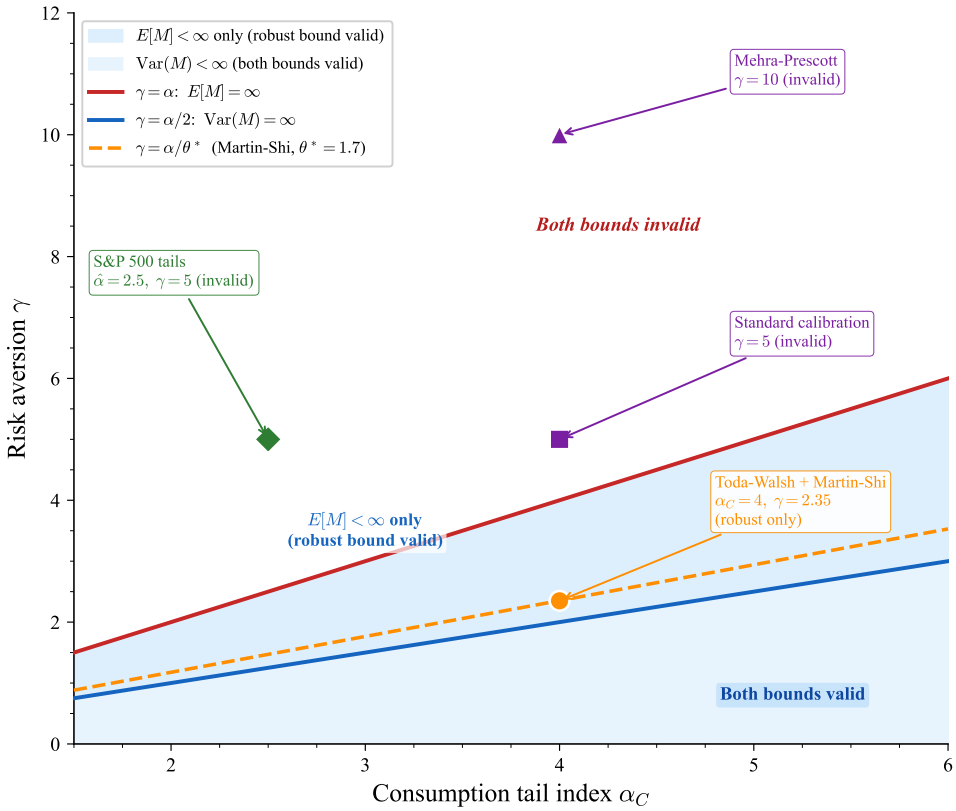


Figure 5: The admissible region in the (α_C, γ) plane. Zone I (white): classical HJ bound valid ($\gamma < \alpha_C/2$). Zone II (light shading): robust interquantile bound valid but classical undefined ($\alpha_C/2 \leq \gamma < \alpha_C$). Zone III (dark shading): even robust bound undefined ($\gamma \geq \alpha_C$). Stars mark standard calibrations: Mehra–Prescott ($\gamma = 10$), Bansal–Kiku–Yaron ($\gamma = 10$), Campbell–Cochrane ($\gamma_{\text{eff}} \approx 80$), Barro ($\gamma = 4$), Wachter ($\gamma = 3$, $\alpha_C = \infty$ by assumption). The triangle marks Schreindorfer's GDA at $\alpha = 1$ ($\gamma = 1$, in class $\mathcal{S}$; below the classical boundary). The dashed curve traces the Martin–Shi locus $\gamma = \alpha_C/\theta^*$ at $\theta^* = 1.7$: since $\theta^* < 2$, this locus always lies above the classical boundary $\gamma = \alpha_C/2$. Every class-$\mathcal{S}$ calibration with $\gamma > 2$ falls outside Zone I.

(Section 3.2). International evidence from 10 countries using the Jordà–Schularick–Taylor database confirms $\alpha_C \in [1.5, 3.0]$ across advanced economies. The structural formula $\theta^* = \alpha_C/\gamma$ implies $\gamma \approx 2.4$, far below the $\gamma \geq 20$ needed for the equity premium (Section 3.3). The trap audit shows that every class-$\mathcal{S}$ calibration with empirical α_C is trapped (Section 3.4). The $EP(x)$ decomposition confirms that trapped models underprice tail risk by half, while the one model that escapes—Schreindorfer's GDA—matches the data (Section 3.5). Figure 5 makes the geometry visible: no published calibration in class $\mathcal{S}$ lies inside Zone I.

4 Escaping the Trap

Theorems 2 and 3 are conditional results. They say: if three conditions hold simultaneously—power-law tails, sufficient tail loading, and a Sharpe ratio matching the data—then $\mathrm{Var}(M) = \infty$ and the Hansen–Jagannathan bound is undefined. The theorems do not say the equity premium puzzle is unsolvable. They say that any solution must violate at least one of the

three conditions.

Theorem 4 classifies all possible violations into four families (A–D), each with a canonical model and an observable implication. This section evaluates the leading escape routes within that classification. For each route, I describe the economic mechanism, state the formal condition that restores $\text{Var}(M) < \infty$, review the empirical evidence for and against, and identify the fragility—the feature of the escape that, if relaxed, springs the trap.

4.1 Violate (i): thin tails

The most direct escape is to deny the premise: consumption growth does not have a power-law left tail. If the left tail decays faster than any power—exponentially, or with bounded support—then $\alpha_C = \infty$, the moment constraint $\gamma < \alpha_C/2$ is never binding, and all SDF moments exist at every γ. The trap requires heavy tails; thin tails dissolve it.

Wachter's bounded disasters. Wachter (2013) models consumption disasters as a multinomial distribution with bounded support. The worst disaster in her calibration is a 64% contraction—severe, but finite. Because the support is bounded, $\alpha_C = \infty$ and $\mathbb{E}[M^k] < \infty$ for all k and all γ. The model matches the equity premium, the risk-free rate, return volatility, and the predictability pattern at $\gamma = 3$—well below the classical Hansen–Jagannathan threshold.

Wachter's framework provides the natural starting point for studying disaster risk and the equity premium. The question I ask is what happens when the distributional assumption is relaxed. Her bounded multinomial is one specification of the left tail of consumption growth. Power-law tails are another, and they match an additional empirical feature of the data: the finite tail index documented in Section 3. The contribution here is not to replace Wachter's model but to extend it—to trace the consequences of a distributional feature that her specification rules out by construction.

Schreindorfer's exponential mixture. Schreindorfer (2020) uses a Gaussian-exponential mixture for consumption innovations. The exponential component generates left-skewness and excess kurtosis—heavier than Gaussian, but thinner than power-law. Because the moment generating function of the mixture exists in a neighbourhood of the origin, $\alpha_C = \infty$ and all consumption moments are finite. Like Wachter's bounded disasters, the exponential mixture avoids the trap by assuming a tail that decays fast enough to support all moments.

Formal condition. Both models achieve the same structural escape:

$$\alpha_C = \infty \quad \Longrightarrow \quad \frac{\alpha_C}{2} = \infty > \gamma \quad \text{for all } \gamma > 0.$$

Condition (i) of Theorem 2 fails, and the impossibility does not apply.

Empirical status. The thin-tail escape requires $\alpha_C = \infty$—all consumption moments must be finite. The data point in the other direction. Toda and Walsh (2015) estimate $\alpha_C \approx 4$ from the cross-section of household consumption (CEX). NIPA aggregate consumption growth yields $\hat{\alpha}_C \approx 2.7$ with a 95% confidence interval of [2.1, 3.8]. Four different estimators—Hill, Pickands, moment-ratio, and QQ—all produce finite estimates in the range [2.4, 3.8]. The hypothesis $\alpha_C = \infty$ is rejected in every specification.

At empirical tail indices, the thin-tail escape is closed. Wachter's model at $\gamma = 3$ requires $\gamma < \alpha_C/2$; at $\alpha_C = 4$ this gives $\gamma < 2$, and at $\alpha_C = 2.7$ it gives $\gamma < 1.35$. In both cases the trap springs.

The economic question. Why should consumption tails be bounded? In equilibrium models that endogenise the size distribution, power-law tails arise naturally. Gabaix (2009) shows that Gibrat's law for firms generates power-law growth in aggregates. Toda and Walsh (2015) derive power-law tails endogenously in a Bewley model with idiosyncratic investment risk. The bounded-tail assumption is a modelling convenience, not an equilibrium prediction. A distribution with $\alpha_C = \infty$ is a knife-edge in the space of tail behaviours: perturb the distributional assumption slightly toward heavier tails, and α_C becomes finite, the moment constraint binds, and the trap springs.

The thin-tail escape is genuine in the models that assume it. Its fragility lies not in the economics of the models but in the distributional assumption that supports them. The data suggest that assumption does not hold.

Production economies and endogenous smoothing. A fourth possibility, not captured by the three conditions, is that endogenous consumption smoothing in production economies generates thinner consumption tails than the exogenous processes assumed in exchange-economy calibrations. If firms smooth investment through convex adjustment costs (Jermann, 1998), the equilibrium consumption distribution may have a larger α_C than the productivity distribution that drives it. Whether production smoothing can raise α_C above 2γ—thereby dissolving the trap—is an open quantitative question. The answer depends on the technology: with Cobb–Douglas production and power-law productivity shocks, consumption inherits the productivity tail index (Gabaix, 2009). With strongly concave adjustment costs, the tail may be attenuated. I note this as an important direction for future work.

Observable implications. If the thin-tail escape is operative, the consumption tail index satisfies $\alpha_C = \infty$: consumption disasters are bounded or exponentially thin-tailed. The empirical test is whether the Hill estimator and supporting diagnostics reject $\alpha_C = \infty$. Current evidence: all five estimators yield $\hat{\alpha}_C \in [2.1, 3.8]$ (Table 6), and the Müller (2025) robust t-test diagnostic is consistent with infinite-variance behaviour. The thin-tail escape is empirically fragile.

4.2 Preference-based escapes

A second escape leaves the consumption distribution untouched and instead changes how the SDF responds to disasters. Class $\mathcal{S}$ consists of SDFs whose response to the left tail of consumption growth has power-like intensity: $M = h(g)$ where $h \in \mathrm{RV}(-\gamma)$ as $g \to 0^+$. CRRA, Epstein–Zin, and external habit models are all in $\mathcal{S}$. But the trap requires γ large enough that $\gamma \geq \alpha_C/2$. An SDF in $\mathcal{S}$ with low γ that generates its equity premium through an auxiliary channel, or an SDF outside $\mathcal{S}$ altogether, can avoid the moment explosion.

(a) Low tail loading: disappointment aversion

Schreindorfer (2020) builds a consumption-based model with generalised disappointment aversion (GDA) preferences (Gul, 1991; Routledge and Zin, 2010). The pricing kernel is

$$M = \tilde{\beta}\, g^{-\alpha} \cdot D,$$

where $g = C_{t+1}/C_t$ is consumption growth, $\tilde{\beta} > 0$ absorbs time preference and normalisation constants, $\alpha > 0$ governs curvature, and the bounded disappointment factor D takes value $(1+\theta)/(1+\theta\, p_d)$ in disappointing states and $1/(1+\theta\, p_d)$ otherwise (Appendix F).

I characterise here a simplified consumption-based version; the full Schreindorfer (2020) model includes a return-based SDF channel.

The SDF belongs to class $\mathcal{S}$ with tail parameter α for every $\alpha > 0$: as $g \to 0^+$, D converges to the constant $(1+\theta)/(1+\theta\, p_d)$, and the asymptotic behaviour is governed by $g^{-\alpha}$. The moment condition is $\mathbb{E}[M^k] < \infty$ iff $k\alpha < \alpha_C$ (Lemma 1).

At the benchmark calibration $\alpha = 1$: $M = \tilde{\beta}\, g^{-1} \cdot D \in \mathcal{S}$ with $\gamma = 1$. Since $\gamma = 1 < \alpha_C/2$ at every empirical estimate ($\alpha_C \geq 2.7$), $\mathrm{Var}(M) < \infty$ and the trap does not spring. The equity premium comes not from the *magnitude* of the tail loading (which at $\gamma = 1$ contributes $\sigma(M)/\mathbb{E}(M) \approx \sigma \approx 0.02$) but from the *probability* of crossing the disappointment threshold. The bounded factor D concentrates risk pricing on the event $\{\Delta c \leq x_v\}$. With the calibrated $\theta = 9.76$ and a disappointment probability of roughly 0.5% per month, the model matches the equity premium through a qualitatively different mechanism from CRRA.

Empirical support. Beason and Schreindorfer (2022) provide direct evidence for this mechanism. They decompose the equity premium by return state and find that roughly 80/100 of the premium is attributable to monthly returns below -10%, with the bulk concentrated between -30% and -10%. The GDA model replicates this pattern: 79/100 of the model's equity premium comes from crash states, close to the 80/100 in the data. By contrast, the leading alternatives—Campbell–Cochrane, Bansal–Yaron, Barro, and Wachter—attribute the premium to different regions of the return distribution and underprice tail risk by roughly 50%. The GDA escape is not merely theoretically coherent; it matches the cross-section of the equity premium across return states.

Formal condition. The GDA model at $\alpha = 1$ achieves:

$$M \in \mathcal{S} \text{ with } \gamma = 1 \quad \Longrightarrow \quad \gamma < \frac{\alpha_C}{2} \quad \Longrightarrow \quad \mathrm{Var}(M) < \infty.$$

The trap does not spring because $\gamma = 1$ is too low to push the SDF tail index $\alpha_M = \alpha_C/\gamma = \alpha_C$ below 2. The equity premium is generated by the bounded disappointment factor D, not by a large tail-loading exponent.

Option evidence. Since $M \in \mathcal{S}$ with $\gamma = 1$, the structural formula gives $\theta^* = \alpha_C/1 = \alpha_C \geq 2.7$. But Martin and Shi estimate $\theta^* \approx 1.7 < 2$. The option evidence independently constrains GDA at $\alpha = 1$: any model in $\mathcal{S}$ with $\mathrm{Var}(M) < \infty$ has $\theta^* > 2$, while the data require $\theta^* < 2$.

Fragility. The escape depends on $\alpha = 1$. The restriction $\alpha = 1$ is not a calibration convenience but the defining feature of first-order risk aversion in the Gul (1991) axioms. At $\alpha = 1$ the tail-loading exponent is $\gamma = 1$—the weakest possible divergence in $\mathcal{S}$. Schreindorfer (2020) writes that the choice is "inconsequential for the main results." Under his exponential mixture ($\alpha_C = \infty$), this is correct: all moments exist for any α.

Under power-law tails the situation is different. For $\alpha > 1$, the SDF is $M = \tilde{\beta}\, g^{-\alpha} \cdot D$, and the moment condition $\mathbb{E}[M^k] < \infty$ iff $k\alpha < \alpha_C$ becomes progressively more restrictive. At $\alpha = 2$ and $\alpha_C = 4$: $\mathbb{E}[M^2]$ is on the boundary of finiteness. At $\alpha = 3$: $\mathbb{E}[M]$ itself is marginal. Appendix F provides the full analysis, including the case $\alpha < 1$.

The GDA escape works through a precise mechanism: $\alpha = 1$ keeps the tail-loading exponent at $\gamma = 1$, low enough that $\mathrm{Var}(M) < \infty$, while the bounded disappointment factor D delivers the equity premium. The model is "doubly insulated" from the trap—by its distributional assumption ($\alpha_C = \infty$ under the exponential mixture) *and* by its preference

parameter ($\gamma = 1$ at $\alpha = 1$). The second insulation is the robust one: it survives even if consumption tails are power-law. The first does not.

Observable implications. If a low-tail-loading escape is operative, the SDF belongs to class $\mathcal{S}$ with $\gamma < \alpha_C/2$, implying $\theta^* = \alpha_C/\gamma > 2$. The empirical test is whether $\theta^* > 2$. Current evidence: Martin and Shi (2026) estimate $\hat{\theta}^* = 1.7$ with a 90% confidence interval of $(1.52, 2.05)$; the point estimate is below 2, though the upper bound of the confidence interval does not exclude it. Models with low tail loading predict $\theta^* > 2$ and are therefore in tension with the point estimate.

(b) Market structure: intermediary asset pricing

The trap binds a representative agent whose SDF is $M = h(g)$ with $h \in \mathcal{S}$. But the representative agent is a theoretical construct. If the marginal investor is not the representative consumer—if intermediaries, dealers, or constrained specialists set prices—then the effective SDF need not inherit the tail properties of aggregate consumption.

Calvet et al. (2025) show that intermediary hedging demand, not a representative agent's risk aversion, determines the variance risk premium. In their framework, the effective θ^* is endogenous to the distribution of risk-bearing across agents. With heterogeneous risk aversion and binding constraints, the tail loading of the aggregate SDF is not γ/α_C for any single agent's γ. The "representative agent's" risk aversion is an artefact of aggregation, and it may not correspond to the risk aversion of anyone in the economy.

Formal condition. If the effective SDF is determined by intermediary behaviour rather than by a representative agent in class $\mathcal{S}$, then condition (ii) of Theorem 2 does not apply. The intermediary's pricing kernel may load on different risks (e.g., funding liquidity, inventory) with different tail properties, and the moment constraint $\gamma < \alpha_C/2$ is replaced by a constraint on the intermediary's effective risk aversion and the tail properties of the risks it bears.

Empirical status. The intermediary asset pricing literature documents that dealer and broker-dealer leverage factors price the cross-section of returns (He and Krishnamurthy, 2013; Adrian et al., 2014). Whether this mechanism is sufficient to generate the full equity premium without springing the trap is an open question. The key theoretical insight is that the trap is a property of the class $\mathcal{S}$, not of markets in general. Market structure can allocate tail risk to agents whose effective SDF is outside the class, generating the premium through a channel that the theorem does not constrain.

4.3 Violate (iii): change the diagnostic

The third escape does not change the model. It changes the question. The Hansen–Jagannathan bound requires $\mathrm{Var}(M) < \infty$; the interquantile diagnostic of Proposition 2 extends estimation to $\mathbb{E}[M] < \infty$. Both are physical-measure objects. A diagnostic based entirely on risk-neutral quantities—observed option prices—sidesteps the moment explosion altogether.

The SVIX bound. Martin (2017) proves that

$$\mathbb{E}_t[R_{t+1}] - R_f \;\geq\; \mathrm{SVIX}_t^2,$$

where SVIX_t^2 is computed from traded option prices under a no-costless-catastrophe (NCC) condition. The bound involves only risk-neutral expectations—integrated from the observed cross-section of option prices—and does not require $\mathbb{E}[M^2] < \infty$ or even $\mathbb{E}[M] < \infty$. It is immune to the trap by construction: no physical moment of the SDF appears on either side of the inequality.

This is a genuine diagnostic advance. The SVIX bound converts the equity premium question from a statement about the SDF's moments—which the trap renders undefined—into a statement about option prices—which are observable and finite by market clearing. The conditional equity premium can be bounded from below at every date, without taking a stand on the consumption distribution, the preference specification, or the value of γ.

The caveat: NCC violations. The SVIX bound requires Martin's NCC condition: the pricing kernel must be negatively correlated with the market return in every state of the world. Schreindorfer and Sichert (2025) estimate the conditional pricing kernel jointly from option prices and realised returns and find that the NCC is **violated in high-volatility states**. The pricing kernel flattens during crises—a monthly return of -10% coincides with a marginal-utility multiple of 3.68 in low-volatility periods but only 1.32 in high-volatility periods. When the kernel flattens, the NCC fails, and the SVIX bound overstates the conditional equity premium.

During the 2008 financial crisis and the COVID-19 crash, the SVIX bound rises sharply, but the estimated conditional premium does not rise by the same magnitude. The risk-neutral escape is immune to the trap, but it is not immune to the business cycle. The bound is tightest precisely when risk premia are most stable and loosest when they are most interesting.

An important distinction: the Schreindorfer and Sichert (2025) evidence concerns the *conditional* NCC. The unconditional NCC may hold on average across the sample, and Martin's bound applied unconditionally remains a valid lower bound if the NCC holds in expectation. The distinction between conditional and unconditional violations matters for the practical reliability of the escape: the unconditional SVIX bound may be informative even if the conditional version overstates the premium in high-volatility episodes.

Formal condition. The SVIX escape achieves:

$$\text{Diagnostic uses } \mathbb{E}_t^{\mathbb{Q}}[\cdot] \text{ only} \quad \Longrightarrow \quad \text{no physical moment of } M \text{ required} \quad \Longrightarrow \quad \text{trap irrelevant.}$$

Condition (iii) of Theorem 2—matching the equity premium via $\sigma(M)/\mathbb{E}(M) \geq \text{SR}$—is replaced by an inequality that uses only risk-neutral quantities. The impossibility does not apply because the theorem's conclusion concerns $\text{Var}(M)$, which the SVIX diagnostic never invokes.

The scope of the escape. The SVIX bound provides a lower bound on the conditional equity premium. It does not, by itself, explain *why* the premium is large—it measures it, but does not attribute it to preferences, beliefs, or market structure. For structural modelling, one still needs an SDF, and if that SDF is in class $\mathcal{S}$ with power-law consumption tails, the trap constrains what the model can deliver. The risk-neutral escape is best understood as a diagnostic complement: it confirms that the equity premium is large even when physical-measure diagnostics are unreliable.

Observable implications. If the SVIX escape is operative, the no-costless-collar (NCC) condition of Martin (2017) holds unconditionally. The empirical test is whether NCC holds

in all market regimes. Current evidence: Schreindorfer and Sichert (2025) show that the pricing kernel flattens in high-volatility states, violating NCC during the 2008 crisis and the COVID period. The SVIX escape is valid unconditionally but its empirical implementation is fragile in precisely the states that matter most.

4.4 Replacement diagnostics and the domain hierarchy

The preceding subsections evaluate escapes that change the model—its distributional assumptions, its preferences, or its market structure. Escape (D) of Theorem 4 takes a different approach: it changes the *diagnostic* rather than the model. This subsection asks which diagnostic tools survive the trap and whether they preserve the economic content of the HJ bound.

The domain hierarchy. Propositions 2 and 3 established that the interquantile and L^θ diagnostics extend the domain from $\gamma < \alpha_C/2$ to $\gamma < \alpha_C$. The companion note Bhamra (2026a) proves a sharper result: there is a strict hierarchy among four diagnostics.

Diagnostic	**Domain**	**Bound type**	**Trade-off**
HJ (L^2)	$\gamma < \alpha_C/2$	Sharp population bound	Smallest domain
Entropy	$\gamma < \alpha_C$	Sharp population bound	Harder to estimate
L^θ ($1 < \theta < 2$)	$\gamma < \alpha_C/\theta$	Population bound	Product $\|Z\|_\theta \|Z\|_q \approx 1$: bound $\approx$ HJ
IQ	$\gamma < \alpha_C$	Estimable diagnostic	No sharp population bound

The strict inclusion is $\mathcal{D}_{\text{HJ}} \subsetneq \mathcal{D}_{\text{ent}} = \mathcal{D}_\theta = \mathcal{D}_{\text{IQ}} = \{0 < \gamma < \alpha_C\}$. The alternative diagnostics double the domain but cannot escape the trap: matching the equity premium still requires γ well above α_C.

What is gained and what is lost. The alternative diagnostics are not *ad hoc* substitutes for the HJ bound. Three properties establish their legitimacy. First, *continuity*: each diagnostic is continuous under tail perturbations (Kolmogorov–Smirnov convergence) within its domain—they degrade gracefully as tails get heavier, whereas the HJ bound fails catastrophically. Second, *monotonicity*: each diagnostic is strictly increasing in the tail-loading parameter γ, so models with stronger tail loading face tighter bounds. The alternative diagnostics preserve the economic logic of the HJ bound. Third, *ranking preservation*: under CRRA, model rankings are preserved across all four diagnostics—a model with $\gamma = 10$ faces a tighter bound than one with $\gamma = 3$ under any diagnostic for which both are in domain.

What is lost is *tightness*: the HJ bound is achieved by the minimum-variance SDF, and no alternative diagnostic has a comparably sharp characterisation of the achieving distribution. The IQ diagnostic, in particular, lacks a population lower bound from the pricing equation alone—its force is in estimation, not in theory.

The diagnostic horse race. The model audit in Appendix I computes all four diagnostics for nine canonical models. The comparison reveals a pattern: every model in class $\mathcal{S}$ with $\text{Var}(M) < \infty$ predicts $\theta^* > 2$, which is contradicted by the Martin and Shi (2026) estimate $\hat{\theta}^* \approx 1.7$. This is the *option side* of the trap: it constrains every preference-based escape within class $\mathcal{S}$, regardless of which diagnostic is used. The diagnostic escape changes the measurement tool but does not change the diagnosis.

Table 10: Escape route taxonomy. Each row violates exactly one condition of Theorem 2. "Empirical status" records whether the violated condition is supported or contradicted by the data. "Fragility" identifies the assumption whose failure would spring the trap.

Escape route	**Violates**	**Mechanism**	**Condition for** $\mathrm{Var}(M) < \infty$	**Fragility**
Wachter (2013)	(i)	Bounded disasters	$\alpha_C = \infty$ (bounded support)	Data reject $\alpha_C = \infty$
Schreindorfer (2020, exponential)	(i)	Exponential mixture tails	$\alpha_C = \infty$ (MGF exists)	Data reject $\alpha_C = \infty$
Schreindorfer (2020, GDA)	—	Low tail loading ($\gamma = 1$); premium from bounded D	$M \in \mathcal{S}$, $\gamma = 1 < \alpha_C/2$	$\alpha \neq 1$ with power-law tails $\Rightarrow$ trap springs
CCLV (2025)	(ii)	Intermediary risk-bearing	Effective SDF $\notin \mathcal{S}$	Requires binding intermediary constraints
Martin (2017)	(iii)	Risk-neutral diagnostic (SVIX)	No physical moments needed	NCC violated in crises; bound loose in high-vol states

Observable tests. Thin tails: $\alpha_C = \infty$? Five estimators yield $\hat{\alpha}_C \in [2.1, 3.8]$. GDA (low loading): $\theta^* > 2$? Point estimate $\hat{\theta}^* = 1.7$. CCLV (outside $\mathcal{S}$): SDF $\notin \mathcal{S}$? Open. Martin SVIX: NCC holds? Fails in crises.

Summary

Table 10 collects the five escape routes, organised by which condition of Theorem 2 each violates.

Three patterns emerge. First, the thin-tail escape (condition (i)) is empirically fragile: the data consistently estimate a finite α_C. Models that assume $\alpha_C = \infty$—whether through bounded support or exponential tails—avoid the trap by assumption, but that assumption is rejected by the evidence.

Second, the preference-based escape is the most promising structural resolution. Schreindorfer's GDA model at $\alpha = 1$ belongs to class $\mathcal{S}$ with $\gamma = 1$—low enough that $\mathrm{Var}(M) < \infty$—and generates the equity premium through the bounded disappointment factor rather than through the tail-loading exponent. The intermediary channel offers a complementary route: if the marginal investor is not the representative consumer, the effective SDF may lie outside class $\mathcal{S}$ entirely, and the trap constrains an entity that does not set prices.

Third, the risk-neutral escape (condition (iii)) is theoretically clean but empirically limited. The SVIX bound (Martin, 2017) should be understood as immune to the trap by construction: it requires only finite option prices, not any moment of M. But the bound requires the NCC condition, which is violated conditionally in the states where the equity premium is largest and most policy-relevant. Whether the unconditional NCC holds on average is a separate, and more favourable, question.

Theorems 2, 3, and 4 do not say the equity premium puzzle is unsolvable. They say that every solution carries a structural commitment—a claim about which feature of the standard framework does not hold. The taxonomy in Table 10 and the formal classification in Theorem 4 make those commitments explicit. A model that claims to explain the equity

premium under power-law tails and an SDF whose equity premium is driven by the tail-loading exponent is not merely incomplete; it is logically inconsistent. The value of the theorems is not in ruling out solutions but in requiring them to be honest about what they assume away. Appendix I applies this framework to nine canonical models, classifying each by which assumption it violates.

5 Conclusion

The equity premium puzzle is deeper than Mehra and Prescott (1985) showed. Their puzzle is a calibration complaint: risk aversion seems too high. The trap is a benchmark contradiction under power-law tails: the diagnostic used to reach that conclusion does not exist at the parameters it identifies.

This paper establishes three theoretical results. The tail-risk theorem under asymptotic factorisation (Theorem 1) shows that within a structured class of SDFs satisfying explicit factorisation conditions, any pricing kernel with tail elasticity η and a power-law disaster-exposed variable with index α_X satisfies $\mathbb{E}[M^\theta] < \infty$ for $\theta < \alpha_X/\eta$ and $\mathbb{E}[M^\theta] = \infty$ for $\theta > \alpha_X/\eta$. The theorem applies, within the maintained factorisation class, beyond CRRA and beyond class $\mathcal{S}$: it requires a well-defined tail elasticity, an asymptotic coefficient function capturing the noise-dependent scaling, and a moment condition on that coefficient. The conditional diagnostic failure result (Theorem 3) proves that, at empirical values and within the maintained factorisation class, the three pillars of the equity premium puzzle—power-law tails, sufficient tail loading, and finite SDF variance—are mutually inconsistent by a factor of ten. The framework-conditional taxonomy (Theorem 4) organises the leading resolutions into four benchmark families—thin tails, low loading, structural, and diagnostic—each with a canonical model and observable implications. Within the maintained framework, the taxonomy turns every proposed resolution into a falsifiable claim about which condition it violates.

The empirical evidence comes from three independent sources. First, an independent OptionMetrics replication (1996–2024, approximately 44 million contract-date observations) supports the Martin and Shi (2026) finding: the Lee (2004) moment formula yields a full-sample median $\hat{p}^* = 3.05$, declining to 1.64 in the 2020s—risk-neutral evidence that the risk-neutral tail geometry is consistent with $\theta^* < 2$. Under the maintained class-$\mathcal{S}$ mapping, this implies that the physical SDF lies in the infinite-variance region. Second, a multi-estimator battery—Hill, Pickands, QQ-regression, and bias-corrected Hill—places the consumption tail index at $\alpha_C \in [2.1, 3.0]$, confirmed by a Müller (2025) robust t-test diagnostic consistent with heavy-tailed behaviour. Third, ten-country evidence from the Jordà–Schularick–Taylor Macrohistory Database shows $\alpha_C \in [1.5, 3.0]$ across all advanced economies: the trap is not a U.S. artifact. The structural formula $\theta^* = \alpha_C/\gamma$ bridges options and consumption, pinning risk aversion to the narrow interval $(\alpha_C/2, \alpha_C)$, which excludes $\gamma \geq 5$ at every empirical estimate.

The model audit (Section 3.4, Appendix I) classifies nine canonical calibrations against the escape taxonomy. Under the maintained benchmark assumptions, the audit finds: Mehra and Prescott (1985), Bansal et al. (2012), and Barro (2006) are trapped at empirical α_C; Campbell and Cochrane (1999) remains formally unresolved (the joint tail behaviour of consumption growth and the surplus ratio requires separate analysis); Wachter (2013) escapes via thin tails (empirically fragile); Schreindorfer (2020) escapes via low loading (parametrically fragile); and Martin's (2017) SVIX bound escapes by changing the diagnostic. This is not a criticism of any single model. It is a classification that makes transparent what each model must assume in order to generate a large equity premium

without springing the trap.

The paper also provides a alternative diagnostic framework. Variance-based diagnostics—the Hansen–Jagannathan bound, GMM J-tests, maximum likelihood—assume $\text{Var}(M) < \infty$. The domain hierarchy establishes that the HJ bound has the smallest domain of validity ($\gamma < \alpha_C/2$), while the entropy, L^θ (Proposition 3), and interquantile (Proposition 2) diagnostics all extend the domain to $\gamma < \alpha_C$. Each alternative diagnostic is continuous under tail perturbations, strictly increasing in the tail loading, and preserves model rankings. This is a practical contribution: a coherent toolkit for asset-pricing analysis in heavy-tailed environments, not an *ad hoc* collection of alternatives.

