Convexity Bounds for the Stochastic Discount Factor: Theory and Evidence

Ian Martin Ran Shi*

September 2025

Abstract

We derive new entropy and moment bounds for the stochastic discount factor (SDF). Our results generalize existing bounds which exploit risk-adjusted measures of investment opportunities—such as Sharpe ratios or expected log returns—that are maximized in the cross-section, across assets. By contrast, we can fix a single asset and optimally exploit information in its true and risk-neutral return distributions. Applying the framework to the S&P 500 index, we find that the θth SDF moment grows extremely rapidly when $\theta > 1$, and appears to diverge to infinity before $\theta = 2$. But entropy measures and the θth moments with $\theta \in (0, 1)$ are well-behaved theoretically and empirically, and can be related to measures of market risk aversion and of the attractiveness of investment opportunities.

*Martin: London School of Economics, i.w.martin@lse.ac.uk; Shi: University of Colorado Boulder, ran.shi@colorado.edu.

Asset prices are widely used to assess market expectations. Policymakers and practitioners talk of forward rates as indicators of expected future interest rates, commodity futures prices as indicators of expected future commodity prices, breakeven inflation as an indicator of expected future inflation, CDS rates as indicators of expected future default rates, implied volatility as an indicator of expected future volatility, and so on. The appeal of doing so is that these quantities, as asset prices, are almost continuously observable; they do not lean on economists' models; and they embody the collective beliefs of market participants. But they may be distorted by risk: the indicators described above measure *risk-neutral* expected future interest rates, *risk-neutral* expected future inflation, and so on, rather than the true expectations that a forecaster would ideally like to know.

Hansen and Jagannathan (1991) introduced the idea that the importance of risk considerations in asset pricing can be captured in a general way via the stochastic discount factor (SDF), whose variability can be understood as an indication of the size of the gap between true and risk-neutral probabilities, and hence as summarizing the importance of risk considerations in pricing. Their work showed that the volatility of the SDF (rescaled to have unit mean) is at least as large as the Sharpe ratio of any asset or strategy.

Hansen and Jagannathan also suggested, as an "important direction" for future research, moving beyond means and variances to characterize the properties of SDFs more fully. As a step in this direction, Snow (1991) derived, for arbitrary $\theta > 1$, lower bounds on the θth moment of the SDF in terms of the $\frac{\theta}{\theta-1}$th moments of assets' returns.[1] But $\frac{\theta}{\theta-1}$ is very large when θ is close to 1, so that in this case the Snow bound depends on extremely high moments of asset returns which are hard to measure in practice. (It may be for this reason that the Snow bound has had limited influence.) This might seem surprising: If a riskless asset is traded each period, we can perfectly infer the conditional mean of the SDF; it should therefore be straightforward to estimate the unconditional first moment of the SDF using the time series of interest rates. Why then is it so hard to restrict nearby moments?

In this paper, we introduce new bounds on the θth moment of the SDF, for arbitrary $\theta \in \mathbb{R}$, and on SDF entropy. The bounds are completely general: we do not impose any restrictions on the form of the SDF, nor do we assume that the market is complete. Like the prior literature, we exploit the true distribution of returns, which we infer from the time series of realized returns. Unlike the prior literature, we also exploit the risk-neutral distribution, which is observable given a suitably rich collection of options traded on the asset

[1] Bansal and Lehmann (1997) and Alvarez and Jermann (2005) have also introduced variations on the approach based on the concept of entropy. We connect to their work below.

in question. As we noted above, the first moment of the SDF is revealed by observation of the riskless rate, that is, of the risk-neutral first moment of returns. Our approach exploits the fact that we can observe arbitrary moments of the risk-neutral distribution. This gives us a crucial degree of freedom that enables us to derive stronger bounds than the prior literature—bounds that are useful even when θ is close to one.

The approach leads us to stronger conclusions than have been reached by the prior literature, and the θ-close-to-one regime turns out to have surprising properties. In particular, our empirical results show that the moments of the SDF rise exceedingly rapidly as θ rises above one. Hansen and Jagannathan's result is often described as showing that the SDF is surprisingly volatile; our findings suggest that the SDF may have *infinite* volatility.[2]

This fact strikes at the heart of the vast empirical literature in financial economics based on mean-variance analysis, dating back to the foundational theories of Markowitz (1952), Sharpe (1964), and Lintner (1965), with early empirical tests by Black, Jensen, and Scholes (1972) and Fama and MacBeth (1973). The more recent literature[3] adopts a multifactor mean-variance approach motivated by the work of Merton (1973) and Ross (1976) and underpinned by the framework of Chamberlain and Rothschild (1983) and Hansen and Richard (1987). If the variance of the SDF is unbounded then this framework breaks down, and with it the statistical foundation for testing whether a collection of factors is mean-variance efficient, either in its traditional form (Gibbons, Ross, and Shanken, 1989) or in more recent high-dimensional variants (Fan, Liao, and Yao, 2015; Bryzgalova, Huang, and Julliard, 2023; Chernov, Kelly, Malamud, and Schwab, 2025).

We show that the bounds on the θth moment of the SDF have particularly good properties when $\theta \in (0, 1)$. These have a natural economic interpretation—they supply bounds on the willingness-to-pay (WTP) to participate in risky financial markets of an investor with power utility and relative risk aversion at least one—and we find lower bounds on WTP that take plausible values in the data.

Analogously, the Hansen–Jagannathan bound relates another measure of the attractiveness of the investment opportunity set—the maximum attainable Sharpe ratio—to the vari-

[2]We write "suggest" and "may have" because we show, as a matter of theory, that bounds on the θth moment of the SDF have bad statistical properties when $\theta < 0$ or $\theta > 1$. Empirically, we find that the lower bound diverges when θ exceeds about 1.7. This is true even though we are careful to make conservative choices in our empirical analysis. Plausible alternatives to our approach (such as using mid-market prices rather than bid and offer prices) give even more extreme results than these. Sections 3.1, 3.2, and 3.3 describe the steps we take to ensure that our empirical approach is conservative.

[3]For example, Fama and French (2015), Hou, Xue, and Zhang (2015), Campbell, Giglio, Polk, and Turley (2018), and Kozak, Nagel, and Santosh (2020).

ance of the SDF. Sharpe ratios make sense from the perspective of investors with quadratic utility[4] because such investors care only about mean and variance, and not about any higher moments of returns. But the quadratic utility assumption is well known to have strange implications: for example, an investor with quadratic utility has increasing absolute risk aversion, and so invests fewer dollars in risky assets as he or she becomes wealthier. This is, to put it mildly, an implausible depiction of investor behavior.[5] And, aside from this issue, our theoretical results reveal a sense in which variance bounds on the SDF are inherently unstable—even, in principle, if we are given an arbitrarily long time series of data.

Our results are related to a literature that has argued that short positions in options have extremely high Sharpe ratios on average.[6] One could try to incorporate the information in option prices by plugging in the returns on option strategies into the Hansen–Jagannathan or Snow bounds, as Liu (2021) does, for example. But, as Broadie, Chernov, and Johannes (2009) emphasize, option returns are unstable, highly skewed and fat-tailed, sensitive to outliers. Given the relatively short time series of observed option returns, these facts make them poorly suited for empirical work: for example, over the 1996–2022 sample period for which we observe option prices, the average realized net return on a one-month index put option struck 10% out of the money is −17.6%, but the standard error on this quantity is huge, at 51.6%.

We do not use options in this way. Instead, we study the risk-neutral distribution directly. Rather than using time series averages of option returns, we exploit time series averages of (appropriately rescaled) option prices. These are much less noisy, and they directly reveal the conditional risk-neutral distribution: for example, based on the 26 years of available option price data, we estimate that the time-series average risk-neutral probability of a 10% decline in the market over the next month is 6.56%, with a standard error of 0.22%.

By contrast, realized returns on the market only provide a noisy measure of its true *un*conditional distribution (and on their own they tell us nothing at all about the true conditional distribution). Indeed, the primary empirical challenge we face is not the short time

[4]Alternatively, Sharpe ratios could be justified if returns followed a distribution in the elliptical family. But elliptical distributions are symmetrical about their means, unlike empirically observed return distributions.

[5]For a comprehensive discussion of the demerits of mean–variance preferences, see Borch (1969), Feldstein (1969), Samuelson (1970), Tsiang (1972), and Levy (1974).

[6]See, for example, Jackwerth (2000), Coval and Shumway (2001), Bondarenko (2003), Jones (2006), Driessen and Maenhout (2007), Goetzmann, Ingersoll, Spiegel, and Welch (2007), and Santa-Clara and Saretto (2009). In a somewhat different direction, Bates (1991) is an influential early study of tail information in option prices.

series of option prices, but the short time series of realized market returns. Even when we use our longest time series of realized returns, we find that the uncertainty associated with estimating the true distribution is greater than that associated with the risk-neutral distribution. In our longest dataset, which starts in 1872, the time-series average true probability of a 10% decline in the market over the next month is 1.82%—but the standard error on this estimate, at 0.32%, is larger than that for the risk-neutral calculation of the previous paragraph, even with 150 years of data.

A century and a half later, it is not the shortness of the option price series but the shortness of the realized return series that remains the central empirical challenge; and it is one that is inherent to any approach that uses the historical time series to make inferences about population moments.

Structure of the paper. We find it helpful to derive and discuss our results using the language of cumulant-generating functions (CGFs). We introduce these, discuss their general properties, and provide some concrete examples of CGFs in equilibrium models, in Section 1. We derive our main theoretical results, including moment bounds, entropy bounds, and their connection to measures of market risk aversion and of the attractiveness of investment opportunities, in Section 2. We discuss the data, and certain important issues that arise in finite samples, in Section 3. We take the theory to the data in Section 4, and conclude in Section 5.

1 Cumulant-generating functions

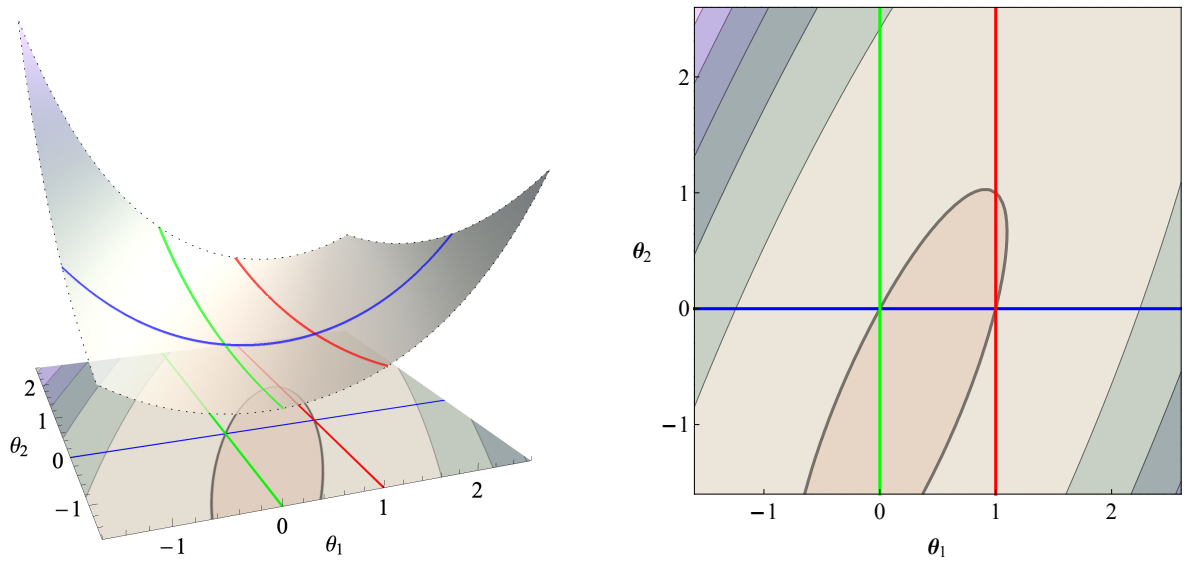
This section introduces CGFs in their conditional and unconditional forms.

Fix some asset, whose gross return from time t to time $t+1$ we write as R_{t+1}. (For the purposes of the theory, R_{t+1} can be an arbitrary return; when we turn to the data, we will set it equal to the return on a broad stock market index.) The gross riskless rate from time t to $t+1$ is $R_{f,t+1}$, which is known at time t. Finally, write M_{t+1} for a stochastic discount factor (SDF) that provides the time t prices of payoffs made at time $t+1$, so that in particular $\mathbb{E}_t\left(M_{t+1}R_{t+1}\right)=1$ and $\mathbb{E}_t\left(M_{t+1}R_{f,t+1}\right)=1$.

We define the conditional CGF

$$\boldsymbol{\kappa}_t(\theta_1,\theta_2)\equiv\log\mathbb{E}_t\left[\left(M_{t+1}R_{f,t+1}\right)^{\theta_1}\left(R_{t+1}/R_{f,t+1}\right)^{\theta_2}\right]. \tag{1}$$

The curved surface shown in the left panel of Figure 1 illustrates how a CGF varies with θ_1 and θ_2. By definition, the CGF goes through the origin: $\boldsymbol{\kappa}_t(0,0)=0$. The properties of



(a) The CGF surface (b) Contour plot

Figure 1: The cumulant-generating function (1). The thick contour in the right panel is the zero contour.

the SDF further restrict the form of the CGF. As $\mathbb{E}_t\left(M_{t+1}R_{f,t+1}\right) = 1$, we have $\boldsymbol{\kappa}_t(1,0) = 0$; and as $\mathbb{E}_t\left(M_{t+1}R_{t+1}\right) = 1$, we have $\boldsymbol{\kappa}_t(1,1) = 0$. Thus the CGF is pinned to zero at the origin, at $(1,0)$, and at $(1,1)$. The contour plot underneath and in the right panel of Figure 1 conveys the same information. The thick contour indicates the set of points at which the CGF equals zero.

Three lines are indicated on the surface and on the contour plot. Two of them—the red and the green—are observable from the data. Each represents a convex function. Indeed, the entire CGF surface is convex: we exploit this important and general property below to infer properties of the third, blue, line. Whereas equilibrium models of financial markets specify the entire surface $\boldsymbol{\kappa}_t(\theta_1,\theta_2)$ for arbitrary θ_1 and θ_2, the goal of this paper is to focus on aspects of the surface that can be determined from observable data without positing a particular functional and distributional form for M_{t+1} and R_{t+1}.

Return cumulants. The green line traces out the curve

$$\boldsymbol{\kappa}_t(0,\theta_2) = \log \mathbb{E}_t\left[\left(R_{t+1}/R_{f,t+1}\right)^{\theta_2}\right] \tag{2}$$

as θ_2 varies. This function encodes the cumulants of the log risky return. For example, by differentiating the right-hand side of equation (2) with respect to θ_2, it is easy to check that the mean and variance of the log excess return reflect, respectively, the slope and curvature

of the CGF at the origin:

$$\mathbb{E}_t \log\left(R_{t+1}/R_{f,t+1}\right) = \frac{\partial\, \boldsymbol{\kappa}_t}{\partial \theta_2}(0,0) \quad \text{and} \quad \operatorname{var}_t \log\left(R_{t+1}/R_{f,t+1}\right) = \frac{\partial^2\, \boldsymbol{\kappa}_t}{\partial \theta_2^2}(0,0)\,.$$

While the cumulants (and hence moments) of log returns are captured by the behaviour of the CGF local to the origin, the cumulants of the simple return depend on its shape away from the origin. For example, $\boldsymbol{\kappa}_t(0,1) = \log \mathbb{E}_t\left(R_{t+1}/R_{f,t+1}\right)$ is the equity premium and $e^{\boldsymbol{\kappa}_t(0,2)} - e^{2\,\boldsymbol{\kappa}_t(0,1)}$ is the variance of $R_{t+1}/R_{f,t+1}$.

Risk-neutral return cumulants. The red line traces out the curve

$$\boldsymbol{\kappa}_t(1,\theta_2) = \log \mathbb{E}_t\left[M_{t+1} R_{f,t+1}\left(R_{t+1}/R_{f,t+1}\right)^{\theta_2}\right] \tag{3}$$

as θ_2 varies. This is observable if R_{t+1} is the return on an asset on which options are traded. To see this, it will be convenient to define the risk-neutral expectation operator $\mathbb{E}_t^*$, which is defined by the property that

$$\frac{1}{R_{f,t+1}}\,\mathbb{E}_t^*\, X_{t+1} \equiv \mathbb{E}_t\left(M_{t+1} X_{t+1}\right)$$

for any tradable payoff X_{t+1} that is received at time $t+1$. The price of the payoff X_{t+1} can be expressed either in terms of the true expectation operator $\mathbb{E}_t$ and the SDF, or in terms of the risk-neutral expectation $\mathbb{E}_t^*$ and the riskless interest rate.

Equation (3) can be rewritten in this notation as

$$\boldsymbol{\kappa}_t\left(1,\theta_2\right) = \log \mathbb{E}_t^*\left[\left(R_{t+1}/R_{f,t+1}\right)^{\theta_2}\right]\,. \tag{4}$$

Thus $\boldsymbol{\kappa}_t(0,\theta_2)$ measures the moments of the excess return under the true probability measure, as in equation (2), and $\boldsymbol{\kappa}_t(1,\theta_2)$ measures the moments of the excess return under the risk-neutral measure, as in equation (4).

We can evaluate $\boldsymbol{\kappa}_t(1,\theta)$ at different values of θ using observable option prices. As we show in the appendix (following Carr and Madan (2001)),

$$\boldsymbol{\kappa}_t\left(1,\theta\right) = \log\left\{1 + \theta(\theta-1)\left[\int_0^1 K^{\theta-2}\,\mathrm{put}_t(KR_{f,t+1})\,\mathrm{d}K + \int_1^\infty K^{\theta-2}\,\mathrm{call}_t(KR_{f,t+1})\,\mathrm{d}K\right]\right\}, \tag{5}$$

where $\mathrm{put}_t(K)$ is the time t price of a European put on the risky return R_{t+1} with strike K, expiring at time $t+1$, and $\mathrm{call}_t(K)$ is the corresponding call price. This formula allows us to incorporate the information in option prices across the full range of traded strikes.

Throughout the paper, we think in terms of options on returns rather than options on prices. Rescaling in this way means that the option prices we work with do not drift with the

level of the underlying asset: they will be stationary if the underlying returns are stationary. Specifically, we divide the prices of options on the level of the S&P 500 index—in the case of call options, these are the prices of assets with payoffs $\max\{0, S_{t+1} - KS_t\}$ for a range of K—by the spot price of the asset, S_t, to infer the price of an option on the return (with payoff $\max\{0, \frac{S_{t+1}}{S_t} - K\}$ in the call option case). As is customary in the literature, we neglect the influence of dividends, so that $R_{t+1} = S_{t+1}/S_t$. The average S&P 500 dividend yield over our option sample period is below 2% annually, so this is a minor assumption at the one-year horizon and even more minor at the one-month horizon.

If θ is large and positive in equation (5), the prices of deep-out-of-the-money calls (that is, calls with high strikes) acquire particular significance; conversely, if θ is large in magnitude and negative, the prices of deep-out-of-the-money puts (that is, puts with low strikes) are important.

On any given date t, we take the range of observed call and put option prices and calculate the above function for any θ of interest. Having done so, the derivatives of $\boldsymbol{\kappa}_t(1, \theta_2)$ at $\theta_2 = 0$ reveal the risk-neutral expectations of the log excess return $\log(R_{t+1}/R_{f,t+1})$,

$$\mathbb{E}_t^* \log(R_{t+1}/R_{f,t+1}) = \frac{\partial \boldsymbol{\kappa}_t}{\partial \theta_2}(1, 0)\,.$$

Similarly, the second derivative of $\boldsymbol{\kappa}_t(1, \theta_2)$ at $\theta_2 = 0$ reveals the risk-neutral variance of the log excess return:

$$\mathrm{var}_t^* \log(R_{t+1}/R_{f,t+1}) = \frac{\partial^2 \boldsymbol{\kappa}_t}{\partial \theta_2^2}(1, 0)\,. \tag{6}$$

Other notions of the (risk-neutral) variability of the return can likewise be interpreted as measures of the convexity of the CGF surface along the line $\theta_1 = 1$.[7] For example, if R_{t+1} is the return on the S&P 500 index, then the level of the VIX index at time t satisfies

$$\mathrm{VIX}_t^2 = -2\,\mathbb{E}_t^* \log(R_{t+1}/R_{f,t+1}) = -2\frac{\partial \boldsymbol{\kappa}_t}{\partial \theta_2}(1, 0), \tag{7}$$

[7] This is obvious in the case of (6), but it is also true for (7), (8), and (9) because of the fact that $\boldsymbol{\kappa}_t(1, 0)$ and $\boldsymbol{\kappa}_t(1, 1)$ are both equal to zero. In more detail: as $\boldsymbol{\kappa}_t(1, 2) = [\boldsymbol{\kappa}_t(1, 2) - \boldsymbol{\kappa}_t(1, 1)] - [\boldsymbol{\kappa}_t(1, 1) - \boldsymbol{\kappa}_t(1, 0)]$, the right-hand side of (8) is a convexity measure that compares the average slope of $\boldsymbol{\kappa}_t(1, \theta_2)$ between $\theta_2 = 1$ and $\theta_2 = 2$ to its average slope between $\theta_2 = 0$ and $\theta_2 = 1$. Meanwhile, the right-hand sides of (7) and (9) are each proportional to weighted averages of the curvature of the risk-neutral CGF evaluated in the range from $\theta_2 = 0$ to $\theta_2 = 1$, because

$$-2\frac{\partial \boldsymbol{\kappa}_t}{\partial \theta_2}(1, 0) = \int_0^1 2(1 - \theta_2)\frac{\partial^2 \boldsymbol{\kappa}_t}{\partial \theta_2^2}(1, \theta_2)\, d\theta_2 \quad \text{and} \quad \frac{\partial \boldsymbol{\kappa}_t}{\partial \theta_2}(1, 1) = \int_0^1 \theta_2 \frac{\partial^2 \boldsymbol{\kappa}_t}{\partial \theta_2^2}(1, \theta_2)\, d\theta_2.$$

Each of these equations follows by integrating by parts and using the fact that $\boldsymbol{\kappa}_t(1, 0) = \boldsymbol{\kappa}_t(1, 1) = 0$.

the level of the SVIX index (Martin, 2017) satisfies

$$\log(1+\text{SVIX}_t^2) = \log\left[1+\text{var}_t^*\left(R_{t+1}/R_{f,t+1}\right)\right] = \boldsymbol{\kappa}_t(1,2), \tag{8}$$

and the LVIX index (Gao and Martin, 2021; Gandhi, Gormsen, and Lazarus, 2025) satisfies

$$\text{LVIX}_t = \mathbb{E}_t^*\left[\left(R_{t+1}/R_{f,t+1}\right)\log\left(R_{t+1}/R_{f,t+1}\right)\right] = \frac{\partial\,\boldsymbol{\kappa}_t}{\partial\theta_2}(1,1). \tag{9}$$

SDF cumulants. Lastly, the blue line in Figure 1 traces out the curve

$$\boldsymbol{\kappa}_t(\theta_1,0) = \log\mathbb{E}_t\left[\left(M_{t+1}R_{f,t+1}\right)^{\theta_1}\right] \tag{10}$$

as θ_1 varies. It summarizes the moments of the SDF. It is not directly observed but, as noted above, it must pass through zero when θ_1 equals zero or one.

