

A Tale of two Tails: the Quantile Approach to Understanding Market Risk-Premia

Tjeerd de Vries *

This Version: September 11, 2023

Abstract

This paper studies the misspecification of asset pricing models by examining the local difference between the physical and risk-neutral measure. Using quantile regression, I find that the conditional physical distribution differs most from the risk-neutral distribution in the left-tail, while the right tails of both distributions are almost identical. These local differences cause the stochastic discount factor to be highly volatile and suggest that the equity premium is driven by disaster risk. I also propose a risk-adjustment term that captures much of the physical and risk-neutral quantile difference in the left-tail. I interpret the risk-adjustment term as a real time measure of the Peso problem and find that it fluctuates significantly over time. At the height of the 2008 financial crisis and 2020 Covid crisis, the risk-adjustment term predicts that a market return of -28% or lower has a 5% probability. Both findings provide model-free evidence of time varying disaster risk.

Keywords: Asset pricing, Model misspecification, Quantile methods

JEL Codes: G13, G17, C14, C22

*Department of Economics, University of California San Diego. Email: tjdevrie@ucsd.edu. I would like to thank Alexis Toda, Allan Timmermann, James Hamilton, Xinwei Ma, Yixiao Sun, Rossen Valkanov, Brendan Beare, Caio Almeida, Ali Uppal, Rob Engle (discussant), Andrew Lo (discussant), Lerby Ergun, Casper de Vries and seminar participants of the 2022 SoFiE conference, 2023 JOIM conference and ESI Brown Bag Seminar for useful feedback. All errors are my own.

1 Introduction

Asset pricing models are essential for understanding the dynamics of financial markets and predicting the behavior of asset prices. Evaluating the misspecification of these models is crucial for understanding why a particular modeling approach falls short in explaining certain characteristics of the data. For example, [Hansen and Jagannathan \(1991\)](#) show that the Sharpe ratio puts a tight constraint on the volatility of the stochastic discount factor, which can be used to sieve out potential asset pricing models. However, a single statistic such as the Sharpe ratio may not give a comprehensive depiction of the underlying data, thereby allowing for diverse interpretations regarding the cause of its high level in U.S. historical data.

In this paper, I propose a new way to analyze model misspecification in asset pricing based on *local differences* between the physical and risk-neutral distribution. The basic insight is that any statistic of interest, such as the equity or variance premium, is determined by differences between the physical and risk-neutral measure. Different asset pricing models imply distinct restrictions on the shape of these distributions, and analyzing this difference empirically provides a new way to discriminate between models. As a leading example, I consider the class of disaster risk models, which predict that almost all differences are concentrated in the left-tail. The main difficulty in assessing this prediction empirically is that the conditional physical distribution is unobserved, thus making inference robust to model misspecification challenging.

The contribution of this paper is to use quantiles to quantify the difference between the physical and risk-neutral measure across various parts of the return distribution. This approach provides a more comprehensive understanding of the requirements for an asset pricing model to explain the data, in contrast to a conventional approach solely based on the estimated equity or variance premium. Empirically, I quantify this conditional difference through quantile regression, using the risk-neutral quantile function as a regressor. This approach is model-free and solely based on forward looking information. The main results highlight the following findings (i) the majority of the conditional difference is concentrated in the left-tail; (ii) quantile regression sheds light on model misspecification and numerous asset pricing puzzles, including the equity premium puzzle, pricing kernel monotonicity, first-order stochastic dominance and belief recovery; (iii) local differences between the physical and risk-neutral measures impact the SDF; and (iv) a model-free predictor variable can real-time forecast the quantile gap and serve as a measure of the Peso problem.

I start by providing empirical evidence that the conditional physical and risk-neutral distribution of the S&P500 index differ most in the left-tail. This differ-

ence cannot be measured directly since the physical distribution is unobserved, even though the risk-neutral distribution can be identified from option prices ([Breeden and Litzenberger, 1978](#)). To proceed, I use quantile regression (QR) to estimate

$$\underbrace{Q_{t,\tau}(R_{m,t \rightarrow N})}_{\text{Unobserved}} = \beta_0(\tau) + \beta_1(\tau) \underbrace{\tilde{Q}_{t,\tau}(R_{m,t \rightarrow N})}_{\text{Observed}} \quad \tau \in (0, 1), \quad (1.1)$$