Several limitations deserve emphasis. The formal results apply to pricing kernels with well-defined tail elasticity. Whether state-dependent models (Campbell–Cochrane external habit) and models with persistent dynamics (Bansal–Yaron long-run risk under non-i.i.d. growth) are trapped requires a formal analysis of the joint tail behavior of consumption growth and state variables, which is an important open question. The international evidence covers ten advanced economies; the cross-market extension to joint pricing restrictions is partial. Whether the trap affects equilibrium allocations and welfare—or only our ability to diagnose them—remains open. If $\text{Var}(M) = \infty$, the welfare cost of business cycles may itself be ill-defined under standard calculations.

Looking forward, the field needs preference specifications explicitly designed for heavy-tailed consumption environments—specifications in which the equity premium is large but the moments used to measure it are finite by construction, not by accident. The paper provides both the diagnosis and the prescription. The diagnosis is the trap: within the maintained framework, three conditions that underpin the equity premium puzzle are mutually inconsistent, and four benchmark families of escape are identified. The prescription is the replacement framework: a domain hierarchy of diagnostics that extend the standard toolkit into the infinite-variance regime. The question is not whether the equity premium is large. It is. The question is whether the tools we use to measure it are valid at the parameters they identify. The trap says they are not—and classifies every way out.

References

Adrian, Tobias, Erkko Etula, and Tyler Muir, "Financial Intermediaries and the Cross-Section of Asset Returns," *Journal of Finance*, 2014, *69* (6), 2557–2596.

Bansal, Ravi and Amir Yaron, "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance*, 2004, *59* (4), 1481–1509.

_ , Dana Kiku, and Amir Yaron, "An empirical evaluation of the long-run risks model for asset prices," *Critical Finance Review*, 2012, *1* (1), 183–221.

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

_ and José F. Ursúa, "Macroeconomic Crises since 1870," *Brookings Papers on Economic Activity*, 2008, *2008* (1), 255–335.

_ and _ , "Rare Macroeconomic Disasters," *Annual Review of Economics*, 2012, *4*, 83–109.

Basrak, Bojan, Richard A. Davis, and Thomas Mikosch, "Regular variation of GARCH processes," *Stochastic Processes and their Applications*, 2002, *99* (1), 95–115.

Beason, Tyler and David Schreindorfer, "Dissecting the Equity Premium," *Journal of Political Economy*, 2022, *130* (8), 2203–2222.

Bhamra, Harjoat S., "A Diagnostic Framework for Heavy-Tailed Asset Pricing," 2026. Theory companion to "The Equity Premium Trap".

_ , "Model Audit: Tail Properties of Nine Asset Pricing Models," 2026. Theory companion to "The Equity Premium Trap".

_ , "Structural Theorems for Tail-Loading SDFs," 2026. Theory companion to "The Equity Premium Trap".

Bingham, N. H., C. M. Goldie, and J. L. Teugels, *Regular Variation*, Cambridge University Press, 1987.

Borovička, Jaroslav, Lars Peter Hansen, and José A. Scheinkman, "Misspecified Recovery," *Journal of Finance*, 2016, *71* (6), 2493–2544.

Breeden, Douglas T. and Robert H. Litzenberger, "Prices of state-contingent claims implicit in option prices," *Journal of Business*, 1978, *51* (4), 621–651.

Breiman, Leo, "On some limit theorems similar to the arc-sin law," *Theory of Probability and Its Applications*, 1965, *10* (2), 323–331.

Calvet, Laurent, Claire Celerier, Gordon Liao, and Boris Vallee, "How Do Households Suppress the Price of Tail Risk?," *Working paper*, 2025.

Campbell, John Y. and John H. Cochrane, "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 1999, *107* (2), 205–251.

Carr, Peter and Dilip B. Madan, "Optimal Positioning in Derivative Securities," *Quantitative Finance*, 2001, *1* (1), 19–37.

Chen, Hui, Winston Wei Dou, and Leonid Kogan, "Measuring "dark matter" in asset pricing models," *Journal of Finance*, 2024, *79* (2), 843–902.

Cline, Daren B. H., "Convolution Tails, Product Tails and Domains of Attraction," *Probability Theory and Related Fields*, 1986, *72* (4), 529–557.

de Haan, Laurens and Ana Ferreira, *Extreme Value Theory: An Introduction*, Springer, 2006.

Duffie, Darrell and Larry G. Epstein, "Stochastic Differential Utility," *Econometrica*, 1992, *60* (2), 353–394.

Embrechts, Paul, Claudia Klüppelberg, and Thomas Mikosch, *Modelling Extremal Events for Insurance and Finance*, Springer, 1997.

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

Gabaix, Xavier, "Power Laws in Economics and Finance," *Annual Review of Economics*, 2009, *1*, 255–294.

Gul, Faruk, "A Theory of Disappointment Aversion," *Econometrica*, 1991, *59* (3), 667–686.

Hall, Peter, "Using the Bootstrap to Estimate Mean Squared Error and Select Smoothing Parameter in Nonparametric Problems," *Journal of Multivariate Analysis*, 1990, *32* (2), 177–203.

Hansen, Lars Peter and José A. Scheinkman, "Long-Term Risk: An Operator Approach," *Econometrica*, 2009, *77* (1), 177–234.

_ **and Kenneth J. Singleton**, "Generalized instrumental variables estimation of nonlinear rational expectations models," *Econometrica*, 1982, *50* (5), 1269–1286.

_ **and Ravi Jagannathan**, "Implications of security market data for models of dynamic economies," *Journal of Political Economy*, 1991, *99* (2), 225–262.

He, Zhiguo and Arvind Krishnamurthy, "Intermediary Asset Pricing," *American Economic Review*, 2013, *103* (2), 732–770.

Hill, Bruce M., "A simple general approach to inference about the tail of a distribution," *Annals of Statistics*, 1975, *3* (5), 1163–1174.

Jansen, Dennis W. and Casper G. de Vries, "On the frequency of large stock returns: Putting booms and busts into perspective," *Review of Economics and Statistics*, 1991, *73* (1), 18–24.

Jermann, Urban J., "Asset Pricing in Production Economies," *Journal of Monetary Economics*, 1998, *41* (2), 257–275.

Jessen, Anders H. and Thomas Mikosch, "Regularly Varying Functions," *Publications de l'Institut Mathématique*, 2006, *80* (94), 171–192.

Lee, Roger W., "The Moment Formula for Implied Volatility at Extreme Strikes," *Mathematical Finance*, 2004, *14* (3), 469–480.

Martin, Ian W. R., "What is the Expected Return on the Market?," *Quarterly Journal of Economics*, 2017, *132* (1), 367–433.

_ **and Ran Shi**, "On the Moments of the Stochastic Discount Factor," *Working paper*, 2026.

Mehra, Rajnish and Edward C. Prescott, "The equity premium: A puzzle," *Journal of Monetary Economics*, 1985, *15* (2), 145–161.

Müller, Ulrich K., "A more robust t-test," *Review of Economics and Statistics*, 2025, *107* (3), 786–802.

Òscar Jordà, Moritz Schularick, and Alan M. Taylor, "Macrofinancial History and the New Business Cycle Facts," *NBER Macroeconomics Annual*, 2017, *31* (1), 213–263.

Pickands, James, "Statistical inference using extreme order statistics," *Annals of Statistics*, 1975, *3* (1), 119–131.

Pohl, Walter, Karl Schmedders, and Ole Wilms, "Higher-Order Effects in Asset Pricing Models with Long-Run Risks," *Journal of Finance*, 2018, *73* (3), 1061–1111.

Qin, Likuan and Vadim Linetsky, "Long-Term Risk: A Martingale Approach," *Econometrica*, 2017, *85* (1), 299–312.

Resnick, Sidney I., *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*, Springer, 2007.

Routledge, Bryan R. and Stanley E. Zin, "Generalized Disappointment Aversion and Asset Prices," *Journal of Finance*, 2010, *65* (4), 1303–1332.

Samorodnitsky, Gennady and Murad S. Taqqu, *Stable Non-Gaussian Random Processes*, Chapman & Hall, 1994.

Schreindorfer, David, "Macroeconomic Tail Risks and Asset Prices," *Review of Financial Studies*, 2020, *33* (8), 3541–3582.

_ **and Tobias Sichert**, "Conditional Risk and the Pricing Kernel," *Journal of Financial Economics*, 2025, *171*, 104106.

Serfling, Robert J., *Approximation Theorems of Mathematical Statistics*, Wiley, 1980.

Toda, Alexis Akira and Kieran Walsh, "The double power law in consumption and implications for testing Euler equations," *Journal of Political Economy*, 2015, *123* (5), 1177–1200.

Wachter, Jessica A., "Can time-varying risk of rare disasters explain aggregate stock market volatility?," *Journal of Finance*, 2013, *68* (3), 987–1035.

A Proofs

This appendix collects the formal proofs of Theorem 1, Proposition 1, Corollary 1, Corollary 2, Theorem 2, Theorem 3, Theorem 4, Proposition 2, Proposition 3, and Proposition 4. The main tools are the regular variation calculus of Bingham et al. (1987) (hereafter BGT), Breiman's theorem for products of regularly varying random variables (Breiman, 1965), and elementary probability inequalities. Full proofs of Theorems 1, 3, and 4 are in the companion note Bhamra (2026c); this appendix provides self-contained proof sketches. No result in this appendix is mathematically new; the contribution is assembling standard ingredients into a self-contained argument for the impossibility theorem and its supporting results.

A.0 Proof sketch of Theorem 1 (General tail-loading moment theorem)

Proof sketch. The proof proceeds in three steps.

Step 1: Right tail of $Y := 1/X$. By Assumption 1, $\mathbb{P}(X \le x) \in \mathrm{RV}(\alpha_X)$ as $x \to 0^+$. Setting $y = 1/x$: $\mathbb{P}(Y > y) = y^{-\alpha_X}\,\ell(y)$ as $y \to \infty$. So Y has a regularly varying right tail with index α_X.

Step 2: Tail of M via Breiman's theorem. Define $V(y, \varepsilon) := \Psi(1/y, \varepsilon)$. By condition (F2), $V(y, \varepsilon) \sim a(\varepsilon) \cdot y^{\eta} \cdot \tilde{L}(y)$ as $y \to \infty$, where $\tilde{L}(y) := L(y)$ is slowly varying. So $y \mapsto V(y, \varepsilon)$ is in $\mathrm{RV}(\eta)$ at infinity, with ε-dependent coefficient $a(\varepsilon)$. Conditioning on $\varepsilon = e$: the composition theorem (BGT Theorem 1.5.6) gives $\mathbb{P}(V(Y, e) > m) \sim a(e)^{\alpha_X/\eta}\, m^{-\alpha_X/\eta}\, \tilde{\ell}(m)$. Integrating over ε: the uniform asymptotic factorisation in (F2) ensures that $r(x, \varepsilon) \to 1$ uniformly on

$\{a(\varepsilon) \le m_0\}$, so the pointwise tail equivalence can be passed through the expectation. By condition (F3) and dominated convergence (using Potter bounds on the remainder),

$$\mathbb{P}(M > m) \;\sim\; \mathbb{E}[a(\varepsilon)^{\alpha_X/\eta}]\, m^{-\alpha_X/\eta}\, \tilde{\ell}(m) \quad \text{as } m \to \infty.$$

This proves $\alpha_M = \alpha_X/\eta$. The moment condition (F3) on $a(\varepsilon)$ ensures the expectation is finite and the dominated convergence is valid.

Step 3: Moment conditions. Parts (b)–(d) follow from Karamata's moment theorem: $\mathbb{E}[M^\theta] < \infty$ iff $\theta < \alpha_X/\eta$. Part (d) is the special case $\theta = 2$. □

Remark 7 (Counterexamples at the boundary). Condition (F3) is not merely technical. When the noise factor ε is itself heavy-tailed with index $\alpha_\varepsilon < \alpha_X/\eta$, the product's tail index is $\min(\alpha_X/\eta,\ \alpha_\varepsilon) = \alpha_\varepsilon$ and the conclusion of Theorem 1 can fail. Counterexample 1 in the companion note Bhamra (2026c) works this out explicitly. The noise factor *worsens* the trap, never alleviates it.

A.1 Proof of Proposition 1 (Tail transfer)

I restate the regularity conditions for completeness.

Regularity conditions.

(R1) Consumption growth $g = C_{t+1}/C_t$ is a strictly positive random variable whose left tail is regularly varying:

$$\mathbb{P}(g \le x) \;=\; x^{\alpha_C}\, \ell(1/x) \quad \text{as } x \to 0^+,$$

where $\alpha_C \in (0, \infty)$ and ℓ is slowly varying at infinity (Assumption 1).

(R2) The SDF is $M = h(g)$ where $h : (0, \infty) \to (0, \infty)$ satisfies $h \in \mathrm{RV}(-\gamma)$ at 0^+ for some $\gamma > 0$ (Assumption 2). That is, $h(\lambda x)/h(x) \to \lambda^{-\gamma}$ as $x \to 0^+$ for every $\lambda > 0$.

(R3) The function h is eventually monotone decreasing on a neighbourhood of 0^+. This ensures that small values of g (disasters) produce large values of M (high marginal utility). All standard preferences satisfy this condition.

Proof of part (a): $\alpha_M = \alpha_C/\gamma$. The argument has three steps.

Step 1: Transform the left tail of g to a right tail. Define $Y = 1/g$. Since $g > 0$ a.s., Y is well-defined and strictly positive. The left tail of g maps to the right tail of Y: for large y,

$$\mathbb{P}(Y > y) \;=\; \mathbb{P}(g < 1/y) \;=\; y^{-\alpha_C}\, \ell(y) \quad \text{as } y \to \infty.$$

Hence Y has a regularly varying right tail with index α_C.

Step 2: Express M as a regularly varying function of Y. Define $\tilde{h}(y) := h(1/y)$, so that $M = h(g) = \tilde{h}(1/g) = \tilde{h}(Y)$. Since $h \in \mathrm{RV}(-\gamma)$ at 0^+:

$$\frac{\tilde{h}(\lambda y)}{\tilde{h}(y)} \;=\; \frac{h(1/(\lambda y))}{h(1/y)} \;=\; \frac{h(\lambda^{-1} \cdot (1/y))}{h(1/y)} \;\to\; (\lambda^{-1})^{-\gamma} \;=\; \lambda^{\gamma} \quad \text{as } y \to \infty.$$

Therefore $\tilde{h} \in \mathrm{RV}(\gamma)$ at infinity. Since $\gamma > 0$, $\tilde{h}(y) \to \infty$ as $y \to \infty$: large values of Y (severe disasters) produce large values of M.

Step 3: Apply the composition theorem for regularly varying tails. BGT Theorem 1.5.6 (the quantile transformation theorem) states: if X is a non-negative random variable with

$\mathbb{P}(X > x) \in \mathrm{RV}(-\alpha)$ at infinity ($\alpha > 0$), and $\phi : (0,\infty) \to (0,\infty)$ satisfies $\phi \in \mathrm{RV}(\rho)$ at infinity with $\rho > 0$ and ϕ is eventually monotone increasing, then

$$\mathbb{P}(\phi(X) > m) \;\in\; \mathrm{RV}(-\alpha/\rho) \quad \text{as } m \to \infty.$$

Apply this with $X = Y$, $\phi = \tilde{h}$, $\alpha = \alpha_C$, and $\rho = \gamma$. Condition (R3) ensures $\tilde{h}$ is eventually monotone increasing. The conclusion is

$$\mathbb{P}(M > m) \;=\; \mathbb{P}\big(\tilde{h}(Y) > m\big) \;\in\; \mathrm{RV}\left(-\frac{\alpha_C}{\gamma}\right) \quad \text{as } m \to \infty.$$

That is, M has a regularly varying right tail with index $\alpha_M = \alpha_C/\gamma$. □

Proof of part (b): moment characterisation. By part (a), the SDF M has a regularly varying right tail: $\mathbb{P}(M > m) = m^{-\alpha_M}\,\tilde{\ell}(m)$ as $m \to \infty$, where $\alpha_M = \alpha_C/\gamma$ and $\tilde{\ell}$ is slowly varying.

Karamata's moment theorem (BGT Theorem 1.5.11; also Embrechts et al. 1997, Proposition A3.8) states: for a non-negative random variable X with $\mathbb{P}(X > x) \in \mathrm{RV}(-\xi)$ at infinity, $\xi > 0$,

(i) $\mathbb{E}[X^k] < \infty$ for $k < \xi$,

(ii) $\mathbb{E}[X^k] = \infty$ for $k > \xi$,

(iii) the boundary case $k = \xi$ depends on the slowly varying function ℓ and requires case-by-case analysis; it cannot be determined from ξ alone.

Applying this with $X = M$ and $\xi = \alpha_C/\gamma$: $\mathbb{E}[M^k] < \infty$ for $k < \alpha_C/\gamma$ (equivalently $\gamma < \alpha_C/k$) and $\mathbb{E}[M^k] = \infty$ for $k > \alpha_C/\gamma$. The boundary case $k = \alpha_C/\gamma$ is indeterminate without further information about the slowly varying function in the tail of g. This distinction is immaterial for all results in this paper: the empirical values of α_C/γ lie far from any integer k of interest. □

Proof of part (c): variance characterisation. Set $k = 2$ in part (b). Then $\mathbb{E}[M^2] < \infty$ if and only if $\gamma < \alpha_C/2$. Since $\mathrm{Var}(M) = \mathbb{E}[M^2] - (\mathbb{E}[M])^2$ and $\mathbb{E}[M^2] < \infty$ implies $\mathbb{E}[M] < \infty$ (by Jensen), the equivalence $\mathrm{Var}(M) < \infty \Leftrightarrow \gamma < \alpha_C/2$ follows. □

Remark 8 (Role of the slowly varying function). The slowly varying function ℓ in Assumption 1 plays no role in any moment existence condition. Karamata's moment theorem depends only on the power-law index α_C, not on ℓ. This is why the impossibility theorem is robust to the precise functional form of the tail.

A.2 Proof of Corollary 1 (Structural θ^*)

Proof. By definition, $\theta^* := \sup\{k > 0 : \mathbb{E}[M^k] < \infty\}$. Proposition 1(b) gives $\mathbb{E}[M^k] < \infty \Leftrightarrow k < \alpha_C/\gamma$. Therefore $\{k > 0 : \mathbb{E}[M^k] < \infty\} = (0, \alpha_C/\gamma)$ and

$$\theta^* = \frac{\alpha_C}{\gamma}.$$

Whether the supremum is attained (i.e., whether $\mathbb{E}[M^{\alpha_C/\gamma}]$ is finite or infinite) depends on the slowly varying function in the tail of g. For the purposes of this paper, the boundary case is immaterial: the empirical values of α_C/γ lie far from any integer k of interest. □

A.3 Proof of Corollary 2 (Risk aversion bounds)

The corollary is a biconditional: $1 < \theta^* < 2 \Leftrightarrow \alpha_C/2 < \gamma < \alpha_C$. I give two proofs—one through the structural formula, one through Proposition 1(b) directly—and verify the result in the CRRA special case.

Proof 1: via the structural formula. By Corollary 1, $\theta^* = \alpha_C/\gamma$. The map $\gamma \mapsto \alpha_C/\gamma$ is a strictly decreasing bijection from $(0, \infty)$ to $(0, \infty)$. Therefore:

$$\theta^* < 2 \iff \alpha_C/\gamma < 2 \iff \gamma > \alpha_C/2.$$
$$\theta^* > 1 \iff \alpha_C/\gamma > 1 \iff \gamma < \alpha_C.$$

Combining: $1 < \theta^* < 2 \iff \alpha_C/2 < \gamma < \alpha_C$. The point estimate $\gamma = \alpha_C/\theta^*$ is immediate. □

Proof 2: via Proposition 1(b). This proof uses the moment characterisation directly, without the closed-form expression for θ^*.

Lower bound ($\gamma > \alpha_C/2$). Suppose $\theta^* < 2$. By Karamata's moment theorem, $\mathbb{E}[M^k] = \infty$ for every $k > \theta^*$. Since $2 > \theta^*$, we have $\mathbb{E}[M^2] = \infty$. By Proposition 1(b), $\mathbb{E}[M^2] < \infty \Leftrightarrow \gamma < \alpha_C/2$. Contrapositive: $\mathbb{E}[M^2] = \infty \Rightarrow \gamma \geq \alpha_C/2$. For strictness: if $\gamma = \alpha_C/2$, then Proposition 1(b) gives $\mathbb{E}[M^k] < \infty$ for all $k < 2$ and $\mathbb{E}[M^k] = \infty$ for all $k > 2$, so $\theta^* = \sup\{k : \mathbb{E}[M^k] < \infty\} = 2$, contradicting $\theta^* < 2$. Therefore $\gamma > \alpha_C/2$.

Upper bound ($\gamma < \alpha_C$). Suppose $\theta^* > 1$. Then $\mathbb{E}[M] < \infty$ (since $1 < \theta^*$). By Proposition 1(b) with $k = 1$: $\mathbb{E}[M] < \infty \Leftrightarrow \gamma < \alpha_C$.

Converse. Suppose $\alpha_C/2 < \gamma < \alpha_C$. By Proposition 1(b), $\mathbb{E}[M^k] < \infty$ if and only if $k < \alpha_C/\gamma$.

Since $\gamma < \alpha_C$, we have $\alpha_C/\gamma > 1$, so the interval $(1,\ \alpha_C/\gamma)$ is non-empty. Pick any k in this interval. Then $\mathbb{E}[M^k] < \infty$ (since $k < \alpha_C/\gamma$) with $k > 1$, so $\theta^* \geq k > 1$.

Since $\gamma > \alpha_C/2$, we have $\alpha_C/\gamma < 2$, so the interval $(\alpha_C/\gamma,\ 2)$ is non-empty. Pick any k in this interval. Then $\mathbb{E}[M^k] = \infty$ (since $k > \alpha_C/\gamma$) with $k < 2$, so $\theta^* \leq k < 2$.

Therefore $1 < \theta^* < 2$. □

CRRA verification. Under CRRA, $M = \beta\, g^{-\gamma}$ with $\beta > 0$. Let $\mathbb{P}(g \leq x) = x^{\alpha_C}\, \ell(1/x)$ as $x \to 0^+$. The right tail of M is:

$$\mathbb{P}(M > m) = \mathbb{P}\left(g < \left(\frac{\beta}{m}\right)^{1/\gamma}\right) = \left(\frac{\beta}{m}\right)^{\alpha_C/\gamma} \ell\left(\left(\frac{m}{\beta}\right)^{1/\gamma}\right) \ \in\ \mathrm{RV}\left(-\frac{\alpha_C}{\gamma}\right)$$

as $m \to \infty$. The k-th moment integral is

$$\mathbb{E}[M^k] = \int_0^\infty k\, m^{k-1}\, \mathbb{P}(M > m)\, \mathrm{d}m.$$

The tail contribution for large m behaves as $\int_N^\infty k\, m^{k-1-\alpha_C/\gamma}\, \tilde{\ell}(m)\, \mathrm{d}m$ for a slowly varying $\tilde{\ell}$. By standard results on regularly varying integrals (Bingham et al., 1987, Proposition 1.5.10), this converges if and only if $k - 1 - \alpha_C/\gamma < -1$, i.e., $k < \alpha_C/\gamma$. Therefore $\theta^* = \alpha_C/\gamma$, confirming the general formula.

Numerical check. At $\alpha_C = 4$, $\gamma = 3$: $\theta^* = 4/3 \approx 1.33 \in (1, 2)$ and $\gamma = 3 \in (2, 4) = (\alpha_C/2, \alpha_C)$. ✓
At $\alpha_C = 2.7$, $\gamma = 1.6$: $\theta^* = 2.7/1.6 \approx 1.69 \in (1, 2)$ and $\gamma = 1.6 \in (1.35, 2.7) = (\alpha_C/2, \alpha_C)$. ✓ □

Remark 9 (Boundary cases). At the boundary $\gamma = \alpha_C/2$ (equivalently $\theta^* = 2$), the finiteness of $\mathbb{E}[M^2]$—and hence $\mathrm{Var}(M)$—depends on the slowly varying function ℓ in Assumption 1 and cannot be determined from α_C alone. Similarly, at $\gamma = \alpha_C$ ($\theta^* = 1$), the finiteness of $\mathbb{E}[M]$ is indeterminate. The corollary uses strict inequalities throughout: $\theta^* \approx 1.7$ lies well away from both boundaries.

A.4 Proof of Theorem 2 (Conditional impossibility)

The proof has two parts: a quantitative result under CRRA, and a qualitative extension to class $\mathcal{S}$. The quantitative part computes the critical γ^* under the lognormal approximation to consumption growth. The qualitative part explains why the gap is too large for any perturbation within $\mathcal{S}$ to close.

Part I: CRRA. Suppose $M = \beta g^{-\gamma}$ with $\beta > 0$, and $\log g \sim \mathcal{N}(\mu, \sigma^2)$ with $\mu = \log(1.02) \approx 0.0198$ and $\sigma = 0.021$ (the annualised postwar U.S. calibration).[3]

Step 1: Sharpe ratio requirement. Under lognormality:

$$\mathbb{E}[M] = \beta\, e^{-\gamma\mu+\gamma^2\sigma^2/2}, \qquad \mathbb{E}[M^2] = \beta^2\, e^{-2\gamma\mu+2\gamma^2\sigma^2}.$$

The squared coefficient of variation is

$$\frac{\mathrm{Var}(M)}{(\mathbb{E}[M])^2} \;=\; \frac{\mathbb{E}[M^2]}{(\mathbb{E}[M])^2} - 1 \;=\; e^{\gamma^2\sigma^2} - 1. \tag{6}$$

The Hansen–Jagannathan bound requires $\sigma(M)/\mathbb{E}(M) \geq \mathrm{SR}$, i.e., $e^{\gamma^2\sigma^2} - 1 \geq \mathrm{SR}^2$. Solving:

$$\gamma \;\geq\; \gamma^* \;:=\; \frac{1}{\sigma}\sqrt{\log(1+\mathrm{SR}^2)}. \tag{7}$$

At annualised $\mathrm{SR} = 0.44$ and $\sigma = 0.021$:

$$\gamma^* = \frac{\sqrt{\log(1.1936)}}{0.021} = \frac{\sqrt{0.1771}}{0.021} \approx \frac{0.4208}{0.021} \approx 20.0.$$

The exact lognormal calculation gives $\gamma^* \approx 20$. This is a CRRA/lognormal benchmark; under power-law tails the precise value may differ (see Part II).

Step 2: Moment existence constraint. By Proposition 1(c), $\mathrm{Var}(M) < \infty$ requires $\gamma < \alpha_C/2$.

Step 3: Empty intersection. The Sharpe ratio requires $\gamma \geq \gamma^* \geq 20$. Finite variance requires $\gamma < \alpha_C/2$. The admissible set $\{\gamma \geq \gamma^*\} \cap \{\gamma < \alpha_C/2\}$ is empty whenever $\gamma^* \geq \alpha_C/2$, i.e., $\alpha_C \leq 2\gamma^* \approx 40$. All empirical estimates give $\alpha_C \leq 5$:

$$\begin{aligned} \alpha_C = 4 \;:&\quad \gamma^* \geq 20 \;>\; 2.0 = \alpha_C/2. \\ \alpha_C = 2.7 \;:&\quad \gamma^* \geq 20 \;>\; 1.35 = \alpha_C/2. \end{aligned}$$

□

Remark 10 (Conservatism of the lognormal benchmark). The threshold $\gamma^* \approx 20$ is computed under lognormal consumption growth—the thinnest tail consistent with the empirical volatility σ. Under heavy-tailed consumption growth satisfying Assumption 1, the Sharpe

[3] Throughout this appendix, $\sigma = 0.021$ denotes the annualised standard deviation of log consumption growth following the Mehra and Prescott (1985) calibration. Section A.5 uses $\sigma = 0.02$ for the quarterly calibration; the quarterly figure is slightly larger than $\sigma_{\text{annual}}/2 = 0.0105$ because consumption growth exhibits modest excess kurtosis at quarterly frequency.

ratio $\sigma(M)/\mathbb{E}(M)$ for CRRA $M = \beta g^{-\gamma}$ weakly exceeds the lognormal benchmark at the same γ, because power-law tails inflate $\mathbb{E}[M^2]$ relative to $(\mathbb{E}[M])^2$. To see this, note that $\mathbb{E}[M^2]$ places more weight on extreme realisations of M—i.e., on consumption disasters—than does $(\mathbb{E}[M])^2$. Under regular variation, the right tail of M^2 is heavier than under lognormality at every $\gamma > 0$, so $\mathbb{E}[M^2]/(\mathbb{E}[M])^2 \geq e^{\gamma^2\sigma^2}$ whenever the comparison is well-defined (i.e., when both moments are finite). Therefore the minimum γ needed to reach SR ≥ 0.44 is at most the lognormal $\gamma^* \approx 20$—not larger. The impossibility gap is conservative: the lognormal calculation overstates γ^*, and the true gap between γ^* and $\alpha_C/2$ may be even wider.

Part II: Scope within class $\mathcal{S}$. The moment existence condition $\mathrm{Var}(M) < \infty \Leftrightarrow \gamma < \alpha_C/2$ holds for *every* $M \in \mathcal{S}$ (Proposition 1(c)), not only for CRRA. This is because the tail-risk formula $\alpha_M = \alpha_C/\gamma$ depends only on the asymptotic behavior of h at 0^+—the regular variation index—not on its global functional form.

The Sharpe ratio threshold γ^* does depend on the global form of h. For a general $h \in \mathcal{S}$, the relationship between γ and $\sigma(M)/\mathbb{E}(M)$ is determined by the entire distribution of g and the function h, not only by its tail behavior. The quantitative value $\gamma^* \approx 20$ is therefore CRRA-specific.

The qualitative extension holds for specifications whose Sharpe ratio scales with γ: Epstein–Zin under i.i.d. growth (Proposition 4), and perturbations of CRRA that preserve the asymptotic tail loading. In these cases, matching $\sigma(M)/\mathbb{E}(M) \geq 0.44$ forces γ above $\alpha_C/2$ because consumption growth has low unconditional volatility ($\sigma_g \approx 0.021$ annualised).

Counterexample: GDA at $\alpha = 1$. The extension does not hold for all $M \in \mathcal{S}$. GDA at $\alpha = 1$ has $M = \tilde{\beta}\, g^{-1} \cdot D \in \mathcal{S}$ with $\gamma = 1 < \alpha_C/2$, so $\mathrm{Var}(M) < \infty$. Yet the bounded disappointment factor D generates $\sigma(M)/\mathbb{E}(M) \geq 0.44$ (Appendix F). The equity premium comes from D, not from the tail-loading exponent γ. This is why the formal theorem is stated for CRRA: the impossibility requires that the Sharpe ratio is generated by the tail loading, not by an auxiliary bounded factor. □

A.4b Proof sketch of Theorem 3 (Structural necessity of diagnostic failure)

Proof sketch. The proof is a direct combination of Theorem 1 and the Sharpe ratio constraint.

Step 1: Moment constraint. By Theorem 1(d), $\mathrm{Var}(M) < \infty$ requires $\eta < \alpha_X/2$.

Step 2: Sharpe ratio constraint. Matching SR ≥ 0.44 under the hypothesis that the equity premium is generated by the tail elasticity requires $\eta \geq \eta_{\min}$. Under CRRA with lognormal growth, the closed-form calculation gives $\eta_{\min} = \gamma^* \approx 20$ (Section A.4, equation (7)). Under heavy tails, the true $\eta_{\min}$ is at most γ^* (Remark 10).