1.1 Conditional and unconditional CGFs

While the conditional risk-neutral CGF $\boldsymbol{\kappa}_t(1,\theta)$ is observable given suitable option price data at time t, as in (5), the corresponding conditional true CGF $\boldsymbol{\kappa}_t(0,\theta)$ is not. We therefore work unconditionally in our empirical work below. We drop subscripts t to indicate unconditional CGFs:

$$\boldsymbol{\kappa}(\theta_1,\theta_2) \equiv \log\mathbb{E}\left[\left(M_{t+1}R_{f,t+1}\right)^{\theta_1}\left(R_{t+1}/R_{f,t+1}\right)^{\theta_2}\right].$$

By the law of iterated expectations, the properties $\boldsymbol{\kappa}(0,0) = \boldsymbol{\kappa}(1,0) = \boldsymbol{\kappa}(1,1) = 0$ also hold unconditionally. The unconditional true CGF of (log) returns is

$$\boldsymbol{\kappa}(0,\theta) = \log\mathbb{E}\left[\left(R_{t+1}/R_{f,t+1}\right)^{\theta}\right]. \tag{11}$$

The unconditional counterpart of the risk-neutral CGF is

$$\begin{aligned}\boldsymbol{\kappa}(1,\theta) &= \log\mathbb{E}\left[M_{t+1}R_{f,t+1}\left(R_{t+1}/R_{f,t+1}\right)^{\theta}\right]\\ &= \log\mathbb{E}\left(\mathbb{E}_t\left[M_{t+1}R_{f,t+1}\left(R_{t+1}/R_{f,t+1}\right)^{\theta}\right]\right)\\ &= \log\mathbb{E}\left(\mathbb{E}_t^*\left[\left(R_{t+1}/R_{f,t+1}\right)^{\theta}\right]\right).\end{aligned} \tag{12}$$

To streamline the notation, we define an operator $\mathbb{E}^*$ which satisfies[8]

$$\mathbb{E}^*\,X \equiv \mathbb{E}\left(\mathbb{E}_t^*\,X\right)$$

[8]The operator $\mathbb{E}^*$ appears, in different notation, in unconditional moment restriction 1 of Hansen and Jagannathan (1991, p. 231), which applies to an unconditional expectation of prices, i.e., to a unconditional expectation of conditional risk-neutral expectations.

for any suitable random variable X. This is notationally convenient as it will allow us to state the unconditional counterparts to our conditional results simply by dropping subscripts t. For example, we can write the unconditional risk-neutral CGF (12) in this notation as

$$\boldsymbol{\kappa}(1,\theta) = \log \mathbb{E}^* \left[\left(R_{t+1}/R_{f,t+1} \right)^\theta \right] . \tag{13}$$

1.2 CGFs in equilibrium

Equilibrium models pin down the form of the CGF. In our empirical work, we will not want to make any assumptions about the shape of the CGF; nonetheless, it may be helpful to see some examples of equilibrium CGFs to make things concrete. Derivations of the CGFs discussed in this section are provided in the appendix.

Example 1. If M_{t+1} and R_{t+1} are jointly lognormal, then the CGF surface is quadratic. To see this, note that we can write $M_{t+1}R_{f,t+1} = e^{-\frac{1}{2}\lambda^2 - \lambda Z}$ and $R_{t+1}/R_{f,t+1} = e^{\mu - \frac{1}{2}\sigma^2 + \sigma W}$, where W and Z are standard Normal with correlation ρ and λ can be interpreted[9] as the maximal attainable Sharpe ratio. (The constants λ, μ, and σ can depend on time-t information, but we suppress subscripts t to simplify the notation.) The CGF is

$$\boldsymbol{\kappa}_t(\theta_1,\theta_2) = \mu\theta_2\left(1-\theta_1\right) + \frac{1}{2}\lambda^2\theta_1\left(\theta_1 - 1\right) + \frac{1}{2}\sigma^2\theta_2\left(\theta_2 - 1\right) \tag{14}$$

after imposing the condition $\mu = \rho\sigma\lambda$, which is implied by $\boldsymbol{\kappa}_t(1,1) = 0$ (that is, by the fact that $\mathbb{E}_t(M_{t+1}R_{t+1}) = 1$).

The two observable slices of the CGF surface are therefore

$$\boldsymbol{\kappa}_t(0,\theta_2) = \mu\theta_2 + \frac{1}{2}\sigma^2\theta_2(\theta_2 - 1) \tag{15}$$

and

$$\boldsymbol{\kappa}_t(1,\theta_2) = \frac{1}{2}\sigma^2\theta_2(\theta_2 - 1)\,. \tag{16}$$

Knowledge of the true distribution of returns (15) amounts to knowing μ and σ, while the information in option prices supplies only σ, via the function (16). Thus in this lognormal case, one learns nothing new from option prices. This property fails in general, but it may help to explain the lack of interest in option prices in much of the mainstream empirical finance literature.

Example 2. If M_{t+1} and R_{t+1} follow jump-diffusions, useful information regarding the SDF can be lost without option prices. Suppose that $M_{t+1}R_{f,t+1} = e^{-\frac{1}{2}\lambda^2 - \lambda Z - \omega J_1}\left(1 + J_1\right)^N$

[9] This is the conventional interpretation: it is exact in continuous time and approximate in discrete time.

and that $R_{t+1}/R_{f,t+1} = e^{\mu-\frac{1}{2}\sigma^2+\sigma W-\omega J_2}\left(1+J_2\right)^N$, where W and Z are standard Normal with correlation ρ and the jump arrival rate is ω, so that the number of jumps, N, has a Poisson distribution with parameter ω. Assume, for simplicity, that the jump size is deterministic, so that J_1 and J_2 represent the proportional impact of a jump on the SDF and risky asset, respectively. (If jumps represent bad news, then we might have J_1 positive and J_2 negative, for example; in any case, we always require $1+J_1$ and $1+J_2$ to be positive.) The CGF is then

$$\begin{aligned}\boldsymbol{\kappa}_t(\theta_1,\theta_2) = \mu\theta_2(1-\theta_1) + \frac{1}{2}\lambda^2\theta_1(\theta_1-1) + \frac{1}{2}\sigma^2\theta_2(\theta_2-1) + \\ + \omega\left[\left(1+J_1\right)^{\theta_1}\left(1+J_2\right)^{\theta_2} - (1+J_1\theta_1)(1+J_2\theta_2)\right]\end{aligned} \tag{17}$$

after imposing the condition that $\rho\sigma\lambda = \mu + \omega J_1 J_2$, which is implied by $\boldsymbol{\kappa}_t(1,1) = 0$. The presence of jumps is reflected in the terms which are exponential in θ_1 and θ_2, but the CGF is defined for all θ_1 and θ_2. This property—which also holds in disaster models such as Barro (2006)—will fail in our next two examples.

The two observable slices of the CGF surface are now

$$\boldsymbol{\kappa}_t(0,\theta_2) = \mu\theta_2 + \frac{1}{2}\sigma^2\theta_2(\theta_2-1) + \omega\left[\left(1+J_2\right)^{\theta_2} - \theta_2 J_2 - 1\right] \tag{18}$$

and

$$\boldsymbol{\kappa}_t(1,\theta_2) = \frac{1}{2}\sigma^2\theta_2(\theta_2-1) + \omega(1+J_1)\left[\left(1+J_2\right)^{\theta_2} - \theta_2 J_2 - 1\right]. \tag{19}$$

In this case, the true distribution of returns, as embodied in (18), reveals μ, σ, ω, and J_2, but not the parameter J_1 which measures the size of an SDF jump shock. We can however infer J_1 from observing option prices, and hence the function (19).

Example 3. Geweke (2001) and Weitzman (2007) have shown that the predictions of conventional models can change dramatically if agents must learn model parameters. To illustrate this point, we adapt Example 2. As our goal is illustrative, we simplify matters by assuming that uncertainty is driven solely by jumps with no diffusion component, that is, $\sigma = \lambda = 0$, so that in the absence of learning the CGF (17) would take the form

$$\boldsymbol{\kappa}_t(\theta_1,\theta_2) = \mu\theta_2(1-\theta_1) + \omega\left[\left(1+J_1\right)^{\theta_1}\left(1+J_2\right)^{\theta_2} - (1+J_1\theta_1)(1+J_2\theta_2)\right], \tag{20}$$

which is defined for all θ_1 and θ_2, so that all moments of the risky return and SDF exist.

To show how parameter uncertainty changes things,[10] we now suppose that agents are

[10] We focus on learning about jump intensities for tractability. Geweke (2001) and Weitzman (2007) have exhibited even more dramatic examples in lognormal models without jumps in which a representative agent performs Bayesian updating about the variance of log consumption growth. In this case every moment of the SDF may be unbounded, with even the risk-free rate undefined.

uncertain about the value of ω: they perceive it as distributed according to an exponential distribution with mean $\overline{\omega}$. We can then write[11] $M_{t+1}R_{f,t+1} = (1-\overline{\omega}J_1)(1+J_1)^N$ and $R_{t+1}/R_{f,t+1} = e^{\mu}(1-\overline{\omega}J_2)(1+J_2)^N$, and the CGF is

$$\boldsymbol{\kappa}_t(\theta_1,\theta_2) = \theta_1\log(1-\overline{\omega}J_1)+\theta_2(\mu+\log(1-\overline{\omega}J_2))-\log\left(1-\overline{\omega}\left[(1+J_1)^{\theta_1}(1+J_2)^{\theta_2}-1\right]\right). \tag{21}$$

We require that $\overline{\omega}J_1 < 1$, $\overline{\omega}J_2 < 1$, and $\overline{\omega}\left[(1+J_1)(1+J_2)-1\right] < 1$ so that the relevant expectations are well defined.

Suppose, for example, that a jump represents bad news, so that $J_1 > 0$ and $-1 < J_2 < 0$. Then equation (21) implies that the true and risk-neutral CGFs, $\boldsymbol{\kappa}_t(0,\theta)$ and $\boldsymbol{\kappa}_t(1,\theta)$, diverge for sufficiently negative values of θ, so the corresponding true and risk-neutral moments of the risky return are unbounded. Consideration of the function $\boldsymbol{\kappa}_t(\theta,0)$ shows that sufficiently positive moments of the SDF also diverge: the θth moment is finite if and only if θ is less than $\frac{\log(1+1/\overline{\omega})}{\log(1+J_1)}$. The requirement that $\overline{\omega}J_1 < 1$ implies that $\frac{\log(1+1/\overline{\omega})}{\log(1+J_1)} > 1$, so that the θth moment of the SDF is well-defined for $\theta \leq 1$ but diverges once θ exceeds some critical value that is above 1. (Alternatively, if jumps are good news, $-1 < J_1 < 0$ and $J_2 > 0$, then the θth moment of the SDF is well-defined for $\theta \geq 0$ but diverges below some critical value of θ that is less than zero.)

Example 4. We now exhibit an example in which the higher moments of the SDF are unbounded even though all return moments are well-behaved under both the true and risk-neutral measures. Specifically, we consider the Brownian limit of the heterogeneous-belief equilibrium model of Martin and Papadimitriou (MP, 2022, Section III). We assume that the median agent at time $t = 0$ has correct beliefs, so calculate the CGF surface from this agent's perspective. The two observable slices are

$$\boldsymbol{\kappa}_t(0,\theta_2) = \frac{1+\delta}{2\delta}\sigma^2\theta_2 + \frac{1}{2}\sigma^2\theta_2^2 \tag{22}$$

and

$$\boldsymbol{\kappa}_t(1,\theta_2) = \frac{1+\delta}{2\delta}\sigma^2\theta_2(\theta_2-1). \tag{23}$$

Here σ is return volatility and $\delta > 0$ is a parameter that controls the amount of disagreement, and hence speculation.[12] Smaller values of δ correspond to greater disagreement: the cross-sectional standard deviation of expected returns across agents equals $\sigma/\sqrt{\delta}$.

[11] The factor $1-\overline{\omega}J_1$ differs from the corresponding factor $e^{-\omega J_1}$ in the case without learning: they are determined in each case by the requirement that $\mathbb{E}_t(M_{t+1}R_{f,t+1}) = 1$. Similarly, the factor $1-\overline{\omega}J_2$ in $R_{t+1}/R_{f,t+1}$ ensures that $\mathbb{E}_t R_{t+1}/R_{f,t+1}$ equals e^{μ}, just as it does in the case without learning. The constant μ is determined in terms of the other parameters by the equilibrium requirement that $\boldsymbol{\kappa}_t(1,1) = 0$.

[12] The parameter δ is referred to as θ by MP. We focus on the Brownian limit to emphasize that option

As in Example 1, the observable slices (22) and (23) are each individually quadratic, indicating that returns are lognormal under both the (perceived) true and risk-neutral distributions. Unlike Example 1, however, they have different amounts of curvature, $\frac{\partial^2 \boldsymbol{\kappa}_t}{\partial \theta_2^2}(0, \theta_2) \neq \frac{\partial^2 \boldsymbol{\kappa}_t}{\partial \theta_2^2}(1, \theta_2)$, reflecting the fact that true and risk-neutral volatility differ (that is, the MP model generates a variance risk premium). This is, therefore, a setting in which an econometrician who neglects option prices is throwing away useful information. In Example 1, the SDF M_{t+1} and return R_{t+1} were *jointly* lognormal, and as a result the entire CGF surface was a quadratic form. Here, by contrast, the CGF surface takes the form

$$\boldsymbol{\kappa}_t(\theta_1, \theta_2) = \frac{1}{2}\left[\frac{1+\delta}{\delta}\sigma^2(\theta_2 - \theta_1) + \frac{1+\delta}{1+\delta-\theta_1}\sigma^2(\theta_2-\theta_1)^2 + \log\frac{1+\delta}{1+\delta-\theta_1} - \theta_1 \log\frac{1+\delta}{\delta}\right]. \tag{24}$$

It follows from (24) that the moments of the SDF are given by

$$\boldsymbol{\kappa}_t(\theta, 0) = \frac{(1+\delta)^2\theta(\theta-1)\sigma^2}{2\delta(1+\delta-\theta)} + \frac{1}{2}\log\frac{1+\delta}{1+\delta-\theta} - \frac{1}{2}\theta\log\frac{1+\delta}{\delta}, \tag{25}$$

so the $(1+\delta)$th and higher moments of the SDF are unbounded. In particular, the second moment is unbounded if $\delta \leq 1$: in this case MP show that arbitrarily large Sharpe ratios can be attained by aggressively shorting out-of-the-money options.

Summing up, these examples make two important points. First, option prices are useful. Other than in the lognormal special case, they convey information that is embodied in the risk-neutral slice of the CGF surface and not revealed by the true distribution of returns. Second, it is possible for the SDF to have unbounded higher moments in equilibrium models with parameter learning or heterogeneous beliefs. In the MP model, for example, there are calibrations in which the volatility of the SDF is unbounded and in which strategies can be constructed with arbitrarily high Sharpe ratios (Hansen and Jagannathan, 1991). As MP emphasize, however, these strategies need not be remotely attractive to investors with conventional utility functions, due to their unappealing higher-moment properties.

In such models, we have the luxury of knowing the full equilibrium and hence the CGF *surface* (e.g., the function (24) in the heterogeneous-agent example). The econometrician, however, can only observe two *slices*—(22) and (23) in this example—of this surface. We now show how to derive empirically implementable bounds on the moments of the SDF that only exploit information in these two observable slices.

prices can be informative even in the benign case in which returns are lognormal under the risk-neutral measure and under every investor's perceived true measure (albeit different investors perceive different mean log returns). MP also study a non-lognormal Poisson limit in which there is a volatility smile and the SDF has infinite variance in every calibration.

2 New bounds for the stochastic discount factor

We introduce our main results, the moment and entropy bounds, in Sections 2.1 and 2.2. We show that they have a natural economic interpretation in Section 2.3.

2.1 Moment bounds

Our first result exploits the convexity of the CGF surface to generalize the Hansen and Jagannathan (1991) and Snow (1991) bounds.

Result 1. *For $\theta < 0$ or $\theta > 1$, we have*

$$\mathbb{E}_t\left[(M_{t+1}R_{f,t+1})^\theta\right] \geq \sup_{R_{t+1}} \sup_{y\in\mathbb{R}} \left\{\mathbb{E}_t^*\left[\left(\frac{R_{t+1}}{R_{f,t+1}}\right)^y\right]\right\}^\theta \left\{\mathbb{E}_t\left[\left(\frac{R_{t+1}}{R_{f,t+1}}\right)^{\frac{\theta}{\theta-1}y}\right]\right\}^{1-\theta}. \tag{26}$$

For $\theta \in (0,1)$, we have

$$\mathbb{E}_t\left[(M_{t+1}R_{f,t+1})^\theta\right] \leq \inf_{R_{t+1}} \inf_{y\in\mathbb{R}} \left\{\mathbb{E}_t^*\left[\left(\frac{R_{t+1}}{R_{f,t+1}}\right)^y\right]\right\}^\theta \left\{\mathbb{E}_t\left[\left(\frac{R_{t+1}}{R_{f,t+1}}\right)^{\frac{\theta}{\theta-1}y}\right]\right\}^{1-\theta}. \tag{27}$$

Equivalently,

$$\boldsymbol{\kappa}_t(\theta,0) \geq \sup_{R_{t+1}} \sup_{y\in\mathbb{R}} \theta\,\boldsymbol{\kappa}_t(1,y) + (1-\theta)\,\boldsymbol{\kappa}_t\left(0, \frac{\theta}{\theta-1}y\right) \text{ for } \theta < 0 \text{ or } \theta > 1 \tag{28}$$

and

$$\boldsymbol{\kappa}_t(\theta,0) \leq \inf_{R_{t+1}} \inf_{y\in\mathbb{R}} \theta\,\boldsymbol{\kappa}_t(1,y) + (1-\theta)\,\boldsymbol{\kappa}_t\left(0, \frac{\theta}{\theta-1}y\right) \text{ for } \theta \in (0,1). \tag{29}$$

These inequalities also hold unconditionally (that is, dropping the ts on the true and risk-neutral expectation operators).

Proof. Let R_{t+1} be an arbitrary gross return. We will prove (26) and (27) by considering the function $\boldsymbol{\kappa}_t(\theta_1,\theta_2)$ evaluated at the points $(\theta,0)$, $(1,y)$, and $(0,x)$. Having done so, the unconditional statements follow by applying the same proof to $\boldsymbol{\kappa}(\theta_1,\theta_2)$ rather than $\boldsymbol{\kappa}_t(\theta_1,\theta_2)$.

We regard y as arbitrary and choose x so that the three points lie on a line, to enable us to invoke convexity of the CGF. This requires that $x = \frac{\theta}{\theta-1}y$. We adopt this notation throughout the proof. The identity of the middle one of the three points depends on whether $\theta < 0$, $\theta > 1$, or $\theta \in (0,1)$, so we handle these cases separately.

$\underline{\theta > 1}$. We can write $(1, y)$ as the convex combination $(1, y) = \frac{1}{\theta}(\theta, 0) + \left(1 - \frac{1}{\theta}\right)(0, x)$. By convexity of the CGF, this implies that

$$\boldsymbol{\kappa}_t(1, y) \leq \frac{1}{\theta}\,\boldsymbol{\kappa}_t(\theta, 0) + \left(1 - \frac{1}{\theta}\right)\boldsymbol{\kappa}_t(0, x)$$

or (rearranging and using the fact that $x = \frac{\theta}{\theta-1}y$)

$$\boldsymbol{\kappa}_t(\theta, 0) \geq \theta\,\boldsymbol{\kappa}_t(1, y) + (1 - \theta)\,\boldsymbol{\kappa}_t\left(0, \frac{\theta}{\theta - 1}y\right). \tag{30}$$

Equivalently, exponentiating and using the definition of $\boldsymbol{\kappa}_t(\cdot, \cdot)$, we have

$$\mathbb{E}_t\left[(M_{t+1}R_{f,t+1})^\theta\right] \geq \left\{\mathbb{E}_t^*\left[\left(\frac{R_{t+1}}{R_{f,t+1}}\right)^y\right]\right\}^\theta \left\{\mathbb{E}_t\left[\left(\frac{R_{t+1}}{R_{f,t+1}}\right)^{\frac{\theta}{\theta-1}y}\right]\right\}^{1-\theta}. \tag{31}$$

As y and R_{t+1} are arbitrary, inequality (26) follows.

$\underline{\theta < 0}$. We write $(0, x)$ as the convex combination $(0, x) = \frac{1}{1-\theta}(\theta, 0) + \left(1 - \frac{1}{1-\theta}\right)(1, y)$. It follows by convexity of the CGF that

$$\boldsymbol{\kappa}_t(0, x) \leq \frac{1}{1 - \theta}\,\boldsymbol{\kappa}_t(\theta, 0) + \left(1 - \frac{1}{1 - \theta}\right)\boldsymbol{\kappa}_t(1, y). \tag{32}$$

This inequality can be rearranged to give (30) above, and hence (31), from which the result follows because y is arbitrary.