where $Q_{t,\tau}$, $\tilde{Q}_{t,\tau}$ denote the physical and risk-neutral τ -quantile respectively of the market return $R_{m,t \rightarrow N}$ from time t to $t + N$. The parameters in (1.1) can be estimated from a quantile regression of $R_{m,t \rightarrow N}$ on $\tilde{Q}_{t,\tau}(R_{m,t \rightarrow N})$, since the realized market return is drawn from the physical distribution. Any deviation from the risk-neutral benchmark, $[\beta_0(\tau), \beta_1(\tau)] = [0, 1]$, signifies a local difference between the physical and risk-neutral measure at the τ -quantile. From the QR estimates, I find that (i) the risk-neutral benchmark cannot be rejected in the right-tail (τ close to 1) but is rejected in the left-tail (τ close to 0); and (ii) the explanatory power in- and out-of-sample is much higher in the right-tail compared to the left-tail. Hence, in order for an asset pricing model to be consistent with empirical data, it must produce a right-tail of the physical distribution that is close to risk-neutral, since investors only get compensated for bearing downside risk. While disaster risk models incorporate this feature, it cannot be achieved by models which assume that the market return follows a conditional lognormal distribution. As a byproduct of the QR estimates, I also clarify the role of stochastic dominance and pricing kernel monotonicity, which are typically embedded in asset pricing models.

To understand the influence of local differences between the physical and risk-neutral measure on the pricing kernel, I introduce a *quantile bound* on the SDF volatility that is closely related to the [Hansen and Jagannathan \(1991\)](#) bound (HJ bound henceforth). The quantile bound depends on the following measure of dispersion:

$$\tau - \phi(\tau), \quad \tau \in (0, 1),$$

where $\phi(\tau) = F(\tilde{Q}_\tau(R))$ and $F(\cdot)$, $\tilde{Q}_\tau(\cdot)$ denote the (unconditional) physical cumulative distribution function (CDF) and risk-neutral quantile of a generic return R . The quantity $\phi(\tau)$ is referred to as the ordinal dominance curve (ODC) of the physical and risk-neutral measure, and $\tau - \phi(\tau)$ can be thought of as a quantile analogue of the equity premium. The quantile bound shows how any pointwise difference between the physical and risk-neutral measure, gauged by $\tau - \phi(\tau)$, leads to a volatile SDF. In the data, the quantile bound peaks around the 5th percentile ($\tau = 0.05$), which implies an SDF volatility of at least 30% in monthly units. In contrast, the Sharpe ratio renders a much looser bound on the SDF volatility, of about 13%. Not

only does this finding put a tighter constraint on a model's ability to match the SDF volatility, it also diagnoses that volatility is mainly driven by the left-tail. In the realm of asset pricing models, I show that these empirical restrictions underscore the necessity of incorporating jumps in disaster risk models.

Based on my empirical findings that the physical and risk-neutral measures differ most in the left-tail, I propose a method for gauging this conditional difference using a risk-adjustment term observed from option prices. This approach provides a model-free way to assess whether disaster risk changes over time. Using some elementary results from functional analysis, I show that up to first-order, the difference between the physical and risk-neutral quantile is directly proportional to $\tau - \phi_t(\tau)$, where $\phi_t(\tau)$ denotes the conditional ODC. Under mild economic constraints: $\tau - \phi_t(\tau) \geq LB_{t,\tau}$, where $LB_{t,\tau}$ is a lower bound that can be computed from option prices. This result is reminiscent of the lower bounds on the equity premium from [Martin \(2017\)](#) and [Chabi-Yo and Loudis \(2020\)](#). Using quantile regression, I show that the risk-adjustment term captures most of the (unobserved) wedge between the physical and risk-neutral quantile function. Since the left-tail of the physical quantile function is a good measure of disaster risk, I obtain a market observed proxy for disaster risk. In the data, this disaster risk proxy correlates well with events associated to high market uncertainty such as the 2008 financial crisis or the 2020 Covid pandemic. During these periods' peak, my measure predicts that with 5% probability, the (monthly) market return would drop by 28% or more. This probability is 57 times higher than what historical data indicate. Hence, under rational expectations, the forward looking probability of a crash perceived by an investor is rather different from historical estimates, representing an instance of the Peso problem. Additionally, I find that the disaster risk measure fluctuates significantly over time, which is model-free evidence of time varying disaster risk.