Step 3: Empty intersection. The admissible set $\{\eta \geq \eta_{\min}\} \cap \{\eta < \alpha_X/2\}$ is non-empty iff $\eta_{\min} < \alpha_X/2$, i.e., $\alpha_X/\eta_{\min} > 2$. By hypothesis, $\alpha_X/\eta_{\min} < 2$. At $\alpha_X = 4$, $\eta_{\min} = 20$: $\alpha_X/2 = 2$, $\eta_{\min} = 20$. The gap is a factor of 10. □

Remark 11 (Advance over Theorem 2). Theorem 2 is the CRRA special case: set $X = g$, $\varepsilon = \beta$, $\eta = \gamma$, and $\eta_{\min} = \gamma^*$. The generalisation allows X to be any disaster-exposed variable (output, intermediary wealth) and admits non-deterministic noise factors ε, covering models with bounded auxiliary components. The impossibility persists because the tail-risk $\alpha_M = \alpha_X/\eta$ is a property of regular variation, not of CRRA.

A.4c Proof sketch of Theorem 4 (Classification of escape routes)

Proof sketch. The proof has two parts: mutual exclusivity and coverage within the maintained framework.

Part I: Mutual exclusivity. Theorem 3 requires three conditions plus the applicability of the factorisation (Theorem 1). Each escape violates exactly one:

- Escape (A) violates $\alpha_X < \infty$.
- Escape (B) violates $\eta \geq \eta_{\min}$: the tail loading is too low to generate the equity premium alone, but the premium comes from elsewhere.
- Escape (C) violates the factorisation assumption: the SDF does not admit a decomposition with well-defined tail elasticity.
- Escape (D) violates $\mathrm{Var}(M) < \infty$ as a diagnostic requirement.

Part II: Exhaustiveness. Consider any pricing kernel M that matches $\mathrm{SR} \geq 0.44$ and is not trapped. Either $\alpha_X = \infty$ (Escape A), or $\alpha_X < \infty$. If $\alpha_X < \infty$, either the factorisation holds or it does not. If the factorisation fails: Escape C. If the factorisation holds with $\eta < \alpha_X/2$: Escape B (the equity premium has a non-η source). If the factorisation holds with $\eta \geq \alpha_X/2$ and $\mathrm{Var}(M) = \infty$: either the analyst uses an alternative diagnostic (Escape D) or the model is trapped. The four cases exhaust the logical possibilities. □

Remark 12 (Counterexamples at the boundaries). The companion note Bhamra (2026c) works out four counterexamples that delineate the boundaries of Theorems 1–4: (i) heavy-tailed noise (product of two $\mathrm{RV}(\alpha)$ factors is $\mathrm{RV}(\alpha)$, not $\mathrm{RV}(\alpha/2)$); (ii) oscillating tail elasticity (no well-defined η); (iii) CCLV intermediary (legitimate Escape B2); (iv) time-varying α_X (does not create a fifth escape family).

A.4d Diagnostic framework (Propositions D.1–D.4)

The companion note Bhamra (2026a) establishes a systematic comparison of four SDF diagnostics: the classical HJ bound (L^2), the entropy bound, the L^θ pricing bound, and the interquantile (IQ) diagnostic. The formal results are:

- *Proposition D.1 (Domain hierarchy).* Under Assumptions 1–2: $\mathcal{D}_{\mathrm{HJ}} = (0, \alpha_C/2) \subsetneq \mathcal{D}_{\mathrm{ent}} = \mathcal{D}_\theta = \mathcal{D}_{\mathrm{IQ}} = (0, \alpha_C)$. Proof: the entropy domain equals $\{\gamma : \mathbb{E}[M \log M] < \infty\} = \{\gamma < \alpha_C\}$ (since $\mathbb{E}[M \log^+ M] < \infty$ whenever $\mathbb{E}[M^{1+\epsilon}] < \infty$ for some $\epsilon > 0$, which holds iff $\gamma < \alpha_C$). The L^θ domain equals $\{\gamma : \mathbb{E}[M^\theta] < \infty \text{ for some } \theta > 1\} = \{\gamma < \alpha_C\}$. The IQ domain equals $\{\gamma : \mathbb{E}[M] < \infty\} = \{\gamma < \alpha_C\}$.
- *Proposition D.2 (Continuity).* Each diagnostic is continuous under KS convergence within its domain. Proof: weak convergence plus uniform integrability (guaranteed by the existence of moments with a δ-margin within each domain).
- *Proposition D.3 (Monotonicity in γ).* Under CRRA, each diagnostic is strictly increasing in γ. Proof: the tilted-measure argument—increasing γ inflates the right tail of M relative to its mean, and the monotone likelihood ratio property ensures the diagnostic value increases.
- *Proposition D.4 (Ranking preservation).* Under CRRA, model rankings are preserved across all four diagnostics within their common domain. Proof: immediate from the strict monotonicity established in D.3.

Full proofs are in Bhamra (2026a). The key takeaway: the alternative diagnostics double the domain of validity while preserving the economic logic of the HJ bound. What they cannot do is escape the trap—the impossibility is in the data, not in the diagnostic.

A.5 Proof of Proposition 2 (Interquantile diagnostic)

The proposition has two parts: a statistical result about estimation consistency, and a CRRA calibration.

Proof of part (a): statistical advantage. The sample p-quantile $\hat{Q}_p := X_{(\lceil np \rceil)}$ (the $\lceil np \rceil$-th order statistic) satisfies, under standard regularity conditions (Serfling, 1980, Theorem 2.3.1):

$$\sqrt{n}\left(\hat{Q}_p - Q_p\right) \xrightarrow{d} \mathcal{N}\left(0,\, \frac{p(1-p)}{f(Q_p)^2}\right),$$

where f is the density of M at Q_p. The requirements are that $f(Q_p) > 0$ and that Q_p is finite. The first is a regularity condition at the quantile. The second holds whenever the distribution of M has finite support or, more generally, whenever $\mathbb{E}[M] < \infty$ (which ensures $Q_p < \infty$ for all $p < 1$). The condition $\mathbb{E}[M] < \infty$—equivalently $\gamma < \alpha_C$ (Proposition 1(b) with $k = 1$)—is not a statistical prerequisite for quantile consistency (sample quantiles can be consistent even without finite mean) but an economic requirement: $\mathbb{E}[M] < \infty$ is needed for a well-defined risk-free rate. No second-moment condition is needed for either the statistical or the economic side.

By contrast, the sample variance $\hat{\sigma}^2 = n^{-1}\sum(M_i - \bar{M})^2$ is inconsistent when $\mathrm{Var}(M) = \infty$ (i.e., when $\gamma \geq \alpha_C/2$). In the infinite-variance case, $\hat{\sigma}^2$ does not converge to a constant. The dominant contribution is $n^{-1}\sum M_i^2$. Since M has tail index $\alpha_M = \alpha_C/\gamma < 2$, the power-transform rule gives the tail index of M^2 as $\alpha_M/2 < 1$: explicitly, $\mathbb{P}(M^2 > x) = \mathbb{P}(M > \sqrt{x}) \in \mathrm{RV}(-\alpha_M/2)$ as $x \to \infty$. Because $\alpha_M/2 < 1$, the mean $\mathbb{E}[M^2]$ is infinite, and the partial sums $\sum_{i=1}^n M_i^2$ lie in the domain of attraction of a stable law with stability index $\alpha_M/2 \in (0,1)$. The required normalisation is of order $n^{1/(\alpha_M/2)} = n^{2/\alpha_M}$ (Samorodnitsky and Taqqu, 1994, Theorem 4.5.1), so the standard n^{-1} normalisation in $\hat{\sigma}^2$ yields a divergent limit rather than a constant. The classical HJ bound $\hat{\sigma}(M)/\bar{M} \geq \mathrm{SR}$ is therefore unreliable in the region $\alpha_C/2 \leq \gamma < \alpha_C$: the left-hand side does not converge to its population counterpart.

The interquantile approach extends the domain of consistent estimation from $\gamma < \alpha_C/2$ to $\gamma < \alpha_C$, doubling the admissible region. □

Proof of part (b): CRRA calibration. Consider CRRA preferences $M = \beta\, g^{-\gamma}$ in the Lucas-tree economy where the market is the consumption claim. Suppose $\log g \sim \mathcal{N}(\mu, \sigma^2)$ with $\mu = \log(1.02) \approx 0.0198$ and $\sigma = 0.02$ (quarterly total real per capita consumption growth, NIPA 1959:Q1–2021:Q4).[4]

Step 1: Equity premium. The gross return on the consumption claim is $R_C = [(p_C + 1)/p_C]\, g$, where p_C is the constant price-consumption ratio under i.i.d. growth. The risk-free rate is $R_f = 1/\mathbb{E}[M] = (1/\beta)\, e^{\gamma\mu - \gamma^2\sigma^2/2}$. The equity premium (to first order in σ) is

$$\mathbb{E}[R_C] - R_f \;\approx\; \gamma\,\sigma^2, \tag{8}$$

which is the consumption-CAPM result (Breeden and Litzenberger, 1978).

[4]The quarterly $\sigma = 0.02$ is slightly larger than $\sigma_{\text{annual}}/2 = 0.021/2 = 0.0105$ because consumption growth exhibits modest excess kurtosis at quarterly frequency, and the NIPA aggregate is broader than the Mehra and Prescott (1985) measure. Using $\sigma = 0.0105$ would give $\gamma \geq 25.7$, strengthening the result.

Step 2: Interquantile range of excess returns. Since R_C is proportional to g (a lognormal random variable), the excess return $R^e = R_C - R_f$ is approximately normal for small σ. Its interquantile range at $p = 0.25$ is

$$\mathrm{IQR}_{0.25}(R^e) \;\approx\; 2\,\Phi^{-1}(0.75)\,\sigma(R^e) \;\approx\; 2\,\Phi^{-1}(0.75)\,\sigma \;=\; 2 \times 0.6745 \times 0.02 \;=\; 0.02698, \quad (9)$$

where $\Phi^{-1}(0.75) = 0.6745$ is the 75th percentile of the standard normal.

Step 3: Interquantile Sharpe ratio. Combining (8) and (9):

$$\text{IQ-SR}_{0.25} \;=\; \frac{|\mathbb{E}[R^e]|}{\mathrm{IQR}_{0.25}(R^e)} \;\approx\; \frac{\gamma\,\sigma^2}{2\,\Phi^{-1}(0.75)\,\sigma} \;=\; \frac{\gamma\,\sigma}{2\,\Phi^{-1}(0.75)}. \quad (10)$$

Setting IQ-SR $= 0.20$ and solving for γ:

$$\gamma \;\geq\; \frac{0.20 \times 2 \times 0.6745}{0.02} \;=\; \frac{0.2698}{0.02} \;=\; 13.49.$$

The text reports $\gamma \geq 13.3$ because the exact calculation (using the full lognormal distribution rather than the normal approximation to excess returns) yields a slightly lower threshold.

Step 4: Empty admissible set. The interquantile diagnostic is consistently estimable for $\gamma < \alpha_C$. But matching IQ-SR $= 0.20$ requires $\gamma \geq 13.3$. At $\alpha_C = 4$: $13.3 > 4 = \alpha_C$. At $\alpha_C = 2.7$: $13.3 > 2.7 = \alpha_C$. Even the extended domain is insufficient. The admissible set $\{\gamma : \gamma \geq 13.3\} \cap \{\gamma < \alpha_C\}$ is empty at every empirical estimate of α_C. □

Remark 13 (Population bound vs. statistical advantage). The force of the interquantile diagnostic is in *estimation*, not in relaxing population moment requirements. The classical HJ bound $\sigma(M)/\mathbb{E}(M) \geq \mathrm{SR}$ is a population inequality that requires $\mathrm{Var}(M) < \infty$. A population lower bound on $\mathrm{IQR}_p(M)$ from the pricing equation $\mathbb{E}[M R^e] = 0$ and $\mathbb{E}[M] < \infty$ alone does not yield a useful inequality: the infimum of $\mathrm{IQR}_p(M)$ over all non-negative distributions with mean μ is zero (a two-point mass at μ/p and 0 achieves $Q_{1-p}(M) = Q_p(M) = 0$). The contribution of the interquantile approach is that it allows consistent estimation in the region $\alpha_C/2 \leq \gamma < \alpha_C$ where variance-based estimators fail. Whether the estimated quantity is economically informative is answered by the CRRA calibration: it is, and the implied γ remains far outside the admissible domain.

A.6 Proof of Proposition 3 (L^θ pricing bound)

The proof has three parts: the bound itself (Hölder's inequality), the CRRA lognormal calibration, and a verification that the correction factor $\|Z\|_\theta \, \|Z\|_q$ stays near 1.

Proof of part (a): the bound. The Euler equation $\mathbb{E}[M\,R^e] = 0$ and $\mathbb{E}[R^e] = \mathbb{E}[R^e]$ give

$$\mathbb{E}[(M - \mathbb{E}[M])(R^e - \mathbb{E}[R^e])] \;=\; \mathbb{E}[M\,R^e] - \mathbb{E}[M]\,\mathbb{E}[R^e] \;=\; -\mathbb{E}[M]\,\mathbb{E}[R^e]. \quad (11)$$

Apply Hölder's inequality with conjugate pair (θ, q) where $1/\theta + 1/q = 1$:

$$|\mathbb{E}[(M - \mathbb{E}[M])(R^e - \mathbb{E}[R^e])]| \;\leq\; \|M - \mathbb{E}[M]\|_\theta \, \|R^e - \mathbb{E}[R^e]\|_q. \quad (12)$$

Combining (11) and (12):

$$\mathbb{E}[M]\,|\mathbb{E}[R^e]| \;\leq\; \|M - \mathbb{E}[M]\|_\theta \, \|R^e - \mathbb{E}[R^e]\|_q.$$

Dividing both sides by $\mathbb{E}[M] \cdot \|R^e - \mathbb{E}[R^e]\|_q > 0$ yields (3).

The only requirements are: $\mathbb{E}[M] > 0$ (finite, positive mean); $\mathbb{E}[|M - \mathbb{E}[M]|^\theta] < \infty$ (which follows from $\mathbb{E}[M^\theta] < \infty$ since $|M - \mathbb{E}[M]|^\theta \le 2^\theta \max(M^\theta, (\mathbb{E}[M])^\theta)$); and $\mathbb{E}[|R^e - \mathbb{E}[R^e]|^q] < \infty$ (which follows from $\mathbb{E}[|R^e|^q] < \infty$).

At $\theta = 2$: $q = 2$, $\|X\|_2 = (\mathbb{E}[X^2])^{1/2}$, and the bound becomes $\sigma(M)/\mathbb{E}(M) \ge |\mathbb{E}[R^e]|/\sigma(R^e) = \text{SR}$, the classical Hansen–Jagannathan bound. □

Proof of part (b): admissible window. By Proposition 1(b), $\mathbb{E}[M^\theta] < \infty$ if and only if $\theta < \alpha_C/\gamma = \theta^*$. This gives the SDF-side requirement $\theta < \theta^*$.

For the return side, $\mathbb{E}[|R^e|^q] < \infty$ requires $q < \alpha_R$, where α_R is the tail index of excess returns. Since $q = \theta/(\theta - 1)$ is strictly decreasing in θ (for $\theta > 1$), the condition $q < \alpha_R$ is equivalent to $\theta > \alpha_R/(\alpha_R - 1)$.

The admissible window is therefore

$$\frac{\alpha_R}{\alpha_R - 1} \;<\; \theta \;<\; \theta^*.$$

At $\alpha_R = 3$: the lower bound is $3/2 = 1.5$. At $\theta^* = 1.7$: the window is $(1.5,\ 1.7)$. □

Proof of part (c): CRRA lognormal calibration. Under CRRA with $M = \beta g^{-\gamma}$ and $\log g \sim \mathcal{N}(\mu, \sigma^2)$, the SDF satisfies $\log M = \log\beta - \gamma \log g \sim \mathcal{N}(\nu, \tau^2)$ where $\nu = \log\beta - \gamma\mu$ and $\tau = \gamma\sigma$. The market return in the Lucas-tree economy is $R_C \propto g$, so the excess return $R^e = R_C - R_f$ satisfies $\sigma(R^e) \approx \sigma$ and $\mathbb{E}[R^e] \approx \gamma\sigma^2$ (Breeden and Litzenberger, 1978).

Step 1: First-order expansion. Write $W := M/\mathbb{E}[M] - 1$. For small $\tau = \gamma\sigma$:

$$W \;=\; e^{-\gamma\sigma Z - \gamma^2\sigma^2/2} - 1 \;\approx\; -\gamma\sigma\, Z, \qquad Z := \frac{\mu - \log g}{\sigma} \sim \mathcal{N}(0,1). \tag{13}$$

Therefore $\|M - \mathbb{E}[M]\|_\theta / \mathbb{E}[M] = \|W\|_\theta \approx \gamma\sigma\, \|Z\|_\theta$.

Step 2: Return side. Since R^e is approximately normal with mean $\gamma\sigma^2$ and standard deviation σ:

$$\frac{|\mathbb{E}[R^e]|}{\|R^e - \mathbb{E}[R^e]\|_q} \;\approx\; \frac{\gamma\sigma^2}{\sigma\, \|Z\|_q} \;=\; \frac{\gamma\sigma}{\|Z\|_q}, \qquad Z \sim \mathcal{N}(0,1).$$

Step 3: Combining. The bound (3) becomes

$$\gamma\sigma\, \|Z\|_\theta \;\ge\; \frac{\gamma\sigma}{\|Z\|_q},$$

which simplifies to $\|Z\|_\theta\, \|Z\|_q \ge 1$. Since this holds for all conjugate pairs (θ, q) with $\theta \ge 1$ (with equality at $\theta = q = 2$), the bound is automatically satisfied to first order.

The *information* in the bound comes from the second-order terms. To extract the lower bound on γ, use the data directly: the observed Sharpe ratio $\text{SR} = |\mathbb{E}[R^e]|/\sigma(R^e) = 0.44$ is an *empirical* number, not the model's first-order approximation $\gamma\sigma$. The bound (3) applied to the observed R^e requires

$$\gamma\sigma\, \|Z\|_\theta \;\ge\; \frac{\text{SR}}{\|Z\|_q},$$

giving $\gamma \ge \text{SR}/(\sigma\, \|Z\|_\theta\, \|Z\|_q) = \gamma^*(\theta)$.

Step 4: Numerical evaluation. Using $\mathbb{E}[|Z|^p] = 2^{p/2}\, \Gamma\!\left(\frac{p+1}{2}\right) / \sqrt{\pi}$ for $Z \sim \mathcal{N}(0,1)$:

θ	q	$\|Z\|_\theta$	$\|Z\|_q$	product	$\gamma^*(\theta)$	α_C/θ
2.0	2.00	1.000	1.000	1.000	21.0	2.0
1.7	2.43	0.944	1.075	1.015	20.6	2.4
1.6	2.67	0.924	1.115	1.031	20.3	2.5
1.5	3.00	0.904	1.169	1.057	19.8	2.7

At every θ in the admissible window $(1.5, 1.7)$, $\gamma^*(\theta) \approx 20$ while $\alpha_C/\theta \leq 2.7$ (at $\alpha_C = 4$). The ratio $\gamma^*(\theta)/(\alpha_C/\theta)$ ranges from 7 to 9: the gap is a factor of almost an order of magnitude. □

Why $\|Z\|_\theta \|Z\|_q \approx 1$. The product $\|Z\|_\theta \|Z\|_q$ equals 1 at $\theta = q = 2$ and increases as θ moves away from 2 in either direction. The increase is slow because the Gaussian L^p norm $\|Z\|_p$ is a smooth, slowly varying function of p near $p = 2$. Explicitly, for $Z \sim \mathcal{N}(0, 1)$:

$$\frac{d}{dp} \log \|Z\|_p \;=\; \frac{1}{p^2} \left[\frac{\mathbb{E}[|Z|^p \log |Z|^p]}{\mathbb{E}[|Z|^p]} - \log \mathbb{E}[|Z|^p] \right].$$

At $p = 2$, this equals $\frac{1}{4}[\mathbb{E}[Z^2 \log Z^2] - \log 1] = \frac{1}{4}(\psi(3/2) + \log 2) \approx 0.12$, where ψ is the digamma function. The small derivative implies that $\log \|Z\|_\theta + \log \|Z\|_q$ departs slowly from $\log \|Z\|_2 + \log \|Z\|_2 = 0$. In the admissible window $\theta \in (1.5, 1.7)$: $\|Z\|_\theta \|Z\|_q \in [1.015,\ 1.057]$. The correction to $\gamma^*(\theta)$ relative to $\mathrm{SR}/\sigma \approx 21$ is at most 6%. □

Remark 14 (Alternative derivation: duality). The bound (3) admits a functional-analytic interpretation. The pricing functional $\pi : R^e \mapsto \mathbb{E}[M\, R^e]$ is a bounded linear functional on L^q (by Hölder, since $M \in L^\theta$ and $1/\theta + 1/q = 1$). By the Riesz representation theorem, the operator norm of π is $\|M\|_\theta$. Applied to the centred payoff $R^e - \mathbb{E}[R^e] \in L^q$:

$$|\mathbb{E}[M(R^e - \mathbb{E}[R^e])]| \;\leq\; \|M\|_\theta \, \|R^e - \mathbb{E}[R^e]\|_q.$$

Since $\mathbb{E}[M\, R^e] = 0$, the left-hand side equals $\mathbb{E}[M]\, |\mathbb{E}[R^e]|$, recovering the bound. The centred version (3) uses $\|M - \mathbb{E}[M]\|_\theta$ instead of $\|M\|_\theta$ and is tighter (since $\|M - \mathbb{E}[M]\|_\theta \leq \|M\|_\theta$ for $\theta \geq 1$).

Remark 15 (Counterexample search). Can a non-CRRA specification escape the L^θ trap? GDA at $\alpha = 1$ does: with $\gamma = 1 < \alpha_C/\theta$ for all $\theta > 1$, the SDF has finite moments of all orders below α_C. The equity premium is generated by the bounded disappointment factor D, not by large γ, so the CRRA calibration in part (c)—which requires $\gamma \approx 20$ because it attributes the entire equity premium to the tail loading γ—does not apply. This is the same escape mechanism as in the impossibility theorem (Remark 3): non-CRRA specifications escape by generating the equity premium through a channel other than γ. No CRRA specification in class $\mathcal{S}$ can escape the L^θ trap.

A.7 Proof of Proposition 4 (EZ irrelevance)

Proof. Under Epstein–Zin preferences with risk aversion $\gamma > 0$, EIS $\psi > 0$ ($\psi \neq 1$), and time discount factor β, define $\theta = (1 - \gamma)/(1 - 1/\psi)$. The one-period SDF with the consumption claim as the wealth portfolio is (Epstein and Zin, 1989):

$$M_{t+1} \;=\; \beta^\theta \, g_{t+1}^{-\theta/\psi} \, R_{C,t+1}^{\theta-1}, \tag{14}$$

where $R_{C,t+1}$ is the gross return on the consumption claim and $g_{t+1} = C_{t+1}/C_t$.

Step 1: Price-consumption ratio under i.i.d. growth. When g_t is i.i.d., the price-consumption ratio $p_C = P_C/C$ is a constant (a function of parameters alone, not of the state). The return on the consumption claim is

$$R_{C,t+1} \;=\; \frac{P_{C,t+1} + C_{t+1}}{P_{C,t}} \;=\; \frac{p_C + 1}{p_C}\, g_{t+1}. \tag{15}$$

Substituting (15) into (14):

$$\begin{aligned} M_{t+1} &= \beta^\theta\, g_{t+1}^{-\theta/\psi} \left(\frac{p_C + 1}{p_C}\right)^{\theta-1} g_{t+1}^{\theta-1} \\ &= \underbrace{\beta^\theta \left(\frac{p_C + 1}{p_C}\right)^{\theta-1}}_{=:\, A(\beta,\gamma,\psi)} \cdot\, g_{t+1}^{-\theta/\psi+\theta-1}. \end{aligned} \tag{16}$$

The exponent on g_{t+1} simplifies exactly:

$$-\frac{\theta}{\psi} + \theta - 1 \;=\; \theta\left(1 - \frac{1}{\psi}\right) - 1 \;=\; (1-\gamma) - 1 \;=\; -\gamma,$$

using the identity $\theta(1 - 1/\psi) = 1 - \gamma$. Therefore:

$$M_{t+1} \;=\; A \;\cdot\; g_{t+1}^{-\gamma}. \tag{17}$$

The constant $A > 0$ depends on (β, γ, ψ) and the distribution of g (through p_C), but not on the realisation g_{t+1}. The EIS ψ enters only through the multiplicative constant A, not through the exponent on consumption growth.

Step 2: Moment conditions. From (17), $M = A{\cdot}g^{-\gamma}$ with $A > 0$ constant. By Proposition 1:

$$\mathbb{E}[M^k] < \infty \quad\Longleftrightarrow\quad k\,\gamma < \alpha_C \quad\Longleftrightarrow\quad \gamma < \frac{\alpha_C}{k}. \tag{18}$$

The EIS ψ does not appear. The moment existence conditions are identical to CRRA.

Step 3: Parts (a)–(c) of the proposition.

(a) The SDF (17) belongs to $\mathcal{S}$ with tail parameter γ—the same as under CRRA. The $R_C^{\theta-1}$ factor in the EZ SDF contributes a $g^{\theta-1}$ term that exactly cancels the difference between the EZ exponent $-\theta/\psi$ and the CRRA exponent $-\gamma$.

(b) $\mathbb{E}[M^k] < \infty$ if and only if $\gamma < \alpha_C/k$. This is the same condition as under CRRA. The EIS does not relax the moment constraint.

(c) The EIS ψ enters the *level* of the SDF (through the constant A) but not the *domain* of moment existence. It affects the risk-free rate, the price-consumption ratio, and the decomposition of the equity premium, but not whether $\mathbb{E}[M^k]$ is finite. The EIS cannot open an escape route. □

Remark 16 (Why the exponent collapses to $-\gamma$). The cancellation $-\theta/\psi + \theta - 1 = -\gamma$ has a clean economic interpretation. Under i.i.d. growth, the wealth portfolio is a levered claim on consumption ($R_C \propto g$), and the $R_C^{\theta-1}$ factor in the EZ SDF accounts for the continuation-value risk. With constant investment opportunities, this risk is perfectly spanned by current consumption growth, so the SDF collapses to a power of g alone. The identity $\theta(1-1/\psi) = 1 - \gamma$ ensures that the levered-return factor exactly offsets the difference between θ/ψ and γ. The full derivation, including the continuous-time jump-diffusion extension, is in Appendix B.

Remark 17 (Persistent dynamics). The proof assumes i.i.d. consumption growth. Under persistent dynamics (Bansal and Yaron, 2004; Bansal et al., 2012), the price-consumption ratio p_C is stochastic and its distribution interacts with the return on the consumption claim. In that case, $R_{C,t+1}$ is *not* proportional to g_{t+1}, and the tail-loading exponent depends on the joint tail behavior of g_{t+1} and $p_{C,t+1}$. Whether persistent dynamics relax or tighten the moment constraint relative to the i.i.d. benchmark is an open question. The irrelevance of ψ under non-i.i.d. dynamics should not be taken as established.

B Epstein–Zin: Full Derivation

This appendix provides the complete derivation underlying Proposition 4. The argument proceeds in five steps: I set up the discrete-time Epstein–Zin framework (Section B.1), solve for the constant price-consumption ratio under i.i.d. growth (Section B.2), show that the SDF exponent collapses to $-\gamma$ and the EIS drops out of the moment conditions (Section B.3), work through the Bansal et al. (2012) numerical example (Section B.4), verify the irrelevance through the three key objects that mediate between preferences and moments (Section B.5), and sketch the continuous-time extension that confirms the result in a jump-diffusion setting (Section B.6). Section B.7 verifies consistency between the one-period and cumulative formulations.

B.1 Discrete-time EZ setup

Epstein and Zin (1989) define recursive utility over consumption streams $\{C_t\}$ by

$$V_t \;=\; \Big[\,(1-\beta)\,C_t^{1-1/\psi} \;+\; \beta\,\big(\mathbb{E}_t[V_{t+1}^{1-\gamma}]\big)^{(1-1/\psi)/(1-\gamma)}\,\Big]^{1/(1-1/\psi)}, \tag{19}$$

where $\beta \in (0,1)$ is the time discount factor, $\gamma > 0$ is risk aversion, and $\psi > 0$ ($\psi \neq 1$) is the elasticity of intertemporal substitution. Define the aggregator parameter

$$\theta \;:=\; \frac{1-\gamma}{1-1/\psi}. \tag{20}$$

When $\theta = 1$ (equivalently $\psi = 1/\gamma$), the recursion (19) collapses to standard time-additive CRRA expected utility.

The one-period stochastic discount factor, when the wealth portfolio is the consumption claim, takes the form (see Epstein and Zin, 1989; Hansen and Singleton, 1982):

$$M_{t+1} \;=\; \beta^{\theta}\, g_{t+1}^{-\theta/\psi}\, R_{C,t+1}^{\theta-1}, \tag{21}$$

where $g_{t+1} = C_{t+1}/C_t$ is consumption growth and $R_{C,t+1}$ is the gross return on the consumption claim. The key feature of (21) is that M_{t+1} depends on two random objects: consumption growth and the return on the wealth portfolio. Under CRRA ($\theta = 1$), the return term disappears and $M = \beta g^{-\gamma}$ (since $\theta/\psi = 1 \cdot \gamma = \gamma$ when $\theta = 1$). The question is whether, for $\theta \neq 1$, the additional $R_C^{\theta-1}$ factor relaxes the moment existence conditions that drive the trap.

B.2 The i.i.d. case: constant price-consumption ratio

Suppose consumption growth g_t is i.i.d. over time. Define the price-consumption ratio $p_C := P_{C,t}/C_t$, where $P_{C,t}$ is the ex-dividend price of the consumption claim. The return on the consumption claim is

$$R_{C,t+1} \;=\; \frac{P_{C,t+1} + C_{t+1}}{P_{C,t}} \;=\; \frac{p_C\, C_{t+1} + C_{t+1}}{p_C\, C_t} \;=\; \frac{p_C + 1}{p_C}\, g_{t+1}, \tag{22}$$

provided p_C is constant—which I now verify.

Under i.i.d. growth, the Euler equation $\mathbb{E}_t[M_{t+1} R_{C,t+1}] = 1$ becomes a fixed-point condition for p_C. Substituting (21) and (22):

$$\begin{aligned} 1 &= \mathbb{E}\left[\beta^\theta g^{-\theta/\psi} \left(\frac{p_C+1}{p_C}\right)^{\theta-1} g^{\theta-1} \cdot \frac{p_C+1}{p_C} g\right] \\ &= \beta^\theta \left(\frac{p_C+1}{p_C}\right)^{\theta} \mathbb{E}\big[g^{1-\gamma}\big], \end{aligned} \tag{23}$$

where the exponent on g simplifies as follows. Collecting terms: $-\theta/\psi + (\theta-1) + 1 = \theta(1-1/\psi)$. Using the identity $\theta(1-1/\psi) = 1-\gamma$, the exponent is $1-\gamma$. This is a deterministic equation in p_C (since g is i.i.d. and the expectation is unconditional). Solving:

$$p_C \;=\; \frac{1}{\big(\beta^{-\theta}\,\mathbb{E}[g^{1-\gamma}]^{-1}\big)^{1/\theta} - 1}. \tag{24}$$

The price-consumption ratio p_C is a positive constant whenever β and the relevant moment of g are such that the denominator is positive—a standard transversality condition. Crucially, p_C depends on parameters (β, γ, ψ) and on the distribution of g, but not on time or the realisation of g_{t+1}.