$\underline{\theta \in (0, 1)}$. We write $(\theta, 0)$ as the convex combination $(\theta, 0) = (1 - \theta)(0, x) + \theta(1, y)$. By convexity of the CGF, and the fact that $x = \frac{\theta}{\theta-1}y$, it follows that

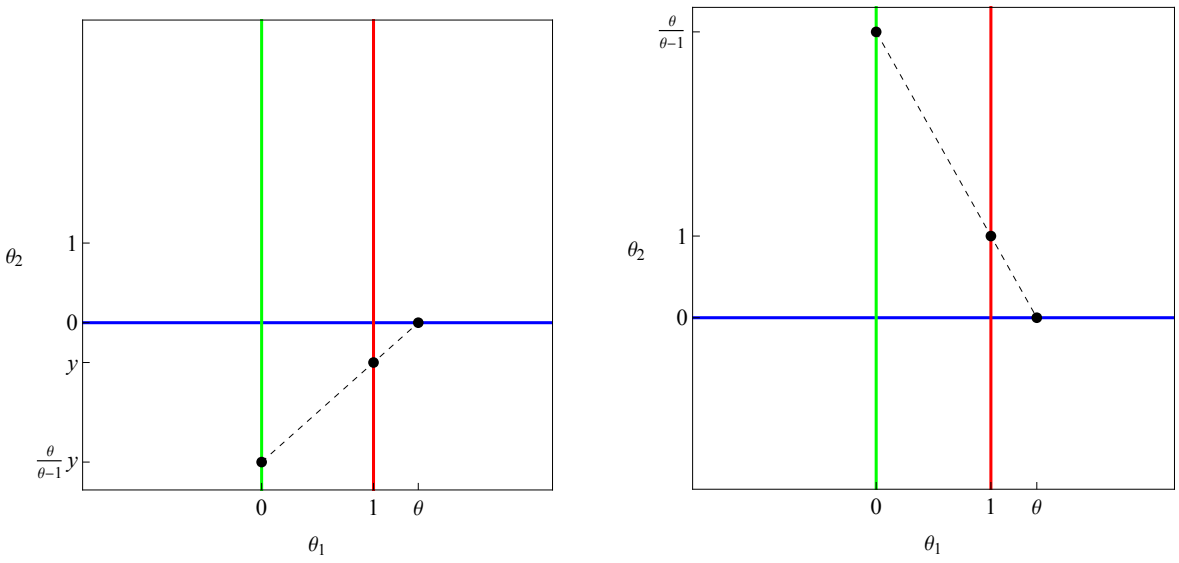
$$\boldsymbol{\kappa}_t(\theta, 0) \leq \theta\,\boldsymbol{\kappa}_t(1, y) + (1 - \theta)\,\boldsymbol{\kappa}_t\left(0, \frac{\theta}{\theta - 1}y\right). \tag{33}$$

Exponentiating, it follows that

$$\mathbb{E}_t\left[(M_{t+1}R_{f,t+1})^\theta\right] \leq \left\{\mathbb{E}_t^*\left[\left(\frac{R_{t+1}}{R_{f,t+1}}\right)^y\right]\right\}^\theta \left\{\mathbb{E}_t\left[\left(\frac{R_{t+1}}{R_{f,t+1}}\right)^{\frac{\theta}{\theta-1}y}\right]\right\}^{1-\theta}.$$

As y and R_{t+1} are arbitrary, inequality (27) follows. □

The left panel of Figure 2 illustrates the idea behind the proof of Result 1. We imagine ourselves standing on the blue line at the point $(\theta, 0)$ and looking across the CGF surface towards the points $(0, \frac{\theta}{\theta-1}y)$ and $(1, y)$, which lie on the green and red lines respectively. Empirically, we will work unconditionally; the convexity of the CGF surface then supplies bounds on $\boldsymbol{\kappa}(\theta, 0)$ in terms of $\boldsymbol{\kappa}(0, \frac{\theta}{\theta-1}y)$ (which is observable based on the time series of realized returns) and $\boldsymbol{\kappa}(1, y)$ (which is observable based on option prices).



(a) The case of general y.

(b) The case $y = 1$ recovers the Snow bound, which fails to exploit information in the risk-neutral distribution because $\boldsymbol{\kappa}_t(1,1) \equiv 0$.

Figure 2: The use of convexity in the proof of Result 1.

In particular, if we set $y = 1$ in inequalities (26) and (27) then we recover the Snow (1991) bounds:[13]

$$\mathbb{E}\left[(M_{t+1}R_{f,t+1})^{\theta}\right] \geq \sup_{R_{t+1}} \left\{\mathbb{E}\left[\left(\frac{R_{t+1}}{R_{f,t+1}}\right)^{\frac{\theta}{\theta-1}}\right]\right\}^{1-\theta} \quad \text{for } \theta < 0 \text{ or } \theta > 1 \tag{34}$$

and the reverse inequality (with an infimum rather than a supremum) for $\theta \in (0,1)$. This is the special case, illustrated in the right panel of Figure 2, in which, by looking in a fixed direction through the point (1,1), we fail to use any nontrivial information in the risk-neutral distribution. (Recall that $\boldsymbol{\kappa}(1,1) = 0$ in any economy because $\mathbb{E}(M_{t+1}R_{f,t+1}) = 1$.) In the case $y = 1$ and $\theta = 2$, for example, inequality (34) simplifies to

$$\mathbb{E}\left[(M_{t+1}R_{f,t+1})^{2}\right] \geq \sup_{R_{t+1}} \left\{\mathbb{E}\left[\left(\frac{R_{t+1}}{R_{f,t+1}}\right)^{2}\right]\right\}^{-1}. \tag{35}$$

The Hansen and Jagannathan (1991) bound follows by applying this inequality to the gross return R_{t+1} with minimum second moment.[14]

[13] Orłowski, Sali, and Trojani (2018) point out that convexity of the joint CGF of M_{t+1} and R_{t+1} can be used in this way to derive existing bounds in the literature. The idea of using CGF convexity to derive bounds also appears, in a different context, in Martin (2013a,b,c).

[14] This return satisfies $\left\{\mathbb{E}\left[\left(\frac{R_{t+1}}{R_{f,t+1}}\right)^{2}\right]\right\}^{-1} = 1 + \text{max. Sharpe ratio}^2$ (Hansen and Richard, 1987). The

The two panels of Figure 2 make two points. First, they illustrate the novel feature of Result 1 that, by using the information in option prices, we can "look around", viewing y as a free parameter that can be optimized to deliver sharp bounds on $\boldsymbol{\kappa}(\theta, 0)$. When y is very large in magnitude, for example, the bounds exploit information about the extreme tails of the risk-neutral distribution. (Turning this round, an econometrician may wish to restrict the range of values of y under consideration in order to avoid using extreme tail information, if this is felt to be unreliable for some reason. We discuss this possibility further below.)

Second, the right panel shows that constraining $y = 1$—that is, "looking through the point (1,1)", as the prior literature does—is particularly problematic when θ is close to 1, because in this case the Snow bound (34) relates the θth moment of the SDF to a very large positive (if $\theta > 1$) or large negative (if $\theta < 1$) moment of returns, which (in either case) is hard to estimate empirically. As our bound does not require this constraint, it is well-behaved when θ is close to one, and we will show below that the θth moment of the SDF exhibits interesting properties in this regime. More generally, however, the ability to choose y optimally supplies new insights even in the case $\theta = 2$, as we will now see.

2.1.1 Instability of volatility bounds

Setting $\theta = 2$ in Result 1 and taking logs, we have a lower bound on the second moment of the SDF:

$$\boldsymbol{\kappa}(2, 0) \geq \sup_{y \in \mathbb{R}} 2\,\boldsymbol{\kappa}(1, y) - \boldsymbol{\kappa}\,(0, 2y)\,. \tag{36}$$

The right-hand side of (36) features the difference of two convex functions.[15] The maximization problem on the right-hand side is therefore not in general well-behaved. This issue is not special to the case $\theta = 2$: the same problem arises for any $\theta > 1$ or $\theta < 0$, as the bound (28) shows.

It is often assumed that the risky return and SDF are conditionally jointly lognormal. In this special case, the bound (36) *is* (conditionally) well-behaved: with the CGF given by (14), inequality (36) becomes

$$\boldsymbol{\kappa}_t(2, 0) \geq -\sigma^2 y^2 - 2\mu y, \tag{37}$$

and as there is a negative coefficient on the quadratic term, the right-hand side has an interior maximum in y.

Hansen–Jagannathan bound follows: $\text{var}[M_{t+1}R_{f,t+1}] = \mathbb{E}[(M_{t+1}R_{f,t+1})^2] - 1 \geq \text{max. Sharpe ratio}^2$.

[15] Each of $\boldsymbol{\kappa}(1, y)$ and $\boldsymbol{\kappa}(0, 2y)$ is a CGF, and CGFs are always convex.

But this does not hold in general, even if true and risk-neutral returns are each individually known to be lognormally distributed so that $\boldsymbol{\kappa}_t(1, y)$ and $\boldsymbol{\kappa}_t(0, y)$ are each quadratic. For example, in the Martin and Papadimitriou (2022) model with CGF (24), inequality (36) becomes

$$\boldsymbol{\kappa}_t(2, 0) \geq \frac{1-\delta}{\delta}\sigma^2 y^2 - \frac{2(1+\delta)}{\delta}\sigma^2 y\,. \tag{38}$$

If $\delta \leq 1$, the right-hand side of (38) can be made arbitrarily large by sending y to minus infinity. This implies that SDF volatility is infinite. Within the MP model, the arbitrarily high Sharpe ratios attainable via option-based strategies reveal this fact without needing to know the full equilibrium or the precise form of the CGF given in (24).

While the MP model is merely a proof of concept—which we mention here, together with the learning-based models of Example 3, to encourage readers to question their assumptions about the finiteness of SDF variance—our empirical results below are consistent with a substantial empirical literature that has argued that short positions in deep-out-of-the-money options can indeed earn very high Sharpe ratios (Jackwerth, 2000; Coval and Shumway, 2001; Bondarenko, 2003; Jones, 2006; Driessen and Maenhout, 2007; Goetzmann et al., 2007; Santa-Clara and Saretto, 2009; Broadie et al., 2009).

The next result, whose proof is in the Appendix, shows that the situation is different for moments between zero and one.

Result 2. *The bounds in Result 1 are well-behaved when $\theta \in (0, 1)$, in the sense that the minimization problem over y on the right-hand side of inequality (27) has a unique interior minimum.*

These intermediate moments, with $\theta \in (0, 1)$, are therefore a reliable object for empirical study. We can continue to think of such bounds as saying that the SDF must be sufficiently variable in some sense.[16] For example, an upper bound on $\boldsymbol{\kappa}(1/2, 0)$, and hence on $\mathbb{E}\sqrt{M_{t+1}R_{f,t+1}}$, gives a *lower* bound on $\text{var}\sqrt{M_{t+1}R_{f,t+1}} = 1 - \left(\mathbb{E}\sqrt{M_{t+1}R_{f,t+1}}\right)^2$. In Section 2.3, we show how to interpret these intermediate moments as measures of the attractiveness of investment opportunities.

The bounds are more sensitive to right-tail behavior of the SDF when $\theta \in (0, 1)$ takes relatively larger values, and to left-tail behavior of the SDF when θ takes relatively smaller values. At either end of the interval $(0, 1)$, Result 1 becomes trivial (because the zeroth

[16] When the SDF is highly variable, its CGF—the blue line in Figure 1—is highly convex. As the CGF is constrained to equal zero when θ equals 0 or 1, a highly convex CGF is associated with *small* (that is, more negative) values of SDF moments when $\theta \in (0, 1)$.

and first moment of the SDF are each pinned down). Nonetheless, it is possible to derive useful bounds on the gradient of $\boldsymbol{\kappa}(\theta, 0)$ at the two end points, and these gradients can be interpreted in terms of familiar entropy measures.

2.2 Entropy bounds

Given a random variable, X, Theil (1967) introduced two entropy measures. The first is defined as

$$L_t^{(1)}(X) \equiv \mathbb{E}_t\left(X \log X\right) - \left(\mathbb{E}_t X\right) \log\left(\mathbb{E}_t X\right) ,$$

and the second measure is defined as

$$L_t^{(2)}(X) \equiv \log\left(\mathbb{E}_t X\right) - \mathbb{E}_t\left(\log X\right) .$$

We define the corresponding unconditional entropy measures by dropping subscripts t on the expectation operators.

Each of these can be interpreted as a measure of variability, as they each take the form $\mathbb{E}_t f(X) - f(\mathbb{E}_t X)$ for appropriately chosen convex functions f: the first entropy measure sets $f(x) = x \log x$ and the second sets $f(x) = -\log x$. Like variance—which sets $f(x) = x^2$—they are strictly positive (unless X is constant, in which case they are zero).

The first entropy measure was proposed by Stutzer (1995) as a measure of SDF variability. The second entropy measure was used by Alvarez and Jermann (2005); it satisfies the Bansal and Lehmann (1997) bound

$$L_t^{(2)}(M_{t+1}R_{f,t+1}) \geq \sup_{R_{t+1}} \mathbb{E}_t \log \frac{R_{t+1}}{R_{f,t+1}}. \tag{39}$$

Our next result provides a new bound for the first entropy measure and generalizes the bound (39) on the second entropy measure.

Result 3. *The entropy measures have geometrical interpretations:* $L_t^{(1)}(M_{t+1}R_{f,t+1})$ *is the slope of* $\boldsymbol{\kappa}_t(\theta, 0)$ *at* $\theta = 1$*, and* $L_t^{(2)}(M_{t+1}R_{f,t+1})$ *is the absolute value of the slope at* $\theta = 0$*.*

$L_t^{(1)}(M_{t+1}R_{f,t+1})$ *satisfies*

$$L_t^{(1)}(M_{t+1}R_{f,t+1}) \geq \sup_{R_{t+1}} \sup_{y \in \mathbb{R}} y\, \mathbb{E}_t^* \log \frac{R_{t+1}}{R_{f,t+1}} - \log \mathbb{E}_t \left[\left(\frac{R_{t+1}}{R_{f,t+1}}\right)^y\right]. \tag{40}$$

$L_t^{(2)}(M_{t+1}R_{f,t+1})$ *satisfies*

$$L_t^{(2)}(M_{t+1}R_{f,t+1}) \geq \sup_{R_{t+1}} \sup_{y \in \mathbb{R}} y\, \mathbb{E}_t \log \frac{R_{t+1}}{R_{f,t+1}} - \log \mathbb{E}_t^* \left[\left(\frac{R_{t+1}}{R_{f,t+1}}\right)^y\right] . \tag{41}$$

The same bounds hold unconditionally (that is, dropping subscripts t*).*

Proof. The interpretation of the first (resp., second) entropy measures in terms of the slope of the CGF follows directly, by differentiating $\boldsymbol{\kappa}_t(\theta, 0)$ with respect to θ and setting θ equal to one (resp., zero).

As the CGF is a differentiable convex function, we have

$$\boldsymbol{\kappa}_t(\boldsymbol{\alpha}) \geq \boldsymbol{\kappa}_t(\boldsymbol{\beta}) + \{\nabla \boldsymbol{\kappa}_t(\boldsymbol{\beta})\}^\top \cdot (\boldsymbol{\alpha} - \boldsymbol{\beta})$$

for all $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. Applying this property with $\boldsymbol{\alpha} = (0, y)$ and $\boldsymbol{\beta} = (1, 0)$, we find $\boldsymbol{\kappa}_t(0, y) \geq y\, \boldsymbol{\kappa}_t^{(2)}(1, 0) - \boldsymbol{\kappa}_t^{(1)}(1, 0)$, or equivalently,

$$\boldsymbol{\kappa}_t^{(1)}(1, 0) \geq y\, \boldsymbol{\kappa}_t^{(2)}(1, 0) - \boldsymbol{\kappa}_t(0, y), \tag{42}$$

where

$$\boldsymbol{\kappa}_t^{(1)}(1, 0) = \frac{\partial\, \boldsymbol{\kappa}_t}{\partial \theta_1}(1, 0) = \mathbb{E}_t\left[(M_{t+1} R_{f,t+1}) \log\left(M_{t+1} R_{f,t+1}\right)\right]$$

and

$$\boldsymbol{\kappa}_t^{(2)}(1, 0) = \frac{\partial\, \boldsymbol{\kappa}_t}{\partial \theta_2}(1, 0) = \mathbb{E}_t\left[(M_{t+1} R_{f,t+1}) \log\left(R_{t+1}/R_{f,t+1}\right)\right].$$

Rewriting the second of these in terms of risk-neutral expectations, we have

$$\mathbb{E}_t\left[M_{t+1} R_{f,t+1} \log\left(M_{t+1} R_{f,t+1}\right)\right] \geq y\, \mathbb{E}_t^* \log \frac{R_{t+1}}{R_{f,t+1}} - \log \mathbb{E}_t\left[\left(\frac{R_{t+1}}{R_{f,t+1}}\right)^y\right].$$

As y is arbitrary, inequality (40) follows.

Applying the same convexity property with $\boldsymbol{\alpha} = (1, y)$ and $\boldsymbol{\beta} = (0, 0)$, and then rearranging, we have

$$-\boldsymbol{\kappa}_t^{(1)}(0, 0) \geq y\, \boldsymbol{\kappa}_t^{(2)}(0, 0) - \boldsymbol{\kappa}_t(1, y), \tag{43}$$

where

$$\boldsymbol{\kappa}_t^{(1)}(0, 0) = \frac{\partial\, \boldsymbol{\kappa}_t}{\partial \theta_1}(0, 0) = \mathbb{E}_t \log\left(M_{t+1} R_{f,t+1}\right),$$

$$\boldsymbol{\kappa}_t^{(2)}(0, 0) = \frac{\partial\, \boldsymbol{\kappa}_t}{\partial \theta_2}(0, 0) = \mathbb{E}_t \log\left(R_{t+1}/R_{f,t+1}\right).$$

This can be rewritten as

$$-\mathbb{E}_t \log\left(M_{t+1} R_{f,t+1}\right) \geq y\, \mathbb{E}_t \log \frac{R_{t+1}}{R_{f,t+1}} - \log \mathbb{E}_t^*\left[\left(\frac{R_{t+1}}{R_{f,t+1}}\right)^y\right].$$

As y is arbitrary, inequality (41) follows.

The unconditional version of the result has the same proof in each case, after dropping subscripts t. □

The bounds in Result 3 are well-behaved, as they each feature suprema over the difference between a linear function of y and a convex function of y. From an empirical perspective, this means that the lower bounds we estimate using realized returns and option price data behave stably across data sources and sampling horizons.

The lower bound (41) is dual to the lower bound (40) in that the roles of the true and risk-neutral distributions are interchanged on the right-hand sides. If we set $y = 1$ on the right-hand side of (41), the dependence on the risk-neutral distribution disappears, because $\mathbb{E}_t^* R_{t+1}/R_{f,t+1} = 1$, and we recover the Bansal and Lehmann (1997) bound (39).

By contrast, the bound (40) on the first entropy measure does not relate to any known bound in the literature. The only way to avoid using information in the risk-neutral distribution of R_{t+1}, when implementing it, is to set $y = 0$ on the right-hand side. But in this case the bound reduces to the vacuous statement $L_t^{(1)}(M_{t+1}R_{f,t+1}) \geq 0$. So an essential feature of the bound (40) is that it plays off the true and risk-neutral distributions of R_{t+1} against one another: the bound is non-trivial only when it exploits *both* the true and risk-neutral distributions of R_{t+1}. This may explain why it has not been discovered in the prior literature, which has focused only on the true distributions of returns.

The difference between the two entropy measures can be expressed[17] in terms of the higher cumulants of the log SDF: writing κ_n for the nth cumulant of $\log(M_{t+1}R_{f,t+1})$,

$$L^{(1)}(M_{t+1}R_{f,t+1}) - L^{(2)}(M_{t+1}R_{f,t+1}) = \sum_{n=3}^{\infty} \frac{n-2}{n!}\kappa_n = \frac{\kappa_3}{6} + \frac{\kappa_4}{12} + \frac{\kappa_5}{40} + \frac{\kappa_6}{180} + \cdots . \quad (44)$$

Hence the gap between the first and second measure is large if the log SDF is right-skewed, has excess kurtosis, and so on. Conversely, if the SDF is lognormal then the higher cumulants of its logarithm are all equal to zero, $\kappa_n = 0$ for $n > 2$: in this case, the two entropy measures are equal to each other (and they each equal $\frac{1}{2}\operatorname{var}\log(M_{t+1}R_{f,t+1})$).

Of the well-behaved bounds studied in this paper—that is, the moment bounds for $\theta \in (0, 1)$, and the two entropy bounds—the bound (40) on the first entropy measure is the most sensitive to the right tail of the SDF distribution, and hence, in equilibrium models, to states in which marginal utility is high. The first entropy measure summarizes the rate at which the SDF moments rise as θ passes through 1, and it is well-behaved even though the SDF moments strictly above 1 are not.

[17] To see this, note that $\boldsymbol{\kappa}(\theta, 0) = \sum_{n=1}^{\infty} \kappa_n \frac{\theta^n}{n!}$. By Result 3, the first and second entropy measures represent the absolute value of the slope of this function at $\theta = 1$ and $\theta = 0$, respectively, so $L^{(1)}(M_{t+1}R_{f,t+1}) = \sum_{n=1}^{\infty} \frac{\kappa_n}{(n-1)!}$ and $L^{(2)}(M_{t+1}R_{f,t+1}) = -\kappa_1$. Moreover, the fact that $\mathbb{E}(M_{t+1}R_{f,t+1}) = 1$ implies that $\boldsymbol{\kappa}(1, 0) = 0$, so $\kappa_1 = -\sum_{n=2}^{\infty} \frac{\kappa_n}{n!}$. Putting these facts together, equation (44) follows.