I conclude by proposing several out-of-sample exercises to test the robustness of my findings. The literature on predicting excess returns highlights the importance of out-of-sample performance ([Campbell and Thompson, 2008](#); [Welch and Goyal, 2008](#)), but many of the performance measures, such as the out-of-sample R^2 , are tailored to OLS and cannot be applied to quantile regression. To address this issue, I propose a clean substitute for the out-of-sample R^2 in the context of quantile regression. Using this measure, I find that the risk-adjustment term better predicts the physical quantile than competing benchmarks. Unlike the equity premium, there is less guidance in the literature on variables that should predict quantiles. As a natural benchmark, I consider the VIX index since it is available daily and known to be correlated with negative market sentiment ([Bekaert and Hoerova, 2014](#)). Sur-

prisingly, I find that the risk-adjustment term outperforms the VIX predictor in the left-tail, even though the VIX uses in-sample information. This fact strengthens my interpretation of the risk-adjustment term as a measure of disaster risk since this term does not require any parameter estimation. Using the risk-neutral quantile as a predictor, I obtain the same result in the right-tail of the return distribution.

1.1 Related Literature

My approach, which uses quantile regression to measure local dispersion between the physical and risk-neutral distribution, is related to a larger body of literature that estimates the pricing kernel from returns and option data (Aït-Sahalia and Lo, 2000; Jackwerth, 2000; Rosenberg and Engle, 2002; Beare and Schmidt, 2016; Linn et al., 2018; Cuesdeanu and Jackwerth, 2018). However, estimating the pricing kernel from returns and options can be challenging, especially in the tails of the distribution, where the ratio of densities that defines the pricing kernel can become unstable. In addition, using historical returns to estimate the physical density can lead to inconsistent results (Linn et al., 2018). In contrast, QR can be used to draw inference on the pricing kernel indirectly, by leveraging the *observed* realized return and risk-neutral distribution, which avoids the estimation of a density ratio. Furthermore, QR can account for changes in the shape and scale of the underlying SDF over time, while the approach of Cuesdeanu and Jackwerth (2018) renders an estimate of the SDF that only allows the normalizing constant to be time varying, since the shape and scale are time invariant (see Section 3.4).

Complimentary to the QR estimates, I derive a nonparametric bound on the SDF volatility closely related to the bound of Hansen and Jagannathan (1991). They argue that the SDF is necessarily volatile and use this observation to screen asset pricing models. Several papers have built on this insight using higher-order moment bounds (Snow, 1991; Almeida and Garcia, 2012; Liu, 2021) and entropy bounds (Stutzer, 1995; Bansal and Lehmann, 1997; Alvarez and Jermann, 2005; Backus et al., 2014). These bounds all provide a measure of how much the risk-neutral distribution differs from the physical distribution. Unlike the quantile bound, all of these measures are global in that they rely on averages over the entire state space. The quantile bound in this paper is a function rather than a single statistic and can be considered an intermediate approach between a single bound and a complete estimate of the SDF.

This paper is also related to the growing literature on using options to estimate forward-looking equity premiums (Martin, 2017; Martin and Wagner, 2019; Chabi-Yo and Loudis, 2020). However, unlike those papers that focus on the conditional

expectation of excess returns, this paper uses option data to predict conditional return quantiles. The relationship between option prices and expected market return shocks has been extensively studied in the literature (Bates, 1991, 2000, 2008; Coval and Shumway, 2001; Bollerslev and Todorov, 2011; Backus et al., 2011; Ross, 2015). Similar to Bollerslev and Todorov (2011), this paper obtains a nonparametric measure of fear or disaster risk. However, the approach differs in that it only uses risk-neutral information and is motivated by the interplay between the physical and risk-neutral quantile function. This approach also complements the recovery literature as it derives forward-looking approximations to the left- and right-tail of the physical distribution using option data (Ross, 2015; Borovička et al., 2016; Qin and Linetsky, 2017; Bakshi et al., 2018; Qin et al., 2018; Schneider and Trojani, 2019; Jackwerth and Menner, 2020). The time variation in the approximation for the left-tail quantile documented in this paper is consistent with the time-varying disaster risk models of Gabaix (2012), Wachter (2013), Isoré and Szczerbowicz (2017), Farhi and Gourio (2018) and Seo and Wachter (2019).