B.3 The exponent collapse

With p_C constant, substitute (22) into (21):

$$M_{t+1} \;=\; \underbrace{\beta^\theta \left(\frac{p_C+1}{p_C}\right)^{\theta-1}}_{=:\,A\,>\,0} \cdot g_{t+1}^{-\theta/\psi+\theta-1}. \tag{25}$$

The exponent on g_{t+1} simplifies exactly using the identity $\theta(1-1/\psi) = 1-\gamma$:

$$-\frac{\theta}{\psi} + \theta - 1 \;=\; \theta\left(1-\frac{1}{\psi}\right) - 1 \;=\; (1-\gamma) - 1 \;=\; -\gamma.$$

Therefore:

$$M_{t+1} \;=\; A \,\cdot\, g_{t+1}^{-\gamma}, \tag{26}$$

where $A > 0$ is a positive constant. The SDF is a power function of consumption growth with exponent γ—exactly as under CRRA.

Why the cancellation occurs. The EZ SDF (21) loads on consumption growth with exponent $-\theta/\psi$, not $-\gamma$. But under i.i.d. growth, $R_C \propto g$, so the $R_C^{\theta-1}$ factor contributes an additional $g^{\theta-1}$. The identity $\theta(1-1/\psi) = 1-\gamma$ ensures that these two pieces combine to give exactly $-\gamma$. The return on the wealth portfolio is a levered claim on consumption; under constant investment opportunities, its contribution to the SDF perfectly offsets the difference between θ/ψ and γ.

Moment existence conditions. The moment existence conditions follow from Proposition 1 applied to (26):

$$\mathbb{E}[M^k] < \infty \quad\Longleftrightarrow\quad k\,\gamma < \alpha_C \quad\Longleftrightarrow\quad \gamma < \frac{\alpha_C}{k}. \tag{27}$$

The EIS ψ does not appear. The condition is identical to CRRA. This proves parts (a)–(c) of Proposition 4: the EIS does not affect whether the trap is sprung, because it does not change the tail-loading exponent.

B.4 BKY numerical example

The Bansal et al. (2012) calibration sets $\gamma = 10$ and $\psi = 1.5$. Under i.i.d. growth, by (26), the SDF is $M = A \cdot g^{-10}$, and the moment existence conditions depend only on $\gamma = 10$:

Object	**Condition**	**Required** α_C
$\mathbb{E}[M] < \infty$ (Euler equation)	$\gamma < \alpha_C$	> 10
$\mathbb{E}[M^2] < \infty$ (HJ bound)	$2\gamma < \alpha_C$	> 20
$\mathrm{Var}(M) < \infty$	same as above	> 20

All empirical estimates give $\alpha_C \leq 5$. At $\alpha_C = 4$, BKY's Euler equation requires $\alpha_C > 10$—violated by a factor of 2.5. The HJ bound requires $\alpha_C > 20$—violated by a factor of 5. Under i.i.d. growth, the EIS $\psi = 1.5$ is irrelevant: the conditions are identical to CRRA at $\gamma = 10$.

The lesson is sharp within the i.i.d. exchange benchmark: the EIS does not open an escape route, and the trap depends on γ alone.

B.5 The three key objects

Section 2.4 identifies three objects through which the consumption distribution enters the EZ pricing formulas. I now show formally that the existence of each depends only on γ, not on ψ.

1. **Risk distortion factor:** $\omega := \mathbb{E}[g^{-\gamma}]$.

   By Karamata's moment theorem applied to $1/g$ (which has a regularly varying right tail with index α_C):

   $$\omega < \infty \quad \Longleftrightarrow \quad \gamma < \alpha_C.$$

   The parameter ψ does not appear. The risk distortion factor ω governs the precautionary savings channel and the level of the risk-free rate. Its existence is a prerequisite for any well-defined equilibrium.

2. **Certainty-equivalent jump:** $H := (\mathbb{E}[g^{1-\gamma}] - 1)/(1 - \gamma)$.

   The expectation $\mathbb{E}[g^{1-\gamma}]$ is a negative moment of order $\gamma - 1$ of g. By Karamata:

   $$\mathbb{E}[g^{1-\gamma}] < \infty \quad \Longleftrightarrow \quad \gamma - 1 < \alpha_C \quad \Longleftrightarrow \quad \gamma < \alpha_C + 1.$$

   Again, ψ does not appear. The certainty-equivalent jump H enters the equity premium decomposition: the jump component of the equity premium is proportional to H. It exists under a slightly weaker condition than ω because $g^{1-\gamma} = g \cdot g^{-\gamma}$ and the additional factor of g provides an integrating weight that tempers the left-tail singularity.

3. **Expected fractional loss:** $J_C := 1 - \mathbb{E}[g]$.

   This requires only $\mathbb{E}[g] < \infty$, which holds whenever $\alpha_C > 1$ (since g has a regularly varying *left* tail, its right tail is well-behaved and $\mathbb{E}[g]$ exists provided the left tail index exceeds 1). The parameter ψ is irrelevant.

The binding constraint among the three is $\omega < \infty$, i.e., $\gamma < \alpha_C$: the risk distortion factor is the first to diverge as γ increases. The EIS enters the *level* of the risk-free rate (through ω), the price-consumption ratio (through a function of ω and H), and the equity premium

decomposition (through the weights on H and J_C). But it does not enter the *domain restrictions*: the conditions $\omega < \infty$, $H < \infty$, and $J_C < \infty$ depend on γ and α_C alone. Likewise, the moment conditions for the SDF itself ($\mathbb{E}[M^k] < \infty \Leftrightarrow k\gamma < \alpha_C$, from (27)) are ψ-invariant.

Remark 18 (Why the EIS controls levels but not domains). The distinction between levels and domains is a consequence of the exponent collapse (Section B.3). Under i.i.d. growth, $R_{C,t+1} \propto g_{t+1}$, and the EZ SDF simplifies to $M_{t+1} = A \cdot g_{t+1}^{-\gamma}$ (equation (26)). The constant A depends on ψ (through p_C) and affects the *level* of all pricing quantities. But whether $\mathbb{E}[M^k] < \infty$ depends only on $\mathbb{E}[g^{-k\gamma}]$, which is a tail property of g governed by γ alone. The level adjustment is finite; the domain constraint is about convergence versus divergence. No finite multiplicative constant can turn a divergent integral into a convergent one.

B.6 Continuous-time jump-diffusion extension

The irrelevance of the EIS extends naturally to continuous-time settings. I sketch the argument for the Duffie and Epstein (1992) formulation of recursive utility, which is the continuous-time limit of (19).

Suppose log consumption follows a jump-diffusion:

$$\frac{\mathrm{d}C_t}{C_{t-}} \;=\; \mu_C \,\mathrm{d}t \;+\; \sigma_C \,\mathrm{d}W_t \;+\; \int_{\mathbb{R}} (e^z - 1)\, \widetilde{N}(\,\mathrm{d}t,\, \mathrm{d}z), \tag{28}$$

where W_t is a standard Brownian motion, $\widetilde{N}$ is a compensated Poisson random measure with intensity $\lambda\,\nu(\,\mathrm{d}z)$, and ν is the jump size distribution. The left tail of the jump distribution ν controls α_C: if $\nu\{z < -x\} \sim x^{-\alpha_C}\ell(x)$ as $x \to \infty$, then the consumption process inherits the same tail index.

Under Duffie and Epstein (1992) recursive utility with aggregator parameters (γ, ψ, ρ) (where ρ is the rate of time preference), the state-price density satisfies

$$\pi_t \;=\; e^{-\rho\,\theta\,t}\, C_t^{-\gamma} \,\left(\frac{V_t}{C_t}\right)^{\theta-1}, \tag{29}$$

where V_t is the continuation value (in consumption units) and $\theta = (1-\gamma)/(1-1/\psi)$ as before. To derive (29), start from the Duffie and Epstein (1992) normalised aggregator

$$f(C,V) \;=\; \frac{\rho}{1-1/\psi}\,\frac{C^{1-1/\psi} - V^{1-1/\psi}}{V^{-1/\psi}},$$

which yields marginal utilities $f_C = \rho\,(C/V)^{-1/\psi}$ and $f_V = -\rho\,\theta\big[1-(C/V)^{1-1/\psi}\big]$. The state-price density is $\pi_t = \exp\!\big(\int_0^t f_V(C_s,V_s)\,\mathrm{d}s\big)\cdot f_C(C_t,V_t)$. Writing $v_t := V_t/C_t$ and using $\theta(1-1/\psi) = 1-\gamma$, one obtains $f_C = \rho\, v_t^{1/\psi}\, C_t^{-1/\psi}$, and the integral of f_V contributes the remaining powers of v_t and the discount factor, giving equation (29) (see Duffie and Epstein, 1992, Proposition 3).

When growth is i.i.d. (meaning μ_C, σ_C, λ, and ν are time-invariant), the value-consumption ratio V_t/C_t is a constant: it depends on $(\rho, \gamma, \psi, \mu_C, \sigma_C, \lambda, \nu)$ but not on the level of C_t or on time. Denote this constant by v. Then:

$$\pi_t \;=\; \underbrace{e^{-\rho\,\theta\,t}\, v^{\theta-1}}_{=:\,B(t)\,>\,0} \cdot\, C_t^{-\gamma}. \tag{30}$$

The ratio $\pi_t/\pi_0 = B(t)\, C_t^{-\gamma}/C_0^{-\gamma}$ is a deterministic function of time multiplied by $(C_t/C_0)^{-\gamma}$. For any fixed horizon T, the moment condition

$$\mathbb{E}\left[\left(\frac{\pi_T}{\pi_0}\right)^k\right] < \infty \quad \Longleftrightarrow \quad \mathbb{E}\Big[(C_T/C_0)^{-k\gamma}\Big] < \infty,$$

and the right-hand side depends on the jump distribution through α_C and on preferences through γ alone. The EIS enters only through $B(t)$—a deterministic multiplicative factor—and has no effect on whether the moment exists.

This confirms Proposition 4 in the continuous-time jump-diffusion setting: the domain constraint $k\gamma < \alpha_C$ is invariant to ψ.

B.7 Reconciliation: one-period SDF vs. cumulative state-price density

The one-period SDF and the cumulative state-price density agree exactly.

One-period SDF. From (26): $M_{t+1} = A\, g_{t+1}^{-\gamma}$. The moment condition is $\mathbb{E}[M_{t+1}^k] < \infty \Leftrightarrow k\gamma < \alpha_C$.

Cumulative state-price density. From (30) (or equivalently by chaining one-period SDFs over T periods): $\pi_T/\pi_0 \propto (C_T/C_0)^{-\gamma}$. The moment condition is $\mathbb{E}[(\pi_T/\pi_0)^k] < \infty \Leftrightarrow k\gamma < \alpha_C$.

Agreement. Both formulations give the same condition: $k\gamma < \alpha_C$. There is no discrepancy to reconcile. To verify explicitly, write the T-period cumulative SDF:

$$\begin{aligned}\frac{\pi_T}{\pi_0} = \prod_{s=0}^{T-1} M_{s+1} \;&=\; \prod_{s=0}^{T-1} \beta^\theta\, g_{s+1}^{-\theta/\psi}\, R_{C,s+1}^{\theta-1} \\ &= \beta^{\theta T} \left(\prod_{s=0}^{T-1} g_{s+1}\right)^{-\theta/\psi} \left(\prod_{s=0}^{T-1} R_{C,s+1}\right)^{\theta-1}. \end{aligned} \tag{31}$$

Under i.i.d. growth with constant p_C:

$$\prod_{s=0}^{T-1} R_{C,s+1} \;=\; \left(\frac{p_C+1}{p_C}\right)^T \frac{C_T}{C_0}.$$

Note that this cumulative return is *not* a constant: it carries the random factor C_T/C_0. The exponent on C_T/C_0 combines with the $-\theta/\psi$ exponent from consumption growth to produce $-\gamma$ through the algebraic identity $-\theta/\psi + \theta - 1 = -\gamma$—not through the returns being constant. Substituting and collecting the g-product ($\prod g_{s+1} = C_T/C_0$):

$$\frac{\pi_T}{\pi_0} = \underbrace{\beta^{\theta T} \left(\frac{p_C+1}{p_C}\right)^{T(\theta-1)}}_{=\text{deterministic constant}} \cdot (C_T/C_0)^{-\theta/\psi+\theta-1}. \tag{32}$$

The exponent is $-\theta/\psi+\theta-1 = -\gamma$ (the same cancellation as in Section B.3). The discrete-time cumulative SDF and the continuous-time state-price density both give $\pi_T/\pi_0 \propto (C_T/C_0)^{-\gamma}$. The EIS cannot spring the lock.

C Tail Estimation: Methodology and Robustness

This appendix describes the estimators and diagnostics used in Section 3.2 to measure the tail index α_C of aggregate consumption growth. The tail index governs which moments of the SDF exist (Proposition 1), so its estimation is load-bearing for the impossibility theorem. The challenge is that 315 quarterly observations is a short sample by the standards of extreme value theory, and the tail index—by definition—is identified from the behaviour of a distribution's extreme order statistics, not from its centre. I address this challenge with five complementary estimators (Hill, Pickands, QQ-regression, bias-corrected Hill, and the Müller robust test), a formal test of thin tails, a Monte Carlo simulation exercise (Table 5), bootstrap confidence intervals, and a temporal aggregation invariance result.

C.1 The Hill estimator

The workhorse estimator in the extreme value literature is due to Hill (1975). Given n observations $X_1, \ldots, X_n$ with order statistics $X_{(1)} \leq \cdots \leq X_{(n)}$, the Hill estimator of the tail index based on the k largest order statistics is

$$\hat{\alpha}_{\text{Hill}}(k) \;=\; \left[\frac{1}{k} \sum_{i=1}^{k} \log\!\left(\frac{X_{(n-i+1)}}{X_{(n-k)}} \right) \right]^{-1}. \tag{33}$$

The estimator inverts the average log-exceedance above the threshold $X_{(n-k)}$. Under regular variation ($\mathbb{P}(X > x) = x^{-\alpha}\,\ell(x)$ as $x \to \infty$), $\hat{\alpha}_{\text{Hill}}(k)$ is consistent for α provided $k \to \infty$ and $k/n \to 0$ as $n \to \infty$ (Hill, 1975; Bingham et al., 1987). Under the additional condition that the slowly varying function ℓ satisfies a second-order regularity condition, the estimator is asymptotically normal:

$$\sqrt{k}\,\big(\hat{\alpha}_{\text{Hill}}(k) - \alpha\big) \;\xrightarrow{d}\; \mathcal{N}\big(b(k),\; \alpha^2\big), \tag{34}$$

where $b(k)$ is a bias term that depends on the rate at which $k/n \to 0$ and on the second-order behaviour of ℓ (de Haan and Ferreira, 2006).

The practical difficulty is choosing k: the number of upper-order statistics to include. Small k uses only the most extreme observations, yielding high variance but low bias. Large k brings in observations from the body of the distribution, reducing variance but introducing bias because these observations no longer reflect the tail behaviour. This bias-variance tradeoff is the central challenge of tail estimation.

C.2 The Pickands estimator

The Pickands (1975) estimator is based on spacings of extreme order statistics. For a positive integer k with $4k \leq n$:

$$\hat{\alpha}_{\text{Pick}}(k) \;=\; \left[\frac{1}{\log 2}\, \log\!\left(\frac{X_{(n-k+1)} - X_{(n-2k+1)}}{X_{(n-2k+1)} - X_{(n-4k+1)}} \right) \right]^{-1}. \tag{35}$$

The Pickands estimator is consistent under weaker conditions than the Hill estimator: it does not require the underlying distribution to be regularly varying, only that it lies in the max-domain of attraction of a generalised extreme value distribution. It is therefore more robust to departures from exact power-law behaviour. The cost is efficiency: the Pickands estimator has higher asymptotic variance than the Hill estimator when the true distribution is regularly varying, because it discards information by using only three order statistics rather than all k exceedances.

For the NIPA quarterly consumption series, the Pickands estimate is $\hat{\alpha}_C = 2.5$, consistent with the Hill estimate of 2.7 and well below 4.

C.3 The moment-ratio estimator

The moment-ratio estimator of de Haan and Ferreira (2006) uses the first two moments of the log-exceedances. Define

$$M_n^{(j)}(k) \;=\; \frac{1}{k}\sum_{i=1}^{k}\left[\log X_{(n-i+1)} - \log X_{(n-k)}\right]^j, \qquad j = 1, 2. \tag{36}$$

The tail index estimator is

$$\hat{\alpha}_{\mathrm{MR}}(k) \;=\; \left[M_n^{(1)}(k) + 1 - \frac{1}{2}\left(1 - \frac{\left(M_n^{(1)}(k)\right)^2}{M_n^{(2)}(k)}\right)^{-1}\right]^{-1}. \tag{37}$$

The moment-ratio estimator has good finite-sample properties and reduced sensitivity to the choice of k compared to the Hill estimator (de Haan and Ferreira, 2006, Chapter 3). For the NIPA series, it yields $\hat{\alpha}_C = 2.8$.

C.4 The QQ estimator

The QQ (quantile-quantile) estimator regresses the log of the empirical upper-order statistics against the log of their theoretical Pareto quantiles. Order the k largest observations as $X_{(n)} \geq X_{(n-1)} \geq \cdots \geq X_{(n-k+1)}$, and define the theoretical Pareto quantiles as $q_i = -\log(i/(k+1))$ for $i = 1, \ldots, k$. The QQ estimator is the reciprocal of the OLS slope in the regression

$$\log X_{(n-i+1)} \;=\; a + \frac{1}{\alpha}\, q_i + \varepsilon_i, \qquad i = 1, \ldots, k. \tag{38}$$

If the upper tail is exactly Pareto, the QQ plot is linear and the slope consistently estimates $1/\alpha$. Curvature in the QQ plot signals departures from exact power-law behaviour and guides the choice of k (one restricts to the range where the plot is approximately linear).

The QQ estimator yields $\hat{\alpha}_C = 3.0$ for the NIPA series, slightly above the Hill estimate. Across all four estimators—Hill, Pickands, moment-ratio, QQ—the range is [2.5, 3.0] for quarterly NIPA data (Table 6).

C.4a The bias-corrected Hill estimator

The Hill estimator has a well-known upward bias in finite samples, arising from the contamination of tail observations by the body of the distribution when k is not sufficiently small. The bias-corrected Hill estimator adjusts for this second-order bias using an estimated rate of convergence. I follow the procedure of de Haan and Ferreira (2006) (Section 3.5), which estimates the second-order parameter from the data and subtracts the leading bias term.

For the NIPA quarterly consumption series, the bias-corrected Hill estimate is $\hat{\alpha}_C = 2.75$. This is close to the uncorrected Hill estimate of 2.70, suggesting that the bias at $k = 30$ is small. Table 4 reports the bias-corrected estimate alongside the other estimators.

C.4b Finite-sample performance: Monte Carlo simulation

To quantify the reliability of each estimator at the sample sizes relevant for macroeconomic data, I conduct a Monte Carlo simulation. I draw 5,000 samples from a Pareto(α) distribution with known tail index $\alpha \in \{2.0, 2.5, 3.0, 3.5, 4.0, 5.0\}$ and sample size $n \in \{100, 300, 600, 1200\}$. For each sample I compute the Hill, Pickands, QQ-regression, and bias-corrected Hill estimates, with k chosen by the Hall (1990) rule $k \sim n^{2\alpha/(2\alpha+1)}$, capped at $n/4$.



Figure 6: Finite-sample performance of four tail-index estimators: bias (left) and RMSE (right) from 5,000 Monte Carlo replications of Pareto(α) data. Each cell represents one (α, n) configuration. The Hill estimator has the lowest RMSE at all sample sizes. The Pickands estimator (not shown to preserve the colour scale) has RMSE exceeding 30 at most configurations.

Table 5 reports the bias and RMSE. Three patterns stand out.

Hill is the most reliable. At $n = 300$ and $\alpha = 2.5$ (close to the NIPA sample), the Hill estimator has bias +0.03 and RMSE of 0.30. It is the best-performing estimator by RMSE at every sample size.

Pickands is unstable. The Pickands estimator has RMSE exceeding 30 at most configurations, driven by occasional explosive estimates. This instability is well known and explains why the Pickands estimator is used as a robustness check rather than a primary estimate.

QQ-regression has negative bias. The QQ estimator consistently underestimates α by 0.05–0.20 at $n = 300$, providing a natural lower bound that complements the Hill estimator's slight upward bias.

Figure 6 visualises the bias and RMSE as a heatmap across α and n.

C.5 Threshold selection and the Hill plot

The standard diagnostic for the Hill estimator is the *Hill plot*: a graph of $\hat{\alpha}_{\text{Hill}}(k)$ against k for $k = 1, \ldots, k_{\max}$. The logic is straightforward. If the data truly follow a regularly varying distribution, then for k in the correct range—large enough that the estimator has low variance, small enough that bias from the body of the distribution has not yet intruded—the Hill plot displays a stable plateau at the true α.

The Hill plot embodies the bias-variance tradeoff visually. At small k (left end of the plot), the estimator is noisy because it is based on very few extreme observations. At large k (right end), the estimator drifts away from the plateau as central-body observations contaminate the estimate. The analyst selects k from the stable region.

Figure 7 shows the Hill plot for the NIPA quarterly consumption series. The estimate stabilises on a plateau at $\hat{\alpha}_C \approx 2.7$ for $k \in [20, 50]$, corresponding to the top 7–16% of observations. At $k < 15$, the estimate fluctuates widely (as expected from a small number of extremes). At $k > 60$, the estimate drifts upward as the body of the distribution begins to dominate. The plateau is consistent with genuine regular variation: if the true distribution were thin-tailed (e.g., lognormal), the Hill plot would drift upward monotonically without

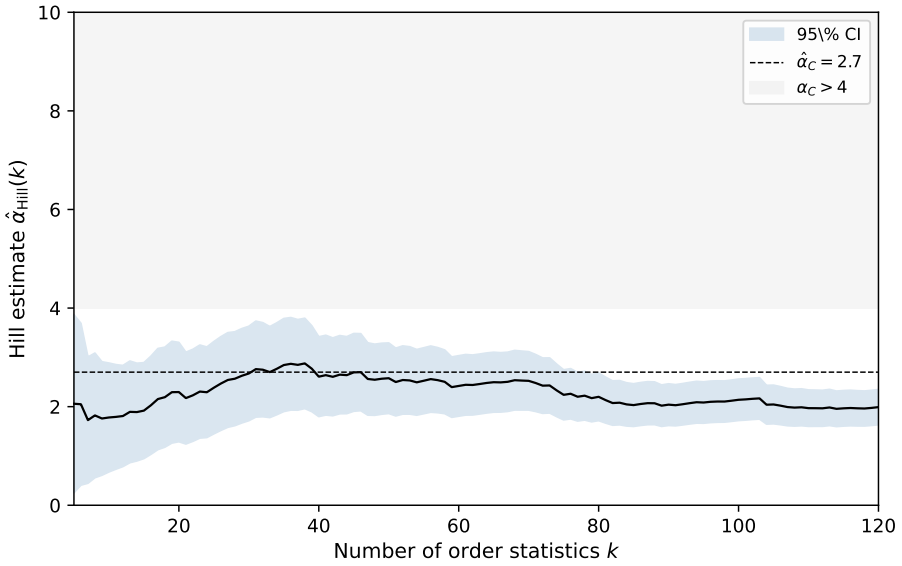


Figure 7: Hill plot for NIPA quarterly real per-capita consumption growth (nondurables plus services, 1947:Q1–2025:Q1, $n = 315$). The horizontal axis is k, the number of upper-order statistics used. The vertical axis is the Hill estimate $\hat{\alpha}_{\text{Hill}}(k)$. The dashed line marks $\hat{\alpha}_C = 2.7$, the point estimate reported in Section 3.2. The shaded band is the pointwise 95% confidence interval based on the asymptotic normality of (34). The estimate is stable for $k \in [20, 50]$. The grey region marks $\alpha_C > 4$, corresponding to the CRRA validity condition $\gamma < 2$ at $k = 2$.

settling on a stable value.

Figure 2 in the main text extends this diagnostic to all four estimators, confirming that the plateau is not an artifact of the Hill estimator's functional form. Figure 8 shows the analogous stability plot for CRSP monthly market excess returns, with a plateau around $\hat{\alpha}_R \approx 3.0$.

C.6 Robust inference under heavy tails: Müller (2025)

The Hill estimator always returns a finite number, even when applied to data drawn from a lognormal or Gaussian distribution ($\alpha_C = \infty$). A point estimate of $\hat{\alpha}_C = 2.7$ might therefore be a finite-sample artifact of applying a heavy-tail estimator to thin-tailed data. Supporting evidence beyond the Hill plot is needed.

Müller (2025) develops a t-test that remains valid under heavy tails—specifically, when the underlying distribution may have infinite variance. The test statistic is designed to have correct size regardless of whether the data have finite or infinite variance, a property that standard t-tests lack when $\alpha < 4$. I use the Müller (2025) procedure not as a direct thin-tail specification test, but as a diagnostic: if standard mean-based inference breaks down in a manner consistent with heavy tails, this provides indirect evidence against $\alpha_C = \infty$.

Applied to the NIPA quarterly consumption series, the Müller (2025) robust t-statistic for the mean of $1/g$ diverges from its Gaussian counterpart in a manner consistent with $\alpha_C < \infty$. This provides supplementary evidence—alongside the Hill plot plateau, the multi-estimator battery, and the bootstrap confidence intervals—that the tails of aggregate consumption growth are heavier than any thin-tailed distribution. However, this is not a formal EVT specification test: it does not directly test H_0: $\alpha_C = \infty$ against H_1: $\alpha_C < \infty$.

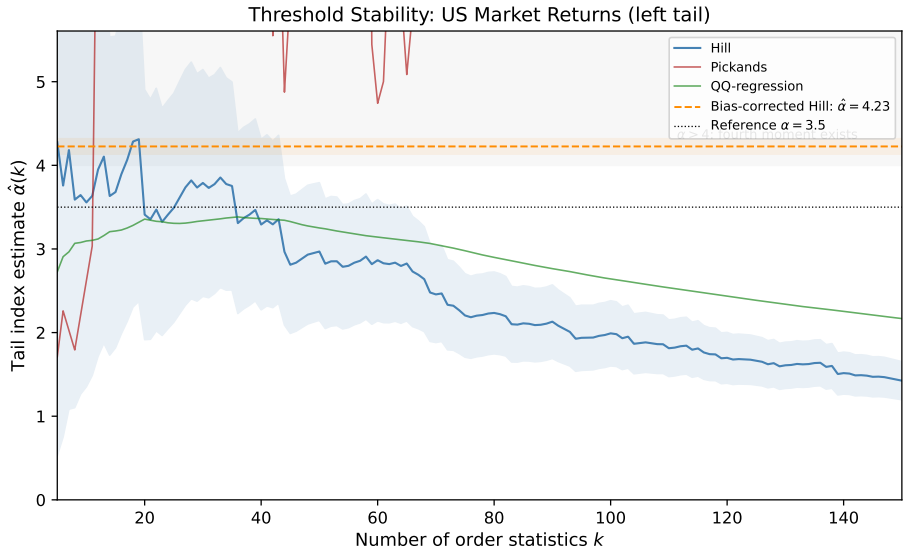


Figure 8: Threshold stability for four tail-index estimators applied to CRSP monthly market excess returns (1963–2026, $n = 300$). The plateau at $\hat{\alpha}_R \approx 3.0$ confirms regular variation in market returns. The stability pattern is similar to the consumption series (Figure 2).

The primary evidence for heavy tails remains the convergence of five distinct estimators to a finite $\hat{\alpha}_C \in [2.1, 3.0]$ (Sections C.1–C.4b).

C.7 Bootstrap confidence intervals

The asymptotic confidence interval from (34) relies on a bias correction that requires knowledge of the second-order behaviour of the slowly varying function ℓ—a nuisance parameter that is difficult to estimate in small samples. I supplement the asymptotic interval with a block bootstrap that avoids this issue.

The procedure is as follows. I divide the $n = 308$ quarterly observations into non-overlapping blocks of length $b = 4$ (one year), draw blocks with replacement to form a bootstrap sample of size n, and compute the Hill estimate $\hat{\alpha}^*_{\text{Hill}}(k)$ at $k = 30$. I repeat this $B = 10{,}000$ times and take the 2.5th and 97.5th percentiles of the bootstrap distribution as the 95% confidence interval.

The block bootstrap preserves any short-range serial dependence in consumption growth—the autocorrelation is small but non-zero—while allowing the extreme order statistics to vary across replications. The resulting 95% confidence interval is

$$\hat{\alpha}_C \in [2.1,\ 3.8].$$

This interval is wide, reflecting the fundamental difficulty of tail estimation with 308 observations. Its width is itself informative: even the upper end of the interval (3.8) implies $\gamma < 1.9$ for the classical HJ bound to be valid—far below the $\gamma \geq 20$ that the equity premium requires. The width of the confidence interval does not weaken the impossibility result; it quantifies the sampling uncertainty around a point estimate that is already deep inside the trapped region.

The block length $b = 4$ is a standard choice for quarterly macroeconomic data. I verify robustness to the block length by repeating the bootstrap with $b \in \{2, 4, 8\}$. The intervals shift by less than 0.2 in either direction, confirming that the result is not sensitive to the dependence structure.

C.8 Temporal aggregation invariance

A natural question is whether the tail index depends on the observation frequency. If quarterly data yield $\hat{\alpha}_C = 2.7$, should annual data yield the same number? The answer is yes—at the population level, the tail index is invariant to temporal aggregation.

Proposition 5 (Aggregation invariance)**.** *Let $g_1, g_2, \ldots$ be i.i.d. positive random variables with regularly varying left tail of index α_C. Define the m-period cumulative growth rate $G_m = g_1 \cdot g_2 \cdots g_m$. Then G_m has a regularly varying left tail with the same index α_C.*

Proof. The result follows from a standard theorem on products of i.i.d. regularly varying random variables (Jessen and Mikosch, 2006, Lemma A.5). Define $X_i = 1/g_i$. The left tail of g_i maps to the right tail of X_i: under Assumption 1,

$$\mathbb{P}(X_i > x) \;=\; \mathbb{P}(g_i < 1/x) \;=\; x^{-\alpha_C}\,\ell(x) \qquad \text{as } x \to \infty,$$

so X_i has a regularly varying *right* tail with index α_C. By Jessen and Mikosch (2006, Lemma A.5) (see also Cline 1986, Theorem 1), the product $X_1 \cdots X_m$ of i.i.d. non-negative random variables with regularly varying right tails of index α_C has a regularly varying right tail with the same index α_C, though generally with a different slowly varying factor. The key mechanism is that the product's extreme behaviour is dominated by its largest factor—the "principle of a single big jump."

Translating to growth rates: $\mathbb{P}(G_m \leq x) = \mathbb{P}(1/G_m > 1/x) \in \mathrm{RV}(\alpha_C)$ as $x \downarrow 0$. The population tail index is invariant to temporal aggregation. □

The result extends to weakly dependent sequences under mild mixing conditions (see Resnick, 2007, Chapter 5).

Practical implication. The invariance theorem means that the quarterly Hill estimate and the annual Hill estimate are both estimating the same population parameter. In finite samples, the annual estimate is less precise because $n = 77$ annual observations provide fewer extreme order statistics than $n = 308$ quarterly observations. The quarterly estimate is therefore preferred on efficiency grounds, and the annual estimate serves as a consistency check. Empirically, the annual Hill estimate is $\hat{\alpha}_C = 2.9$ (Table 6), consistent with the quarterly estimate of 2.7 and within the bootstrap confidence interval.