2.2.1 Measuring risk aversion

A classical result of Merton (1969) and Samuelson (1969) shows that if an investor with constant relative risk aversion, γ, chooses to invest fully in an asset whose returns are iid lognormal, then γ can be inferred from the ratio of the asset's risk premium to its return variance. The next result (which we implement empirically in Table 4 below) shows that the optimizing values of y in Result 3 reveal γ in a more general setting.

Result 4 (Merton–Samuelson Redux)**.** *If a myopic investor with constant relative risk aversion, γ, chooses to invest fully in the asset with return R_{t+1}, then the optimizing values of y in inequalities (40) and (41) for the SDF entropy measures are $-\gamma$ and γ, respectively.*

Result 4 replaces the iid lognormal assumption of Merton and Samuelson with the strictly weaker assumption[18] that the investor behaves myopically, so that the SDF is proportional to $R_{t+1}^{-\gamma}$. It is easy to check that Result 4 recovers the Merton–Samuelson result in the lognormal special case: with a lognormal CGF, as in equation (14), it implies that risk aversion satisfies $\gamma = \mu/\sigma^2$, where μ is the risk premium and σ^2 is return variance.[19]

2.3 An economic interpretation of the bounds

The variance of the SDF has the appealing property that it can be related to measures of the attractiveness of investment opportunities without any assumptions on the underlying stochastic processes, if one is prepared to adopt the perspective of a one-period investor with quadratic utility so that the Sharpe ratio can be taken as an index of investment opportunities. Specifically, the Hansen–Jagannathan bound states that the volatility of $M_{t+1}R_{f,t+1}$ must exceed the maximum attainable Sharpe ratio. In a similar vein, the Bansal–Lehmann bound relates the (second) entropy measure of the SDF to the maximum attainable expected log return, which measures the attractiveness of investment opportunities from the perspective of an investor with log utility.

We now show that similar properties hold for the intermediate moments $\mathbb{E}\left[(M_{t+1}R_{f,t+1})^{\theta}\right]$, where $\theta \in (0,1)$.

Our analysis adopts the perspective of a one-period investor with constant relative risk aversion equal to γ (of which log utility is a special case). The attractiveness of investment

[18] Merton and Samuelson show that a CRRA investor behaves myopically if returns are iid.

[19] Under lognormality, return variance is identical under both the true and risk-neutral distributions. Our approach allows true variance and risk-neutral variance to differ, as documented by a large empirical literature on the variance risk premium.

opportunities can be then quantified using the willingness-to-pay (WTP) to trade.[20] This is the proportional fraction of wealth, g_γ, that the investor would be prepared to pay in order to be allowed to trade risky assets (as opposed to being allowed to trade only the riskless asset). This satisfies

$$\frac{1}{1-\gamma}\,\mathbb{E}\left(\left[e^{-g_\gamma}W_t\right]^{1-\gamma}R_{\gamma,t+1}^{1-\gamma}\right)=\frac{1}{1-\gamma}\,\mathbb{E}\left(W_t^{1-\gamma}R_{f,t+1}^{1-\gamma}\right), \tag{45}$$

where we write $R_{\gamma,t+1}$ for the return on the investor's optimal strategy. It follows that

$$g_\gamma=\frac{1}{1-\gamma}\log\mathbb{E}\left[\left(\frac{R_{\gamma,t+1}}{R_{f,t+1}}\right)^{1-\gamma}\right]. \tag{46}$$

This equation motivates an asset-level performance measure,[21]

$$\phi(R_{t+1})=\frac{1}{1-\gamma}\log\mathbb{E}\left[\left(\frac{R_{t+1}}{R_{f,t+1}}\right)^{1-\gamma}\right]. \tag{47}$$

If offered a choice between a collection of different investment opportunities with returns $\{R_{i,t+1}\}_{i=1,\ldots,N}$, a myopic power utility investor would choose the opportunity with highest $\phi(R_{i,t+1})$. (In the log utility case, i.e., the limit as $\gamma\to 1$, $\phi(R_{i,t+1})$ equals $\mathbb{E}\log R_{i,t+1}-\log R_{f,t+1}$: the investor chooses the opportunity with the highest expected log return.)

The next result shows that our bounds translate into bounds on g_γ. To state the result concisely, we write $B(\theta)$ for the bound on the right-hand side of inequalities (26) and (27) in Result 1 (as illustrated in Figures 4 and 6), written as a function of θ. That is,

$$B(\theta)=\inf_{y\in\mathbb{R}}\theta\,\boldsymbol{\kappa}(1,y)+(1-\theta)\,\boldsymbol{\kappa}\left(0,\frac{\theta}{\theta-1}y\right)\quad\text{when }\theta\in(0,1).$$

When $\theta\notin(0,1)$, we define $B(\theta)$ similarly, but with a supremum over $y\in\mathbb{R}$.

Result 5 (The relationship between SDF moments and WTP)**.** *The lower bounds are informative about the attractiveness of the investment opportunity set:*

$$g_\gamma\geq\left|\frac{B(1-1/\gamma)}{1-1/\gamma}\right|\quad\textit{for }\gamma>0,\ \gamma\neq 1. \tag{48}$$

[20] Černý (2003) points out the unattractiveness of the Sharpe ratio as a performance measure and proposes a related approach based on generalised Sharpe ratios, which are defined in such a way that they reduce to the true Sharpe ratio as the Sharpe ratio tends to zero. Our WTP measure is a monotonic function of Černý's generalised Sharpe ratio in the CRRA case.

[21] This performance measure is intentionally easy to calculate. Real-world investors have long horizons, background risk, and so on, but we choose not to incorporate such considerations into the performance measure as the details of horizon length and background risk vary from person to person.

The right-hand side of this inequality is the absolute value of the slope of the line that joins the origin to $B(\theta)$ at the value $\theta = 1 - 1/\gamma$.

In the log utility case, we have $g_1 \geq |B'(0)|$. The right-hand side of this inequality is the absolute value of the tangent to $B(\theta)$ at $\theta = 0$.

Writing $\tau = 1/\gamma$, we can define WTP as a function of risk tolerance as opposed to risk aversion, $g(\tau) = g_{1/\tau}$. Then we have $g'(0) = L^{(1)}(M_{t+1}R_{f,t+1})$, so that the first entropy measure[22] *reveals the marginal increase in WTP when someone with zero risk tolerance becomes slightly risk tolerant.*

We use this result below to translate our moment bounds into more easily interpretable WTP numbers.

3 From theory to data

In our empirical work, we apply our bounds to the return on the S&P 500 index. For the rest of the paper, we use R_{t+1} to denote this return.

As noted above, we work unconditionally throughout our empirical work, in order to avoid restrictive assumptions about which state variables are relevant for calculating conditional moments. We estimate the unconditional CGF (11) via a time-series average over a long sample of realized returns,

$$\widehat{\boldsymbol{\kappa}}(0,\theta) = \log \frac{1}{T}\sum_{t=0}^{T-1} \left(R_{t+1}/R_{f,t+1}\right)^{\theta} . \tag{49}$$

Similarly, we replace $\mathbb{E}$ with $1/T\sum_{t=1}^{T}$ in (13) to give our empirical implementation of the unconditional risk-neutral CGF,

$$\widehat{\boldsymbol{\kappa}}(1,\theta) = \log \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}_t^*\left[\left(R_{t+1}/R_{f,t+1}\right)^{\theta}\right] = \log \frac{1}{T}\sum_{t=1}^{T} \exp\left(\boldsymbol{\kappa}_t(1,\,\theta)\right), \tag{50}$$

where the term inside the sum can be computed from option prices on any given date, as shown in equation (5).

The quantities on the right-hand sides of equations (49) and (50) are CGFs calculated with respect to particular measures—the empirical measure in the case of (49) and a time average of conditional risk-neutral measures in the case of (50)—so they have the usual properties of CGFs: they are convex and they pass through the origin.

[22] Stutzer (1995) offers a similar gain-from-trade interpretation for this entropy measure from the perspective of an investor with constant *absolute* risk aversion.

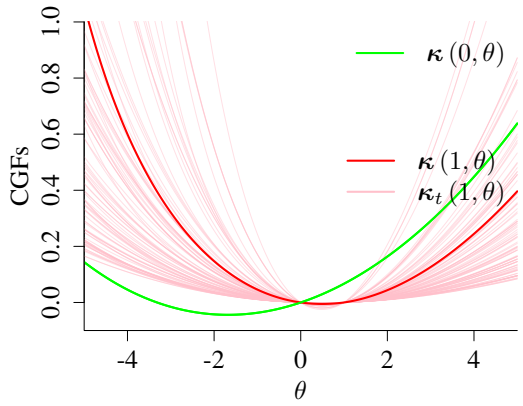


Figure 3: The CGFs observable from realized S&P 500 returns and SPX index option prices. Two observable slices of the CGF surface in the data. The solid green line is the unconditional true CGF calculated using realized one-month returns on the S&P 500 index from 1872 to 2022. Thin red lines indicate conditional risk-neutral CGFs for one-month S&P 500 index returns on 100 randomly selected trading days drawn from the 1996–2022 sample, calculated as in (5); the solid red line is the unconditional risk-neutral CGF $\boldsymbol{\kappa}(1,\theta)$, calculated according to (50).

Figure 3 illustrates these quantities in our dataset using S&P 500 index returns from 1872 to 2022 and option price data from 1996 to 2022. (A detailed description of our data is deferred to Section 3.3.)

3.1 Conservative measurement of risk-neutral quantities

Our results exploit the fact that risk-neutral expectations of powers of returns are observable via option prices, as shown in equation (5). It is conventional in the literature to treat mid-market prices as exact measures of risk-neutral quantities. In reality, the existence of a bid-offer spread means that we only observe a *region* in which the risk-neutral quantity lies. To be conservative, we use bid and offer prices throughout our empirical work. If we used mid-market prices, our results would become more extreme: the lower bounds would increase, the upper bounds would decrease, and the singularity would move closer to one.

More specifically, we use whichever of bid or offer prices gives the conservative choice in any given situation. Suppose, for example, we are interested in the θth moment of the SDF for some $\theta > 1$. In this case, the right-hand side of inequality (26) in Result 1 represents a lower bound on the θth moment, and the conservative choice is therefore to make $\mathbb{E}_t^* \left[(R_{t+1}/R_{f,t+1})^y \right]$—equivalently, $\boldsymbol{\kappa}_t(1,y)$—small. Recalling equation (5) and the subsequent

discussion, this is achieved by using the offer prices of options if the optimizing value of y is between zero and one, and bid prices of options if it is less than zero or greater than one.

If we are interested in measuring the θth moment of the SDF for some $\theta < 0$, there is a sign flip: the right-hand side of inequality (26) continues to represent a lower bound, but because θ is now negative, the conservative choice is to make $\boldsymbol{\kappa}_t(1, y)$ large. We therefore use bid prices when y is between zero and one, and offer prices when y is less than zero or greater than one.

There is a sign flip of a different form when $\theta \in (0, 1)$: we now use inequality (27), which supplies an *upper* bound on the θth SDF moment. With θ positive, the conservative choice is to make $\boldsymbol{\kappa}_t(1, y)$ large. As in the previous paragraph—but for a different reason—we therefore use bid prices when y is between zero and one, and offer prices when y is less than zero or greater than one.

3.2 Finite sample considerations

In the data, it often happens that options with strikes so far out of the money that the corresponding return has never been realized in sample have positive bid prices. The worst monthly return in the sample, of -29%, occurred in September 1931; and put options with strikes more than 29% out of the money have positive *bid* prices on 73% of days in our sample (or 16% of days after we apply our filters).[23] These options are so far out of the money that they would never have paid off in the sample period, so that selling them at a positive price looks like an arbitrage opportunity to a naive econometrician. Our version of the data-mining problem, then, is that by shorting sufficiently far out-of-the-money options we can construct strategies with superficially very attractive in-sample risk-reward characteristics.

Recall that when $\theta < 0$ or $\theta > 1$, the bound of Result 1 is equivalent to the inequality (30). We implement this bound (in its unconditional form) in the conventional way, replacing expectations with time-series averages, so that the estimated lower bound on the θth moment of the SDF is

$$\boldsymbol{\kappa}(\theta, 0) \geq \sup_{y} \; \theta\, \widehat{\boldsymbol{\kappa}}(1, y) + (1 - \theta)\, \widehat{\boldsymbol{\kappa}}\left(0, \frac{\theta}{\theta - 1} y\right) \tag{51}$$

where $\widehat{\boldsymbol{\kappa}}(0, \cdot)$ and $\widehat{\boldsymbol{\kappa}}(1, \cdot)$ are the empirical estimates of the true and risk-neutral CGFs, as defined in equations (49) and (50).

[23] The most extreme example is that we observe a put option with strike $K/S_0 = 0.33$ and a positive bid price on March 19, 2020. This option would only pay out if the market dropped by more than 67% over a month.

Result 6. *Fix $\theta > 1$. Suppose that one (or both) of the following conditions hold: (i) The most extreme put strike (of a put with positive bid price) in the dataset is lower than the lowest observed return in sample,* $\min_t K_{\min,t}/R_{f,t+1} < \min_t R_{t+1}/R_{f,t+1}$*; (ii) The most extreme call strike (of a call with positive bid price) in the dataset is higher than the highest observed return in sample,* $\max_t K_{\max,t}/R_{f,t+1} > \max_t R_{t+1}/R_{f,t+1}$*. Then the lower bound in equation (51) can be made arbitrarily large.*

We have already seen examples of models that feature divergent SDF moments, and we have shown that the bounds in Result 1 need not be well-behaved for θth moments if $\theta > 1$ or $\theta < 0$. Result 6 sharpens the issue: it shows (under conditions that hold in the data) that if $\theta > 1$ the empirical lower bound on the θth moment of the SDF will *always* diverge in a finite sample.

The only truly robust response—and our preferred response—to this issue is to restrict attention to θth moment bounds with $\theta \in (0, 1)$ and to entropy bounds. These are well-behaved, as we showed in Section 2, and below we will see that these bounds have a natural economic interpretation.

The prior literature has focussed attention outside this range, however. To connect to this literature by allowing $\theta \notin (0, 1)$, one can take a more pragmatic approach by constraining the range of risk-neutral moments that are exploited in the bounds of Result 1: that is, by constraining y to lie in some range $[-B, B]$ where $B > 0$ is a fixed parameter. This avoids the problems associated with apparent in-sample arbitrage opportunities, but it creates the new problem that B is arbitrary.

For bounds on the lower moments of the SDF—when θ is greater than, but sufficiently close to, 1, or when θ is negative—we find, however, that there is an interior optimum which remains constant as we change B within some range. This implies that the data-implied bound is independent of B within this range, as we would wish (and in practice we find, encouragingly, that the interior optima $y^*(\theta)$ that arise in our empirical study tend to be small in magnitude).

In contrast, when we seek bounds on higher moments of the SDF—for values of θ above, say, two—we find that the optimizing value of y is always a corner solution: for any given B, the optimizing value of y equals $-B$ or B. In the case $\theta = 2$, for example, this is analogous to finding that deeper and deeper out-of-the-money options have higher and higher Sharpe ratios.

In short, we can distinguish three different ranges of values of θ. When $\theta \in (0, 1)$, there is always an interior optimum, so that the lower bound is guaranteed to be finite both in

population and in finite samples. Using a traffic light coloring scheme, we can think of this as the *green* range, and we indicate it as such in our figures. Only this range is truly robust in general. There is also a *yellow* range in which $\theta \notin (0,1)$ but there are interior optima in y. Finally, the *red* range encompasses the remaining values of $\theta \notin (0,1)$ for which there are no interior optima, so that the bounds diverge monotonically to infinity. While the boundaries between the green and yellow ranges always occur at $\theta = 0$ and $\theta = 1$, the locations of the boundaries between yellow and red are an empirical question. We refer to the boundary value of $\theta > 0$ that divides the yellow and red ranges as the *singularity*.

3.3 Data

3.3.1 Equity market returns

Our monthly stock market return data is drawn from three sources: the US stock returns compiled by (Schwert, 1990, `https://www.billschwert.com/gws_data.htm`), the S&P 500 index from the Global Financial Data (GFD), and the Center for Research in Security Prices (CRSP). As the return time-series before 1872 are aggregated from a handful of bank and railroad stocks, we focus on the post-1872 sample. We download annual US stock market return data from the Macrohistory database introduced in Jordà, Knoll, Kuvshinov, Schularick, and Taylor (2019, JKKST) covering 1872 through 2020. Lastly, we use daily return data to construct realized monthly returns—and returns over other horizons when needed—in some of our analysis for robustness. The daily data is consolidated from Schwert (1990) and CRSP, covering February 17, 1885 through December 30, 2022.

3.3.2 Riskless rates

We also combine multiple data sources for the riskless rate observations. Before 1926, we use the US long-term bond yield data from the GFD, assuming a flat term structure. Between January 1926 and May 1961, we combine the monthly Fama–Bliss discount bond prices and the risk-free rate series from the CRSP US treasury database. Missing rates at specific maturities are interpolated linearly from the observed data. After June 1961, we use the yield curve data constructed and maintained by Liu and Wu (2021), which is available at daily frequency.

3.3.3 Option prices

We collect daily Cboe SPX index option prices from OptionMetrics, from January 4, 1996 to December 30, 2022. After March 6, 2008, OptionMetrics reports closing bid and offer prices at 3:59pm Eastern Time, synchronized with the equity market. Before that date, OptionMetrics reports the last quotes of the trading day, at 4pm for the equity market and 4:15pm for the index option market. In a robustness check in Section 4.1, we confirm that our results are almost identical if we restrict the option sample to the period following March 6, 2008.

We apply various filters to the options sample. We exclude contracts that expire within the next week. For each maturity-date combination, we require an available risk-free rate. Each option record must satisfy several conditions to remain in our sample: the implied volatility must be available; the best offer price must exceed the best bid price; the best bid price must be positive; and the contract must have positive open interest and must have been traded on the current day. When multiple puts (or calls) share the same strike price and maturity on a given date, we keep the observation with the highest trading volume. Table A1 summarizes the effect of imposing these filters on the size of the sample.

On each day in sample, we first linearly interpolate option prices between quoted strikes. We consider several different approaches to extrapolating outside the range of quoted strikes. They all give broadly similar results and, in particular, the finding that SDF moments explode rapidly when $\theta > 1$ emerges in every specification. In this range, the critical inputs to the bounds are the bid prices of options: the single most important empirical fact for us is that these are surprisingly high.

It is therefore particularly important to make conservative choices when extrapolating bid prices.[24] We exploit a classical argument based on the absence of so-called butterfly arbitrage trades, taking account of the cost of "crossing the spread". This argument allows us to extrapolate bid prices linearly outside the quoted range of strikes, using the *bid* price of the deepest out-of-the-money option and the *offer* price of the next deepest out-of-the-money option. This procedure applies for puts and for calls, and it relies only on the absence of static arbitrage opportunities.

Offer prices are less important for our key findings. In our baseline analysis, we extrapolate offer prices outside the range of strikes by linearly extrapolating the associated Black–Scholes (offer) implied volatilities, terminating once the corresponding option prices

[24] Figure A5, in the Appendix, shows that our results are broadly unchanged if we do not extrapolate at all, setting prices to zero outside the range of traded quotes.

fall below one cent.[25] In a robustness exercise in Section 4.1, we entertain the possibility that offer prices behave as badly as possible, subject to the absence of arbitrage. This has a modest effect on the moment bounds for values of θ less than about 0.5, but almost no effect above this range.

SPX options are available only on certain expiration dates, so we use time-weighted interpolation to create risk-neutral CGFs at fixed maturities (such as one month or one year). This is the approach used by Cboe to construct the VIX index. Table A2 explains the procedure in detail and reports the resulting sample sizes, for both daily and monthly observations, of the conditional risk-neutral CGFs.

4 Empirical results

Figure 4 plots the bounds on moments of the SDF for θ between -3 and 3 at the one-month horizon; we annualize the monthly bounds by multiplying by 12. Figure A2, in the appendix, reports similar results at the one-year horizon. The left panel shows results based on the full realized return sample (1872–2022). The right panel presents results using realized returns for the postwar period (1946–2022). In both panels, our sample of option prices (expiring in one month, and observed at month end) runs from 1996 to 2022.

The bounds are highly asymmetric, growing far more rapidly for positive than for negative θ. For comparison, recall that in the lognormal case (14) the SDF, $\boldsymbol{\kappa}(\theta, 0) = \frac{1}{2}\lambda^2\theta(\theta-1)$, is quadratic and symmetric around $\theta = 1/2$. In particular, the bounds diverge for values of θ below two, both in our long sample and in the post-war sample; as discussed in Section 3.2, we refer to the point of divergence as the singularity. The singularity is closer to one in the post-war sample, reflecting the fact that the post-war sample contains fewer very extreme market declines of the type that occurred during the Great Depression years.[26]

Recall that we use option bid and offer prices to calculate our bounds, and we extrapolate conservatively outside the range of traded strikes. Figure A3 shows that our results would appear more dramatic—the upper bounds would be lower and the lower bounds would be higher—if we used mid-market prices, extrapolated assuming a flat volatility smile.

[25] To be conservative, we impose a constraint that the slope of the linearly extrapolated implied volatilities should be at most zero for puts, and at least zero for calls.

[26] If, hypothetically, we were able to observe a long sample of option prices, including data from the Great Depression, it is likely that our results would become more dramatic, because measures of implied volatility would plausibly have been even higher during that period than is suggested by the unconditional mean of our more recent sample.