Finally, the QR approach is similar to conditional mean regressions that are common in the equity premium literature. The performance evaluation of conditional expected return predictors is well established in the literature, with important contributions from Campbell and Thompson (2008) and Welch and Goyal (2008). To evaluate the performance of the QR approach, I draw on earlier work of Koenker and Machado (1999) and extend the evaluation toolkit to the quantile setting, specifically focusing on out-of-sample performance. This paper thus complements the literature on conditional return prediction by extending it to the entire distribution.

The rest of this paper is organized as follows. Section 2 presents the main empirical results from the quantile regressions and its consequences for the SDF are discussed in Section 3. Section 4 introduces the quantile bound, discusses its use in asset pricing models, and presents estimates of the quantile bound from empirical data. Section 5 complements the results from Section 2 and shows how risk-neutral information can be used to measure conditional disaster risk. Finally, Section 6 concludes.

2 Empirical Estimates of Quantile Difference

This section documents empirical estimates of the conditional quantile difference between the physical and risk-neutral measures. I first discuss the notation and then consider an example to clarify the idea and motivate the methodology.

2.1 Notation

Let $R_{m,t \rightarrow N}$ denote the market return from period t to $t + N$, where N can be measured in days or years, depending on the context. The risk-free rate over the same period is denoted by $R_{f,t \rightarrow N}$, which is assumed to be known at time t . In the absence of arbitrage, there exists a positive random variable $M_{t \rightarrow N}$ such that, conditional on all time t information

$$\mathbb{E}_t [M_{t \rightarrow N} R_{m,t \rightarrow N}] = 1. \quad (2.1)$$

The random variable $M_{t \rightarrow N}$ is referred to as the stochastic discount factor (SDF) and the expectation in (2.1) is calculated under the *physical* probability measure \mathbb{P}_t , which is the actual distribution of the market return, i.e. $R_{m,t \rightarrow N} \sim \mathbb{P}_t$. The SDF can potentially depend on many state variables, but these are suppressed from the notation for brevity. It is convenient to restate (2.1) in terms of risk-neutral probabilities:

$$\tilde{\mathbb{E}}_t (R_{m,t \rightarrow N}) = 1 / \mathbb{E}_t [M_{t \rightarrow N}] = R_{f,t \rightarrow N},$$

where the expectation is calculated under the *risk-neutral* measure $\tilde{\mathbb{P}}_t$ induced by $M_{t \rightarrow N}$. Finally, $F_t(x) := \mathbb{P}_t(R_{m,t \rightarrow N} \leq x)$ denotes the physical CDF of the market return conditional on all information available at time t , $f_t(\cdot)$ denotes the conditional PDF and $Q_{t,\tau}$ denotes the conditional τ -quantile. As before, a tilde superscript refers to the risk-neutral measure, so that

$$\tilde{F}_t(\tilde{Q}_{t,\tau}) = \tilde{\mathbb{P}}_t \left(R_{m,t \rightarrow N} \leq \tilde{Q}_{t,\tau} \right) = \tau, \quad \text{for } \tau \in (0, 1).$$

The physical and risk-neutral quantiles depend on the underlying random variable $R_{m,t \rightarrow N}$ (i.e. $\tilde{Q}_{t,\tau} := \tilde{Q}_{t,\tau}(R_{m,t \rightarrow N})$), but I typically omit this dependence as the underlying random variable should be clear from the context.

To clarify my approach of using quantiles to analyze model misspecification, I consider the following asset pricing model that will be used several times in the paper.

Example 2.1 (Disaster risk). Consider the disaster risk model analyzed in [Backus et al. \(2011\)](#). The SDF process M_{t+1} is given by

$$\log M_{t+1} = \log \beta - \gamma \log G_{t+1},$$

where $G_{t+1} = C_{t+1}/C_t$ is consumption growth in period $t + 1$, which follows a two-

component structure

$$\log G_{t+1} = z_{1,t+1} + z_{2,t+1}, \quad z_{1,t+1} \sim \mathcal{N}(\mu, \sigma^2),$$

and $z_{2,t+1}$ is a Poisson mixture of normals to capture jumps, which represent rare shocks to consumption growth that are large in magnitude. The number of jumps, j , take on nonnegative integer values with probability $e^{-\omega} \omega^j / j!$, and conditional on j , the jump term is normal: $z_{2,t+1}|j \sim \mathcal{N}(j\theta, j\delta^2)$. [Backus et al. \(2011\)](#) show that the risk-neutral distribution of consumption growth in a representative agent model is again a normal mixture with parameters:

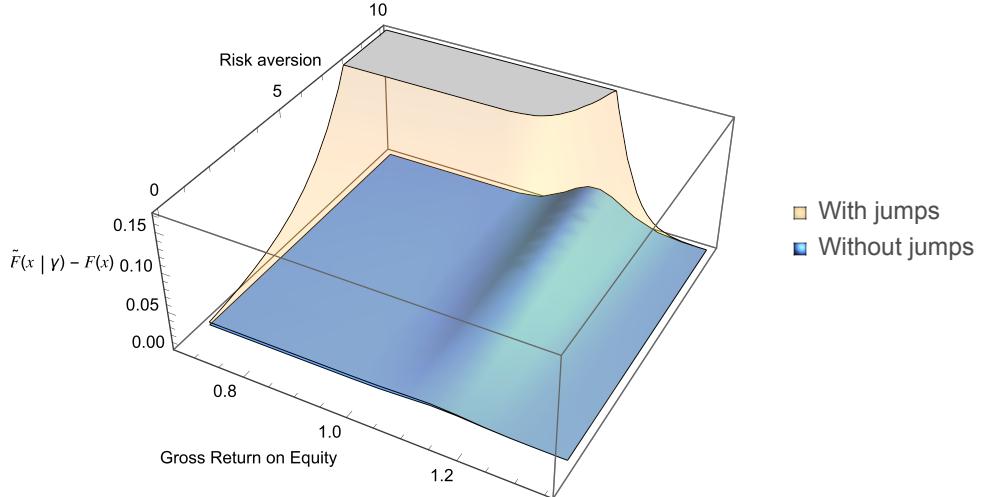
$$\tilde{\mu} = \mu - \gamma\sigma^2, \quad \tilde{\omega} = \omega e^{-\gamma\theta + (\gamma\delta)^2}, \quad \tilde{\theta} = \theta - \gamma\delta^2, \quad (2.2)$$

where γ is the coefficient of risk aversion and δ denotes the time discount factor. In this setup, risk aversion amplifies the jump frequency ($\tilde{\omega} > \omega$ if $\theta < 0$) as well as the jump size ($\tilde{\theta} < \theta$). Letting $\tilde{F}(\cdot|\gamma)$ and $F(\cdot)$ denote the risk-neutral and physical CDF of return on equity respectively, Panel 1a shows the impact of jumps on the difference between these distributions. It is clear that the difference is more pronounced with jumps, and moreover, this difference is almost entirely concentrated in the left-tail, unlike the model without jumps. This result is driven by the impact of jumps on the risk-neutral distribution, and the requirement $\theta < 0$ is crucial to drive a wedge between the physical and risk-neutral measure in the left-tail (see (2.2)). The question is whether these distinct shape restrictions on the physical and risk-neutral distribution are supported by the data.

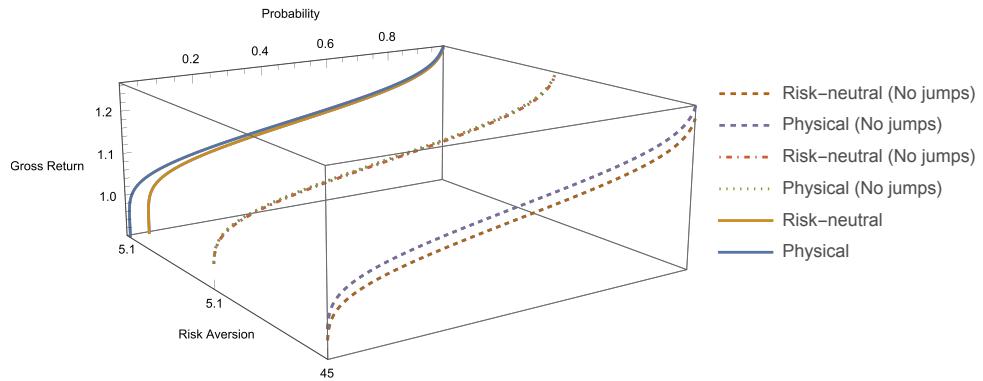
To answer this question, consider the physical and risk-neutral quantile functions with and without jumps in Panel 1b. When jumps are absent, it is much harder to drive a wedge between the physical and risk-neutral distribution using reasonable levels of risk aversion; a manifestation of the equity premium puzzle. Even when the level of risk aversion is sufficiently high to explain, say, the equity premium, the disparity between the physical and risk-neutral quantile functions is mostly concentrated around the median in a model without jumps. To discriminate between these models empirically, I use return data to estimate the quantile wedge at a given probability level.