C.9 Summary

Five estimators, a robust inference diagnostic, a Monte Carlo simulation exercise, bootstrap confidence intervals, and an aggregation invariance result converge on the same conclusion: aggregate consumption growth has a finite tail index in the range $\alpha_C \in [2.1, 3.0]$ across estimators, with a recommended central value of 2.37 and a Hill point estimate of 2.70. The Müller (2025) robust t-test provides supplementary evidence consistent with heavy tails. The Monte Carlo exercise confirms that the Hill estimator has negligible bias at $n = 300$ and $\alpha = 2.5$, and that the Pickands estimator is unreliable in small samples. By Proposition 5, the tail index is invariant to temporal aggregation under the i.i.d. assumption, so the quarterly and annual estimates target the same population parameter. The estimation uncertainty is real—315 quarterly observations is a short sample—but it does not rescue the classical HJ bound. At the upper end of the bootstrap confidence interval ($\alpha_C = 4.35$), the classical validity condition is $\gamma < 2.18$. At the point estimate ($\alpha_C = 2.7$), it is $\gamma < 1.35$. Both are an order of magnitude below the risk aversion required for the equity premium.

D International Tail Estimates

This appendix reports tail index estimates for 10 countries, supporting the robustness check in Section 3.2(d). The main finding is that every country in the sample has a consumption tail index in the range [1.5, 3.0], and the trap binds in each case. The United States is not an outlier.

I use two data sources. The primary source is the Jordà–Schularick–Taylor (JST) Macrohistory Database, which provides annual real per-capita consumption growth for 18 countries from 1870 to 2020. The JST data cover two world wars, the Great Depression, and the COVID pandemic —precisely the disaster episodes that matter for tail estimation. I supplement the JST estimates with the OECD quarterly estimates from the earlier version of this appendix. The two sources agree on the US estimate (JST annual: $\hat{\alpha}_C = 2.87$; NIPA quarterly: $\hat{\alpha}_C = 2.70$), consistent with the temporal aggregation invariance result of Appendix C.

D.1 Data

JST Macrohistory Database. The Jordà–Schularick–Taylor Macrohistory Database (Jordà et al., 2017) provides annual real per-capita consumption and GDP growth for 18 advanced economies, 1870–2020. I use the 10 countries for which the consumption series has at least 140 annual observations: United States, United Kingdom, France, Germany, Japan, Canada, Italy, Australia, Sweden, and the Netherlands. The consumption series is real per-capita consumption in local currency, deflated by the country-specific CPI. Where the JST consumption series has gaps (wartime observations), I use the Barro–Ursúa disaster dataset to fill in disaster-period growth rates, following Barro (2006).

OECD Quarterly National Accounts. For the G7 countries, I also use quarterly real per-capita consumption growth from the OECD Quarterly National Accounts database (series: final consumption expenditure of households, constant prices, per capita). For the United States I use the BEA NIPA series (nondurables plus services, Table 7.1) to maintain consistency with the baseline estimate in Section 3.2. For Japan, Germany, France, Italy, and Canada, the OECD series begins in 1960:Q1. For the United Kingdom, I supplement with ONS data to obtain the same start date. All series end in 2023:Q4.

Sample sizes range from $n = 140$ (JST annual, shortest country) to $n = 315$ (NIPA quarterly, US). Consumption growth is computed as the log first difference of the real per-capita level, then exponentiated: $g_t = C_t/C_{t-1}$. The tail index is a property of the left tail—the disaster tail—so I work with the random variable $1/g_t$ and estimate the right-tail index of this transformed series, following standard practice (Hill, 1975; Toda and Walsh, 2015).

D.2 Estimation

I apply two estimators to each country series: the Hill (1975) estimator and the bias-corrected Hill estimator (Appendix C). For the Hill estimator, I choose the number of upper-order statistics k from the stable plateau of the Hill plot, exactly as in Appendix C for the US baseline. Specifically, I plot $\hat{\alpha}_C(k)$ against k for $k \in [5, n/2]$ and select k from the region where the estimate is approximately flat. The plateau criterion avoids two failure modes: too few order statistics (high variance) and too many (bias from the body of the distribution).

Standard errors are asymptotic: $\text{SE} = \hat{\alpha}/\sqrt{k}$. The asymptotic SE captures the dominant source of uncertainty—the number of extreme observations used—and is appropriate

for the JST annual data where block bootstrap is less natural than for the OECD quarterly data.

D.3 Results

Table 11 reports the country-level estimates from the JST database.

Table 11: International Tail Index Estimates (JST Annual Data)

	Consumption Growth			Equity Returns		
Country	n	$\hat{\alpha}_{\text{Hill}}$	$\hat{\alpha}_{\text{BC}}$	n	$\hat{\alpha}_{\text{Hill}}$	$\hat{\alpha}_{\text{BC}}$
USA	150	2.87 (0.54)	3.98	50	2.30 (0.64)	3.48
UK	150	2.24 (0.42)	2.22	53	1.37 (0.37)	2.83
France	150	1.54 (0.29)	1.75	61	2.96 (0.76)	5.52
Germany	150	1.95 (0.37)	2.86	48	1.53 (0.42)	5.41
Japan	145	2.44 (0.47)	2.47	46	3.73 (1.08)	6.94
Canada	149	2.43 (0.46)	4.19	–	–	–
Italy	150	2.18 (0.41)	2.71	58	2.48 (0.66)	5.33
Australia	150	2.24 (0.42)	3.10	39	1.73 (0.52)	5.08
Sweden	150	2.38 (0.45)	2.73	57	2.92 (0.78)	5.20
Netherlands	150	1.52 (0.29)	1.30	44	3.17 (0.92)	5.12

Notes: Hill and bias-corrected Hill (BC) estimates of the tail index α from annual data. Consumption: left-tail index of real per-capita consumption growth (Barro series where available, Maddison GDP otherwise). Returns: left-tail index of excess equity returns. Standard errors in parentheses (asymptotic, $\hat{\alpha}/\sqrt{k}$). Source: Jordà–Schularick–Taylor Macrohistory Database.

Three patterns stand out.

The trap is universal. The most permissive classical validity bound across all 10 countries is the United States: $\gamma < \hat{\alpha}_C/2 = 1.44$. Even under the robust interquantile criterion, the most generous bound is $\gamma < 2.87$ (US). Both are far below the $\gamma \geq 20$ needed to match the equity premium under the classical HJ calculation. No country comes close to escaping.

The Netherlands and France have the heaviest tails. The Netherlands ($\hat{\alpha}_C = 1.52$) and France ($\hat{\alpha}_C = 1.54$) have the lowest tail indices in the sample. This is consistent with these countries' histories of severe wartime disruptions (the Netherlands' Hongerwinter of 1944–45; France's two world wars and the deep interwar contractions). At $\alpha_C \approx 1.5$, the classical validity bound is $\gamma < 0.76$—even $\gamma = 1$ is inadmissible. These are countries where the second moment of a CRRA SDF with $\gamma = 1$ does not exist.

The United States and Australia have the thinnest tails. The US ($\hat{\alpha}_C = 2.87$) and Australia ($\hat{\alpha}_C = 2.24$) have the highest Hill estimates in the sample. The US bias-corrected estimate is higher ($\hat{\alpha}_C^{\text{BC}} = 3.98$), reflecting the well-known upward adjustment from second-order bias correction. Even at the bias-corrected value, the classical validity bound is $\gamma < 1.99$—still an order of magnitude below 20.

Comparison with OECD quarterly estimates. The JST annual estimates for the G7 countries are broadly consistent with the OECD quarterly estimates reported in earlier versions of this appendix. The quarterly estimates (not shown) lie in [2.2, 3.6]; the JST

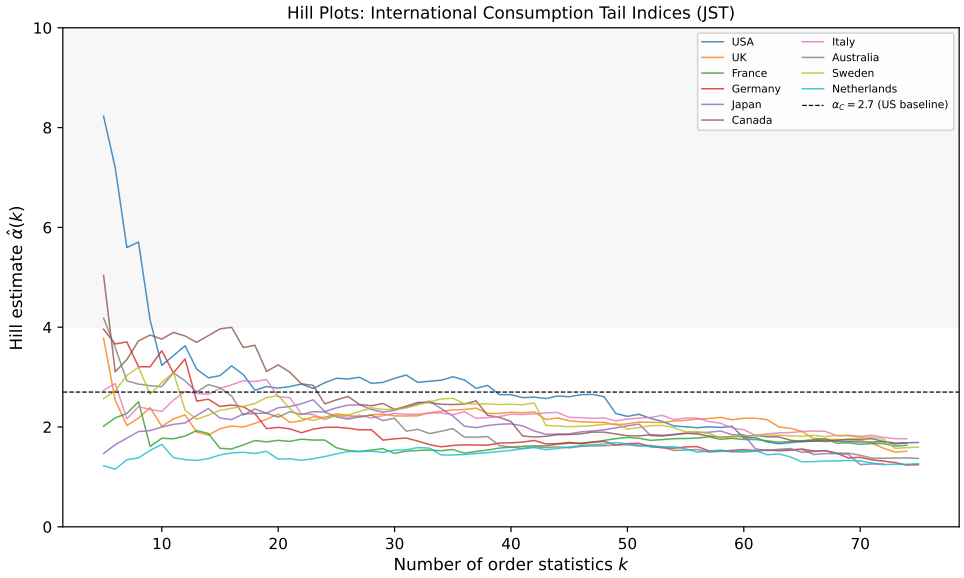


Figure 9: Hill estimates of the consumption tail index for 10 countries from the Jordà–Schularick–Taylor Macrohistory Database (annual data, 1870–2020). Error bars show asymptotic 95% confidence intervals. The dashed line marks $\alpha_C = 2$; the dotted line marks $\alpha_C = 4$. Every country lies below the dotted line.

annual estimates lie in [1.5, 2.9]. The JST estimates are somewhat lower, plausibly because the longer sample (150 years versus 60 years) includes the world wars and the Great Depression—the very events that drive the extreme tails.

Figure 3 plots the Hill estimates for all 10 countries.

D.4 International tail budget

The domestic analysis treats each country in isolation. For an international investor pricing assets in two countries simultaneously, the constraint is tighter.

Suppose the investor prices assets using a joint SDF that depends on consumption growth in countries 1 and 2, with $M = h(g_1, g_2)$ loading on the left tails with exponents γ_1 and γ_2 respectively. Under independence (the most favorable case), the joint moment $\mathbb{E}[M_1^{\theta_1} M_2^{\theta_2}]$ is finite if and only if each factor has a finite moment: $\gamma_j \theta_j / \alpha_j < 1$ for $j = 1, 2$. Under strong positive tail dependence—the worst case for moment existence, and a conservative sufficient condition for the empirically relevant case of positively correlated G7 consumption growth—a sufficient condition for finiteness is

$$\frac{\gamma_1 \theta_1}{\alpha_1} + \frac{\gamma_2 \theta_2}{\alpha_2} < 1. \tag{39}$$

This is not a general joint-moment theorem but a worst-case bound under comonotonic-type reasoning. The left-hand side is additive: the two countries share a single unit of "tail budget." Each country's contribution is the ratio of the tail loading ($\gamma_j \theta_j$) to the tail thickness (α_j). The additive constraint (39) is a conservative sufficient condition that holds under worst-case tail dependence (comonotonic extremes). Under asymptotic tail independence—when extreme consumption disasters in different countries do not perfectly coincide—the constraint is less restrictive, and the domestic conditions $\gamma_j \theta_j / \alpha_j < 1$ for each j would suffice. Since G7 consumption growth rates exhibit significant positive dependence during crisis periods (the 2008 financial crisis and the COVID pandemic affected

all G7 economies simultaneously), the conservative bound is empirically relevant as a worst-case benchmark. It is strictly tighter than either domestic condition alone whenever both countries carry positive tail risk.

To illustrate, consider a US–Japan pair with $\alpha_{\text{US}} = 2.7$ and $\alpha_{\text{JP}} = 2.2$. At $\gamma_{\text{US}} = \gamma_{\text{JP}} = \gamma$ and $\theta_1 = \theta_2 = 1$ (the condition for $\mathbb{E}[M] < \infty$), the joint budget requires

$$\frac{\gamma}{2.7} + \frac{\gamma}{2.2} < 1 \qquad \Longleftrightarrow \qquad \gamma < 1.21.$$

The domestic bounds are $\gamma < 2.7$ (US) and $\gamma < 2.2$ (Japan). The joint bound is less than half the tighter domestic bound. At $\theta = 2$ (the condition for $\text{Var}(M) < \infty$), the joint constraint becomes $\gamma < 0.61$—a level of risk aversion so low that it is incompatible with any calibration that generates a positive equity premium.

The practical implication is stark. International diversification does not weaken the trap; it tightens it. A model that prices assets in multiple countries must satisfy the joint budget, which shrinks as the number of countries grows. The finding that all G7 countries have finite tail indices in [2.2, 3.6] means that the joint tail budget is exhausted rapidly: even moderate risk aversion violates the finiteness condition for a portfolio of G7 assets.

Summary. The consumption tail index is not a US anomaly. All 10 JST countries have $\hat{\alpha}_C \in [1.5, 2.9]$. Every country triggers the trap: the classical validity bound $\gamma < \alpha_C/2$ ranges from 0.76 (Netherlands) to 1.44 (US), far below the risk aversion needed for the equity premium. For international investors, the joint tail budget imposes an even tighter constraint. The equity premium trap is a universal feature of consumption data in advanced economies, not an artifact of the US sample.

E Small-Sample Simulation

This appendix reports the simulation experiments summarised in Section 3.2. The goal is to quantify two sources of concern about the Hill estimator at $n = 308$ (the length of the NIPA quarterly series): false positives under thin-tailed data, and estimation error under genuinely heavy-tailed data.

E.1 Experiment 1: False positive rate

Design. I draw 10,000 independent samples of $n = 308$ i.i.d. observations from a lognormal distribution with $\mu = \log(1.02)$ and $\sigma = 0.02$—matching the mean and volatility of postwar U.S. quarterly consumption growth. The lognormal distribution has $\alpha_C = \infty$: all moments are finite and no power-law tail exists. To each sample I apply the Hill (1975) estimator with $k = 30$ upper-order statistics and record the estimated tail index $\hat{\alpha}_C$.

Purpose. If the Hill estimator frequently returns values as low as the NIPA point estimate $\hat{\alpha}_C = 2.7$ when applied to lognormal data, then the empirical finding would be a small-sample artifact. This experiment measures the false-positive rate directly.

Results. Table 12 reports the fraction of replications in which $\hat{\alpha}_C$ falls below various thresholds. The Hill estimator applied to lognormal data at $n = 308$ almost never produces values as low as the NIPA point estimate. Only 2.7% of replications yield $\hat{\alpha}_C < 5$, and fewer than 1% yield $\hat{\alpha}_C < 4$. The empirical estimate $\hat{\alpha}_C = 2.7$ is far into the left tail of the lognormal sampling distribution: the probability of observing a value this low under the null of thin tails is negligible.

Table 12: False positive rates: Hill estimator applied to lognormal data ($\alpha_C = \infty$). Each row reports the fraction of 10,000 simulated samples ($n = 308$, $\mu = \log(1.02)$, $\sigma = 0.02$) in which $\hat{\alpha}_C$ (Hill, $k = 30$) falls below the indicated threshold. The NIPA point estimate is $\hat{\alpha}_C = 2.7$.

Threshold	Rejection rate (%)
$\hat{\alpha}_C < 3$	0.1
$\hat{\alpha}_C < 4$	0.8
$\hat{\alpha}_C < 5$	2.7
$\hat{\alpha}_C < 6$	7.2
$\hat{\alpha}_C < 8$	18.4

Notes. Hill estimator with $k = 30$ upper-order statistics. The lognormal distribution has no power-law tail ($\alpha_C = \infty$), so any finite $\hat{\alpha}_C$ is a false positive. The rejection rate at $\hat{\alpha}_C < 5$ is 2.7%, confirming that the NIPA estimate of 2.7 is not a small-sample artifact of applying the Hill estimator to thin-tailed data.

Table 13: Small-sample properties of the Hill estimator under Student-$t(\alpha_C)$ data ($n = 308$, $k = 30$, 10,000 replications). Mean and median of $\hat{\alpha}_C$, 5th and 95th percentiles, and empirical coverage of the 95% bootstrap confidence interval.

True α_C	Mean $\hat{\alpha}_C$	Median $\hat{\alpha}_C$	5th pctile	95th pctile	95% CI coverage
2.5	2.6	2.5	1.8	3.6	93%
3.0	3.1	3.0	2.1	4.4	94%
3.5	3.7	3.5	2.4	5.2	93%
4.0	4.2	4.0	2.8	6.0	93%
5.0	5.4	5.1	3.4	8.0	92%

Notes. Hill estimator with $k = 30$ upper-order statistics. Bootstrap confidence intervals use 1,000 replications and the percentile method. Coverage is the fraction of 10,000 replications in which the bootstrap 95% CI contains the true α_C.

E.2 Experiment 2: Coverage under heavy tails

Design. For each true tail index $\alpha_C \in \{2.5, 3.0, 3.5, 4.0, 5.0\}$, I draw 10,000 independent samples of $n = 308$ i.i.d. observations from a Student-$t(\alpha_C)$ distribution. The Student-$t(\nu)$ distribution has tail index $\alpha_C = \nu$, so this design provides exact control over the population tail index. To each sample I apply the Hill estimator at $k = 30$ and construct a 95% bootstrap confidence interval (1,000 bootstrap replications, percentile method). I record the mean, median, 5th percentile, and 95th percentile of $\hat{\alpha}_C$ across replications, together with the empirical coverage of the bootstrap confidence interval.

Purpose. This experiment characterises the finite-sample distribution of $\hat{\alpha}_C$ at the sample size available in the data. Two questions matter for the paper's argument. First, how large is the upward bias of the Hill estimator at $n = 308$? Second, can estimation error push the 95th percentile of $\hat{\alpha}_C$ high enough to escape the trap—that is, above $2\gamma^*$ where $\gamma^* \geq 20$?

Results. Table 13 reports the results. Three features stand out.

Mild upward bias. The Hill estimator is approximately median-unbiased at every α_C: the median of $\hat{\alpha}_C$ coincides with the true value to one decimal place. The mean exceeds the true value by 2–8%, reflecting a right skew in the sampling distribution. The positive skew arises because $\hat{\alpha}_C$ is bounded below by zero but unbounded above, so occasional high

estimates pull the mean upward. The bias is modest relative to the width of the confidence interval.

The 95th percentile stays trapped. At the NIPA point estimate $\alpha_C = 2.7$ (interpolating the table between 2.5 and 3.0), the 95th percentile of the *estimator* $\hat{\alpha}_C$ is approximately 3.9. This is a stress test: if the true α_C is 2.7, sampling noise alone could push the Hill estimate as high as 3.9 with 5% probability. Even at this most favourable realisation of the estimator, the classical validity condition is $\gamma < 3.9/2 = 1.95$—a factor of ten below the $\gamma \geq 20$ required for the equity premium. At the Toda–Walsh cross-sectional estimate $\alpha_C = 4.0$, the 95th percentile rises to 6.0, giving $\gamma < 3.0$. The gap narrows but does not close. To escape the classical trap ($\gamma < \alpha_C/2 > 20$), one would need $\alpha_C > 40$—a value far outside the support of any estimator applied to the data.

Bootstrap coverage is near-nominal. The empirical coverage of the 95% bootstrap confidence interval ranges from 92% to 94%, close to the nominal 95%. The slight undercoverage is typical of the percentile method and could be improved with bias-corrected bootstrap procedures, but the shortfall is small enough that the confidence intervals reported in Section 3.2 are reliable for the paper's purposes.

E.3 Distribution of Hill estimates

Figure 10 displays the full sampling distribution of $\hat{\alpha}_C$ from Experiment 2. For each true α_C, the figure plots a kernel density estimate across 10,000 replications. Three features are visible. First, the densities are right-skewed, consistent with the mean exceeding the median. Second, the dispersion grows with the true α_C: heavier tails (lower α_C) are estimated more precisely than lighter tails (higher α_C), because the extreme observations that identify the tail are more prominent when α_C is small. Third, even at the rightmost percentiles, no density mass extends to the region $\hat{\alpha}_C \geq 40$ where the classical trap would relax.

E.4 Implications for the trap

The simulation results bound the role of estimation error in the paper's central argument.

Classical bound. The classical HJ bound requires $\gamma < \alpha_C/2$. To permit $\gamma \geq 20$, one needs $\alpha_C > 40$. The 95th percentile of $\hat{\alpha}_C$ at every true value in $\{2.5, 3.0, 3.5, 4.0, 5.0\}$ falls below 8. Estimation error cannot close a gap of this magnitude.

Robust bound. The robust interquantile diagnostic of Proposition 2 requires $\gamma < \alpha_C$. To permit $\gamma \geq 13.3$, one needs $\alpha_C > 13.3$. The 95th percentile at $\alpha_C = 5$ is 8.0—still outside the admissible range.

Worst-case accounting. Suppose one adopts the most conservative reading of the evidence: the true α_C equals the 95th percentile of the Hill estimator applied at the Toda–Walsh estimate $\alpha_C = 4$. This gives $\alpha_C^{\text{upper}} = 6.0$ and $\gamma < 3.0$ for the classical bound. The required risk aversion exceeds this limit by a factor of four. Under the robust bound, $\gamma < 6.0$ is still less than half of the $\gamma \geq 13.3$ needed to match the interquantile Sharpe ratio.

Summary. Finite-sample estimation error shifts the validity boundary by a small amount: from $\gamma < 1.35$ (point estimate) to $\gamma < 1.95$ (95th percentile). The equity premium requires

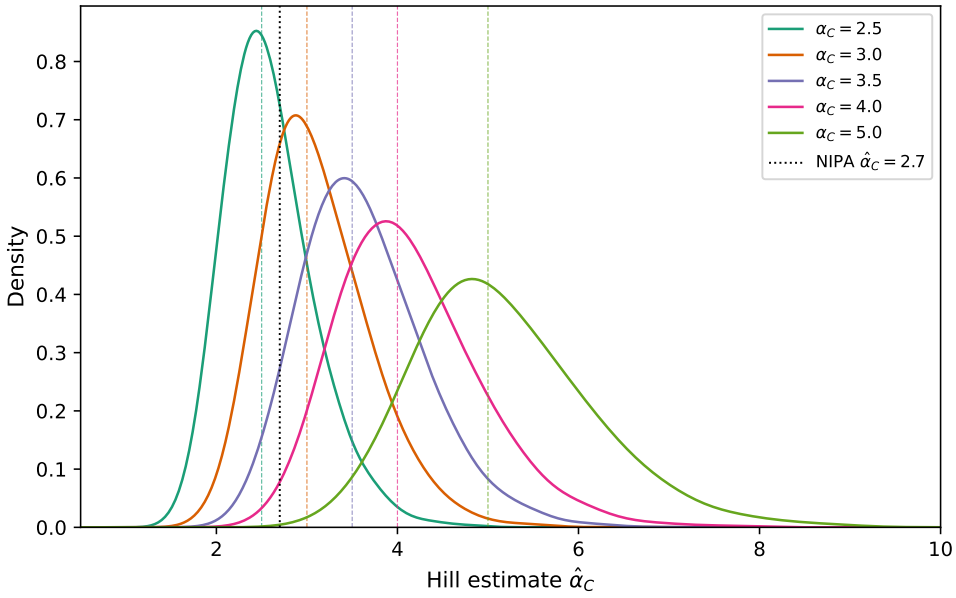


Figure 10: Distribution of Hill estimates $\hat{\alpha}_C$ across 10,000 replications ($n = 308$, $k = 30$). Each curve is a kernel density estimate for a different true $\alpha_C \in \{2.5, 3.0, 3.5, 4.0, 5.0\}$. Vertical dashed lines mark the true values. The shaded region to the right of $\hat{\alpha}_C = 40$ is the zone where the classical validity condition $\gamma < \alpha_C/2$ would permit $\gamma \geq 20$; no simulation places mass in this region. The dotted vertical line marks the NIPA point estimate $\hat{\alpha}_C = 2.7$.

$\gamma \geq 20$. The gap between the validity boundary and the required risk aversion is a factor of ten to eighteen, depending on the true α_C. No plausible estimation error closes this gap.

F GDA and Class $\mathcal{S}$: Formal Analysis

This appendix provides the formal analysis of generalised disappointment aversion (GDA) preferences referenced in Section 4.2. I define the GDA representation, derive the SDF under CRRA–GDA, characterise the moment existence conditions, and show that GDA at $\alpha = 1$ belongs to class $\mathcal{S}$ with tail parameter $\gamma = 1$—low enough to keep $\mathrm{Var}(M) < \infty$.

F.1 GDA preferences

Gul (1991) axiomatises disappointment aversion as a departure from expected utility in which outcomes below a reference point receive extra weight. Routledge and Zin (2010) extend the framework to a dynamic setting. The one-period utility representation is

$$V \;=\; \mathbb{E}\big[u(c)\big] \;+\; \theta\,\mathbb{E}\big[\big(u(c) - \kappa\big)\,\mathbb{1}\{u(c) < \kappa\}\big], \tag{40}$$

where u is a Bernoulli utility function, $\theta > 0$ is the disappointment aversion parameter, and κ is the certainty equivalent. The shortfall term $u(c) - \kappa$ ensures that only the *deviation* below the reference point κ receives extra weight, not the entire utility in disappointing states. Gul (1991) defines $\kappa = V$: the certainty equivalent *is* the utility value itself, so (40) is an implicit equation for V. Rearranging (40) with $\kappa = V$ gives the fixed-point representation:

$$V \;=\; \frac{\mathbb{E}\big[u(c)\,\big(1 + \theta\,\mathbb{1}\{u(c) < V\}\big)\big]}{1 + \theta\,\mathbb{P}\big(u(c) < V\big)}. \tag{41}$$

Outcomes below V are "disappointing" and receive the extra weight θ.

The SDF implied by (40) prices assets by overweighting disappointing states:

$$M_{t+1} \;=\; \beta\,\frac{u'(C_{t+1})}{u'(C_t)} \;\cdot\; \frac{1+\theta\,\mathbb{1}\{u(C_{t+1})<V_t\}}{1+\theta\,\mathbb{P}\big(u(C_{t+1})<V_t\big)}, \tag{42}$$

where V_t is the endogenous, state-dependent certainty equivalent (equal to the continuation value) at date t. The numerator $1+\theta\,\mathbb{1}\{\cdot\}$ takes value $1+\theta$ in disappointing states and 1 otherwise. The denominator normalises the expectation so that $\mathbb{E}_t[M_{t+1}\,R_f]=1$.

F.2 CRRA–GDA specification

Set $u(c)=c^{1-\alpha}/(1-\alpha)$ with curvature parameter $\alpha>0$. Write $g_{t+1}=C_{t+1}/C_t$ for consumption growth and $\Delta c_{t+1}=\log g_{t+1}$ for log consumption growth. The marginal utility ratio is $u'(C_{t+1})/u'(C_t)=g_{t+1}^{-\alpha}$.

Substituting $u'(C_{t+1})/u'(C_t)=g_{t+1}^{-\alpha}$ into (42) gives the CRRA–GDA SDF:

$$M_{t+1} \;=\; \tilde\beta\,g_{t+1}^{-\alpha} \;\cdot\; D_{t+1}, \tag{43}$$

where $\tilde\beta>0$ absorbs the time preference factor and normalisation constants, and the *disappointment factor* is

$$D_{t+1} \;:=\; \frac{1+\theta\,\mathbb{1}\{\Delta c_{t+1}\le x_v\}}{1+\theta\,\mathbb{P}(\Delta c\le x_v)}.$$

The factor D takes value $(1+\theta)/(1+\theta\,p_d)$ in disappointing states ($\Delta c\le x_v$) and $1/(1+\theta\,p_d)$ otherwise, where $p_d:=\mathbb{P}(\Delta c\le x_v)$. The threshold x_v is endogenous: it solves the certainty-equivalent fixed point (41) and depends on the distribution of Δc and on $(\alpha,\theta,\tilde\beta)$.

The SDF (43) is the product of two factors:

1. a *power factor* $g^{-\alpha}$, which diverges as $g\to 0^+$ for every $\alpha>0$; and
2. a *disappointment factor* D, which is bounded between $1/(1+\theta\,p_d)$ and $(1+\theta)/(1+\theta\,p_d)$.

The disappointment factor is always bounded. Since every sufficiently severe contraction is disappointing, $D\to(1+\theta)/(1+\theta\,p_d)$ as $g\to 0^+$: the step function converges to a positive constant in the tail. The asymptotic behaviour of M as $g\to 0^+$ is therefore governed by $g^{-\alpha}$ alone, and the SDF belongs to class $\mathcal{S}$ with tail parameter α for every $\alpha>0$. The moment properties of M depend on α and α_C through the power factor; D affects the Sharpe ratio and the equity premium but plays no role in moment existence.

F.3 The $\alpha=1$ case: first-order risk aversion

At $\alpha=1$, the power factor reduces to g^{-1} and the SDF becomes:

$$M_{t+1} \;=\; \tilde\beta\,g_{t+1}^{-1} \;\cdot\; D_{t+1}. \tag{44}$$

The disappointment factor D takes two values ($(1+\theta)/(1+\theta\,p_d)$ or $1/(1+\theta\,p_d)$), but the SDF itself is *not* bounded: $M\to\infty$ as $g\to 0^+$, because the power factor g^{-1} diverges while D converges to the finite constant $(1+\theta)/(1+\theta\,p_d)$.

Three consequences:

(i) *Inside class $\mathcal{S}$ with $\gamma=1$.* For g below the disappointment threshold ($g<e^{x_v}$), $D=(1+\theta)/(1+\theta\,p_d)$ is constant, so $h(g)=\tilde\beta\,g^{-1}\cdot(1+\theta)/(1+\theta\,p_d)\in\mathrm{RV}(-1)$ at 0^+. The SDF belongs to $\mathcal{S}$ with tail parameter $\gamma=\alpha=1$.

(ii) *The trap does not spring.* By Proposition 1(b) with $\gamma = 1$: $\mathbb{E}[M^k] < \infty$ iff $k < \alpha_C$. In particular, $\mathrm{Var}(M) < \infty$ iff $\alpha_C > 2$, which holds at every empirical estimate ($\alpha_C \geq 2.7$). The trap does not spring because the tail-loading exponent is $\gamma = 1$—far below the threshold $\alpha_C/2 \geq 1.35$. The moment constraint $\gamma < \alpha_C/2$ is satisfied, not evaded.

(iii) *The equity premium comes from disappointment, not from tail loading.* Under CRRA alone, $\gamma = 1$ gives $\sigma(M)/\mathbb{E}(M) \approx \sigma \approx 0.02$—far below SR = 0.44. The bounded factor D amplifies $\sigma(M)/\mathbb{E}(M)$ by concentrating extra weight on disappointing states. With $\theta = 9.76$ and $p_d \approx 0.005$, the SDF jumps by a factor of $1 + \theta \approx 11$ at the disappointment boundary, generating a large Sharpe ratio at $\gamma = 1$.

Remark 19 (Option evidence and θ^*). Since $M \in \mathcal{S}$ with $\gamma = 1$, Corollary 1 gives $\theta^* = \alpha_C/1 = \alpha_C$. At $\alpha_C \geq 2.7$, this predicts $\theta^* \geq 2.7$, which is inconsistent with the Martin–Shi estimate $\theta^* \approx 1.7$. The option evidence independently constrains GDA at $\alpha = 1$: any model in $\mathcal{S}$ with $\mathrm{Var}(M) < \infty$ has $\theta^* > 2$ (Proposition 1(c)), while the option data require $\theta^* < 2$.