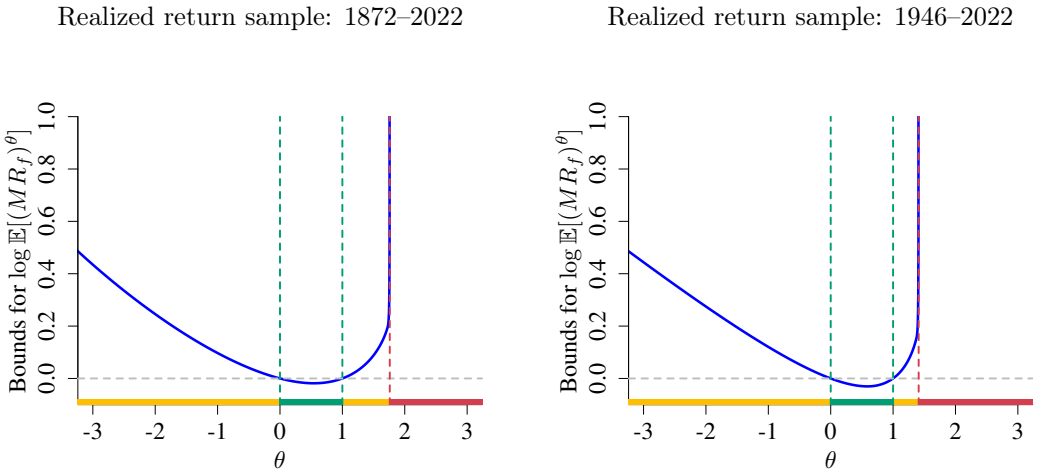


Figure 4: Convexity bounds for the moments of the SDF. Shading along the axis indicates the green, yellow and red regions as discussed in Section 3.2.

Applying Result 1 to the S&P 500 index. The blue curves are lower bounds on SDF moments when $\theta < 0$ or $\theta > 1$ and upper bounds when $\theta \in (0, 1)$.

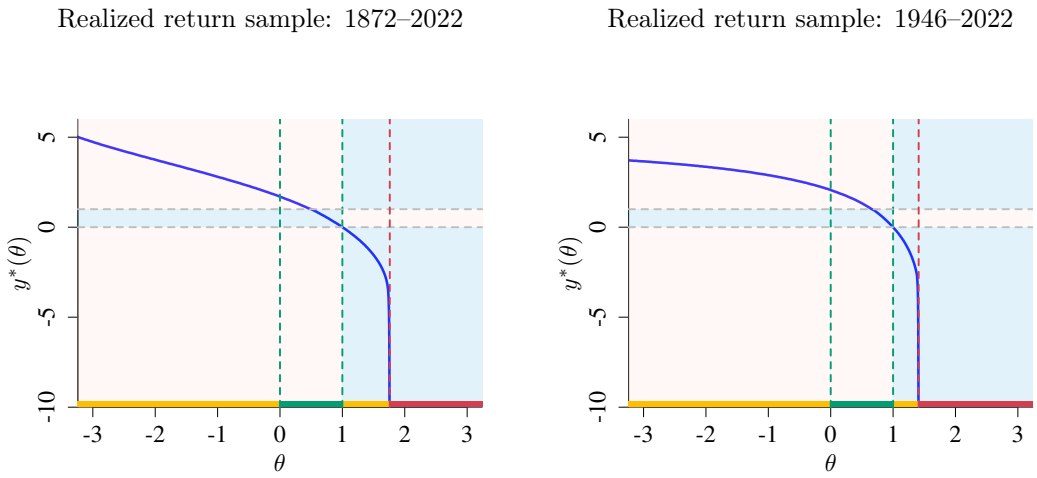


Figure 5: Optimizing values for the convexity bounds.

Applying Result 1 to the S&P 500 index. The blue curves show the values of y where (interior) optima are achieved, for different values of θ. Regions colored in light blue (red) are combinations of (θ, y) values such that we will use the bid (offer) option prices to construct the risk-neutral curves for evaluating our bounds, as discussed in Section 3.1.

Table 1: The locations of the singularity in $\boldsymbol{\kappa}(\theta, 0) = \log \mathbb{E}[(M_{t+1}R_{f,t+1})^{\theta}]$

This table reports the location of the singularity in (the bounds on) $\log \mathbb{E}[(M_{t+1}R_{f,t+1})^{\theta}]$. The estimation uses monthly US stock-market returns and month-end SPX option prices. The row "JKKST annual" in Panel (b) uses the annual realized return series from 1872–2020 in Jordà et al. (2019). The options data cover 1996 through 2022 for all specifications. Column 3 reports 95% confidence intervals using a block bootstrap with block length equal to the return horizon, resampling both series separately. Column 4 ($\mathbb{E}$-bootstrap) resamples only the realized returns used for estimating $\widehat{\boldsymbol{\kappa}}(0, \theta)$ in equation (49). Column 5 ($\mathbb{E}^*$-bootstrap) resamples only the risk-neutral conditional CGFs used for estimating $\widehat{\boldsymbol{\kappa}}(1, \theta)$ in equation (50).

sample	est.	bootstrap CI	$\mathbb{E}$-bootstrap CI	$\mathbb{E}^*$-bootstrap CI
(a) one-month horizon				
1872-2022	1.72	(1.52, 2.05)	(1.50, 1.97)	(1.62, 1.87)
1946-2022	1.38	(1.20, 1.60)	(1.20, 1.56)	(1.34, 1.45)
1996-2022	1.44	(1.23, 1.78)	(1.23, 1.77)	(1.39, 1.52)
(b) one-year horizon				
1872-2022	1.67	(1.27, 2.50)	(1.27, 2.05)	(1.60, 2.10)
1946-2022	1.28	(1.15, 1.50)	(1.14, 1.42)	(1.23, 1.38)
1996-2022	1.36	(1.14, 1.92)	(1.13, 1.73)	(1.32, 1.52)
JKKST annual	1.36	(1.23, 1.56)	(1.21, 1.52)	(1.32, 1.51)

Figure 5 plots, as a function of θ, the optimizing value of y that produces the convexity bounds shown in Figure 4. The same red, amber, and green regions are also shaded over the x-axis. In the data, $y^*(\theta)$ decreases monotonically with θ. It is positive for all $\theta < 1$ but plunges from zero to negative infinity—eliminating any interior maximum—as θ increases from one toward the singularity. As noted earlier, this phenomenon can emerge in a variety of equilibrium models with parameter learning and heterogeneous beliefs (for example, in equation (38) we have $y^*(2) = -\infty$).

Light blue (resp. light red) shading in Figure 5 indicates regions in which we use bid (resp. offer) option prices to estimate the risk-neutral CGFs. We use bid and offer prices appropriately, for given values of θ and y, to ensure that our bounds are conservative, as described in Section 3.1. In practice, as the figure shows, our procedure switches from using offer prices when θ is below about 0.5 to using bid prices above this value.

Table 1 reports estimates of the location of the singularity using one-month (Panel a) and one-year (Panel b) returns in both the full and post-war samples. To address the concern

that the recent time period (when we observe option prices) is somehow different from the earlier sample period, we also report results using the time series of returns from 1996 to 2022 only. All these realized returns are based on our monthly data series. The point estimate of the location of the singularity is similar to that of the post-war sample, though with a somewhat wider confidence interval at both one-month and one-year horizons.

As a further validation, we also consider the annual one-year return series of Jordà et al. (2019), which covers the period from 1872 to 2020 (JKKST annual). This exercise leaves the point estimate virtually unchanged. The confidence interval narrows, despite the smaller sample size, because the positions of the singularity are heavily influenced by extreme observations. Specifically, the JKKST data exclude severe mid-year crashes such as the -65% drop from June 1931 to June 1932: the worst return realization in their series is around -40% for the year 1931.

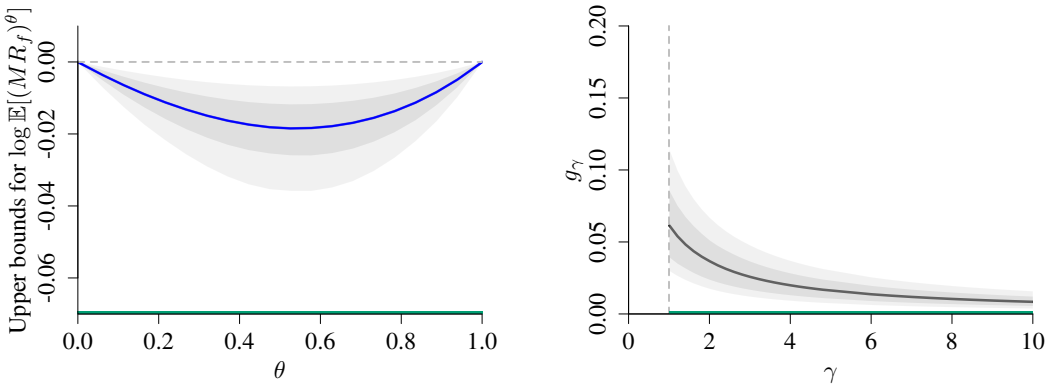
Table 1 also demonstrates that the sampling uncertainty is much larger for the physical distribution than for the risk-neutral one. Column 4 bootstraps realized returns, while holding the risk-neutral CGF fixed, and Column 5 bootstraps the time series of option prices, holding the true CGF fixed. Even with the longest return sample, the confidence intervals based on resampling returns (Column 4) are almost twice as wide as the intervals based on resampling the options (Column 5). Across all specifications, bootstrapping returns alone produces nearly the same intervals as bootstrapping both data sources, indicating that most of the sampling uncertainty is associated with the time series of realized returns.

The left panels of Figure 6 show zoomed-in versions of the two panels of Figure 4, focussing on the well-behaved range $\theta \in (0, 1)$. As moments in this range are unfamiliar, the right panels of the figure convert the moment bounds into bounds on the attractiveness of investment opportunities, as discussed in Section 2.3.

Table A4, in the appendix, reports the values of the bounds at $\theta = 1/2$ together with confidence intervals. The estimates are stable across subsamples, and once again the confidence intervals reveal that the majority of the statistical uncertainty is attributed to uncertainty about the true CGF, not the risk-neutral CGF. Larger return sample sizes consistently tighten the intervals by increasing efficiency, in contrast to the singularity results in Table 1, where extreme observations play a more crucial role and additional data does not always narrow the confidence intervals.

Tables 2 and 3 report results for the two SDF entropy measures introduced in Result 3. As predicted, these bounds are stable across subsamples, and most of the uncertainty stems from the noise in realized returns. The Alvarez–Jermann (A–J) measure is a special case of

(a) Realized return sample: 1872–2022



(b) Realized return sample: 1946–2022

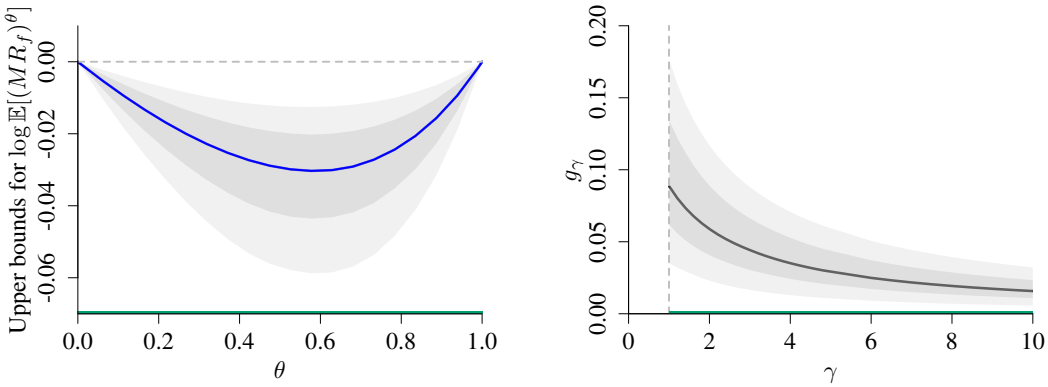


Figure 6: Convexity bounds for $\theta \in (0, 1)$ (left). The attractiveness of investment opportunities, from the perspective of an investor with relative risk aversion γ (right). Dark and light gray shaded areas indicate point-wise 68% and 95% confidence intervals.

our second entropy bound when the right-hand side of (41) is evaluated at $y = 1$ rather than optimized. At the one-month and one-year horizons, our new bounds improve on the A–J measure by at least 15% (30%) in the full (post-war) sample, though these improvements are not always statistically significant.[27]

Table 4 uses Result 4 to calculate measures of risk aversion implied by the optimizing values of y in the entropy bounds of Result 3. The risk aversion estimates are similar, and statistically indistinguishable from each other, across return horizons and for both entropy measures, taking values between 1.6 and 2.4.

[27]Figure A4, in the appendix, shows that these results are extremely stable across a wider range of horizons between 1 and 12 months.

Table 2: Lower bounds for the first entropy metric

This table reports the lower bounds for the first entropy measure of the SDF

$$L^{(1)}(MR_f) \equiv \mathbb{E}[(MR_f)\log(MR_f)] = \mathbb{E}^*[\log(MR_f)]$$

according to (40) in Result 3, and estimated with monthly US stock-market returns in various samples and month-end SPX option prices which always cover 1996 through 2022. The one-month estimates in Panel (a) are annualized by multiplying by twelve. Panel (b) of each table adds an extra row, "JKKST annual," which swaps in the annual return series from 1872–2020 in Jordà et al. (2019) while keeping the same option sample. For all estimates in Column 2, we report 95% confidence intervals using block bootstrap, the block length of which matches each return horizon. Column 3 resamples realized returns and the risk-neutral conditional CGFs. Column 4 ($\mathbb{E}$-bootstrap) resamples only the realized returns. Column 5 ($\mathbb{E}^*$-bootstrap) resamples only the risk-neutral conditional CGFs.

sample	est.	bootstrap CI	$\mathbb{E}$-bootstrap CI	$\mathbb{E}^*$-bootstrap CI
(a) one-month horizon				
1872-2022	0.088	(0.033, 0.175)	(0.033, 0.177)	(0.084, 0.093)
1946-2022	0.173	(0.070, 0.344)	(0.071, 0.348)	(0.165, 0.182)
1996-2022	0.157	(0.016, 0.445)	(0.016, 0.439)	(0.149, 0.165)
(b) one-year horizon				
1872-2022	0.070	(0.026, 0.151)	(0.027, 0.150)	(0.062, 0.079)
1946-2022	0.143	(0.060, 0.301)	(0.058, 0.313)	(0.129, 0.158)
1996-2022	0.122	(0.013, 0.476)	(0.013, 0.478)	(0.110, 0.133)
JKKST annual	0.078	(0.028, 0.178)	(0.028, 0.179)	(0.068, 0.088)

4.1 Robustness

The extreme asymmetry of our SDF moment bounds—and the presence of a singularity—emerges robustly across a range of alternative specifications that we now describe.

Appendix Figure A5 shows that the moment bounds are almost unchanged if we do not extrapolate outside the observed option contracts at all when implementing equation (5).

To capture more extreme return realizations,[28] we have also explored also using daily return data to construct realized monthly returns. Appendix Figure A6 shows that our moment bounds are broadly the same in daily data. For this exercise, we construct rolling one-month

[28]For instance, the worst calendar-month loss was −29% from the close on August 31 to the close on September 30, 1931. Daily return data reveal a larger drop of −31% from February 20 to March 20, 2020, during the COVID-19 crash. The worst one-month decline is −37% from October 14 to November 14, 1929.

Table 3: Lower bounds for the second entropy metric

This table reports the lower bounds for the second entropy measure of the SDF

$$L^{(2)}(MR_f) = -\mathbb{E}[\log(MR_f)]$$

according to inequality (41) in Result 3, and estimated with monthly US stock-market returns in various samples and month-end SPX option prices which always cover 1996 through 2022. The one-month estimates in Panel (a) are annualized by multiplying by twelve. Panel (b) of each table adds an extra row, "JKKST annual," which swaps in the annual return series from 1872–2020 in Jordà et al. (2019) while keeping the same option sample. Columns 2 and 3 report the original Alvarez–Jermann measures, which is a special case of our bounds by fixing $y = 1$ in (41). For all estimates of our bounds in Column 4, we report 95% confidence intervals using block bootstrap, the block length of which matches each return horizon. Column 5 resamples realized returns and the risk-neutral conditional CGFs. Column 6 ($\mathbb{E}$-bootstrap) resamples only the realized returns. Column 7 ($\mathbb{E}^*$-bootstrap) resamples only the risk-neutral conditional CGFs.

	A–J measure ($y = 1$)					
sample	est.	bootstrap CI	est.	bootstrap CI	$\mathbb{E}$-bootstrap CI	$\mathbb{E}^*$-bootstrap CI
(a) one-month horizon						
1872-2022	0.052	(0.023, 0.080)	0.062	(0.023, 0.122)	(0.023, 0.120)	(0.060, 0.065)
1946-2022	0.066	(0.036, 0.100)	0.089	(0.038, 0.174)	(0.038, 0.174)	(0.084, 0.094)
1996-2022	0.067	(0.006, 0.125)	0.091	(0.009, 0.262)	(0.009, 0.258)	(0.087, 0.097)
(b) one-year horizon						
1872-2022	0.049	(0.023, 0.072)	0.057	(0.023, 0.108)	(0.023, 0.105)	(0.053, 0.063)
1946-2022	0.063	(0.034, 0.093)	0.084	(0.035, 0.173)	(0.035, 0.164)	(0.076, 0.096)
1996-2022	0.067	(0.007, 0.118)	0.091	(0.010, 0.273)	(0.010, 0.265)	(0.083, 0.104)
JKKST annual	0.045	(0.018, 0.076)	0.050	(0.018, 0.117)	(0.018, 0.114)	(0.048, 0.055)

returns from daily returns (thereby including more extreme index return realizations), and we include daily index option data.[29]

Appendix Figure A7 shows that the imperfect synchronization between SPX option closing quotes and the closing index level before March 6, 2008 does not affect our results: we find almost identical results when we only consider observations after this date, so that option and index closes are synchronized within one minute.

[29] We do not use daily data in our baseline specification because we would need a block length of at least 250 days to calculate bootstrapped confidence intervals for our one-year analysis. A standard rule of thumb is that the block length should be of order $T^{1/3}$, where T is the sample length. We would therefore need on the order of 250^3 days of data; we only have 150 years, i.e. roughly 150×250 days.

Table 4: Implied risk aversion

This table reports risk-aversion parameters calculated using Result 4. Estimates are based on monthly US stock-market returns (1872–2022) and month-end SPX option prices (1996–2022). 95% confidence intervals are constructed using block bootstrap, with the block length equaling each return horizon. We resample both the realized returns and the conditional risk-neutral CGFs.

	first entropy measure $L^{(1)}$		second entropy measure $L^{(2)}$	
horizon	estimate	bootstrap CI	estimate	bootstrap CI
1	2.36	(1.42, 3.38)	1.69	(1.00, 2.48)
2	2.21	(1.41, 3.11)	1.67	(1.05, 2.40)
3	2.14	(1.38, 3.08)	1.65	(1.02, 2.36)
4	2.17	(1.43, 3.10)	1.63	(1.09, 2.27)
5	2.15	(1.40, 3.14)	1.62	(1.03, 2.35)
6	2.07	(1.34, 3.12)	1.63	(1.03, 2.39)
9	1.84	(1.06, 2.95)	1.66	(1.00, 2.50)
12	1.74	(0.96, 2.99)	1.66	(0.99, 2.65)

Appendix Figure A8 shows our results are little changed when we implement the conservative approach to extrapolating option offer prices described in Figure A1. This procedure accommodates the possibility that (offered) risk-neutral moments may be infinite when $\theta < 0$ or $\theta > 1$, as suggested by Bondarenko, Dillschneider, Schneider, and Trojani (2025). This has a modest effect on the moment bounds for values of θ less than about 0.5, but almost no effect above this range, because the offer prices becomes irrelevant for reasons described in Section 3.1.

Appendix Figure A9 shows that our results are almost unchanged if we drop the deepest out-of-the-money options. Specifically, we define adjusted moneyness for each option contract as

$$m = \frac{\log(K/F(\tau))}{\sigma_{ATM}\sqrt{\tau}}, \tag{52}$$

where τ is the option's maturity, K is the strike price of the options, $F(\tau)$ is the maturity-matched forward price of the index, and σ_{ATM} is the at-the-money implied volatility, measured at the strike equal to the forward price. We keep out-of-the-money put options for which m is between -6 and 0 and out-of-the-money call options for which m is between 0 and 3. This choice is motivated by the fact that the average daily trading volumes (in delta-adjusted dollars) of option contracts within these moneyness ranges are comparable to those of stocks in the Fama-French size-sorted portfolios. (Figure A10 illustrates the time-

series patterns of the dollar volumes and Table A3 reports summary statistics.) If a financial economist is comfortable using the size-sorted portfolio returns in empirical tests, there is no reason to disregard our option panel on the basis that the contracts are too thinly traded.

5 Conclusions

We have derived new bounds for moments of the stochastic discount factor that exploit a comparison between the risk-neutral and true distribution of returns of a specific underlying asset. The bounds can be applied to any risky asset return on which options are traded. We apply them to the return on the S&P 500 index, and find that the θth moment of the SDF rises extremely rapidly as θ rises above one and exhibits a singularity, diverging to infinity, at around $\theta = 1.7$ in our long dataset. In particular, our results undercut the assumption, almost universal in financial economics, that the SDF has finite variance.

This fact conflicts with the intuition that returns with extremely high Sharpe ratios are in some sense too good to be true (Cochrane and Saá-Requejo, 2000). That intuition is justified if one adopts the perspective of a mean–variance investor who cares *only* about means and volatilities of returns (and not about skewness or fat-tailedness, other than to the extent that they are captured in volatility), but, as has been widely noted in the literature, mean–variance preferences generate implausible patterns of investor behavior.

The intuition can be salvaged, however. Just as the maximal attainable Sharpe ratio summarizes the attractiveness of investment opportunities from the point of view of a mean–variance investor—and is related to SDF variance via the Hansen and Jagannathan (1991) bound—our family of bounds can be related to measures of the attractiveness of investment opportunities from the perspective of a myopic investor with constant relative risk aversion γ. From such an investor's point of view, the natural index of SDF variability is not its variance but its $(1 - 1/\gamma)$th moment. In sharp contrast to variance bounds, these moment bounds are well-behaved (and have plausible magnitudes empirically) when $\gamma \geq 1$.

The single most important empirical fact underpinning our moment divergence finding is that index options have high bid prices. Our results therefore relate to a prior literature that has argued that options have low average realized returns. Average returns on options are hard to measure, however, both because they are inherently skewed and fat-tailed and because the observed time series is relatively short. Our approach avoids this issue—and deviates from the prior literature—by exploiting the fact that option prices directly reveal the risk-neutral distribution of the underlying asset. In fact, in our applications we find

that there is more estimation uncertainty associated with estimating the true unconditional distribution of returns (based on the 150-year time series of realized returns) than with estimating the corresponding risk-neutral distribution (based on 26 years of option price data).