2.2 Methodology and Econometric Model

The conditional risk-neutral distribution, and hence the quantile function, can be inferred from option prices without making modeling assumptions ([Breeden and Litzenberger, 1978](#)). However, the physical distribution cannot be recovered in a



(a) Effect of jumps and risk aversion on $\tilde{F}(\cdot | \gamma) - F(\cdot)$



(b) Effect of jumps on physical and risk-neutral quantile function

Figure 1: Effect of risk aversion and jumps on the difference between the physical and risk-neutral distribution. Panel (a) shows the difference between the risk-neutral and physical distribution with and without jumps, following the calibration of [Backus et al. \(2011\)](#). Return on equity is defined as a levered claim on the consumption asset. Panel (b) shows the corresponding quantile functions of equity return with and without jumps. The figure shows the model without jumps for two different values of risk aversion, $\gamma = 5.1$ and $\gamma = 45$ respectively.

model-free way from option prices, unless one makes strong assumptions about the martingale component of the SDF (Ross, 2015; Borovička et al., 2016). The information available about the conditional physical distribution is limited to a single realization of the market return, since the distribution of $R_{m,t \rightarrow N}$ is \mathbb{P}_t conditional on time t . Consequently, the main difficulty is to come up with a method to gauge the conditional difference between the physical and risk-neutral measures without relying on the physical distribution directly.

In order to overcome this difficulty, I estimate the model

$$\underbrace{Q_{t,\tau}(R_{m,t \rightarrow N})}_{\text{Unobserved}} = \beta_0(\tau) + \beta_1(\tau) \underbrace{\tilde{Q}_{t,\tau}(R_{m,t \rightarrow N})}_{\text{Observed}}, \quad \forall \tau \in (0, 1). \quad (2.3)$$

If the world is risk-neutral, $[\beta_0(\tau), \beta_1(\tau)] = [0, 1]$ for all τ . Departures from risk-neutrality at a specific percentile τ are reflected by point estimates of $[\beta_0(\tau), \beta_1(\tau)]$ that are far from the $[0, 1]$ benchmark. Given a sample of T observations $\{R_{m,t \rightarrow N}, \tilde{Q}_{t,\tau}\}_{t=1}^T$, the unknown parameters in (2.3) can be estimated by quantile regression (Koenker and Bassett, 1978):

$$[\hat{\beta}_0(\tau), \hat{\beta}_1(\tau)] = \arg \min_{(\beta_0, \beta_1) \in \mathbb{R}^2} \sum_{t=1}^T \rho_\tau(R_{m,t \rightarrow N} - \beta_0 - \beta_1 \tilde{Q}_{t,\tau}), \quad (2.4)$$

where $\rho_\tau(\cdot)$ is the check function from quantile regression

$$\rho_\tau(x) = \begin{cases} \tau x, & \text{if } x \geq 0 \\ (\tau - 1)x & \text{if } x < 0. \end{cases}$$

Even if (2.3) is misspecified, Angrist et al. (2006) show that QR finds the best linear approximation to the conditional quantile function.¹ From the disaster risk model with jumps in Example 2.1, we expect that $\tilde{Q}_{t,\tau}$ is a poor approximation to $Q_{t,\tau}$ in the left-tail ($\tau \approx 0$), but an excellent approximation in the right-tail ($\tau \approx 1$). The model without jumps should render good approximations in the tail, but poor approximations around the median ($\tau \approx 1/2$). QR allows me to assess these competing hypotheses by testing how close the point estimates from (2.4) are to the $[0, 1]$ benchmark for different values of τ .