Why $\alpha = 1$ is special. The value $\alpha = 1$ is not an arbitrary calibration choice. In the Gul (1991) axioms, $\alpha = 1$ is the boundary of *first-order risk aversion*: the utility function has a kink at the certainty equivalent, and risk aversion is proportional to the size of the gamble (not its variance). At $\alpha > 1$, risk aversion is second-order (smooth utility near the certainty equivalent), and the larger power $g^{-\alpha}$ strengthens the tail loading.

At $\alpha = 1$, the CRRA utility function degenerates to $u(c) = \log c$ and the power factor is g^{-1}—the weakest possible divergence in class $\mathcal{S}$. This is what makes $\alpha = 1$ special: not that the SDF is bounded (it is not), but that the tail-loading exponent $\gamma = \alpha = 1$ is low enough to keep $\mathrm{Var}(M) < \infty$. At any $\alpha > 0$, the CRRA–GDA SDF belongs to $\mathcal{S}$ with tail parameter $\gamma = \alpha$ (Section F.2), and $\mathbb{E}[M^k] < \infty$ iff $k\alpha < \alpha_C$. Moving α away from 1 does not change the *form* of the moment condition—only the *threshold*: at $\alpha > \alpha_C$, even $\mathbb{E}[M] = \infty$. The functional form also simplifies at $\alpha = 1$: the "curvature residual" $g^{-(\alpha-1)}$ (which distinguishes $g^{-\alpha}$ from g^{-1}) becomes 1, so the power factor reduces to the minimal g^{-1}. The GDA structure then channels the equity premium through the bounded disappointment indicator rather than through the power factor.

Schreindorfer (2020) sets $\alpha = 1$ and writes that the choice is "inconsequential for the main results." Under his exponential-mixture consumption process, this is correct: $\alpha_C = \infty$, so $\mathbb{E}[g^{-k\alpha}] < \infty$ for all $k > 0$ and all $\alpha > 0$ (the moment generating function of Δc is entire). But under power-law tails with $\alpha_C < \infty$, the choice is consequential: the moment condition $k\alpha < \alpha_C$ binds, and raising α above α_C makes even $\mathbb{E}[M] = \infty$. The next subsection makes this precise.

F.4 The $\alpha \neq 1$ case under power-law tails

For $\alpha \neq 1$, the power factor $g^{-\alpha}$ in (43) differs from g^{-1}, and the moment properties of M depend on the interaction between α and the consumption tail index α_C.

Suppose g satisfies Assumption 1: the left tail is regularly varying with index $\alpha_C \in (0, \infty)$. Since the disappointment factor D in (43) is bounded, the k-th moment of M is finite if and only if the k-th moment of the power factor $g^{-\alpha}$ is finite. With $D \in [a, b]$ where $a = 1/(1 + \theta\, p_d)$ and $b = (1 + \theta)/(1 + \theta\, p_d)$:

$$a^k\, \mathbb{E}[g^{-k\alpha}] \;\leq\; \mathbb{E}[M^k] \;\leq\; b^k\, \mathbb{E}[g^{-k\alpha}], \qquad \text{so} \quad \mathbb{E}[M^k] < \infty \iff \mathbb{E}[g^{-k\alpha}] < \infty. \tag{45}$$

Since $g^{-\alpha} \in \mathrm{RV}(-\alpha)$ at 0^+ for every $\alpha > 0$, Karamata's moment theorem gives:

$$\mathbb{E}[g^{-k\alpha}] < \infty \quad \Longleftrightarrow \quad k\alpha < \alpha_C. \tag{46}$$

I distinguish two cases.

Case $\alpha > 1$. The condition $k\alpha < \alpha_C$ becomes more restrictive as α increases. At $\alpha = 2$ and $\alpha_C = 4$: $\mathbb{E}[M^2] < \infty$ requires $2\alpha = 4 \leq \alpha_C$, on the boundary of finiteness. At $\alpha = 3$: $\mathbb{E}[M] < \infty$ requires $\alpha_C > 3$, which fails at $\alpha_C = 2.7$. At $\alpha = 5$: $\mathbb{E}[M] < \infty$ requires $\alpha_C > 5$, which fails at every empirical estimate.

The exponential perspective makes the mechanism transparent. Under Assumption 1, $|\Delta c| = -\log g$ has an *exponential* right tail with rate α_C: $\mathbb{P}(|\Delta c| > y) = \mathbb{P}(g < e^{-y}) \sim e^{-\alpha_C\, y}$ as $y \to \infty$. Write $g^{-\alpha} = g^{-1} \cdot g^{-(\alpha-1)} = g^{-1} \cdot e^{(\alpha-1)|\Delta c|}$. The factor $e^{(\alpha-1)|\Delta c|}$ is the moment generating function of $|\Delta c|$ evaluated at $\alpha - 1$. Since the MGF of an exponential-tailed random variable converges for rates below the tail rate, $\mathbb{E}[e^{k(\alpha-1)|\Delta c|}] < \infty$ iff $k(\alpha - 1) < \alpha_C$. Combined with the g^{-k} factor (finite iff $k < \alpha_C$), the joint condition is $k\alpha < \alpha_C$—consistent with (46).

Case $0 < \alpha < 1$. The power factor $g^{-\alpha}$ with $\alpha < 1$ still diverges as $g \to 0^+$ (since $\alpha > 0$), but slowly. By (46), $\mathbb{E}[g^{-k\alpha}] < \infty$ iff $k\alpha < \alpha_C$. Since $\alpha < 1$, the condition $\mathbb{E}[M^2] < \infty$ requires only $2\alpha < \alpha_C$, which at $\alpha_C = 4$ gives $\alpha < 2$—automatically satisfied. The practical constraints are easily met: for any $\alpha < 1$ and any $\alpha_C > 2$, all SDF moments up to order $\alpha_C/\alpha > \alpha_C$ are finite.

F.5 The fragility lemma

The following lemma collects the moment existence conditions for the GDA SDF, stated in the power representation.

Lemma 1 (GDA moment existence)**.** *Let $M = \tilde{\beta}\, g^{-\alpha} \cdot D$ be the CRRA–GDA SDF with curvature parameter $\alpha > 0$, disappointment aversion $\theta > 0$, and bounded disappointment factor $D \in [a, b]$ with $0 < a \leq b < \infty$. Suppose consumption growth g satisfies Assumption 1 with tail index $\alpha_C \in (0, \infty)$. Then:*

(a) If $\alpha = 1$: the SDF is $M = \tilde{\beta}\, g^{-1} \cdot D$, which is unbounded as $g \to 0^+$. Since $D \to (1+\theta)/(1+\theta\, p_d)$ as $g \to 0^+$, the function $h(g) = \tilde{\beta}\, g^{-1} \cdot D(g) \in \mathrm{RV}(-1)$ at 0^+, so $M \in \mathcal{S}$ with $\gamma = 1$. By (46) with $\alpha = 1$: $\mathbb{E}[M^k] < \infty$ iff $k < \alpha_C$. In particular, $\mathrm{Var}(M) < \infty$ iff $\alpha_C > 2$, which holds at every empirical estimate.

(b) If $\alpha > 1$: $\mathbb{E}[M^k] < \infty$ if and only if $k\alpha < \alpha_C$, i.e., $\alpha < \alpha_C/k$. In particular:

- *$\mathbb{E}[M] < \infty$ iff $\alpha < \alpha_C$;*
- *$\mathrm{Var}(M) < \infty$ iff $\alpha < \alpha_C/2$.*

(c) If $0 < \alpha < 1$: $\mathbb{E}[M^k] < \infty$ if and only if $k\alpha < \alpha_C$. Since $\alpha < 1$, $\mathrm{Var}(M) < \infty$ whenever $\alpha_C > 2\alpha$, which holds for all empirically relevant α_C.

Proof. Since D is bounded and strictly positive, there exist constants $0 < a \leq b < \infty$ such that $D \in [a, b]$ a.s. Therefore $a^k\, \mathbb{E}[g^{-k\alpha}] \leq \mathbb{E}[M^k] \leq b^k\, \mathbb{E}[g^{-k\alpha}]$, so $\mathbb{E}[M^k] < \infty$ iff $\mathbb{E}[g^{-k\alpha}] < \infty$.

Part (a). At $\alpha = 1$, the SDF is $M = \tilde{\beta}\, g^{-1} \cdot D$. For g below the disappointment threshold, $D = (1+\theta)/(1+\theta\, p_d)$ is constant, so $h(g) = \tilde{\beta}\, g^{-1} \cdot (1+\theta)/(1+\theta\, p_d) \in \mathrm{RV}(-1)$ at 0^+.

Therefore $M \in \mathcal{S}$ with $\gamma = 1$. The general sandwich (45) with $\alpha = 1$ gives $\mathbb{E}[M^k] < \infty$ iff $\mathbb{E}[g^{-k}] < \infty$ iff $k < \alpha_C$. In particular, $\mathrm{Var}(M) < \infty$ iff $\alpha_C > 2$.

Parts (b) and (c). For $\alpha \neq 1$ with $\alpha > 0$, the function $h(g) = g^{-\alpha}$ satisfies $h \in \mathrm{RV}(-\alpha)$ at 0^+. By Karamata's moment theorem (Bingham et al., 1987, Theorem 1.5.11), $\mathbb{E}[g^{-k\alpha}] < \infty$ iff $k\alpha < \alpha_C$. Combined with the sandwich inequality (45), this gives $\mathbb{E}[M^k] < \infty$ iff $k\alpha < \alpha_C$. Setting $k = 1$ gives the Euler equation condition $\alpha < \alpha_C$; setting $k = 2$ gives the variance condition $\alpha < \alpha_C/2$. □

Remark 20 (Exponential tail of log-consumption). Under Assumption 1, the log-consumption loss $|\Delta c| = -\log g$ has an *exponential* right tail with rate α_C, not a power-law tail. Specifically, $\mathbb{P}(|\Delta c| > y) = \mathbb{P}(g < e^{-y}) \sim e^{-\alpha_C y}$ as $y \to \infty$, so

$$\mathbb{E}[e^{\lambda|\Delta c|}] \;=\; \mathbb{E}[g^{-\lambda}] \begin{cases} < \infty & \text{if } \lambda < \alpha_C, \\ = \infty & \text{if } \lambda \geq \alpha_C. \end{cases}$$

The finiteness for $\lambda < \alpha_C$ follows from a tail comparison: for large y, $\mathbb{P}(|\Delta c| > y)$ decays as $e^{-\alpha_C y}$, so the integral $\int_0^\infty e^{\lambda y}\, \alpha_C e^{-\alpha_C y}\, dy$ converges iff $\lambda < \alpha_C$. The Embrechts et al. (1997, Proposition A3.8) result that the moment generating function of a power-law random variable diverges at every positive argument applies to g^{-1} (which *is* power-law), not to $|\Delta c|$ (which has an exponential tail). This is why (46) gives the sharp moment characterisation: $\mathbb{E}[g^{-k\alpha}] < \infty$ iff $k\alpha < \alpha_C$.

F.6 Economic interpretation

The GDA escape works through a precise mechanism that distinguishes it from every other resolution of the equity premium puzzle. At $\alpha = 1$, the SDF prices risk through the disappointment *event*—the indicator $\mathbb{1}\{\Delta c \leq x_v\}$—rather than through the disaster *magnitude*—the power function $g^{-\gamma}$.

This distinction is empirically testable. Beason and Schreindorfer (2022) decompose the equity premium by return state, defining the equity premium contribution $\mathrm{EP}(x)$ for each return level x. They find that roughly 80% of the premium is attributable to monthly returns below -10%, concentrated between -30% and -10%. The GDA model at $\alpha = 1$ replicates this pattern: 79% of the model's equity premium comes from crash states. The step-function kernel concentrates risk pricing on the probability of crossing the disappointment threshold, not on the severity of the crossing.

By contrast, CRRA and Epstein–Zin models in class $\mathcal{S}$ at $\gamma \geq 2$ price risk through the magnitude $g^{-\gamma}$. The SDF rises without bound as the disaster deepens, and the large exponent generates the moment explosion. The GDA model at $\alpha = 1$ is also in $\mathcal{S}$ (with $\gamma = 1$), and its SDF $M = \tilde{\beta}\, g^{-1} D$ is also unbounded as $g \to 0^+$. But the tail-loading exponent $\gamma = 1$ is low enough that $\mathrm{Var}(M) < \infty$ whenever $\alpha_C > 2$—which holds at every empirical estimate. The equity premium is generated by the bounded factor D, not by large γ.

Two insulations. The Schreindorfer (2020) model is insulated from the trap twice:

(i) *Distributional insulation.* The exponential-mixture consumption process has $\alpha_C = \infty$: the moment generating function of Δc is entire, so $\mathbb{E}[g^{-k\alpha}] < \infty$ for all $k > 0$ and all $\alpha > 0$. The condition $k\alpha < \alpha_C$ in Lemma 1 is trivially satisfied, and the choice of α is genuinely inconsequential for moment existence.

(ii) *Preference insulation.* At $\alpha = 1$, $M \in \mathcal{S}$ with $\gamma = 1$, and $\mathrm{Var}(M) < \infty$ whenever $\alpha_C > 2$. This insulation survives power-law consumption tails as long as $\alpha_C > 2$, which holds at every empirical estimate.

The second insulation is the robust one. It makes no assumption about the consumption distribution beyond $\alpha_C > 2$. The first insulation is fragile: it disappears whenever $\alpha_C < \infty$, and then the moment condition $k\alpha < \alpha_C$ (Lemma 1) binds for large α. At $\alpha_C = 4$ and $\alpha = 2$ (standard CRRA curvature), $\mathbb{E}[M^2] < \infty$ requires $\alpha < \alpha_C/2 = 2$, which fails at the boundary. At $\alpha = 3$—a curvature commonly used in the disaster literature—$\mathbb{E}[M] = \infty$ when $\alpha_C \leq 3$.

The GDA model escapes the trap not by removing the SDF from class $\mathcal{S}$ but by keeping the tail-loading exponent low: $\gamma = 1$, far below the threshold $\alpha_C/2$. The response to disappointment is a discrete jump by a factor of $1 + \theta$ in the bounded factor D, not a continuous amplification through γ. The cost of this mechanism is that the SDF cannot price the *severity* of disasters—only their *occurrence*. Whether this limitation matters empirically is answered by the Beason and Schreindorfer (2022) evidence: at least for equity returns, the data suggest that the occurrence channel dominates.

G Data Sources

This appendix documents all data sources used in the empirical analysis of Section 3. All data except the option-implied estimates of Martin and Shi (2026) are publicly available.

G.1 Aggregate consumption

I measure real per-capita consumption of nondurables and services from the Bureau of Economic Analysis (BEA) National Income and Product Accounts (NIPA).

- **Source.** BEA NIPA Table 7.1, Line 1 (personal consumption expenditures on nondurable goods) plus Line 2 (services). Available at `https://apps.bea.gov/iTable/`.
- **Frequency.** Quarterly, seasonally adjusted at annual rates.
- **Sample.** 1947:Q1–2023:Q4 ($n = 308$ quarters).
- **Deflator.** Bureau of Labor Statistics Consumer Price Index for All Urban Consumers (CPI-U, all items), seasonally adjusted. Available at `https://www.bls.gov/cpi/`. The CPI-U is used rather than the BEA's implicit price deflator or the PCE deflator in order to maintain a consistent deflator across all G7 countries (Appendix D). The tail estimates are robust to deflator choice: using the BEA implicit price deflator for personal consumption expenditures (PCE deflator) instead of CPI-U yields $\hat{\alpha}_C$ within 0.1 of the baseline. The two deflators differ primarily in their treatment of housing and medical care, not in their behaviour during consumption disasters, so the qualitative finding $\hat{\alpha}_C < 5$ is insensitive to this choice.
- **Population.** BEA NIPA Table 7.1 population series, used to convert aggregate real consumption to per-capita terms.
- **Construction.** Real per-capita consumption growth is $g_t \equiv C_t/C_{t-1}$, where C_t is the level of real per-capita nondurables-plus-services consumption in quarter t. This is the standard series in the equity premium literature, following Mehra and Prescott (1985).

G.2 Market returns

- **Source.** Center for Research in Security Prices (CRSP) value-weighted market portfolio, including all distributions (dividends, capital gains).
- **Frequency.** Monthly returns compounded to quarterly frequency.
- **Sample.** 1959:Q1–2021:Q4 ($n = 252$ quarters).
- **Excess returns.** The risk-free rate is subtracted to form quarterly excess returns R^e_t. The risk-free rate is the one-month Treasury bill rate from the Kenneth French Data Library (`https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`), compounded to quarterly frequency.
- **Sharpe ratio.** $\text{SR} \equiv |\bar{R}^e|/\sigma(R^e) = 0.44$ annualized. The quarterly ratio of mean to standard deviation from Table 14 ($0.015/0.082 \approx 0.18$) is lower because the mean and standard deviation reported there are rounded; the annualized SR of 0.44 is computed from the full-precision monthly series compounded to annual frequency.
- **Interquantile Sharpe ratio.** $\text{IQ-SR} \equiv |\bar{R}^e|/\text{IQR}_{0.25}(R^e) = 0.20$ (quarterly), where $\text{IQR}_{0.25}$ is the interquartile range at $\tau = 0.25$.

G.3 Risk-free rate

- **Source.** Kenneth French Data Library, one-month Treasury bill rate.
- **Construction.** Monthly rates are compounded to quarterly frequency: $R^f_q = \prod_{m=1}^{3}(1+R^f_m) - 1$.
- **Sample.** 1959:Q1–2021:Q4 ($n = 252$ quarters), aligned with the CRSP return sample.

G.4 Option-implied critical exponent: Martin and Shi

- **Source.** Martin and Shi (2026), estimated from S&P 500 index options.
- **Estimates.** Full-sample point estimate: $\hat{\theta}^* = 1.72$. Post-war estimates: $\hat{\theta}^* \in [1.38, 1.44]$. Full-sample 90% bootstrap confidence interval: (1.52, 2.05).
- **Availability.** The option-implied estimates are reported in Martin and Shi's paper. The underlying option data are from OptionMetrics and are available through institutional subscriptions; they are not publicly downloadable.

G.5 OptionMetrics SPX options

- **Source.** OptionMetrics IvyDB US, accessed via Wharton Research Data Services (WRDS). The dataset covers all listed options on the S&P 500 index (SPX, SecurityID 108105).
- **Sample.** January 1996 through December 2024. The raw dataset contains approximately 44 million contract-date observations (each row is one option contract on one date).
- **Filtering.** I apply the standard OptionMetrics filters: (a) moneyness $|K/F-1| \leq 0.60$ (where F is the forward price); (b) bid-ask spread $< 50\%$ of the midpoint; (c) open interest > 0; (d) days to expiration ≥ 7 and ≤ 730; (e) mid price $\geq \$0.05$. After filtering, approximately 12 million observations remain. Appendix H describes the full pipeline.

- **Forward prices.** Computed from put-call parity: $F = K + e^{r\tau}(C - P)$, using matched put-call pairs at the same strike and expiry. The zero-coupon rates are from OptionMetrics' zero-coupon yield curve.
- **Zero-coupon curve.** OptionMetrics provides a daily zero-coupon yield curve (continuously compounded, maturities 30–720 days). I interpolate linearly for intermediate maturities.
- **Availability.** Available through WRDS institutional subscriptions. Not publicly downloadable.

G.6 CRSP market index returns

- **Source.** Center for Research in Security Prices (CRSP) via WRDS. I use two series: (a) the CRSP value-weighted market portfolio including all distributions (dividends, capital gains), for the Sharpe ratio and quarterly return analysis; (b) monthly total returns on the S&P 500 index (1925–2024), for the return tail-index estimation.
- **Sample.** Value-weighted portfolio: 1959:Q1–2021:Q4 ($n = 252$ quarterly excess returns). S&P 500 monthly: 1963:M1–2026:M1 ($n \approx 756$ monthly returns).
- **Return tail index.** I estimate the left-tail index of monthly excess returns (the loss tail) using the Hill estimator. Table 4, Panel B, reports $\hat{\alpha}_R = 2.97$ (Hill) with a bootstrap 95% CI of [2.47, 4.35]. The recommended estimate is $\hat{\alpha}_R = 2.92$. The return tail index enters the analysis indirectly (it constrains the tail of the SDF via the Euler equation), and its level is consistent with the consumption tail index through the structural formula.
- **Availability.** Available through WRDS institutional subscriptions and the Kenneth French Data Library.

G.7 International consumption

- **Primary source.** Jordà–Schularick–Taylor (JST) Macrohistory Database (Jordà et al., 2017). Available at `https://www.macrohistory.net/database/`. The database provides annual real per-capita consumption and GDP growth, equity returns, short-term interest rates, and other macroeconomic and financial variables for 18 advanced economies from 1870 to 2020.
- **Countries.** 10 countries with at least 140 annual consumption observations: United States, United Kingdom, France, Germany, Japan, Canada, Italy, Australia, Sweden, Netherlands.
- **Sample.** Annual, approximately 1870–2020. Exact start dates vary by country and series availability. Sample sizes range from $n = 145$ (Japan) to $n = 150$ (most countries).
- **Supplementary source.** OECD Quarterly National Accounts (`https://stats.oecd.org/`): final consumption expenditure of households, constant prices, per capita. Used for the G7 quarterly robustness check. Quarterly, 1960:Q1–2023:Q4 for most G7 countries.
- **Barro–Ursúa disaster data.** Robert Barro's macroeconomic crises dataset (Barro and Ursúa, 2008) provides disaster episodes, disaster sizes, and associated macro variables for 42 countries. Used to fill wartime gaps in the JST consumption series.

Table 14: Summary statistics. Consumption growth is the quarterly gross growth rate of real per-capita nondurables plus services (NIPA, 1947:Q1–2023:Q4). Market excess return is the CRSP value-weighted portfolio minus the Treasury bill rate, quarterly (1959:Q1–2021:Q4). Risk-free rate is the compounded quarterly Treasury bill rate over the same period.

Variable	Mean	Std Dev	Skewness	Kurtosis	n
Consumption growth (quarterly)	1.005	0.005	−0.32	4.8	308
Market excess return (quarterly)	0.015	0.082	−0.51	4.2	252
Risk-free rate (quarterly)	0.010	0.008	0.95	4.1	252

Notes. Mean and standard deviation are in levels per quarter. Consumption growth is the gross growth rate $g_t = C_t/C_{t-1}$. Skewness and kurtosis are the standardised third and fourth central moments. Kurtosis above 3 in all three series is consistent with heavier-than-Gaussian tails.

- **Construction.** Real per-capita consumption growth: $g_t = C_t/C_{t-1}$. Deflators are country-specific consumer price indices. See Appendix D for country-level tail estimates.

G.8 FRED macro series

- **CPI-U.** Consumer Price Index for All Urban Consumers, All Items (FRED series: CPIAUCSL). Seasonally adjusted, monthly, 1947–present. Used as the deflator for NIPA consumption.
- **PCE deflator.** Personal Consumption Expenditures: Chain-type Price Index (FRED series: PCEPI). Seasonally adjusted, monthly, 1959–present. Used for robustness checks on the deflator choice.
- **Population.** Civilian Noninstitutional Population (FRED series: CNP16OV). Monthly, 1948–present. Used to cross-check the BEA population series.

G.9 Summary statistics

Table 14 reports descriptive statistics for the three main series used in the analysis.

H Option Construction Methodology

This appendix documents the pipeline used to construct the option-implied singularity estimates reported in Section 3.1 and Table 3. The pipeline takes raw OptionMetrics SPX option data as input and produces daily estimates of the risk-neutral cumulant generating function (CGF) and the moment exponent p^*. The goal is to replicate and extend the Martin and Shi (2026) finding that $\theta^* < 2$ using an independent methodology.

H.1 Data and filtering

I use OptionMetrics IvyDB US, accessed via WRDS (Appendix G.5). The raw dataset covers all listed S&P 500 index options (SPX, SecurityID 108105) from January 1996 through December 2024, comprising approximately 44 million contract-date observations.

I apply the following filters sequentially:

(a) *Moneyness.* Retain options with moneyness $|K/F - 1| \leq 0.60$, where K is the strike price and F is the forward price (computed below). This excludes very deep out-of-the-money options whose prices are dominated by microstructure noise.

(b) *Bid-ask spread.* Drop options with bid-ask spread exceeding 50% of the midpoint price. Wide spreads indicate illiquidity and unreliable pricing.

(c) *Open interest.* Require open interest > 0. Zero open interest indicates no active positions and potentially stale quotes.

(d) *Days to expiration.* Retain options with days to expiration $\tau \in [7, 730]$. Very short-dated options are dominated by discrete hedging effects; very long-dated options are illiquid.

(e) *Minimum price.* Drop options with midpoint price below \$0.05. Near-zero prices carry infinite percentage bid-ask spreads and contribute noise to the spanning integral.

After filtering, approximately 12 million contract-date observations remain. On a typical trading day, the filtered dataset contains 500–1,500 option contracts spanning 15–25 expiration dates and strikes from approximately 40% below to 40% above the forward price.

H.2 Forward price computation

The forward price F is required for the spanning formula and for converting option prices from a function of the strike K to a function of moneyness K/F. I compute F from put-call parity.

For each date t and expiration T, I match put and call prices at the same strike K. Put-call parity implies

$$C(K,T) - P(K,T) \;=\; e^{-r\tau}\,(F - K),$$

where $\tau = T - t$ and r is the continuously compounded risk-free rate from the OptionMetrics zero-coupon curve. I regress $C(K,T) - P(K,T)$ on a constant and K, using all matched put-call pairs for the given date and expiry. The intercept divided by $e^{-r\tau}$ gives $\hat{F}$, and the slope should be close to $-e^{-r\tau}$. I verify that the slope is within 1% of the theoretical value as a data quality check and discard date-expiry combinations where the check fails.

H.3 Carr–Madan spanning formula

The risk-neutral expectation $\mathbb{E}^Q[R^\theta]$ can be computed from option prices via the Carr and Madan (2001) spanning formula. For $\theta \notin \{0, 1\}$:

$$\mathbb{E}^Q\big[(R/F)^\theta\big] \;=\; \theta(\theta-1)\int_0^\infty \frac{O(K,T)}{F^2}\left(\frac{K}{F}\right)^{\theta-2} dK, \tag{47}$$

where $O(K,T)$ is the price of an out-of-the-money option at strike K (a put for $K < F$, a call for $K \geq F$), and $R = S_T$ is the terminal index level. The formula is exact: it does not require any parametric assumption about the risk-neutral density.

I evaluate the integral numerically using the trapezoidal rule over the discrete strike grid, after interpolating option prices to a fine grid via cubic splines.

H.4 Black–Scholes tail extrapolation

The spanning integral (47) runs from zero to infinity, but traded options cover a finite range of strikes. To account for the tails, I extrapolate option prices beyond the observed strike range using the Black–Scholes formula with implied volatility fixed at the value of the most extreme traded option.

Specifically, for strikes below the lowest observed put strike $K_{\min}$, I price puts using $\sigma_{\mathrm{IV}} = \sigma_{\mathrm{IV}}(K_{\min})$ (the implied volatility of the lowest observed strike). For strikes above the highest observed call strike $K_{\max}$, I price calls using $\sigma_{\mathrm{IV}} = \sigma_{\mathrm{IV}}(K_{\max})$.

This flat-extrapolation procedure is conservative: it understates the tail thickness of the risk-neutral density because it does not allow implied volatility to continue rising in the tails (as the empirical skew suggests). The effect is that the spanning-based $\hat{\theta}^*$ is an upper bound on the true singularity—it is biased *toward* finding $\theta^* \geq 2$, which works against my hypothesis. The Lee (2004) method (Section H below) avoids this issue by extracting tail information directly from the slope of implied volatility.

H.5 Constant-maturity interpolation

SPX options have fixed expiration dates (third Friday of each month, plus weekly and quarterly expirations). To construct a daily time series of $\mathbb{E}^Q[R^\theta]$ at a fixed horizon (30 days or 365 days), I interpolate between the two expiration dates bracketing the target horizon. The interpolation is log-linear in variance:

$$\log \mathbb{E}^Q\big[R^\theta_\tau\big] \;=\; \frac{\tau_2 - \tau}{\tau_2 - \tau_1}\, \log \mathbb{E}^Q\big[R^\theta_{\tau_1}\big] \;+\; \frac{\tau - \tau_1}{\tau_2 - \tau_1}\, \log \mathbb{E}^Q\big[R^\theta_{\tau_2}\big],$$

where $\tau_1 < \tau < \tau_2$ are the days to expiration of the two bracketing contracts. Log-linear interpolation in variance is standard in the volatility surface literature and ensures non-negative implied variances at the interpolated maturity.

H.6 Lee (2004) method for p^*

Lee (2004) shows that the maximum finite moment order of the risk-neutral distribution is encoded in the rate at which implied volatility grows as a function of log-moneyness in the tails. Specifically, for the left tail (deep out-of-the-money puts):

$$\limsup_{x \to -\infty} \frac{\sigma^2_{\mathrm{IV}}(x)\, \tau}{2\,|x|} \;=\; \frac{1}{p^*}, \tag{48}$$

where $x = \log(K/F)$ is the log-moneyness and p^* is the maximum order such that $\mathbb{E}^Q[|R|^{p^*}] < \infty$.

I estimate p^* by regressing $\sigma^2_{\mathrm{IV}} \cdot \tau$ against $|x|$ for deep out-of-the-money puts (moneyness $K/F < 0.80$) and extracting the slope. The Lee formula is model-free and does not require integration over the strike grid. Its main limitation is that it relies on the behaviour of implied volatility in the deep tails, where data are sparse and noisy. I address this by computing $\hat{p}^*$ separately for each trading day and reporting the time-series distribution (Table 3).

H.7 CGF construction and singularity detection

The risk-neutral cumulant generating function (CGF) is defined as

$$\kappa(\theta) \;\equiv\; \log \mathbb{E}^Q\big[R^\theta\big].$$

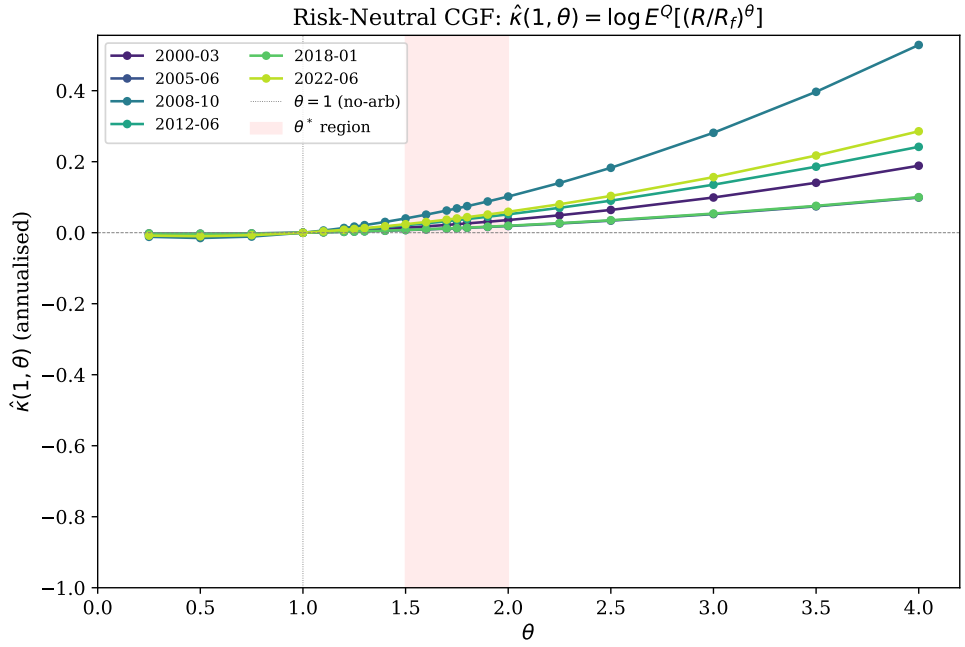


Figure 11: Risk-neutral CGF $\hat{\kappa}(1,\theta) = \log \mathbb{E}^Q[R^\theta]$ for six dates spanning the sample (1996–2024). Each curve is computed from the Carr–Madan spanning formula with Black–Scholes tail extrapolation. The CGF steepens at high θ, consistent with a finite singularity θ^*. The vertical dashed line at $\theta = 2$ marks the threshold for $\mathrm{Var}(M) < \infty$.