We have not attempted to optimize our bounds on the cross-sectional dimension, so that it is likely possible to sharpen the empirical bounds we have derived by expanding the range of assets under investigation. The finance literature has documented a wide range of strategies that appear to have attractive Sharpe ratios. It would be interesting to assess these strategies on the metrics we have suggested, which make sense from the perspective of investors with constant relative risk aversion; this represents a natural direction for future research.

Statistical tests of asset pricing models may still be well specified even if SDF variance is infinite. For example, GMM-based tests of the prediction $\mathbb{E}(M_{t+1}R_{i,t+1}) = 1$ require that $\text{var}(M_{t+1}R_{i,t+1})$ is finite.[30] This can hold, even with infinite SDF variance, if $R_{i,t+1}$ and M_{t+1} are not perfectly (positively or negatively) correlated.

But our results are problematic for (multifactor) mean–variance analysis, one of the central frameworks of empirical financial economics. A superficially plausible reaction is to argue that mean–variance analysis is appropriate so long as the econometrician restricts attention to assets whose payoffs are "not option-like". But this is not an adequate solution, in part because the separation between options and other assets is not a sharp one. A stock with financial or operational leverage *is* an option: to quote Black and Scholes (1973), "almost all corporate liabilities can be viewed as combinations of options." Moreover, a large literature has connected meaures of volatility (and hence option prices) to macroeconomic and financial outcomes.[31] To neglect option prices is to miss out an important part of the story.

More generally, economists who use measures of SDF variability as diagnostics for equilibrium models, along the lines proposed by Hansen and Jagannathan (1991), should confront the fact that SDF moments are unstable, and appear to rise very rapidly, above the first moment.

[30] Such tests rest on central limit theorems, so researchers typically require an assumption that $\mathbb{E}\left[(M_{t+1}R_{i,t+1})^{2+\delta}\right]$ is finite for some $\delta > 0$. (Consistency of GMM may still hold if this fails, but standard errors and conventional test statistics will become invalid.)

[31] See, for example, Bloom (2009), Bollerslev, Tauchen, and Zhou (2009), Bekaert and Hoerova (2014), Adrian and Brunnermeier (2016), Martin (2017), Martin and Wagner (2019), and Martin and Shi (2025).

References

Tobias Adrian and Markus K Brunnermeier. CoVaR. *American Economic Review*, 106(7): 1705–1741, 2016.

Fernando Alvarez and Urban J. Jermann. Using asset prices to measure the persistence of the marginal utility of wealth. *Econometrica*, 73(6):1977–2016, 2005.

Ravi Bansal and Bruce N. Lehmann. Growth-optimal portfolio restrictions on asset pricing models. *Macroeconomic Dynamics*, 1(2):333–354, 1997.

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

David S. Bates. The crash of '87: Was it expected? The evidence from options markets. *Journal of Finance*, 46(3):1009–1044, 1991.

Geert Bekaert and Marie Hoerova. The VIX, the variance premium and stock market volatility. *Journal of Econometrics*, 183(2):181–192, 2014.

Fischer Black and Myron S. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–659, 1973.

Fischer Black, Michael C. Jensen, and Myron S. Scholes. The capital asset pricing model: Some empirical tests. In M.C. Jensen, editor, *Studies in the Theory of Capital Markets*, pages 79–121. Praeger, 1972.

Nicholas Bloom. The impact of uncertainty shocks. *Econometrica*, 77(3):623–685, 2009.

Tim Bollerslev, George Tauchen, and Hao Zhou. Expected stock returns and variance risk premia. *Review of Financial Studies*, 22(11):4463–4492, 2009.

Oleg Bondarenko. Statistical arbitrage and securities prices. *Review of Financial Studies*, 16:875–919, 2003.

Oleg Bondarenko, Yannick Dillschneider, Paul Schneider, and Fabio Trojani. What can you really tell from option prices? Working Paper, 2025.

Karl Borch. A note on uncertainty and indifference curves. *Review of Economic Studies*, 36 (1):1–4, 1969.

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

Mark Broadie, Mikhail Chernov, and Michael Johannes. Understanding index option returns. *Review of Financial Studies*, 22(11):4493–4529, 2009.

Svetlana Bryzgalova, Jiantao Huang, and Christian Julliard. Bayesian solutions for the factor zoo: We just ran two quadrillion models. *Journal of Finance*, 78(1):487–557, 2023.

John Y. Campbell, Stefano Giglio, Christopher Polk, and Robert Turley. An intertemporal capm with stochastic volatility. *Journal of Finance*, 73(4):1571–1631, 2018.

Peter Carr and Dilip Madan. Towards a theory of volatility trading. In E. Jouini, J. Cvitanic, and Marek Musiela, editors, *Option Pricing, Interest Rates and Risk Management*, Handbooks in Mathematical Finance, pages 458–476. Cambridge University Press, 2001.

Aleš Černý. Generalised Sharpe ratios and asset pricing in incomplete markets. *European Finance Review*, 7:191–233, 2003.

Gary Chamberlain and Michael Rothschild. Arbitrage, factor structure, and mean-variance analysis on large asset markets. *Econometrica*, 51(5):1281–1304, 1983.

Mikhail Chernov, Bryan Kelly, Semyon Malamud, and Johannes Schwab. A test of the efficiency of a given portfolio in high dimensions. Working Paper, 2025.

John H. Cochrane and Jesús Saá-Requejo. Beyond arbitrage: Good-deal asset price bounds in incomplete markets. *Journal of Political Economy*, 108(1):79–119, 2000.

Joshua D. Coval and Tyler Shumway. Expected option returns. *Journal of Finance*, 56(3): 983–1009, 2001.

Joost Driessen and Pascal Maenhout. An empirical portfolio perspective on option pricing anomalies. *Review of Finance*, 11:561–603, 2007.

Eugene F. Fama and Kenneth R. French. A five-factor asset pricing model. *Journal of Financial Economics*, 116(1):1–22, 2015.

Eugene F. Fama and James D. MacBeth. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, 81(3):607–636, 1973.

Jianqing Fan, Yuan Liao, and Jiawei Yao. Power enhancement in high-dimensional cross-sectional tests. *Econometrica*, 83(4):1497–1541, 2015.

Martin S. Feldstein. Mean–variance analysis in the theory of liquidity preference and portfolio selection. *Review of Economic Studies*, 36(1):5–12, 1969.

Mihir Gandhi, Niels Gormsen, and Eben Lazarus. Forward return expectations. Working Paper, 2025.

Can Gao and Ian W. R. Martin. Volatility, valuation ratios, and bubbles: An empirical measure of market sentiment. *Journal of Finance*, 2021. forthcoming.

John Geweke. A note on some limitations of CRRA utility. *Economics Letters*, 71(3): 341–345, 2001.

Michael R. Gibbons, Stephen A. Ross, and Jay Shanken. A test of the efficiency of a given portfolio. *Econometrica*, 57(5):1121–1152, 1989.

William Goetzmann, Jonathan Ingersoll, Matthew Spiegel, and Ivo Welch. Portfolio performance manipulation and manipulation-proof performance measures. *Review of Financial Studies*, 20:1503–1546, 2007.

Lars P. Hansen and Scott F. Richard. The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models. *Econometrica*, 55(3):587–613, 1987.

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

Kewei Hou, Chen Xue, and Lu Zhang. Digesting anomalies: An investment approach. *Review of Financial Studies*, 28(3):650–705, 2015.

Jens C. Jackwerth. Recovering risk aversion from option prices and realized returns. *Review of Financial Studies*, 13(2):433–451, 2000.

Christopher S. Jones. A nonlinear factor analysis of S&P 500 index option returns. *Journal of Finance*, 61(5):2325–2363, 2006.

Òscar Jordà, Katharina Knoll, Dmitry Kuvshinov, Moritz Schularick, and Alan M Taylor. The rate of return on everything, 1870–2015. *Quarterly Journal of Economics*, 134(3): 1225–1298, 2019.

Serhiy Kozak, Stefan Nagel, and Shrihari Santosh. Shrinking the cross-section. *Journal of Finance*, 75(3):1183–1233, 2020.

Haim Levy. The rationale of the mean–standard deviation analysis: Comment. *American Economic Review*, 64(3):434–441, 1974.

John Lintner. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The review of economics and statistics*, pages 13–37, 1965.

Yan Liu. Index option returns and generalized entropy bounds. *Journal of Financial Economics*, 139:1015–1036, 2021.

Yan Liu and Jing Cynthia Wu. Reconstructing the yield curve. *Journal of Financial Economics*, 142(3):1395–1425, 2021.

Harry Markowitz. Portfolio selection. *Journal of Finance*, 7:77–91, 1952.

Ian W. R. Martin. Consumption-based asset pricing with higher cumulants. *Review of Economic Studies*, 80(2):745–773, 2013a.

Ian W. R. Martin. The Lucas orchard. *Econometrica*, 81(1):55–111, 2013b.

Ian W. R. Martin. The forward premium puzzle in a two-country world. NBER Working Paper No. 17564, 2013c.

Ian W. R. Martin. What is the expected return on the market? *Quarterly Journal of Economics*, 132(1):367–433, 2017.

Ian W. R. Martin and Dimitris Papadimitriou. Sentiment and speculation in a market with heterogeneous beliefs. *American Economic Review*, 112(8):2465–2517, 2022.

Ian W. R. Martin and Ran Shi. Forecasting crashes with a smile. Working paper, 2025.

Ian W. R. Martin and Christian Wagner. What is the expected return on a stock? *Journal of Finance*, 74(4):1887–1929, 2019.

Robert C. Merton. Lifetime portfolio selection under uncertainty: The continuous-time case. *Review of Economics and Statistics*, 51(3):247–257, 1969.

Robert C. Merton. An intertemporal capital asset pricing model. *Econometrica*, 41(5): 867–887, 1973.

Piotr Orłowski, Andras Sali, and Fabio Trojani. Arbitrage free dispersion. Swiss Finance Institute Research Paper No. 19–20, 2018.

Stephen A. Ross. The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, 13(3):341–360, 1976.

Paul A. Samuelson. Lifetime portfolio selection by dynamic stochastic programming. *Review of Economics and Statistics*, 51(3):239–246, 1969.

Paul A. Samuelson. The fundamental approximation theorem of portfolio analysis in terms of means, variances and higher moments. *Review of Economic Studies*, 37(4):537–542, 1970.

Pedro Santa-Clara and Alessio Saretto. Option strategies: Good deals and margin calls. *Journal of Financial Markets*, 12:391–417, 2009.

G. William Schwert. Indexes of US stock prices from 1802 to 1987. *Journal of Business*, pages 399–426, 1990.

William F. Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19(3):425–442, 1964.

Karl N. Snow. Diagnosing asset pricing models using the distribution of asset returns. *Journal of Finance*, 46(3):955–983, 1991.

Michael Stutzer. A Bayesian approach to diagnosis of asset pricing models. *Journal of Econometrics*, 68(2):367–397, 1995.

Henri Theil. *Economics and Information Theory*. Amsterdam: North-Holland, 1967.

S. C. Tsiang. The rationale of the mean–standard deviation analysis, skewness preference, and the demand for money. *American Economic Review*, 62(3):354–371, 1972.

Martin L. Weitzman. Subjective expectations and asset-return puzzles. *American Economic Review*, 97(4):1102–1130, 2007.

A Proofs

Proof of equation (5). By the logic of Breeden and Litzenberger (1978), as expressed in the Carr–Madan formula (Carr and Madan, 2001), for any smooth function $g(\cdot)$, we have

$$\begin{aligned} g(R) = g(R_f) + g'(R_f)(R - R_f) + \\ + R_f \int_0^1 g''(KR_f) \max\{KR_f - R,\, 0\} \,\mathrm{d}K + \\ + R_f \int_1^\infty g''(KR_f) \max\{R - KR_f,\, 0\} \,\mathrm{d}K. \end{aligned}$$

Let $g(R) = R^\theta$; then we have

$$\begin{aligned} R_{t+1}^\theta = R_{f,t+1}^\theta + \theta R_{f,t+1}^{\theta-1}(R_{t+1} - R_{f,t+1}) + \\ + R_{f,t+1}^{\theta-1} \int_0^1 \theta(\theta-1)K^{\theta-2} \max\{KR_{f,t+1} - R_{t+1},\, 0\} \,\mathrm{d}K + \\ + R_{f,t+1}^{\theta-1} \int_1^\infty \theta(\theta-1)K^{\theta-2} \max\{R_{t+1} - KR_{f,t+1},\, 0\} \,\mathrm{d}K. \end{aligned}$$

Dividing both sides by $R_{f,t+1}^\theta$ and taking risk-neutral expectations,

$$\begin{aligned} \mathbb{E}^*\left[(R_{t+1}/R_{f,t+1})^\theta\right] = 1 + \int_0^1 \theta(\theta-1)K^{\theta-2} \,\mathrm{put}_t(KR_{f,t+1}) \,\mathrm{d}K + \\ + \int_1^\infty \theta(\theta-1)K^{\theta-2} \,\mathrm{call}_t(KR_{f,t+1}) \,\mathrm{d}K. \end{aligned}$$

□

Derivation of the CGF (14). Note that if X is Normally distributed,

$$\log \mathbb{E}_t \exp X = \mathbb{E}_t X + \frac{1}{2} \mathrm{var}_t X. \tag{A1}$$

It follows that

$$\begin{aligned} \boldsymbol{\kappa}_t(\theta_1, \theta_2) &= \log \mathbb{E}_t \left[\exp\left\{\theta_1\left(-\frac{1}{2}\lambda^2 - \lambda Z\right) + \theta_2\left(\mu - \frac{1}{2}\sigma^2 + \sigma W\right)\right\}\right] \\ &= \mu\theta_2 + \frac{1}{2}\lambda^2\theta_1(\theta_1 - 1) + \frac{1}{2}\sigma^2\theta_2(\theta_2 - 1) - \rho\sigma\lambda\theta_1\theta_2. \end{aligned} \tag{A2}$$

As $\boldsymbol{\kappa}_t(1,1) = 0$, we must have $\mu = \rho\sigma\lambda$. Using this fact to eliminate ρ in (A2), we have

$$\boldsymbol{\kappa}_t(\theta_1, \theta_2) = \mu\theta_2(1 - \theta_1) + \frac{1}{2}\lambda^2\theta_1(\theta_1 - 1) + \frac{1}{2}\sigma^2\theta_2(\theta_2 - 1),$$

as required. □

Derivation of the CGF (17). We seek

$$\boldsymbol{\kappa}_t(\theta_1,\theta_2) = \log \mathbb{E}_t\left[\exp\left\{\theta_1\left(-\frac{1}{2}\lambda^2 - \lambda Z - J_1\omega + N\log(1+J_1)\right) + \right.\right.$$
$$\left.\left. + \ \theta_2\left(\mu - \frac{1}{2}\sigma^2 + \sigma W - J_2\omega + N\log(1+J_2)\right)\right\}\right].$$

As N is independent of (W, Z), we can split the expectation to give

$$\boldsymbol{\kappa}_t(\theta_1,\theta_2) = \log \mathbb{E}_t\left[\exp\left\{\theta_1\left(-\frac{1}{2}\lambda^2 - \lambda Z - J_1\omega\right) + \theta_2\left(\mu - \frac{1}{2}\sigma^2 + \sigma W - J_2\omega\right)\right\}\right] +$$
$$+ \log \mathbb{E}_t\left[\exp\left\{\left(\theta_1\log(1+J_1) + \theta_2\log(1+J_2)\right)N\right\}\right]. \tag{A3}$$

The first term on the right-hand side of (A3) can be evaluated using (A1). The second can be calculated using the fact that if N is Poisson distributed with parameter ω,

$$\log \mathbb{E}_t \exp\{\theta N\} = \omega\left(e^{\theta} - 1\right). \tag{A4}$$

We then have

$$\boldsymbol{\kappa}_t(\theta_1,\theta_2) = \mu\theta_2 - \omega J_1\theta_1 - \omega J_2\theta_2 + \frac{1}{2}\lambda^2\theta_1(\theta_1-1) +$$
$$+ \frac{1}{2}\sigma^2\theta_2(\theta_2-1) - \rho\sigma\lambda\theta_1\theta_2 + \omega\left[(1+J_1)^{\theta_1}(1+J_2)^{\theta_2} - 1\right]. \tag{A5}$$

Now, $\boldsymbol{\kappa}_t(1,1) = 0$, so from (A5) we must have $\mu + \omega J_1 J_2 = \rho\sigma\lambda$. Using this fact to substitute for ρ in (A5),

$$\boldsymbol{\kappa}_t(\theta_1,\theta_2) = \mu\theta_2(1-\theta_1) + \frac{1}{2}\lambda^2\theta_1(\theta_1-1) + \frac{1}{2}\sigma^2\theta_2(\theta_2-1) +$$
$$+ \omega\left[(1+J_1)^{\theta_1}(1+J_2)^{\theta_2} - (1+J_1\theta_1)(1+J_2\theta_2)\right], \tag{A6}$$

as required. □

Derivation of the CGF (21). Recall that the jump arrival rate, ω, is distributed according to an exponential distribution with mean $\overline{\omega}$. Thus, $N \mid \omega \sim \text{Poisson}(\omega)$ and $\omega \sim \text{Exp}(1/\overline{\omega})$. The key calculation we require is $\mathbb{E}_t\left(a^N\right)$, where $a > 0$ is an arbitrary constant. We have

$$\begin{aligned}
\mathbb{E}_t\left(a^N\right) &= \mathbb{E}_t\left[\mathbb{E}_t\left(a^N \mid \omega\right)\right] \\
&= \int_{\omega=0}^{\infty} \frac{1}{\overline{\omega}} e^{-\frac{\omega}{\overline{\omega}}} \sum_{n=0}^{\infty} \frac{e^{-\omega}\omega^n}{n!} a^n \ \mathrm{d}\omega \\
&= \int_{\omega=0}^{\infty} \frac{1}{\overline{\omega}} e^{-\omega\left(\frac{1}{\overline{\omega}}+1-a\right)} \ \mathrm{d}\omega \\
&= \frac{\frac{1}{\overline{\omega}}}{\frac{1}{\overline{\omega}}+1-a} \quad \text{if } \frac{1}{\overline{\omega}} + 1 - a > 0.
\end{aligned}$$

That is, $\mathbb{E}_t\left(a^N\right) = 1/\left(1-\overline{\omega}(a-1)\right)$ for $a < 1 + 1/\overline{\omega}$. This implies that

$$\mathbb{E}_t\left[(1+J_1)^{\theta_1 N}(1+J_2)^{\theta_2 N}\right] = \frac{1}{1-\overline{\omega}\left[(1+J_1)^{\theta_1}(1+J_2)^{\theta_2}-1\right]},$$

and the form of the CGF (21) and all other calculations in the example follow. □

Derivation of the CGF (25). In the Martin and Papadimitriou (MP, 2022) model, the gross interest rate is normalized so that $R_{f,t+1} = 1$. We therefore want to find

$$\boldsymbol{\kappa}_t(\theta_1,\theta_2) = \log \mathbb{E}_t\left[\exp\left\{\theta_1 \log M_{t+1} + \theta_2 \log R_{t+1}\right\}\right]. \tag{A7}$$

We take the perspective of the median agent ($z = 0$ in the notation of MP). MP write today's date as time 0 and the terminal horizon as time T. In the present paper, we write today's date as time t and the terminal horizon as time $t+1$, so let us first restate the relevant results of MP in our notation.