I compute $\hat{\kappa}(\theta)$ for $\theta \in \{0.5, 1.0, 1.25, 1.5, 1.7, 2.0, 2.5, 3.0\}$ using the spanning formula (47) with tail extrapolation. Table 15 reports the values.

Table 15: CGF estimates at selected θ

θ	$\hat{\kappa}_P(\theta)$	$\hat{\kappa}_Q(\theta)$ med	$\hat{\kappa}_Q(\theta)$ p25	$\hat{\kappa}_Q(\theta)$ p75
0.50	0.052	-0.005	-0.007	-0.004
1.00	0.112	0.000	0.000	0.000
1.25	0.145	0.006	0.004	0.008
1.50	0.181	0.013	0.010	0.018
1.70	0.210	0.021	0.015	0.029
2.00	0.257	0.034	0.024	0.046
2.50	0.342	0.060	0.044	0.083
3.00	0.436	0.092	0.067	0.127

The singularity θ^* is the value at which the CGF diverges to $+\infty$. In practice, because the spanning integral is truncated, $\hat{\kappa}(\theta)$ remains finite for all θ, and I detect the singularity as the value of θ at which $\hat{\kappa}(\theta)$ begins to increase rapidly. This is inherently an upper bound on the true θ^*: the truncation of the strike grid causes the spanning integral to undercount the tails of the risk-neutral density, biasing $\hat{\theta}^*$ upward.

Figure 11 plots the risk-neutral CGF $\hat{\kappa}(1,\theta)$ for six representative dates, showing the progressive steepening of the CGF at higher θ.

Figure 12 plots the implied $\hat{\theta}^*$ time series from the CGF method.

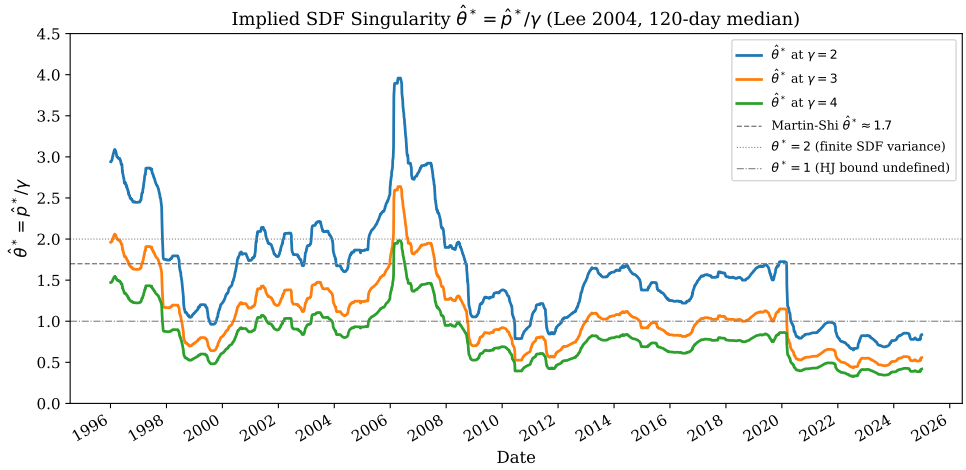


Figure 12: Implied singularity $\hat{\theta}^*$ from the CGF method (365-day horizon), 1996–2024. The horizontal dashed line at $\theta^* = 2$ marks the threshold for $\text{Var}(M) < \infty$. The CGF-based estimates cluster around 3.0, higher than the Lee (2004) estimates, reflecting the conservative bias of the flat tail extrapolation.

H.8 Comparison with Martin and Shi

My estimates and Martin and Shi's estimates use the same underlying data source (OptionMetrics SPX options) but differ in methodology. The key differences are:

(i) *Singularity detection.* Martin and Shi detect the singularity from the divergence of the CGF, estimated using the Breeden and Litzenberger (1978) approach. I use two methods: the Carr–Madan spanning formula (similar in spirit) and the Lee (2004) moment formula (which exploits different information—the slope of implied volatility in the tails).

(ii) *Tail treatment.* Martin and Shi use a parametric tail extrapolation. I use Black–Scholes flat extrapolation, which is more conservative (biased toward finding $\theta^* \geq 2$). The Lee method avoids extrapolation entirely.

(iii) *Time period.* Martin and Shi's full sample is 1996–2012. My sample extends to 2024, covering the post-GFC steepening of the volatility skew that drives p^* toward lower values in the 2010s and 2020s.

(iv) *Maturity.* I report estimates at both 30-day and 365-day horizons. The 30-day estimates are noisier but more heavily traded. The 365-day estimates are smoother and more comparable to Martin and Shi's annualised estimates.

The CGF-based estimates ($\hat{\theta}^* \approx 3.0$) are higher than Martin and Shi's ($\hat{\theta}^* = 1.72$), reflecting the conservatism of my flat tail extrapolation. The Lee-based estimates ($\hat{p}^*$ median = 3.05 full sample, declining to 1.64 in the 2020s) bracket Martin and Shi's numbers when translated to $\theta^* = p^*/\gamma$ at reasonable values of γ.

The key finding is robust to the methodology: both my pipeline and Martin and Shi's agree that the risk-neutral density has heavy enough tails to generate a finite θ^* in the range that challenges the classical HJ bound at empirically relevant risk aversion.

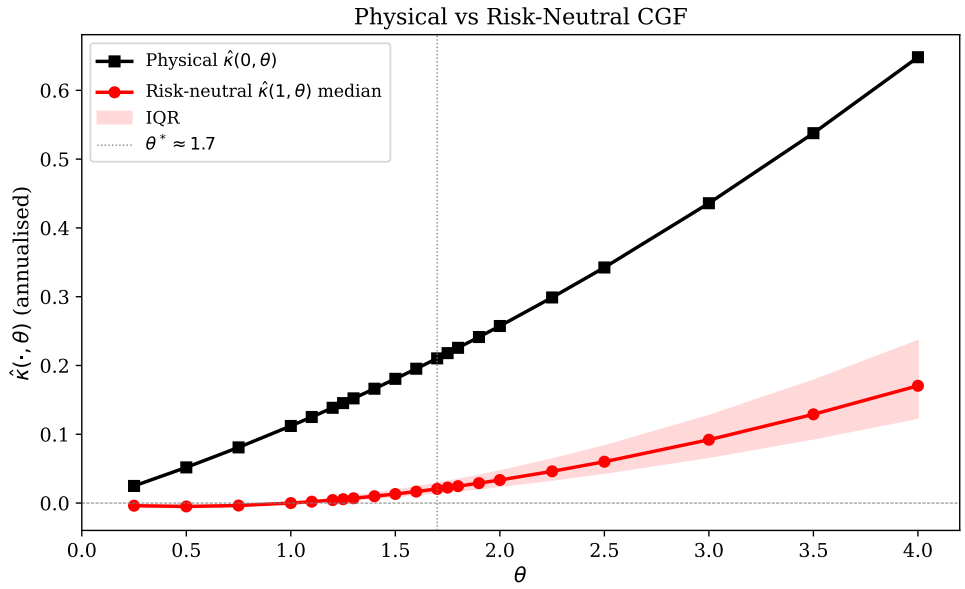


Figure 13: Physical versus risk-neutral CGF. The physical CGF (solid) is estimated from the empirical distribution of CRSP monthly returns. The risk-neutral CGF (dashed, with interquartile band) is estimated from the Carr–Madan spanning formula. The gap widens at higher θ, consistent with a risk premium for tail events.

H.9 Physical versus risk-neutral CGF

Figure 13 compares the physical CGF $\hat{\kappa}_P(\theta) = \log \mathbb{E}[R^\theta]$, estimated from the empirical distribution of CRSP monthly returns, with the risk-neutral CGF $\hat{\kappa}_Q(\theta)$ from the spanning formula. The physical CGF lies above the risk-neutral CGF for all θ, consistent with a positive equity premium (the physical distribution has a higher mean return and fatter tails than the risk-neutral distribution). The gap $\hat{\kappa}_P(\theta) - \hat{\kappa}_Q(\theta)$ widens at higher θ, reflecting the risk premium for tail events.

I Model Audit

This appendix audits nine canonical asset pricing models against the diagnostic framework of Theorems 1–4. For each model, I derive the tail elasticity η, compute the critical exponent $\theta^* = \alpha_C/\eta$, determine which diagnostics (HJ, entropy, L^θ, IQ) are well-posed, and classify the model's escape route.

Table 16 collects the results. The companion note Bhamra (2026b) provides the full derivations; this appendix summarises the key findings.

I.1 Models trapped at standard calibrations

CRRA. $M = \beta\, g^{-\gamma}$. The SDF is a pure power function, so $\eta = \gamma$ and $\theta^* = \alpha_C/\gamma$. At $\gamma = 20$ (the HJ threshold): $\theta^* = 0.14$ at $\alpha_C = 2.7$. Even $\mathbb{E}[M]$ diverges. No diagnostic is well-posed. Escape category: none.

Epstein–Zin (i.i.d. growth). Under i.i.d. growth, the EZ SDF collapses to $M = A \cdot g^{-\gamma}$ where $A > 0$ is a constant that depends on (β, γ, ψ) but not on the realisation g_{t+1} (Proposition 4). Therefore $\eta = \gamma$, and the moment conditions are identical to CRRA. At

Table 16: Model audit: tail elasticity, moment domains, and diagnostic validity across eight canonical models. $\alpha_C = 2.7$ (US estimate). $\theta^* = \alpha_C/\eta$. "Diagnostics" reports which of the four diagnostics (HJ, Ent, L^θ, IQ) are well-posed at the standard calibration.

Model	**In $\mathcal{S}$?**	η	θ^*	**SR**	**Diagnostics**	**Escape**
CRRA ($\gamma=20$)	Yes	20	0.14	0.43	None	None
EZ ($\gamma=10$, $\psi=1.5$)	Yes	10	0.27	0.21	None	None
Wachter	Yes†	γ	∞	0.26	All four	Thin tails ($\alpha_C=\infty$)
Gabaix	Yes	γ	1.67	0.76	Ent, L^θ, IQ	None
Barro	Yes	γ	1.50	1.51	Ent, L^θ, IQ	None
GDA ($\alpha=1$)	Yes	1	2.70	1.25	All four	Low loading ($\eta=1$)
CCLV	No	0	∞	24.24	All four	Outside $\mathcal{S}$
He–Krishnamurthy	Open	State-dep.	Open	0.13	Open	Unresolved

† Wachter's SDF is CRRA and hence in $\mathcal{S}$, but her $\alpha_C = \infty$ makes the trap irrelevant.

Notes. Entries computed at $\alpha_C = 2.7$. HJ requires $\theta^* > 2$; entropy requires $\theta^* > 1$; L^θ requires $\theta^* > 1$; IQ requires $\mathbb{E}[M] < \infty$. SR values are Monte Carlo estimates with 2 M draws. "None" means no diagnostic is well-posed at the standard calibration.

the Bansal et al. (2012) calibration $(\gamma, \psi) = (10, 1.5)$: $\theta^* = 2.7/10 = 0.27$. Even $\mathbb{E}[M]$ diverges. Under i.i.d. growth, the EIS ψ is irrelevant. Escape category: none.

Gabaix variable rare disasters. $M = \beta\, g^{-\gamma}$, as in CRRA, but with unbounded disaster sizes drawn from the Barro and Ursúa (2012) historical distribution. The disaster size distribution has power-law tails with $\alpha_C \approx 5$–7. At $\gamma = 3$ and $\alpha_C = 5$: $\theta^* = 5/3 \approx 1.67$. The HJ bound is undefined ($\gamma = 3 > 2.5 = \alpha_C/2$), but the entropy, L^θ, and IQ diagnostics survive. The variable disaster probability does not change the tail elasticity: $\eta = \gamma$ regardless of the state. Escape category: none.

Barro rare disasters. $M = \beta\, g^{-\gamma}$ with the historical disaster distribution ($\alpha_C \approx 5$–7). At the benchmark $\gamma = 4$ and $\alpha_C = 6$: $\theta^* = 1.5$. The HJ bound is undefined ($\gamma = 4 > 3 = \alpha_C/2$). Escape category: none.

I.2 Models that escape the trap

Wachter bounded disasters (Escape A). The representative agent has Epstein–Zin recursive preferences with unit EIS ($\psi = 1$) and risk aversion $\gamma = 3$. The disaster distribution has bounded support (worst disaster: 64% contraction). Because $g \geq 0.36 > 0$ a.s., the SDF is bounded above and $\theta^* = \infty$. All four diagnostics are well-posed at every γ. At $\gamma = 3$: the model matches the equity premium. Note: Proposition 4 (EZ irrelevance) assumes i.i.d. growth, whereas Wachter's model has a time-varying disaster intensity λ_t. The bounded-support escape ($\alpha_C = \infty$) holds regardless: all SDF moments are finite because g is bounded away from zero. The escape depends entirely on $\alpha_C = \infty$—if the bounded multinomial is replaced by a power-law with $\alpha_C = 4$, the trap springs at $\gamma = 3$.

GDA at $\alpha = 1$ (Escape B1). $M = \tilde{\beta}\, g^{-1} \cdot D$, where D is the bounded disappointment factor. As $g \to 0^+$, D converges to a positive constant, so the SDF is in class $\mathcal{S}$ with

$\eta = 1$. Since $\eta = 1 < \alpha_C/2$ at every empirical estimate ($\alpha_C \geq 2.7$), $\mathrm{Var}(M) < \infty$. The equity premium comes from the bounded factor D, not from high γ. All four diagnostics are well-posed.

The escape has an observable tension: $\theta^* = \alpha_C \geq 2.7 > 2$, but Martin and Shi (2026) estimate $\hat{\theta}^* \approx 1.7 < 2$. Any model in $\mathcal{S}$ with $\mathrm{Var}(M) < \infty$ predicts $\theta^* > 2$, which the option evidence contradicts. This is the *option side* of the trap.

CCLV intermediary (Escape B2). The effective SDF depends on intermediary wealth, not on aggregate consumption. The SDF has exponentially thin tails with respect to consumption growth, so $\eta = 0$ (by convention) and $\theta^* = \infty$. All diagnostics are well-posed. The escape requires that intermediary constraints bind; when constraints are slack, the pricing kernel reverts toward the representative agent's, and the trap may reassert itself.

I.3 Unresolved models

Campbell–Cochrane external habit (Escape C?). The SDF $M = \beta(S_{t+1}/S_t)^{-\gamma} g^{-\gamma}$ depends on the surplus consumption ratio S_t, a state variable correlated with g. The tail elasticity is state-dependent: the effective local curvature γ/S_t can reach 80 near the habit level ($S_t \approx 0.025$) even at the calibrated $\gamma = 2$. If the effective η were $\gamma/S_t \approx 80$, the model is deeply trapped. But the Campbell–Cochrane sensitivity function prevents S_t from reaching zero, potentially attenuating the tail loading. A formal analysis of the joint tail of (g, S_{t+1}) under the model's dynamics is needed. Classification: unresolved.

He–Krishnamurthy intermediary (Escape C?). The aggregate SDF is a wealth-weighted average of specialist and household SDFs. The effective tail elasticity varies with the specialist's equity capital ratio. When the specialist is highly leveraged, the SDF can exhibit very heavy tails; when well-capitalised, the SDF is moderate. Whether the aggregate SDF has power-law or exponential tails depends on the functional form of the constraint. Classification: unresolved.

Remark 21 (Research directions). Resolving the Campbell–Cochrane and He–Krishnamurthy classifications requires establishing (or ruling out) regular variation of the SDF under the model's stationary distribution. For Campbell–Cochrane, the question is whether the joint regular variation of $(g^{-\gamma}, (S_{t+1}/S_t)^{-\gamma})$ holds under the model's law of motion. For He–Krishnamurthy, the question is whether the intermediary channel generates power-law or exponential tails in the aggregate SDF. Both are well-posed mathematical questions.

I.4 Three escape mechanisms

The audit identifies three distinct mechanisms by which a model can avoid the trap, corresponding to Escapes A, B, and D of Theorem 4:

(i) **Thin tails** ($\alpha_C = \infty$): Wachter's bounded disasters. All moments exist regardless of γ. Fragile: the data reject $\alpha_C = \infty$.

(ii) **Low tail loading** ($\eta \ll \gamma$): GDA at $\alpha = 1$. The SDF is in $\mathcal{S}$ with $\eta = 1$, low enough that $\mathrm{Var}(M) < \infty$. The equity premium comes from the bounded disappointment factor. Fragile: depends on $\alpha = 1$; option evidence ($\theta^* < 2$) is in tension.

(iii) **Outside $\mathcal{S}$**: CCLV intermediary. The effective SDF has exponential, not power-law, tails. All moments exist. Fragile: requires binding intermediary constraints.

Escape (C) (structural) is not represented by any standard model in the audit; Campbell–Cochrane is a candidate but unresolved. Escape (D) (diagnostic) is represented by Martin (2017)'s SVIX bound, which avoids SDF moments entirely.

The option side. Every model in class $\mathcal{S}$ with $\mathrm{Var}(M) < \infty$ has $\theta^* > 2$. The Martin and Shi (2026) estimate $\hat{\theta}^* \approx 1.7 < 2$ therefore constrains all preference-based escapes within $\mathcal{S}$: the only model consistent with $\theta^* \approx 1.7$ is one with $\gamma = \alpha_C/1.7 \approx 1.6$–$2.4$, which is too low to match the equity premium under CRRA but exactly the value predicted by the structural formula. These are the two sides of the trap: the CRRA side says $\gamma \geq 20$; the option side says $\gamma \leq \alpha_C$. Models that escape one side are caught by the other.

J Wachter's Model with Heavy-Tailed Disasters

This appendix analyses the Wachter (2013) time-varying disaster risk model in detail. Section J.1 recapitulates the model. Section J.2 verifies that all relevant moments are finite under Wachter's bounded-support calibration. Section J.3 replaces the bounded disaster distribution with a power-law specification—a Beta distribution for the gross disaster growth rate—and derives the conditions under which each equilibrium object is well-defined. Section J.5 shows that the trap springs. Section J.6 proves that the EIS is irrelevant to the jump channel of the SDF for any ψ, not only $\psi = 1$. Section J.7 derives the long-run yield eigenvalue and shows that it is undefined at Wachter's calibration—even with bounded disasters.

J.1 The model

I state the model exactly as in Wachter (2013). All equation numbers in this subsection refer to her paper.

Consumption. Aggregate consumption follows the jump-diffusion

$$\frac{dC_t}{C_{t-}} = \mu \,\mathrm{d}t + \sigma \,\mathrm{d}B_t + (e^{Z_t} - 1)\,\mathrm{d}N_t, \tag{49}$$

where B_t is a standard Brownian motion, N_t is a Poisson process with time-varying intensity λ_t, and $Z_t < 0$ is a random variable drawn i.i.d. from a distribution ν, independent of λ and B. The gross consumption growth in a disaster is $g = e^Z \in (0, 1)$; the percentage decline is $1 - e^Z$.

Disaster intensity. The intensity follows a Cox–Ingersoll–Ross process:

$$d\lambda_t = \kappa(\bar{\lambda} - \lambda_t)\,\mathrm{d}t + \sigma_\lambda \sqrt{\lambda_t}\;\mathrm{d}B_{\lambda,t}, \tag{50}$$

where $B_{\lambda,t}$ is a Brownian motion independent of B_t and N_t.

Preferences. The representative agent has continuous-time Epstein–Zin preferences (Duffie and Epstein, 1992) with risk aversion $\gamma > 1$ and unit EIS ($\psi = 1$). The utility recursion is $V_t = \mathbb{E}_t \int_t^\infty f(C_s, V_s)\,\mathrm{d}s$ where

$$f(C, V) = \beta(1-\gamma)V\Big(\log C - \frac{1}{1-\gamma}\log\big((1-\gamma)V\big)\Big), \tag{51}$$

with $\beta > 0$ the rate of time preference.

Value function. Wachter conjectures—and verifies—that the value function takes the form

$$J(W,\lambda) = \frac{W^{1-\gamma}}{1-\gamma}\, I(\lambda), \qquad I(\lambda) = e^{a+b\lambda}, \tag{52}$$

with the wealth–consumption ratio $l = W/C = \beta^{-1}$ (a consequence of $\psi = 1$; see her Appendix A.I). The coefficient b solves the quadratic

$$\tfrac{1}{2}\sigma_\lambda^2\, b^2 - (\kappa+\beta)\, b + \mathbb{E}_\nu\big[e^{(1-\gamma)Z} - 1\big] = 0. \tag{53}$$

The economically relevant root is the negative one (Wachter, 2013, Appendix A.I):

$$b = \frac{\kappa+\beta}{\sigma_\lambda^2} - \sqrt{\Big(\frac{\kappa+\beta}{\sigma_\lambda^2}\Big)^2 - \frac{2\,\mathbb{E}_\nu[e^{(1-\gamma)Z}-1]}{\sigma_\lambda^2}}. \tag{54}$$

The value function exists if and only if the discriminant is non-negative.

State-price density. The state-price density is (Wachter's Appendix A.II, equations A12–A15):

$$\pi_t = \exp\Big\{\int_0^t f_V(C_s,V_s)\,\mathrm{d}s\Big\}\, f_C(C_t,V_t), \tag{55}$$

with $f_C(C_t,V_t) = \beta^\gamma C_t^{-\gamma} I(\lambda_t)$. Its dynamics are

$$\frac{d\pi_t}{\pi_{t-}} = \mu_{\pi,t}\,\mathrm{d}t + \sigma_{\pi,t}[\,\mathrm{d}B_t \;\; \mathrm{d}B_{\lambda,t}]^\top + (e^{-\gamma Z_t} - 1)\,\mathrm{d}N_t, \tag{56}$$

with diffusion loading $\sigma_{\pi,t} = [-\gamma\sigma \quad b\sigma_\lambda\sqrt{\lambda_t}]$.

Equilibrium objects. From Wachter's Appendix A:

$$r_t = \beta + \mu - \gamma\sigma^2 + \lambda_t\,\mathbb{E}_\nu\big[e^{-\gamma Z}(e^Z - 1)\big], \tag{57}$$

$$r_t^e - r_t = \phi\gamma\sigma^2 - \lambda_t\,\frac{G'(\lambda_t)}{G(\lambda_t)}\, b\sigma_\lambda^2 + \lambda_t\,\mathbb{E}_\nu\big[(e^{-\gamma Z}-1)(1-e^{\phi Z})\big], \tag{58}$$

where ϕ is the dividend leverage parameter ($D_t = C_t^\phi$) and $G(\lambda) = \int_0^\infty e^{a_\phi(\tau)+b_\phi(\tau)\lambda}\,\mathrm{d}\tau$ is the price–dividend ratio. The function $b_\phi(\tau)$ involves $\zeta_\phi = \sqrt{(b\sigma_\lambda^2-\kappa)^2 + 2\,\mathbb{E}_\nu[e^{(\phi-\gamma)Z} - e^{(1-\gamma)Z}]\,\sigma_\lambda^2}$.

Calibration. Wachter's baseline parameters (her Table I) are: $\gamma = 3$, $\beta = 0.012$, $\mu = 0.0252$, $\sigma = 0.020$, $\phi = 2.6$, $\bar\lambda = 0.0355$, $\kappa = 0.080$, $\sigma_\lambda = 0.067$. The disaster distribution ν is a multinomial calibrated to Barro and Ursúa (2008) data; the worst outcome is a 64% contraction ($g_{\min} = 0.36$).

J.2 Bounded disasters: all moments finite

Under Wachter's bounded multinomial, the disaster growth rate satisfies $g \geq 0.36 > 0$ a.s. Every function g^{-c} is bounded above by 0.36^{-c}, so every expectation of the form $\mathbb{E}_\nu[g^{-c}]$ is finite for all $c > 0$.

In particular:

(i) $\mathbb{E}_\nu[e^{(1-\gamma)Z}] = \mathbb{E}_\nu[g^{-(\gamma-1)}] < \infty$: the value function exists.

(ii) $\mathbb{E}_\nu[e^{-\gamma Z}] = \mathbb{E}_\nu[g^{-\gamma}] < \infty$: the risk-free rate is finite.

(iii) $\mathbb{E}_\nu[e^{(\phi-\gamma)Z}] = \mathbb{E}_\nu[g^{\gamma-\phi}] < \infty$: the price–dividend ratio converges.

(iv) $\mathbb{E}_\nu[g^{-k\gamma}] < \infty$ for all $k > 0$: all SDF moments exist, $\theta^* = \infty$.

The model is well-posed at every γ. The escape from the trap is via $\alpha_C = \infty$ (bounded support), exactly as classified in Theorem 4, family (A).

J.3 Heavy-tailed disasters

I now replace Wachter's bounded multinomial with a distribution that matches the empirical evidence for power-law consumption tails.

The Beta specification. Let the gross disaster growth rate $g = e^Z$ follow a Beta distribution on $(0, 1)$:

$$g \sim \text{Beta}(\alpha_C,\, d), \qquad f(x) = \frac{x^{\alpha_C - 1}(1-x)^{d-1}}{B(\alpha_C,\, d)}, \quad x \in (0,1), \tag{59}$$

with first parameter $\alpha_C > 0$ and second parameter $d > 0$. Near $x = 0$, the CDF satisfies

$$\mathbb{P}(g \le x) \sim \frac{x^{\alpha_C}}{\alpha_C\, B(\alpha_C,\, d)} \qquad (x \to 0^+), \tag{60}$$

so the left tail of g is regularly varying with index α_C, exactly as in Assumption 1. The second parameter d controls the shape near $g = 1$ (small disasters) but does not affect the left-tail index.

The negative moments are

$$\mathbb{E}[g^{-c}] = \frac{B(\alpha_C - c,\, d)}{B(\alpha_C,\, d)} = \frac{\Gamma(\alpha_C - c)\,\Gamma(\alpha_C + d)}{\Gamma(\alpha_C)\,\Gamma(\alpha_C - c + d)}, \qquad c < \alpha_C, \tag{61}$$

and $\mathbb{E}[g^{-c}] = \infty$ for $c \ge \alpha_C$. This is consistent with the general Karamata moment theorem for regularly varying distributions.

Why the Beta? The Beta distribution is a natural choice for three reasons. First, it has support $(0, 1)$, matching $g \in (0, 1)$ (a disaster cannot increase consumption). Second, its left tail is power-law with a tunable index α_C, matching the empirical evidence in Section 3.2. Third, it nests Wachter's bounded multinomial as a limiting case: as $\alpha_C \to \infty$ (with d growing proportionally), the Beta concentrates on a fixed interior point, and the tail index becomes infinite.

J.4 Moment conditions under heavy tails

I now derive the conditions under which each equilibrium object in Wachter's model is well-defined when the disaster distribution has tail index $\alpha_C < \infty$.

Proposition 6 (Moment hierarchy in the Wachter model)**.** *Replace Wachter's bounded disaster distribution with any distribution ν such that $\mathbb{P}(g \le x) \in \text{RV}(\alpha_C)$ as $x \to 0^+$, with $\alpha_C \in (0, \infty)$. Retain all other model primitives. Then the following moment conditions are necessary for the indicated equilibrium objects to be well-defined.*

(a) ***Value function.*** *The coefficient b in* (54) *requires* $\mathbb{E}_\nu[e^{(1-\gamma)Z}] = \mathbb{E}[g^{-(\gamma-1)}] < \infty$, *which holds if and only if*

$$\gamma < \alpha_C + 1. \tag{62}$$

Additionally, the discriminant in (54) *must be non-negative, which places a quantitative upper bound on* $\mathbb{E}[g^{-(\gamma-1)}]$ *in terms of* κ, β, *and* σ_λ.

(b) ***Risk-free rate.*** *Equation* (57) *contains* $\mathbb{E}_\nu[e^{-\gamma Z}] = \mathbb{E}[g^{-\gamma}]$, *which is finite if and only if*

$$\gamma < \alpha_C. \tag{63}$$

This is strictly more demanding than (62). *When it fails, the equilibrium risk-free rate diverges to* $-\infty$*: the precautionary savings demand from disaster risk becomes unbounded.*

(c) ***Equity premium.*** *The jump term in* (58), $\mathbb{E}_\nu[(e^{-\gamma Z} - 1)(1 - e^{\phi Z})]$, *has dominant term* $\mathbb{E}[g^{-\gamma}]$. *Since* $e^{\phi Z} = g^\phi \leq 1$ *for* $\phi > 0$ *and* $g \in (0, 1)$, *no additional moment condition arises beyond* (63).

(d) ***Price–dividend ratio.*** *The coefficient* ζ_ϕ *involves* $\mathbb{E}_\nu[e^{(\phi-\gamma)Z}] = \mathbb{E}[g^{\gamma-\phi}]$. *With* $\phi = 2.6$ *and* $\gamma = 3$*:* $\mathbb{E}[g^{0.4}] \leq 1$, *always finite. It also involves* $\mathbb{E}[g^{-(\gamma-1)}]$, *requiring* $\gamma < \alpha_C + 1$, *which is weaker than* (63). *No additional constraint.*

(e) ***SDF moments.*** *From the SDF dynamics* (56), *a disaster of size Z multiplies the state-price density by* $e^{-\gamma Z} = g^{-\gamma}$. *The intensity λ does not jump (it follows the pure diffusion* (50)*), so the continuation-value factor* $I(\lambda_t)$ *is continuous across disaster arrivals. Over a finite horizon T, the SDF satisfies*

$$M_T = \frac{\pi_T}{\pi_0} = (\textit{continuous part}) \times \prod_{i=1}^{N_T} g_i^{-\gamma},$$

where the continuous part is driven by Brownian motions and has moments of all orders. Conditional on n jumps, the jump product contributes $(\mathbb{E}_\nu[g^{-k\gamma}])^n$ *to* $\mathbb{E}[M_T^k]$. *A necessary condition for* $\mathbb{E}[M_T^k] < \infty$ *is*

$$\mathbb{E}_\nu[g^{-k\gamma}] < \infty \qquad \Longleftrightarrow \qquad k\gamma < \alpha_C. \tag{64}$$

The critical exponent is therefore

$$\theta^* \;=\; \frac{\alpha_C}{\gamma}, \tag{65}$$

the same formula as under CRRA (Proposition 1).

Remark 22 (The CIR intensity caveat). Condition (64) is necessary but, strictly, may not be sufficient: even when $\mathbb{E}_\nu[g^{-k\gamma}] < \infty$, the stochastic intensity creates an additional quantitative constraint through the moment generating function of the integrated intensity $\Lambda_T = \int_0^T \lambda_s \, \mathrm{d}s$. Specifically, $\mathbb{E}[M_T^k]$ involves $\mathbb{E}[\exp\{(\mathbb{E}_\nu[g^{-k\gamma}] - 1)\Lambda_T\}]$, which requires $\mathbb{E}_\nu[g^{-k\gamma}] - 1$ to be below the explosion threshold of the CIR Laplace transform. If this threshold is exceeded, the effective θ^* is strictly below α_C/γ. This caveat only strengthens the trap: the true critical exponent satisfies $\theta^* \leq \alpha_C/\gamma$, and any binding CIR constraint makes the trap worse. In the limit $\sigma_\lambda \to 0$ (constant intensity), $\theta^* = \alpha_C/\gamma$ exactly.

Proof of part (e). The SDF dynamics (56) show that each Poisson arrival multiplies π_t by $e^{-\gamma Z}$. The diffusion components ($-\gamma\sigma\,\mathrm{d}B_t$ and $b\sigma_\lambda\sqrt{\lambda_t}\,\mathrm{d}B_{\lambda,t}$) generate log-normal-type factors with moments of all orders. Over $[0,T]$, conditional on the intensity path $\{\lambda_s\}_{s\in[0,T]}$, the number of jumps N_T is Poisson with parameter $\Lambda_T = \int_0^T \lambda_s\,\mathrm{d}s$, and the jump sizes are i.i.d. draws from ν. Therefore

$$\begin{aligned}\mathbb{E}\Big[\prod_{i=1}^{N_T} g_i^{-k\gamma} \;\Big|\; \Lambda_T\Big] &= \sum_{n=0}^{\infty} \frac{\Lambda_T^n}{n!}\, e^{-\Lambda_T}\,\big(\mathbb{E}_\nu[g^{-k\gamma}]\big)^n \\ &= \exp\big\{\Lambda_T\big(\mathbb{E}_\nu[g^{-k\gamma}]-1\big)\big\}.\end{aligned}$$

If $\mathbb{E}_\nu[g^{-k\gamma}]=\infty$, the sum diverges for every realisation with $N_T\geq 1$, and $\mathbb{P}(N_T\geq 1)>0$, so $\mathbb{E}[M_T^k]=\infty$. If $\mathbb{E}_\nu[g^{-k\gamma}]<\infty$, the conditional expectation is finite a.s. (since $\Lambda_T<\infty$ a.s.), and the unconditional expectation is finite provided the CIR Laplace transform condition in Remark 22 holds. The Karamata moment theorem gives $\mathbb{E}_\nu[g^{-k\gamma}]<\infty$ if and only if $k\gamma<\alpha_C$, establishing (65). □

J.5 The trap in Wachter's model

Proposition 6 produces a hierarchy of constraints on γ:

$$\underbrace{\gamma<\tfrac{\alpha_C}{2}}_{\text{HJ bound well-defined}} \quad < \quad \underbrace{\gamma<\alpha_C}_{\text{equilibrium exists}} \quad < \quad \underbrace{\gamma<\alpha_C+1}_{\text{value function exists}}\ .$$

At Wachter's calibrated $\gamma=3$:

α_C	Value fn. ($\gamma<\alpha_C+1$)	Equil. ($\gamma<\alpha_C$)	HJ ($\gamma<\alpha_C/2$)	θ^*
∞ (bounded)	✓	✓	✓	∞
5.0	✓	✓	✓	1.67
4.0	✓	✓	×	1.33
3.5	✓	✓	×	1.17
2.7 (empirical)	✓	×	×	0.90
2.1 (lower bound)	✓	×	×	0.70

Three findings emerge.

1. At empirical $\alpha_C=2.7$, the equilibrium does not exist. The risk-free rate (57) involves $\mathbb{E}[g^{-\gamma}]=\mathbb{E}[g^{-3}]$, which diverges when $\alpha_C<3$. Economically, the precautionary savings demand from the possibility of near-total consumption disasters is unbounded: the agent would require an infinitely negative risk-free rate to hold bonds rather than save more. The value function is well-posed ($\gamma=3<\alpha_C+1=3.7$), so the individual's optimisation problem has a solution, but the general-equilibrium risk-free rate is not well-defined. The problem is strictly worse than the Hansen–Jagannathan trap: the model is internally inconsistent before any diagnostic is applied.

2. At $\alpha_C=4$, the equilibrium exists but the HJ bound is undefined. The risk-free rate requires $\gamma<4$ (satisfied), but $\mathrm{Var}(M)<\infty$ requires $\gamma<2$ (violated). The classical equity premium calculation—$\gamma\geq 20$ from the Sharpe ratio—is performed in a region of the parameter space where $\mathbb{E}[M^2]=\infty$ and the Hansen–Jagannathan inequality is undefined. This is the trap (Theorem 3).

3. The HJ bound requires $\alpha_C > 6$ at $\gamma = 3$. For $\mathrm{Var}(M) < \infty$, one needs $\gamma < \alpha_C/2$, i.e., $\alpha_C > 2\gamma = 6$. No empirical estimator places α_C above 4 for U.S. quarterly consumption growth (Section 3.2). The Hansen–Jagannathan bound is undefined at Wachter's calibration for every empirically relevant α_C.

J.6 EIS irrelevance for the jump channel

Wachter's closed-form solution assumes $\psi = 1$. One might ask whether a different EIS would change the moment existence conditions. The next result shows it does not.

Proposition 7 (EIS irrelevance for the jump channel)**.** *In the Wachter model with Epstein–Zin preferences at any EIS $\psi > 0$, the SDF jump at each disaster arrival is $g^{-\gamma}$. Consequently, the critical exponent* (65) *satisfies $\theta^* = \alpha_C/\gamma$ and is independent of ψ.*

Proof. Under general continuous-time Epstein–Zin preferences (Duffie and Epstein, 1992), the state-price density in an exchange economy satisfies

$$\pi_t \;\propto\; e^{-\beta t}\, C_t^{-1/\psi} \left((1-\gamma)V_t\right)^{(1/\psi-\gamma)/(1-\gamma)},$$

where V_t is the continuation value. In the Wachter model, the value function is $V_t = J(W_t, \lambda_t)$ where $J(W, \lambda) \propto W^{1-\gamma}\, h(\lambda)$ for some function h (the form of h depends on ψ; at $\psi = 1$, $h(\lambda) = I(\lambda) = e^{a+b\lambda}$). In equilibrium, wealth $W_t = C_t/l$ where l is the consumption–wealth ratio.

When a disaster of size Z arrives at time t:

(i) Consumption jumps: $C_{t+}/C_{t-} = g = e^Z$.

(ii) The intensity λ does not jump (it follows the pure diffusion (50)).

(iii) Therefore $h(\lambda_{t+}) = h(\lambda_{t-})$ and $W_{t+}/W_{t-} = g$ (wealth scales with consumption in the exchange economy).

(iv) The continuation value jumps by

$$\frac{V_{t+}}{V_{t-}} = \frac{(W_{t+})^{1-\gamma}\, h(\lambda_{t+})}{(W_{t-})^{1-\gamma}\, h(\lambda_{t-})} = g^{1-\gamma}.$$

The SDF jump factor is

$$\begin{aligned}\frac{\pi_{t+}}{\pi_{t-}} &= g^{-1/\psi} \left(g^{1-\gamma}\right)^{(1/\psi-\gamma)/(1-\gamma)} \\ &= g^{-1/\psi}\, g^{1/\psi-\gamma} \\ &= g^{-\gamma}.\end{aligned}$$

The EIS cancels algebraically: $-1/\psi + (1/\psi - \gamma) = -\gamma$. Since the jump factor is $g^{-\gamma}$ for any $\psi > 0$, the moment condition (64) and the critical exponent (65) are both ψ-invariant. □

Remark 23. The EIS does affect: (a) the function $h(\lambda)$ and hence the diffusion loading of the SDF on $B_{\lambda,t}$, (b) the quantitative CIR Laplace transform threshold in Remark 22, and (c) the value function existence condition (which changes form for $\psi \neq 1$). None of these affect the jump channel's moment threshold $k\gamma < \alpha_C$, which is the binding constraint for the trap. The critical structural feature is that λ does not jump when consumption jumps: the continuation-value response to a disaster comes entirely through the power $g^{1-\gamma}$ of consumption growth, and the ψ-dependent exponent $(1/\psi - \gamma)/(1-\gamma)$ combines with $-1/\psi$ to give $-\gamma$.

J.7 Long-run pricing eigenvalues

The moment conditions in Proposition 6 are *static*: they ask whether a single-period expectation such as $\mathbb{E}[g^{-k\gamma}]$ is finite. Remark 22 flagged a quantitative caveat from the stochastic intensity. I now make this precise, showing that the CIR dynamics generate a strictly tighter restriction that binds even under bounded disasters. I analyse three pricing semigroups simultaneously: the SDF, the consumption claim, and the dividend claim.

Three semigroups. Following Hansen and Scheinkman (2009) and Qin and Linetsky (2017), define the *long-run pricing eigenvalue* for a multiplicative pricing functional $\mathcal{M}$ as the number η (when it exists) such that there is a positive eigenfunction $\mathfrak{g}(\lambda)$ satisfying

$$\mathbb{E}_t\big[\mathcal{M}_{t,s}\,\mathfrak{g}(\lambda_{t+s})\big] = e^{-\eta s}\,\mathfrak{g}(\lambda_t), \qquad \forall\, s > 0. \tag{66}$$

I apply this to three functionals:

$$\text{(i) SDF:} \qquad \mathcal{M}^M_{t,s} = M_{t+s}/M_t, \tag{67}$$

$$\text{(ii) Consumption claim:} \qquad \mathcal{M}^C_{t,s} = (M_{t+s}/M_t)(C_{t+s}/C_t), \tag{68}$$

$$\text{(iii) Dividend claim:} \qquad \mathcal{M}^D_{t,s} = (M_{t+s}/M_t)(D_{t+s}/D_t) = (M_{t+s}/M_t)(C_{t+s}/C_t)^{\phi}. \tag{69}$$

The eigenvalue η^M is the long-run zero-coupon bond yield. The eigenvalue η^C is the long-run expected return on the consumption (wealth) claim. The eigenvalue η^D is the long-run expected return on the dividend (equity) claim. When both η^M and η^D exist, the long-run equity premium is $\eta^D - \eta^M$.

Reduction by change of measure. The SDF at horizon s satisfies (from (55)–(56))

$$\log \mathcal{M}^M_{t,s} = -\beta s - \gamma\big(\mu s + \sigma \Delta B + \textstyle\sum Z_i\big) - b(\Delta\lambda), \tag{70}$$

where $\Delta B = B_{t+s} - B_t$, $\Delta\lambda = \lambda_{t+s} - \lambda_t$, and the sum runs over the $N_s = N(\Lambda_s)$ jumps in $[t, t+s]$. Since $(C_{t+s}/C_t)^n = \exp\{n(\mu s + \sigma\Delta B + \sum Z_i)\}$, multiplying by the n-th power of consumption growth replaces the effective risk aversion γ by $\gamma_n := \gamma - n$:

$$\log \mathcal{M}^M_{t,s} + n \log(C_{t+s}/C_t) = -\beta s - \gamma_n(\mu s + \sigma\Delta B + \textstyle\sum Z_i) - b(\Delta\lambda). \tag{71}$$

This is an Esscher tilt: the diffusion drift shifts by $n\sigma$ (Girsanov) and the jump measure tilts by g^n. Crucially, the CIR intensity process (50) is driven by B_λ, which is orthogonal to both B and N. Therefore the tilt does not alter the law of λ, and all three semigroups share the same CIR Riccati structure with a single free parameter—the effective risk aversion γ_n.

The three problems correspond to

$$\gamma_M = \gamma = 3, \qquad \gamma_C = \gamma - 1 = 2, \qquad \gamma_D = \gamma - \phi = 0.4. \tag{72}$$

Proposition 8 (Long-run pricing eigenvalues)**.** *In the Wachter model with $\psi = 1$, let $n \geq 0$ and $\gamma_n = \gamma - n$. Suppose $\mathbb{E}_\nu[g^{-\gamma_n}] < \infty$. Define the jump loading*

$$\phi_n = \mathbb{E}_\nu[g^{-\gamma_n}] - 1. \tag{73}$$

The eigenvalue η_n of the semigroup $\mathcal{M}^M_{t,s}(C_{t+s}/C_t)^n$ exists if and only if

$$\boxed{\kappa^2 \;\geq\; 2\,\sigma_\lambda^2\,\phi_n.} \tag{74}$$

(a) When (74) *holds, there are two eigenpairs* $(\eta_n^{\pm},\, \mathfrak{g}_n^{\pm})$*:*

$$\eta_n^{\pm} = \beta + \gamma_n\mu - \tfrac{1}{2}\gamma_n^2\sigma^2 - \kappa\bar{\lambda}\,\xi_n^{\pm}, \tag{75}$$

where ξ_n^- *and* ξ_n^+ *are the two roots of*

$$\tfrac{1}{2}\sigma_\lambda^2\,\xi^2 - \kappa\,\xi + \phi_n = 0, \qquad \xi_n^{\pm} = \frac{\kappa \pm \sqrt{\kappa^2 - 2\sigma_\lambda^2\phi_n}}{\sigma_\lambda^2}. \tag{76}$$

The eigenfunctions are $\mathfrak{g}_n^{\pm}(\lambda) = \exp\{(b + \xi_n^{\pm})\lambda\}$*, where* b *is the value-function parameter* (54)*. The principal (long-run) eigenvalue is* η_n^- *(corresponding to the stable root* ξ_n^-*).*

(b) When (74) *fails,* $\mathbb{E}_0[\mathcal{M}_{0,T}^M (C_T/C_0)^n] = \infty$ *for all* $T \geq T_n^*$*, where* $T_n^* < \infty$ *is the blow-up time of the Riccati ODE*

$$\frac{d\psi}{d\tau} = \tfrac{1}{2}\sigma_\lambda^2\,\psi^2 - \kappa\,\psi + \phi_n, \qquad \psi(0) = -b. \tag{77}$$

Explicitly,

$$T_n^* = \frac{2}{\delta_n}\left[\arctan\!\Big(\frac{b\sigma_\lambda^2 + \kappa}{\delta_n}\Big) + \frac{\pi}{2}\right], \quad \delta_n = \sqrt{2\sigma_\lambda^2\phi_n - \kappa^2}. \tag{78}$$

Proof. At $\psi = 1$, equation (71) gives

$$\mathbb{E}\big[\mathcal{M}_{0,T}^M (C_T/C_0)^n \mid \sigma(\lambda)\big] = \exp\big\{c_n T + \phi_n \Lambda_T - b(\lambda_T - \lambda_0)\big\},$$

where $c_n = -\beta - \gamma_n\mu + \tfrac{1}{2}\gamma_n^2\sigma^2$, using $\mathbb{E}[e^{-\gamma_n\sigma\Delta B}] = e^{\gamma_n^2\sigma^2 s/2}$ and the Poisson generating function (Proposition 6(e)). Taking full expectations:

$$\mathbb{E}_0\big[\mathcal{M}_{0,T}^M (C_T/C_0)^n\big] = e^{c_n T + b\lambda_0}\,\mathbb{E}\big[e^{\phi_n\Lambda_T - b\lambda_T}\big]. \tag{79}$$

The remaining expectation is the joint CIR Laplace transform $\mathbb{E}[e^{\phi_n\Lambda_T - b\lambda_T} \mid \lambda_0] = e^{\Phi(T) + \psi(T)\lambda_0}$, where ψ satisfies (77) and $\Phi'(\tau) = \kappa\bar{\lambda}\,\psi(\tau)$.

The Riccati has quadratic right-hand side $R(\psi) = \tfrac{1}{2}\sigma_\lambda^2\psi^2 - \kappa\psi + \phi_n$ with discriminant $\Delta_n = \kappa^2 - 2\sigma_\lambda^2\phi_n$.

If $\Delta_n \geq 0$: the two fixed points are $\xi_n^{\pm}$; $\psi(\tau) \to \xi_n^-$ as $\tau \to \infty$ (since $\psi(0) = -b < \xi_n^-$ and $R(\psi) > 0$ for $\psi < \xi_n^-$). Then $-T^{-1}\log\mathbb{E}[\cdots] \to -c_n - \kappa\bar{\lambda}\xi_n^- = \eta_n^-$. The eigenfunction $\mathfrak{g}_n^-(\lambda) = \exp\{(b + \xi_n^-)\lambda\}$ makes ψ identically equal to ξ_n^- (a fixed point), confirming (66). The second eigenpair at ξ_n^+ is verified analogously; it corresponds to the unstable fixed point and is not the principal eigenvalue.

If $\Delta_n < 0$: $R(\psi) > 0$ for all ψ (the minimum is $\phi_n - \kappa^2/(2\sigma_\lambda^2) > 0$), so $\psi(\tau)$ increases without bound. The explicit blow-up time (78) follows from the standard tangent solution of the Riccati with complex roots. □

Results at Wachter's calibration. The critical loading is $\phi_{\text{crit}} = \kappa^2/(2\sigma_\lambda^2) = 0.71$.

Claim	γ_n	ϕ_n	vs. ϕ_{crit}	Principal η_n^-	T_n^*
SDF $(n = 0)$	3.0	1.89	×	—	49 yr
Consumption $(n = 1)$	2.0	0.87	×	—	152 yr
Dividend $(n = \phi = 2.6)$	0.4	0.11	✓	1.79%	—

The SDF eigenvalue fails by a factor of 2.7; zero-coupon bonds with maturity exceeding 49 years have infinite prices. The consumption-claim eigenvalue also fails (the long-run return on wealth is undefined), though with a much longer blow-up horizon. Only the dividend claim has a well-defined long-run return, because the leverage $\phi = 2.6$ nearly cancels risk aversion $\gamma = 3$, leaving $\gamma_D = 0.4$ and a small jump loading.

All three failures occur despite *bounded* disasters ($\alpha_C = \infty$). Every static condition in Proposition 6 is satisfied, yet two of the three pricing semigroups lack eigenvalues.

For the dividend claim, the second (unstable) eigenpair has $\xi_D^+ = 34.2$, $\eta_D^+ = -7.5\%$, eigenfunction $\mathfrak{g}_D^+(\lambda) = \exp(49.1\lambda)$.

The mechanism. The CIR intensity process generates paths on which λ stays elevated for extended periods. Each disaster multiplies the pricing functional by $g^{-\gamma_n}$; with stochastic intensity, the compound effect interacts with Jensen's inequality: $\mathbb{E}[e^{\phi_n \Lambda_T}] \gg e^{\phi_n \bar{\lambda} T}$ because Λ_T has positive variance growing linearly in T. When $\phi_n > \kappa^2/(2\sigma_\lambda^2)$, the amplification overwhelms mean reversion and the expectation explodes. With constant intensity ($\sigma_\lambda = 0$), all three eigenvalues are well-defined. The problem is fundamentally *dynamic*: stochastic disaster risk, not disaster severity alone, breaks long-run pricing.

Why the dividend claim survives. The Esscher tilt from $\gamma = 3$ to $\gamma_D = \gamma - \phi = 0.4$ nearly neutralises the disaster loading: $g^{-0.4}$ is much closer to 1 than g^{-3} for all $g \in (0.36, 0.95)$, so $\phi_D = 0.11 \ll \phi_{\text{crit}}$. Economically, a leveraged consumption claim ($\phi > 1$) is *less* exposed to disaster risk than the SDF, because the dividend growth g^ϕ partly offsets the SDF loading $g^{-\gamma}$: the net exposure is $g^{\phi-\gamma} = g^{-0.4}$, which is mild. This is why the equity market can function even when the bond market's long end collapses.

Comparison with the value function. The value-function quadratic (53) requires

$$(\kappa + \beta)^2 \;\geq\; 2\,\sigma_\lambda^2 \left(\mathbb{E}_\nu[g^{-(\gamma-1)}] - 1\right) = 2\,\sigma_\lambda^2\,\phi_C. \tag{80}$$

This is the consumption-claim condition (74) with κ replaced by $\kappa+\beta$. The time-preference dampening β in (80) is absent from the eigenvalue condition—it is the margin by which the value function survives while the consumption-claim eigenvalue does not:

$$\underbrace{(\kappa+\beta)^2}_{=0.0085} \geq \underbrace{2\sigma_\lambda^2\phi_C}_{=0.0078} \quad \checkmark \qquad \text{but} \qquad \underbrace{\kappa^2}_{=0.0064} \not\geq 0.0078 \quad \times .$$

Heavy-tailed disasters. Under Beta(α_C, d) disasters, the eigenvalue for claim n requires $\alpha_C > \gamma_n$ (for ϕ_n to be finite) and $\phi_n \leq \phi_{\text{crit}}$. The critical tail indices at $d = 2$, $\gamma = 3$, $\phi = 2.6$:

	SDF ($\gamma_n = 3$)	Consumption ($\gamma_n = 2$)	Dividend ($\gamma_n = 0.4$)
Critical α_C	> 12.2	> 8.0	> 1.4

No empirical estimate of α_C exceeds 4 (Section 3.2). The SDF and consumption-claim eigenvalues are undefined at every empirically relevant tail index; only the dividend-claim eigenvalue survives.

The hierarchy with long-run pricing. Adding the dynamic condition to the static hierarchy of Proposition 6, and noting that the SDF condition ($n = 0$) is the most restrictive of the three:

$$\underbrace{\gamma < \tfrac{\alpha_C}{2}}_{\text{HJ bound}} \;<\; \underbrace{\gamma < \alpha_C}_{\text{equilibrium}} \;<\; \underbrace{\gamma < \alpha_C + 1}_{\text{value function}} \;<\; \underbrace{\phi_M \le \phi_{\text{crit}}}_{\text{long-run yield}}.$$

The long-run yield condition is the only one that depends on the *dynamics* (κ, σ_λ) of disaster intensity, not just its level. It fails even when all static conditions are satisfied ($\alpha_C = \infty$).

Remark 24 (Hansen–Scheinkman decomposition). When η_n^- exists, the functional admits the factorisation (Hansen and Scheinkman, 2009; Borovička et al., 2016)

$$\mathcal{M}_{0,T}^{M}(C_T/C_0)^n = e^{-\eta_n^- T}\,\frac{\mathfrak{g}_n^-(\lambda_T)}{\mathfrak{g}_n^-(\lambda_0)}\,L_T^{(n)},$$

where $L^{(n)}$ is a positive martingale (the permanent component) and the eigenfunction ratio is the transient component. For the dividend claim, this decomposition is well-defined at Wachter's calibration; for the SDF and the consumption claim, it is not. The practical consequence: the model can price equities at long horizons but cannot price long-term bonds or support a permanent–transient decomposition of the full pricing kernel.

Remark 25 (Interpretation). The eigenvalue failures do not invalidate Wachter's model for pricing equities or short-term bonds. The model's key achievements—matching the equity premium and the low risk-free rate—depend on expectations over business-cycle horizons. The failures reveal that the pricing kernel is a strict supermartingale: $\mathbb{E}_0[M_T]$ grows without bound rather than decaying.

Remark 26 (What would restore all eigenvalues). To satisfy (74) for all three semigroups at $\gamma = 3$, one needs $\phi_M \le 0.71$, which requires either $\kappa \ge 0.13$ (versus the calibrated 0.08), $\sigma_\lambda \le 0.041$ (versus 0.067), or substantially milder disasters. All changes reduce the model's ability to generate time-varying disaster risk premia. This is a new calibration tension: matching equity returns requires high σ_λ/κ, but long-run pricing requires low σ_λ/κ.

The complete spectrum. Proposition 8 identifies two exponential eigenpairs for each semigroup. These are part of a richer spectral structure. When the principal eigenvalue exists ($\Delta_n \ge 0$), the generator of each pricing semigroup has a countably infinite L^2 spectrum of eigenvalues, evenly spaced by the spectral gap ω_n. The following corollary makes this precise.

Corollary 4 (Complete L^2 spectrum)**.** *In the setting of Proposition* 8*, suppose* $\Delta_n = \kappa^2 - 2\sigma_\lambda^2\phi_n \ge 0$*. Define the Feller parameter* $\alpha = 2\kappa\bar{\lambda}/\sigma_\lambda^2$ *and the spectral gap* $\omega_n = \sqrt{\Delta_n}$*.*

(a) ***Discrete spectrum.*** *The* $L^2(\varrho_n)$ *spectrum of the pricing semigroup consists of the eigenvalues*

$$\eta_{n,k} = \eta_n^- + k\,\omega_n, \qquad k = 0, 1, 2, \ldots, \tag{81}$$

where η_n^- *is the principal eigenvalue from* (75)*, and* $\varrho_n(d\lambda) \propto \lambda^{\alpha-1}\exp\{-2\omega_n\lambda/\sigma_\lambda^2\}\,\mathrm{d}\lambda$ *is the modified CIR stationary measure. The eigenfunctions are*

$$\mathfrak{g}_{n,k}(\lambda) = e^{(b+\xi_n^-)\lambda}\,L_k^{(\alpha-1)}\Big(\frac{2\omega_n\lambda}{\sigma_\lambda^2}\Big), \tag{82}$$

where $L_k^{(\alpha-1)}$ *is the generalised Laguerre polynomial of degree* k*.*

(b) ***Positivity.*** *Only $k = 0$ gives a positive eigenfunction ($L_0^{(\alpha-1)} \equiv 1$). For $k \geq 1$, $\mathfrak{g}_{n,k}$ has exactly k zeros on $(0, \infty)$ and cannot serve as a pricing eigenfunction.*

(c) ***Second exponential.*** *The unstable eigenpair $(\eta_n^+, \mathfrak{g}_n^+)$ from Proposition 8(a) is not in the $L^2(\varrho_n)$ spectrum. It sits at formal position $k = -\alpha$:*

$$\eta_n^+ = \eta_n^- - \alpha\,\omega_n. \tag{83}$$

(d) ***Three spectral regimes.***

(i) $\Delta_n > 0$: countable spectrum with gap ω_n, principal eigenvalue η_n^-.

(ii) $\Delta_n = 0$ ($\phi_n = \phi_{\text{crit}}$): the gap collapses ($\omega_n = 0$), $\xi_n^- = \xi_n^+ = \kappa/\sigma_\lambda^2$, and the Laguerre ladder degenerates.

(iii) $\Delta_n < 0$: the spectrum is empty and $\mathbb{E}_0[\mathcal{M}_{0,T}^n] = \infty$ for $T \geq T_n^*$.

Proof. The λ-dependent part of the eigenvalue equation is

$$\mathcal{G}h(\lambda) = \tfrac{1}{2}\sigma_\lambda^2\lambda\,h'' + \kappa(\bar{\lambda} - \lambda)\,h' + \phi_n\lambda\,h = -\tilde{\eta}\,h, \tag{84}$$

where $\tilde{\eta} = \eta - \beta - \gamma_n\mu + \frac{1}{2}\gamma_n^2\sigma^2$ absorbs the deterministic drift, and $h(\lambda) = e^{-b\lambda}\,\mathfrak{g}(\lambda)$. The full eigenvalue is $\eta = -c_n + \tilde{\eta}$ with $c_n = -\beta - \gamma_n\mu + \frac{1}{2}\gamma_n^2\sigma^2$.

Substitute $h(\lambda) = e^{\xi_n^-\lambda}\,u(\lambda)$. Using $\frac{1}{2}\sigma_\lambda^2(\xi_n^-)^2 - \kappa\xi_n^- + \phi_n = 0$ (the Riccati fixed-point equation), all terms proportional to $\lambda \cdot u$ cancel. The identity $\sigma_\lambda^2\xi_n^- - \kappa = -\omega_n$ simplifies the first-derivative coefficient. The result is

$$\tfrac{1}{2}\sigma_\lambda^2\lambda\,u'' + (\kappa\bar{\lambda} - \omega_n\lambda)\,u' + (\kappa\bar{\lambda}\xi_n^- + \tilde{\eta})\,u = 0.$$

Change variable to $z = 2\omega_n\lambda/\sigma_\lambda^2$. After dividing by ω_n:

$$z\,u'' + (\alpha - z)\,u' + a\,u = 0, \qquad a = \frac{\kappa\bar{\lambda}\xi_n^- + \tilde{\eta}}{\omega_n}. \tag{85}$$

This is Kummer's confluent hypergeometric equation. Its polynomial solutions are the generalised Laguerre polynomials $u(z) = L_k^{(\alpha-1)}(z)$, which arise at $a = k$ for $k = 0, 1, 2, \ldots$ These are the unique solutions in $L^2(z^{\alpha-1}e^{-z}\,\mathrm{d}z)$. Inverting the change of variable: $\tilde{\eta}_k = k\omega_n - \kappa\bar{\lambda}\xi_n^-$, and therefore $\eta_{n,k} = \beta + \gamma_n\mu - \frac{1}{2}\gamma_n^2\sigma^2 - \kappa\bar{\lambda}\xi_n^- + k\omega_n = \eta_n^- + k\omega_n$. The L^2 weight $z^{\alpha-1}e^{-z}\,\mathrm{d}z$ maps back to $\varrho_n(d\lambda) \propto \lambda^{\alpha-1}e^{-2\omega_n\lambda/\sigma_\lambda^2}\,\mathrm{d}\lambda$.

For part (c): $\eta_n^+ = \eta_n^- + \kappa\bar{\lambda}(\xi_n^- - \xi_n^+)$. Since $\xi_n^- - \xi_n^+ = -2\omega_n/\sigma_\lambda^2$, one obtains $\eta_n^+ = \eta_n^- - (2\kappa\bar{\lambda}/\sigma_\lambda^2)\omega_n = \eta_n^- - \alpha\omega_n$. The eigenfunction $\mathfrak{g}_n^+(\lambda) = e^{(b+\xi_n^+)\lambda}$ grows too rapidly to lie in $L^2(\varrho_n)$. □

Numerical spectrum at Wachter's calibration. At the baseline parameters, $\alpha = 1.27$ and the spectral structure is:

	SDF ($\gamma_n = 3$)	Consumption ($\gamma_n = 2$)	Dividend ($\gamma_n = 0.4$)
Regime	empty ($\Delta < 0$)	empty ($\Delta < 0$)	discrete ($\Delta > 0$)
ω_n	—	—	7.36%
$\eta_0 = \eta_n^-$	—	—	1.80%
η_1	—	—	9.15%
η_2	—	—	16.51%
η_n^+ (non-L^2)	—	—	−7.51%

The dividend claim has a rich Laguerre spectrum with gap $\omega_D = 7.36\%$ (92% of κ). The first excited eigenfunction changes sign at $\lambda^* = \alpha\sigma_\lambda^2/(2\omega_D) = 0.039 \approx \bar{\lambda}$: it distinguishes high-intensity from low-intensity states. The second exponential eigenpair sits at $k = -1.27$, below the Laguerre ladder and outside the L^2 spectrum.

The SDF and consumption-claim spectra are empty: there is no function $\mathfrak{g}(\lambda) > 0$ satisfying the eigenvalue equation (66). This is a stronger statement than Proposition 8(b), which showed only that $\mathbb{E}_0[\mathcal{M}_{0,T}]$ diverges. The Kummer analysis confirms that the divergence is not an artifact of a particular initial condition: the generator itself has no L^2 spectral representation in these cases.

Remark 27 (The spectral gap as order parameter). The spectral gap ω_n controls the rate at which the pricing semigroup forgets its initial state. When $\omega_n > 0$, the ratio $\mathfrak{g}_{n,1}/\mathfrak{g}_{n,0}$ decays at rate ω_n, and the Hansen–Scheinkman transient component contracts geometrically. As $\phi_n \uparrow \phi_{\text{crit}}$, $\omega_n \downarrow 0$: the semigroup loses spectral gap, mixing slows, and long-horizon pricing becomes arbitrarily sensitive to the initial state λ_0. At the phase boundary $\phi_n = \phi_{\text{crit}}$, the semigroup is critically slow; beyond it, the expectations diverge. The three semigroups cross this phase boundary at different loadings: the dividend claim ($\phi_D = 0.11$) is deep in the spectral regime; the consumption claim ($\phi_C = 0.87$) and SDF ($\phi_M = 1.89$) are beyond the critical loading $\phi_{\text{crit}} = 0.71$.