By Result 7 of MP, the market return R_{t+1} is lognormally distributed from the perspective of the median agent, with

$$\mathbb{E}_t \log R_{t+1} = \frac{1+\delta}{2\delta}\sigma^2 \quad \text{and} \quad \operatorname{var}_t \log R_{t+1} = \sigma^2. \tag{A8}$$

As agents have log utility in the MP model, the SDF perceived by the median agent is $M_{t+1} = 1/R^{(0)}_{t+1}$ where $R^{(0)}_{t+1}$ is the return on the median agent's chosen trading strategy. (The MP model features complete markets, in which different agents perceive different SDFs because they have different beliefs, but all see the same prices.) Using Result 9 of MP, this implies that we can write, in our notation,

$$M_{t+1} = \sqrt{\frac{\delta}{1+\delta}}\exp\left\{-\frac{1}{2}\frac{(1+\delta)^2}{\delta}\sigma^2 + \frac{1}{2(1+\delta)\sigma^2}\left(\log R_{t+1} - \frac{(1+\delta)(1+2\delta)}{2\delta}\sigma^2\right)^2\right\}. \tag{A9}$$

By (A8), we can write

$$\log R_{t+1} = \frac{1+\delta}{2\delta}\sigma^2 + \sigma Z \tag{A10}$$

where $Z \sim N(0,1)$ is standard Normal. Substituting (A10) into (A9) we find, after some simplification and rearranging, that

$$M_{t+1} = \sqrt{\frac{\delta}{1+\delta}}\exp\left\{-\frac{(1+\delta)}{2\delta}\sigma^2 - \sigma Z + \frac{1}{2(1+\delta)}Z^2\right\}. \tag{A11}$$

Substituting equations (A10) and (A11) into equation (A7), it follows that

$$\boldsymbol{\kappa}_t(\theta_1,\theta_2) = \log \mathbb{E}_t \exp\left\{\frac{1}{2}\theta_1 \log\frac{\delta}{1+\delta} + (\theta_2-\theta_1)\frac{1+\delta}{2\delta}\sigma^2 + (\theta_2-\theta_1)\sigma Z + \frac{\theta_1}{2(1+\delta)}Z^2\right\}. \tag{A12}$$

We now use the fact that if Z is standard Normal and b, c, and d are constants with $d < 1$, then

$$\log \mathbb{E}_t \exp\left\{b + cZ + \frac{1}{2}dZ^2\right\} = b + \frac{1}{2}\frac{c^2}{1-d} - \frac{1}{2}\log(1-d). \tag{A13}$$

Applied to equation (A12), this implies that

$$\boldsymbol{\kappa}_t(\theta_1, \theta_2) = \frac{1}{2}\left[\frac{1+\delta}{\delta}\sigma^2\left(\theta_2 - \theta_1\right) + \frac{1+\delta}{1+\delta-\theta_1}\sigma^2\left(\theta_2 - \theta_1\right)^2 + \log\frac{1+\delta}{1+\delta-\theta_1} - \theta_1 \log\frac{1+\delta}{\delta}\right],$$

which is equation (25), as required. □

Proof of Result 2. (27) is equivalent to inequality (33). The right-hand side of inequality (33) is the weighted average of two strictly convex functions, and therefore is itself strictly convex. Moreover, $\boldsymbol{\kappa}_t(1, y)$ equals zero when y equals 0 or 1; as it is strictly convex it is therefore unbounded as $y \to -\infty$ or $y \to \infty$. Summarizing, the right-hand side of (33) is a strictly convex function of y that is unbounded as y tends to plus or minus infinity. It therefore has a unique interior minimum. □

Proof of Result 4. If the investor chooses to invest in the asset with return R_{t+1}, then the SDF is proportional to $R_{t+1}^{-\gamma}$. It follows that $M_{t+1}R_{f,t+1} = \frac{R_{t+1}^{-\gamma}}{\mathbb{E}_t R_{t+1}^{-\gamma}}$, so

$$\mathbb{E}_t^* \log\frac{R_{t+1}}{R_{f,t+1}} = \frac{\mathbb{E}_t\left[\left(\frac{R_{t+1}}{R_{f,t+1}}\right)^{-\gamma}\log\frac{R_{t+1}}{R_{f,t+1}}\right]}{\mathbb{E}_t\left[\left(\frac{R_{t+1}}{R_{f,t+1}}\right)^{-\gamma}\right]}.$$

The right-hand side of inequality (40) can then be written

$$\sup_{y\in\mathbb{R}} y\frac{\mathbb{E}_t\left[\left(\frac{R_{t+1}}{R_{f,t+1}}\right)^{-\gamma}\log\frac{R_{t+1}}{R_{f,t+1}}\right]}{\mathbb{E}_t\left[\left(\frac{R_{t+1}}{R_{f,t+1}}\right)^{-\gamma}\right]} - \log\mathbb{E}_t\left[\left(\frac{R_{t+1}}{R_{f,t+1}}\right)^{y}\right].$$

The first-order condition associated with the above expression is

$$\frac{\mathbb{E}_t\left[\left(\frac{R_{t+1}}{R_{f,t+1}}\right)^{-\gamma}\log\frac{R_{t+1}}{R_{f,t+1}}\right]}{\mathbb{E}_t\left[\left(\frac{R_{t+1}}{R_{f,t+1}}\right)^{-\gamma}\right]} = \frac{\mathbb{E}_t\left[\left(\frac{R_{t+1}}{R_{f,t+1}}\right)^{y}\log\frac{R_{t+1}}{R_{f,t+1}}\right]}{\mathbb{E}_t\left[\left(\frac{R_{t+1}}{R_{f,t+1}}\right)^{y}\right]},$$

which is satisfied when $y = -\gamma$, as claimed. Moreover, y is uniquely determined because the objective function is strictly concave, and so has a unique maximum.

Similarly, the assumption on the SDF implies that

$$\mathbb{E}_t^*\left[\left(\frac{R_{t+1}}{R_{f,t+1}}\right)^y\right] = \frac{\mathbb{E}_t\left[\left(\frac{R_{t+1}}{R_{f,t+1}}\right)^{y-\gamma}\right]}{\mathbb{E}_t\left[\left(\frac{R_{t+1}}{R_{f,t+1}}\right)^{-\gamma}\right]}$$

so the right-hand side of inequality (41) can be written

$$\sup_{y\in\mathbb{R}} y\,\mathbb{E}_t \log\frac{R_{t+1}}{R_{f,t+1}} - \log\mathbb{E}_t\left[\left(\frac{R_{t+1}}{R_{f,t+1}}\right)^{y-\gamma}\right] + \log\mathbb{E}_t\left[\left(\frac{R_{t+1}}{R_{f,t+1}}\right)^{-\gamma}\right].$$

The first-order condition is

$$\mathbb{E}_t \log\frac{R_{t+1}}{R_{f,t+1}} = \frac{\mathbb{E}_t\left[\left(\frac{R_{t+1}}{R_{f,t+1}}\right)^{y-\gamma}\log\frac{R_{t+1}}{R_{f,t+1}}\right]}{\mathbb{E}_t\left[\left(\frac{R_{t+1}}{R_{f,t+1}}\right)^{y-\gamma}\right]},$$

and it is satisfied when $y = \gamma$, as claimed. Again, y is uniquely determined because the objective function is strictly concave, so has a unique maximum. □

Proof of Result 5. We can find an SDF induced by the optimal strategy of the marginal investor, $M_{\gamma,t+1}$, defined by

$$M_{\gamma,t+1}R_{f,t+1} = \frac{\left(\frac{R_{\gamma,t+1}}{R_{f,t+1}}\right)^{-\gamma}}{\mathbb{E}\left[\left(\frac{R_{\gamma,t+1}}{R_{f,t+1}}\right)^{1-\gamma}\right]}. \tag{A14}$$

If the market is incomplete, there may be other SDFs, but $M_{\gamma,t+1}$ is *an* SDF; and all SDFs must satisfy Results 1 and 3. We can use equation (A14) to rewrite equation (46) as

$$g_\gamma = \frac{1}{1-\gamma}\log\mathbb{E}\left[\left(M_{\gamma,t+1}R_{f,t+1}\right)^{\frac{\gamma-1}{\gamma}}\mathbb{E}\left[\left(\frac{R_{\gamma,t+1}}{R_{f,t+1}}\right)^{1-\gamma}\right]^{\frac{\gamma-1}{\gamma}}\right].$$

But, using equation (46) once again, this implies that

$$g_\gamma = \frac{1}{1-\gamma}\log\mathbb{E}\left[\left(M_{\gamma,t+1}R_{f,t+1}\right)^{\frac{\gamma-1}{\gamma}}\right] + \frac{\gamma-1}{\gamma}g_\gamma,$$

and hence that

$$g_\gamma = \frac{\gamma}{1-\gamma}\log\mathbb{E}\left[\left(M_{\gamma,t+1}R_{f,t+1}\right)^{\frac{\gamma-1}{\gamma}}\right]. \tag{A15}$$

The result follows by Result 1 when $\gamma \in (0,1)$ or $\gamma > 1$, and hence also in the log utility case by continuity.

For the final result, rewrite (A15) as

$$g(\tau) = \frac{1}{\tau - 1} \log \mathbb{E}\left[(M_{\gamma,t+1} R_{f,t+1})^{1-\tau}\right] = \frac{1}{\tau - 1} \boldsymbol{\kappa}(1 - \tau, 0).$$

It follows that $g'(0) = \boldsymbol{\kappa}^{(1)}(1,\, 0) = L^{(1)}\left(M_{t+1} R_{f,t+1}\right)$, as required. □

Proof of Result 6. We begin by introducing a lemma:

Lemma A1. *For any random variable X, the function $\kappa(\theta) = \log \mathbb{E}[e^{\theta X}]$ satisfies*

$$\lim_{\theta \to -\infty} \frac{\kappa(\theta)}{\theta} = \operatorname{ess\,inf}\, X, \quad \lim_{\theta \to \infty} \frac{\kappa(\theta)}{\theta} = \operatorname{ess\,sup}\, X.$$

Proof. To prove the lemma, let $L = \operatorname{ess\,sup}\, X$. As $X \leq L$ with probability one, $\kappa(\theta) \leq \theta L$ for all $\theta > 0$, that is, $\limsup_{\theta \to \infty} \kappa(\theta)/\theta \leq L$.

Now consider sets $A_n = \{L - 1/n < X \leq L\}$, $n = 1, 2, \ldots$, as $\mathbb{P}(A_n) > 0$ for any n,

$$\frac{\kappa(\theta)}{\theta} \geq \frac{\log\left\{\mathbb{E}[e^{\theta X} \mid X \in A_n] \times \mathbb{P}(A_n)\right\}}{\theta} \geq L - \frac{1}{n} + \frac{\log \mathbb{P}(A_n)}{\theta}, \quad \forall\, n = 1, 2, \ldots.$$

Thus, $\liminf_{\theta \to \infty} \kappa(\theta)/\theta \geq \limsup_{n \to \infty} \liminf_{\theta \to \infty} \{L - 1/n + \log \mathbb{P}(A_n)/\theta\} = L$. Combining the two limiting statements for $\kappa(\theta)/\theta$, $\lim_{\theta \to \infty} \kappa(\theta)/\theta = L$.

The case for $\theta \to -\infty$ follows immediately by applying the same arguments to $-X$. □

It follows that in finite samples, the estimated true and risk-neutral CGFs will be asymptotically linear. More precisely, as $|\theta| \to \infty$, we have

$$\lim_{\theta \to -\infty} \frac{\widehat{\boldsymbol{\kappa}}(0, \theta)}{\theta} = \log \min_t \frac{R_{t+1}}{R_{f,t+1}} \quad \text{and} \quad \lim_{\theta \to +\infty} \frac{\widehat{\boldsymbol{\kappa}}(0, \theta)}{\theta} = \log \max_t \frac{R_{t+1}}{R_{f,t+1}}, \tag{A16}$$

in a finite sample of returns $\{R_{t+1}/R_{f,t+1}\}_{t=1,\ldots,T}$. To see this, consider a random variable X whose distribution follows the empirical measure $\frac{1}{T}\sum_{t=0}^{T-1} \delta_{X_t}$ where $X_t = \log(R_{t+1}/R_{f,t+1})$, so that by the definition (49), $\widehat{\boldsymbol{\kappa}}(0, \theta) = \log \mathbb{E}[e^{\theta X}]$. Equation (A16) follows from Lemma A1.

Similarly, given a finite collection of strikes, where $K_{\min,t}$ and $K_{\max,t}$ denote the smallest and largest strikes observed for out-of-the-money puts and calls at time t, we have

$$\lim_{\theta \to -\infty} \frac{\widehat{\boldsymbol{\kappa}}(1, \theta)}{\theta} = \log \min_t \frac{K_{\min,t}}{R_{f,t+1}} \quad \text{and} \quad \lim_{\theta \to +\infty} \frac{\widehat{\boldsymbol{\kappa}}(1, \theta)}{\theta} = \log \max_t \frac{K_{\max,t}}{R_{f,t+1}}. \tag{A17}$$

(Note that K here indicates a strike expressed as a return, consistent with our notation in equation (5), i.e., the strike of an option on the asset divided by the spot price of the underlying asset.) This follows because we can write $\widehat{\boldsymbol{\kappa}}(1, \theta) = \log \int e^{\theta x}\, q(\mathrm{d}x)$, where the

probability measure $q(\cdot) = \frac{1}{T}\sum_{t=1}^{T} q_t(\cdot)$, and q_t denotes the conditional risk-neutral distribution of $\log(R_{t+1}/R_{f,t+1})$. The claim follows from Lemma A1 because q_t only has nonzero probability density in $[\log(K_{\min,t}/R_{f,t+1}),\, \log(K_{\max,t}/R_{f,t+1})]$.

Now we are in a position to prove the result. Rewrite inequality (51) as

$$\frac{\boldsymbol{\kappa}(\theta,0)}{\theta} \geq \left[\frac{\widehat{\boldsymbol{\kappa}}(1,y)}{y} - \frac{\widehat{\boldsymbol{\kappa}}(0,\frac{\theta}{\theta-1}y)}{\frac{\theta}{\theta-1}y}\right] y. \tag{A18}$$

If condition (i) holds, then we can take the limit as $y \to -\infty$: the term in the square brackets tends to $\log \min_t K_{\min,t} - \log \min_t R_{t+1}$, which is negative, so that the right-hand side of (A18) diverges to $+\infty$. If condition (ii) holds, then we can take the limit as $y \to \infty$: in this case the term in square brackets tends to $\log \max_t K_{\max,t} - \log \max_t R_{t+1}$, which is positive, so that the right-hand side of (A18) diverges to $+\infty$ once again. It follows that $\boldsymbol{\kappa}(\theta,0)/\theta$ can be made arbitrarily large, and hence that $\boldsymbol{\kappa}(\theta,0)$ can be made arbitrarily large, because θ is fixed and positive. □

B Additional Tables and Figures

Table A1: Summary of the option sample

This table summarizes the number of observed SPX option contracts covering January 1996 through December 2022. For each filter applied, we report the fraction of contracts that are deleted and the number of contracts that remain.

	DTM$\leq$ 30		30 <DTM$\leq$ 182		DTM> 182	
	deleted (%)	remaining	deleted (%)	remaining	deleted (%)	remaining
Initial sample		4,384,236		4,261,474		911,732
Call		1,740,265		1,601,206		361,198
Put		2,643,971		2,660,268		550,534
DTM$\geq$ 8	26.34%	3,229,180	0.00%	4,261,474	0.00%	911,732
no missing R_f	0.01%	3,228,757	0.03%	4,259,831	0.06%	911,188
implied vol. exists	1.49%	3,180,755	0.54%	4,236,609	0.71%	904,701
no duplication	6.63%	2,969,538	11.32%	3,757,977	0.55%	899,743
last traded today	0.01%	2,969,124	0.03%	3,756,786	0.07%	899,109
OI > 0	4.76%	2,827,852	6.68%	3,505,516	2.81%	873,848
offer $>$ bid > 0	0.00%	2,827,844	0.00%	3,505,516	0.00%	873,844
out-of-the-money	24.21%	2,142,995	20.08%	2,801,297	21.40%	686,821
Final sample		2,142,995		2,801,297		686,821
Call		697,503		893,373		227,741
Put		1,445,492		1,907,924		459,080
$-6 \leq m \leq 3$	11.72%	1,891,835	6.19%	2,628,005	2.94%	666,616
m-filtered sample		1,891,835		2,628,005		666,616
Call		683,447		889,902		227,553
Put		1,208,388		173,8103		439,063

Table A2: Contracts used for constructing constant-maturity risk-neutral CGFs

This table summarizes the SPX option contracts used to construct the risk-neutral CGFs at constant maturities (or, equivalently, the return horizons) of $\tau = 1, 2, 3, 4, 5, 6, 9, 12$ months (Column 2 gives the calendar days).

At date t, for a target maturity τ, we identify the near-term (T_1) and the next-term expiry (T_2) from our filtered options sample. We then compute the corresponding risk-neutral CGFs, $\boldsymbol{\kappa}_t^{(T_1)}(1,\theta)$ and $\boldsymbol{\kappa}_t^{(T_2)}(1,\theta)$, using equation (5). The CGF at the target maturity τ is then obtain by linearly interpolating in calendar time:

$$\boldsymbol{\kappa}_t^{(\tau)}(1,\theta) = \frac{T_2 - \tau}{T_2 - T_2}\,\boldsymbol{\kappa}_t^{(T_1)}(1,\theta) + \frac{\tau - T_1}{T_2 - T_2}\,\boldsymbol{\kappa}_t^{(T_2)}(1,\theta).$$

Column 1 reports the number of days between January 6, 1996, and December 30, 2022, for which enough options contracts pass our filter to construct both $\boldsymbol{\kappa}_t^{(T_1)}(1,\theta)$ and $\boldsymbol{\kappa}_t^{(T_2)}(1,\theta)$, and thus $\boldsymbol{\kappa}_t^{(\tau)}(1,\theta)$. Columns 3 and 4 report the average values of T_1 and T_2. Column 5 gives the number of month-end dates with $\boldsymbol{\kappa}_t^{(\tau)}(1,\theta)$.

num. obs.	τ in cal. days	avg. T_1	avg. T_2	num. months
5,856	30	23	42	324
6,106	60	49	74	324
5,640	91	77	115	323
5,063	122	99	152	321
4,308	152	120	182	315
3,589	182	148	217	314
2,476	273	234	305	294
1,511	365	325	431	250

Table A3: Comparing the dollar trading volumes of options and stocks: summary statistics

Every day, for each SPX option, we compute two measures. The first measure is the dollar volume, defined as $DV = 100S_0 \times V \times |\Delta|$, where S_0 is the current price of the S&P 500 index, V is the number of contracts traded, and Δ is the option delta. (Note that SPX options use a 100 contract multiplier.) The second measure is the adjusted moneyness, calculated as

$$m = \frac{\log(K/F(\tau))}{\sqrt{\tau\sigma^2_{ATM}}},$$

where τ is the annualized time to maturity, K is the strike price, $F(\tau)$ is the maturity-matched forward price, and σ_{ATM} is the at-the-money implied volatility for that maturity on the same day.
In Panels (a) and (b), we show the average of $\log_{10}(DV)$ for all contracts in our sample from January 4, 1996 to December 30, 2022, grouped by different adjusted moneyness ranges. In comparison, Panel (c) presents the average daily dollar trading volumes of stocks, divided into ten portfolios sorted by size based on NYSE market capitalization.

(a) Out-of-the-money puts

m	mean	med.	min	q25	q75	max
$(-\infty, -7)$	5.26	5.42	2.91	5.07	5.61	7.26
$[-7, -6)$	5.48	5.63	2.90	5.20	5.86	7.30
$[-6, -5)$	5.66	5.80	4.25	5.38	6.00	6.66
$[-5, -4)$	5.94	6.03	4.52	5.66	6.25	6.77
$[-4, -3)$	6.32	6.39	5.27	6.06	6.57	7.24
$[-3, -2)$	6.72	6.77	5.88	6.49	6.94	7.32
$[-2, -1)$	7.17	7.19	6.53	6.98	7.38	7.73
$[-1, 0)$	7.65	7.64	7.07	7.45	7.88	8.25

(b) Out-of-the-money calls

m	mean	med.	min	q25	q75	max
$[4, +\infty)$	5.68	5.75	3.11	5.45	6.03	7.30
$[3, 4)$	5.57	5.65	3.57	5.40	5.89	7.52
$[2, 3)$	5.88	5.92	3.78	5.72	6.08	7.03
$[1, 2)$	6.73	6.72	6.10	6.61	6.85	7.47
$[0, 1)$	7.59	7.57	7.11	7.44	7.73	8.18

(c) Fama-French portfolios

size	mean	med.	min	q25	q75	max
1	5.77	5.83	4.82	5.56	5.99	7.04
2	6.46	6.54	5.62	6.21	6.68	7.75
3	6.77	6.87	6.04	6.51	6.98	7.59
4	7.00	7.12	6.31	6.76	7.22	7.57
5	7.20	7.31	6.51	6.95	7.43	7.71
6	7.36	7.46	6.70	7.11	7.58	7.86
7	7.54	7.63	6.85	7.32	7.75	8.06
8	7.78	7.86	7.09	7.58	7.97	8.29
9	8.03	8.11	7.29	7.85	8.23	8.52
10	8.56	8.62	7.74	8.36	8.73	9.15

Table A4: Upper bounds for $\log \mathbb{E}[(MR_f)^{1/2}]$

This table reports the upper bounds for $\boldsymbol{\kappa}(1/2, 0) = \log \mathbb{E}[(MR_f)^{1/2}]$. The estimation uses monthly US stock-market returns and month-end SPX option prices. The row "JKKST annual" in Panel (b) uses the annual realized return series from 1872–2020 in Jordà et al. (2019). The options data cover 1996 through 2022 for all specifications. The one-month estimates in Panel (a) are annualized by multiplying by twelve. Column 3 reports the 95% confidence intervals using a block bootstrap with block length equal to the return horizon, resampling both series separately. Column 4 ($\mathbb{E}$-bootstrap) resamples only the realized returns used for estimating $\widehat{\boldsymbol{\kappa}}(0, \theta)$ in equation (49). Column 5 ($\mathbb{E}$-bootstrap) resamples only the risk-neutral conditional CGFs used for estimating $\widehat{\boldsymbol{\kappa}}(1, \theta)$ in equation (50).

By Result 5, we can convert the figures in the second column of the table into lower bounds on the WTP of an investor with $\gamma = 2$ by multiplying them by -2.

sample	est.	bootstrap CI	$\mathbb{E}$-bootstrap CI	$\mathbb{E}^*$-bootstrap CI
(a) one-month horizon				
1872-2022	-0.018	$(-0.036, -0.007)$	$(-0.035, -0.007)$	$(-0.018, -0.018)$
1946-2022	-0.029	$(-0.057, -0.012)$	$(-0.057, -0.012)$	$(-0.030, -0.029)$
1996-2022	-0.029	$(-0.081, -0.003)$	$(-0.081, -0.003)$	$(-0.029, -0.029)$
(b) one-year horizon				
1872-2022	-0.015	$(-0.029, -0.006)$	$(-0.029, -0.006)$	$(-0.016, -0.015)$
1946-2022	-0.026	$(-0.052, -0.011)$	$(-0.051, -0.011)$	$(-0.026, -0.025)$
1996-2022	-0.025	$(-0.084, -0.003)$	$(-0.083, -0.003)$	$(-0.026, -0.025)$
JKKST annual	-0.015	$(-0.033, -0.005)$	$(-0.033, -0.005)$	$(-0.015, -0.015)$

(a) OOM puts

(b) OOM calls

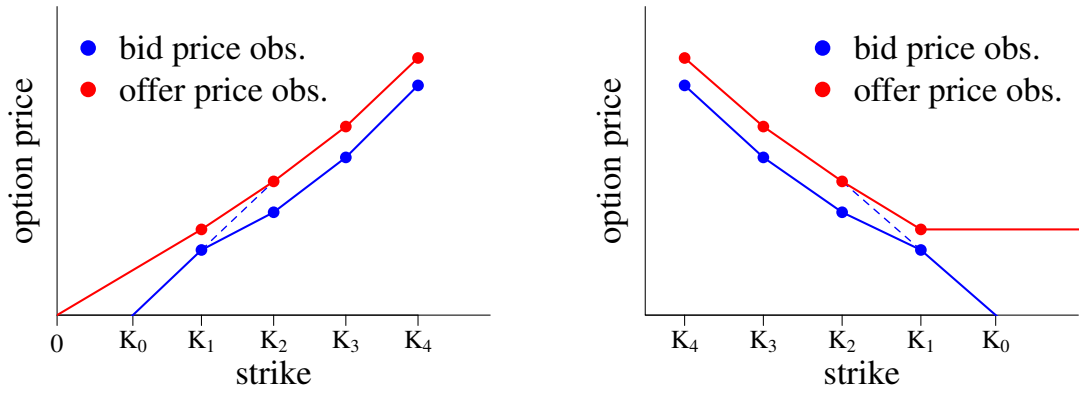


Figure A1: Conservative extrapolation of option prices in Section 4.1.

In these illustrations, we plot increasing out-of-the-money put and call prices at strikes $K_1, K_2, K_3, K_4, \ldots$, marked by blue (bid) and red (offer) dots. We apply linear interpolation between observed strikes, and treat bid and offer price extrapolations separately as follows.

Bid prices. When constructing risk-neutral CGFs from bid prices, we aim to find the lowest arbitrage-free prices. Consider two strikes $K_1 < K_2$ with observed bid-offer quotes for OOM puts, as shown in Panel (a). To rule out butterfly arbitrage, the cost of buying $\frac{K_2-K_1}{K_2-K_0}$ K_0-puts plus $\frac{K_1-K_0}{K_2-K_0}$ K_2-puts must be no less than the proceeds from shorting one put at K_1. Solving this inequality yields a unique maximum arbitrage-free strike K_0 such that $\text{put}(K_0) = 0$. A similar argument yields K_0 for calls in Panel (b).

Offer prices. When constructing risk-neutral CGFs from offer prices, we use the highest possible prices to remain conservative. For OOM puts in Panel (a), since $\text{put}(0) = 0$, linear extrapolation from the lowest observed offer price at K_1 to the origin gives the highest possible extrapolated prices by a standard convexity argument (the function $\text{put}(K)$ must be convex in K). As $\text{put}_t(KR_{f,t+1})$ is then linear in K for K close to zero, $\theta(\theta-1)\int_0^1 K^{\theta-2}\,\text{put}_t(KR_{f,t+1})\,\mathrm{d}K$ in equation (5) will diverge when $\theta < 0$. For OOM calls, in Panel (b), the most conservative extrapolation holds the offer prices constant at and above the highest observed strike K_1. As a result, with $\text{call}_t(KR_{f,t+1})$ constant for large K, $\theta(\theta-1)\int_1^\infty K^{\theta-2}\,\text{call}_t(KR_{f,t+1})\,\mathrm{d}K$ in equation (5) will diverge when $\theta > 1$.

To summarize, when we extrapolate offer prices in this maximally conservative way, the "offer" risk-neutral CGF $\boldsymbol{\kappa}_t(1,\theta)$ becomes infinite for $\theta \notin [0,1]$.

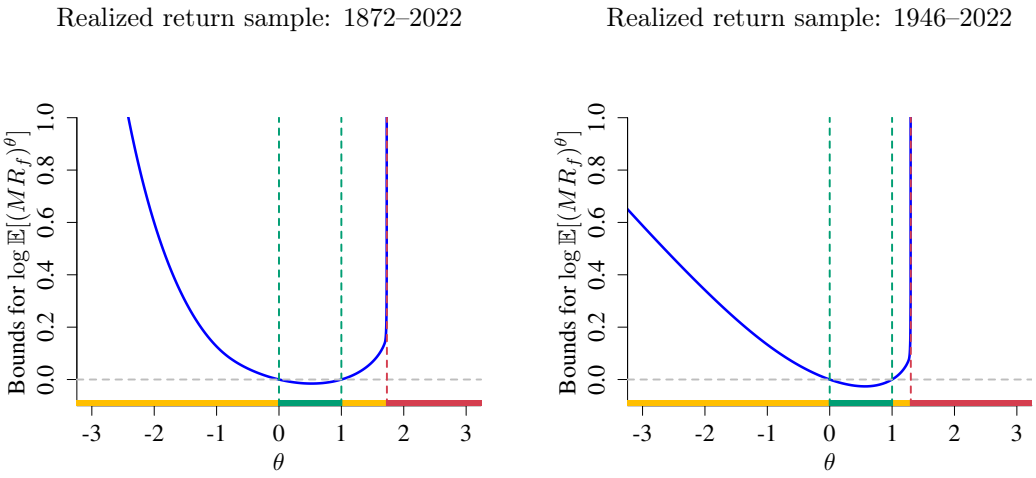


Figure A2: Convexity bounds for the moments of the SDF: one-year horizon.

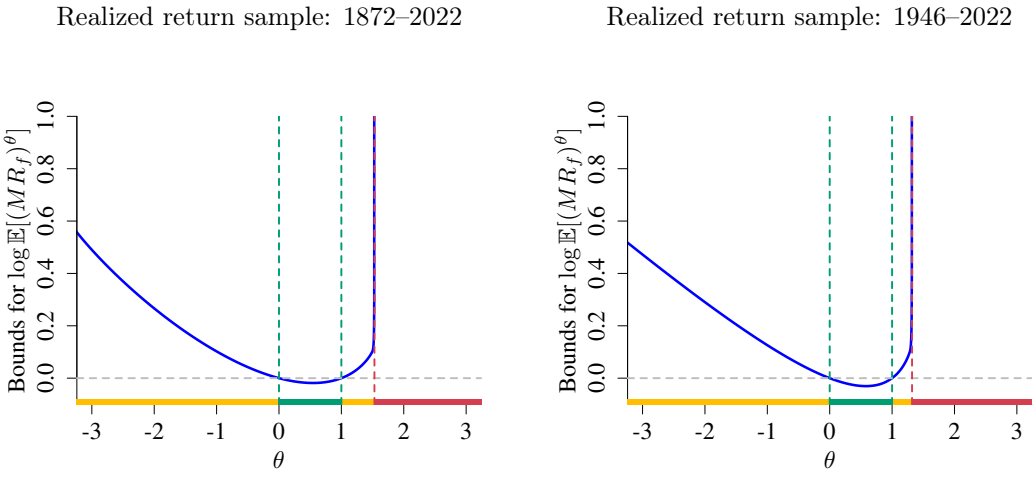


Figure A3: Moment bounds calculated using mid-market option prices, extrapolating using a flat volatility smile.

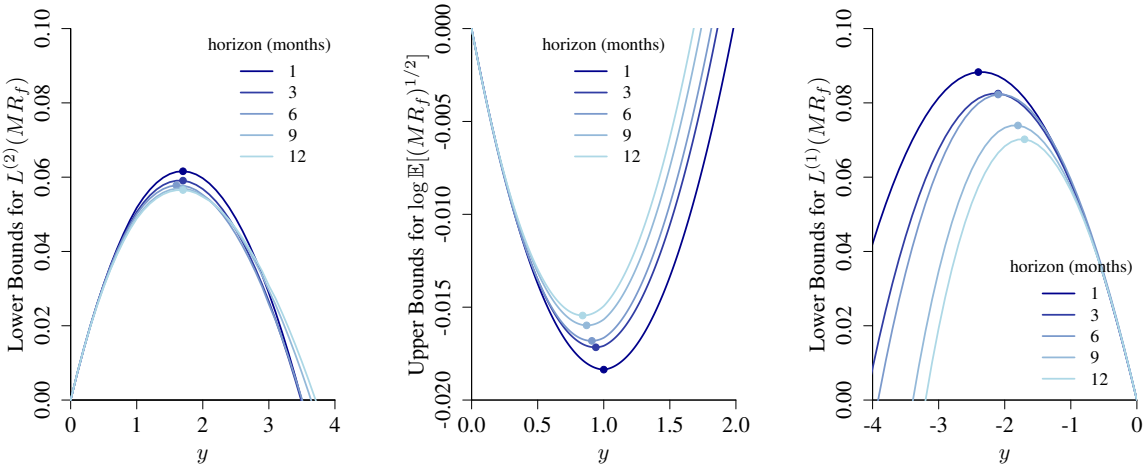


Figure A4: Moment bounds are well-behaved for $\theta \in (0, 1)$ and for the two entropy bounds. This figure shows bounds on the $\theta = 1/2$ moment of the SDF and on the two entropy measures as a function of y, across different horizons. All curves are annualized. The points on the curves represent the optimizing values of y (which, in the left and right panels, supplies the risk aversion measure described in Result 4).

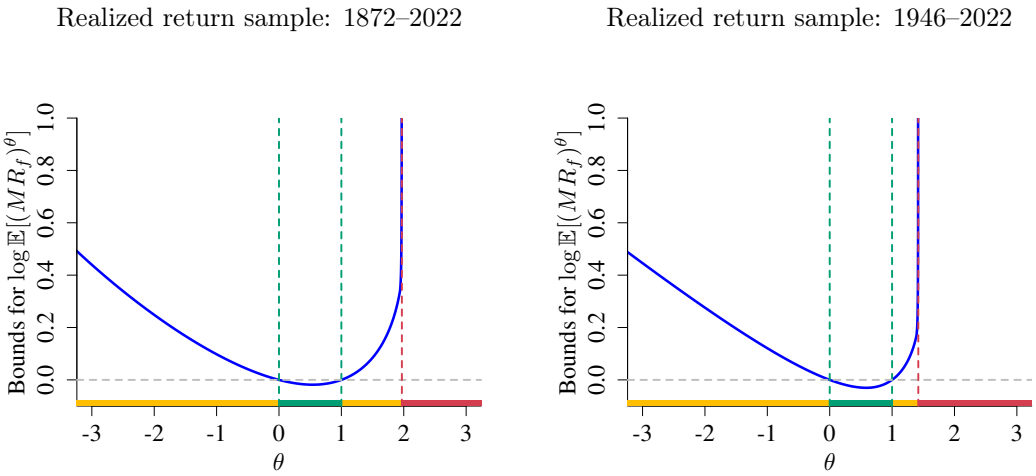


Figure A5: Convexity bounds for the moments of the SDF: no extrapolation outside the range of observed option strikes.

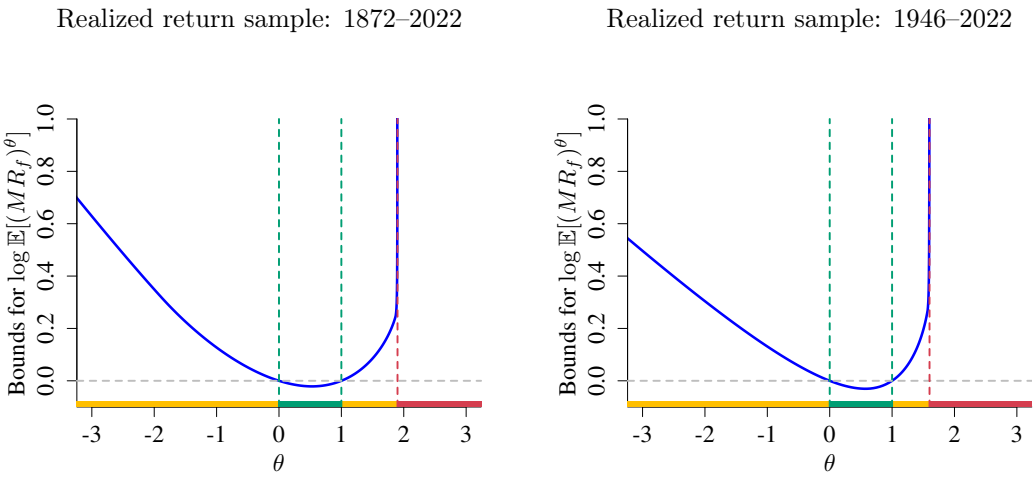


Figure A6: Convexity bounds for the moments of the SDF: using daily observations of option prices and realized returns.

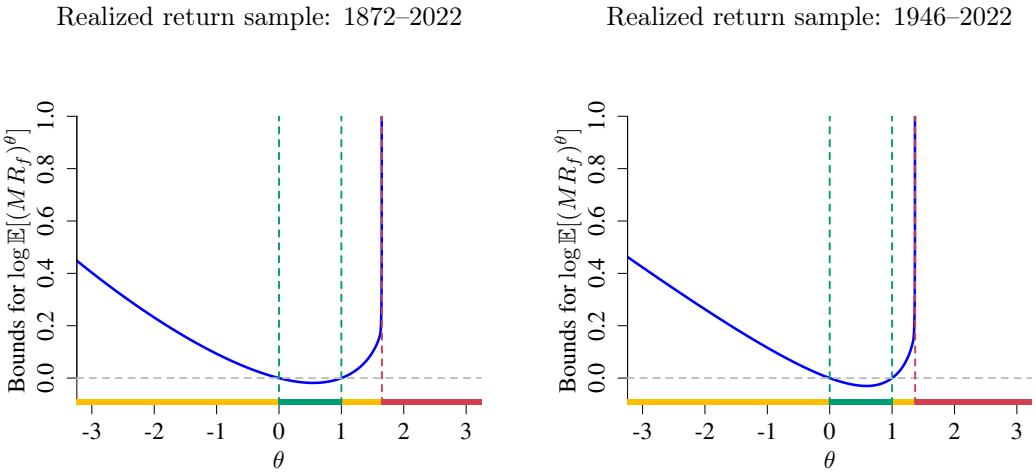


Figure A7: Convexity bounds for the moments of the SDF: option sample starting March 6, 2008.

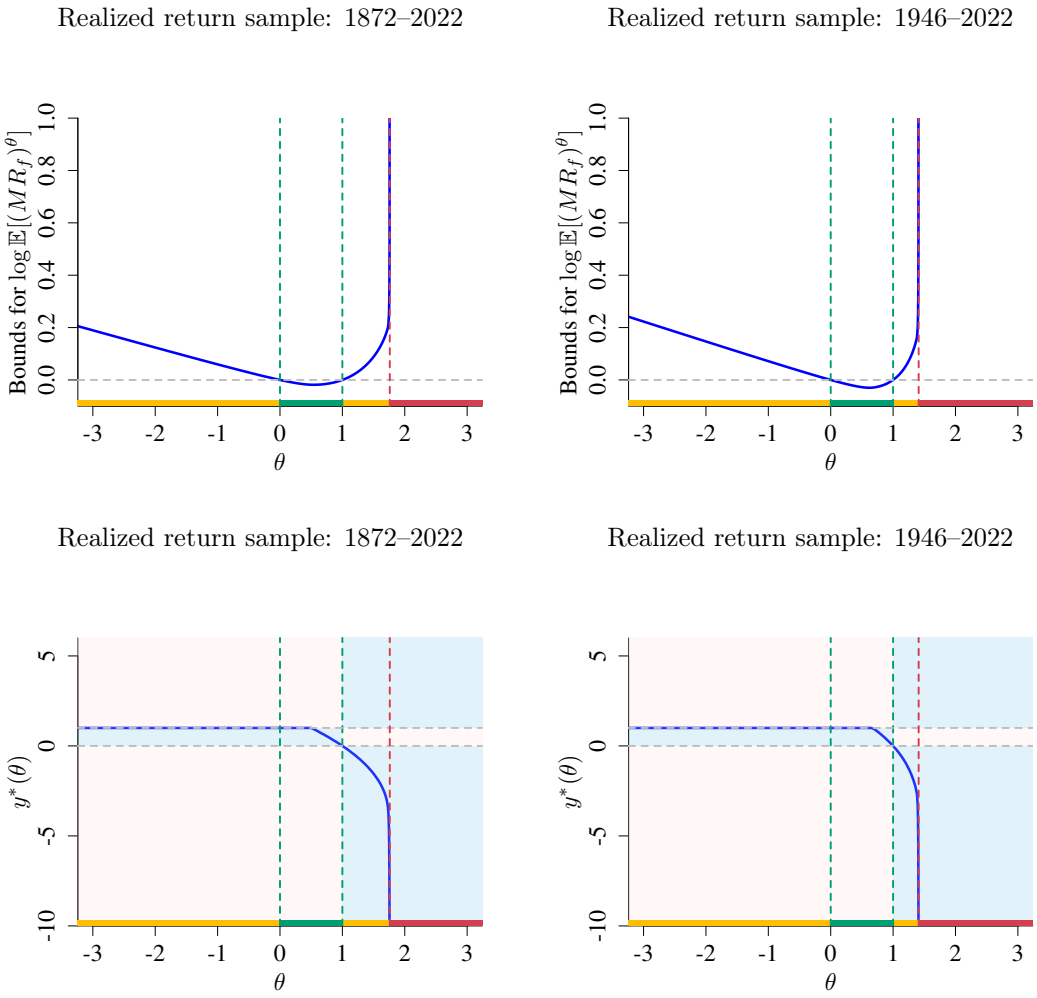


Figure A8: Convexity bounds for the moments of the SDF and their optimizing values based on the most conservative approach to extrapolating option offer prices.

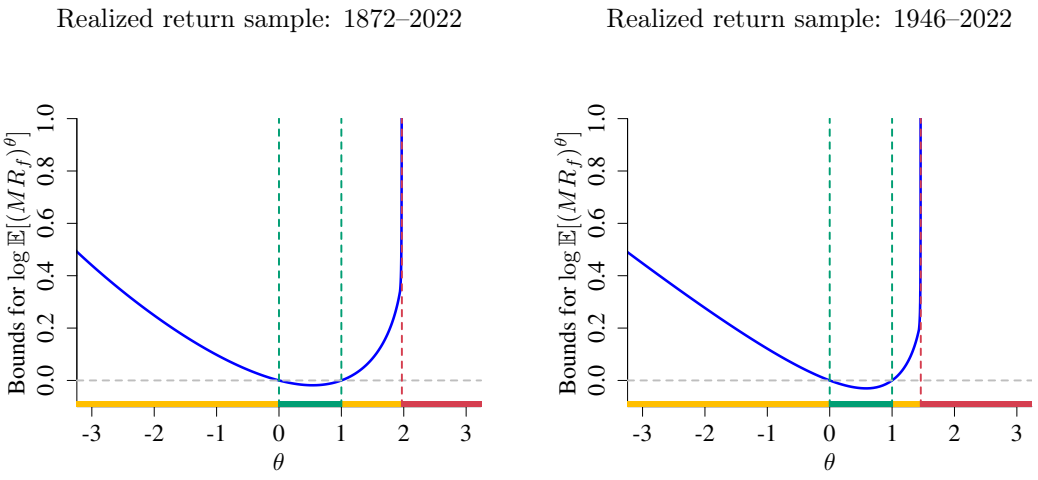
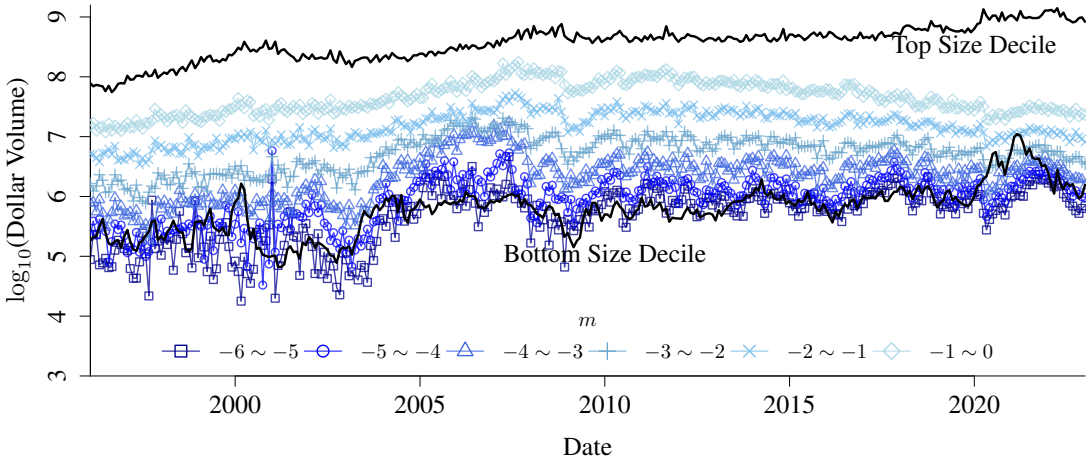


Figure A9: Convexity bounds for the moments of the SDF: moneyness filter $-6 \leq m \leq 3$ applied to the option sample.

(a) Out-of-the-money puts



(b) Out-of-the-money calls

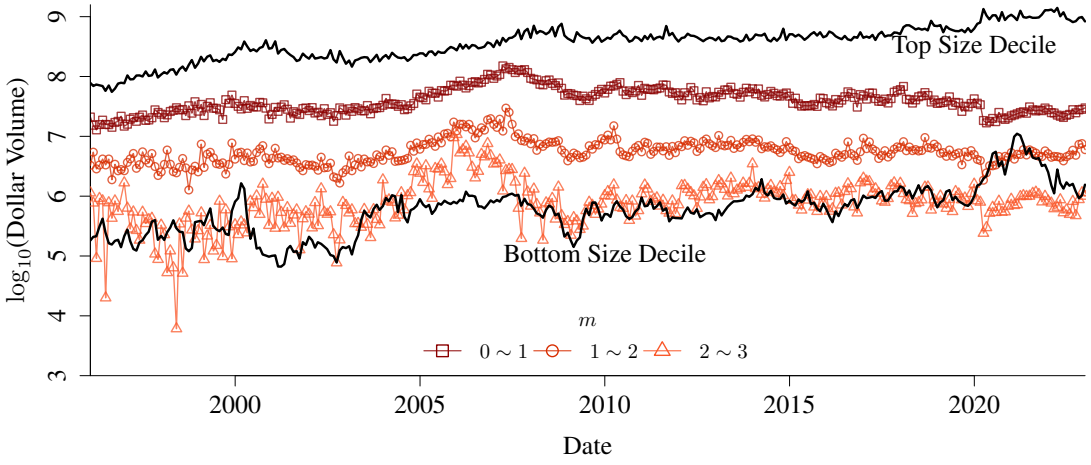


Figure A10: Comparing the dollar trading volumes of options and stocks: time-series patterns

This plot displays the average daily dollar trading volumes for SPX options by month and moneyness groups. Blue lines represent puts, while red lines represent calls. Additionally, two black lines show the average daily dollar trading volumes for stocks, one for the top decile (the largest companies) and one for the bottom decile (the smallest companies) of a size-sorted portfolio.