Remark. As an alternative to quantile regression, one can estimate the SDF non-parametrically as in Aït-Sahalia and Lo (2000), Jackwerth (2000) and Rosenberg and Engle (2002). The quantile difference can then be deduced from the estimated

¹This result is analogous to OLS, which finds the best linear approximation to the conditional expectation function, even if the model is misspecified.

SDF. However, this method uses pooled historical returns, raising concerns about their relevance to estimate a forward-looking distribution (Linn et al., 2018). The QR approach is not prone to this critique, since it leverages directly observed data $\{R_{m,t \rightarrow N}, \tilde{Q}_{t,\tau}\}_{t=1}^T$ that condition on the same information set and are forward-looking. Linn et al. (2018) and Cuesdeanu and Jackwerth (2018) propose an alternative nonparametric approach to estimate the SDF that also accounts for forward-looking information. However, their method can be challenging to implement since (i) the optimization problem is not convex and the objective function might not be well defined due to the small number of existing risk-neutral moments (see Figure F7); (ii) it is generally unclear how to choose the basis functions and how many to select; (iii) the estimated SDF is a projection on the return space, and even in a lognormal model with time-varying volatility, this may lead to incorrect inference about the *conditional* quantile difference. The QR approach does not suffer from either of these difficulties as the optimization problem is convex, and if the market is conditionally lognormal, quantile regression recovers the exact conditional quantile difference between the physical and risk-neutral measures (see Section 3.4).

Based on the quantile regression (2.3), I consider two measures of fit to evaluate how well the risk-neutral distribution locally approximates the physical distribution. The first in-sample measure, $R^1(\tau)$, is defined as²

$$R^1(\tau) := 1 - \frac{\min_{b_0, b_1} \sum_{t=1}^T \rho_\tau(R_{m,t \rightarrow N} - b_0 - b_1 \tilde{Q}_{t,\tau})}{\min_{b_0} \sum_{t=1}^T \rho_\tau(R_{m,t \rightarrow N} - b_0)}. \quad (2.5)$$

This measure of fit was proposed by Koenker and Machado (1999) and is a clean substitute for the OLS R^2 . I also consider an out-of-sample measure of fit

$$R_{oos}^1(\tau) := 1 - \frac{\sum_{t=w}^T \rho_\tau(R_{m,t \rightarrow N} - \bar{Q}_{t,\tau})}{\sum_{t=w}^T \rho_\tau(R_{m,t \rightarrow N} - \bar{Q}_{t,\tau})}, \quad (2.6)$$

where $\bar{Q}_{t,\tau}$ is the historical rolling quantile of the market return from time $t-w+1$ to t , and w is the rolling window length. Notice that (2.6) is a genuine out-of-sample metric since no parameter estimation is used. In the equity premium literature, Campbell and Thompson (2008) stress the importance of out-of-sample predictability; (2.6) is analogous to their out-of-sample R^2 .

2.3 Data and Estimation

To estimate the quantile regression in (2.4), I need data on the market return and the risk-neutral distribution over time. I use overlapping returns on the S&P500 index from WRDS over the period 2003–2021 to represent the market return. I

²It is well known that b_0 in the denominator of (2.5) equals the in-sample τ -quantile.

calculate the market return over a horizon of 30, 60, and 90 days. Second, over the same horizon, I use put and call option prices on the S&P500 on each day t from OptionMetrics to estimate the risk-neutral quantile function $\tilde{Q}_{t,\tau}$ nonparametrically, based on the [Breeden and Litzenberger \(1978\)](#) formula:

$$\tilde{F}_t \left(\frac{K}{S_t} \right) = R_{f,t \rightarrow N} \frac{\partial}{\partial K} \text{Put}_t(K), \quad (2.7)$$

where $\text{Put}_t(K)$ denotes the time t price of a European put option on the S&P500 index with stock price S_t , strike price K and expiration date $t + N$. Due to the lack of a continuum of option prices, interpolation of different maturity options and missing data for option prices far in- and out-of-the money, it is a nontrivial exercise to obtain accurate estimates of \tilde{F}_t (and hence $\tilde{Q}_{t,\tau}$) from (2.7). A detailed description of my approach that overcomes these issues is described in Appendix B.2, which is based on [Filipović et al. \(2013\)](#).³ Finally, I obtain the risk-free rate from Kenneth French's website.⁴

Table 1 shows the QR estimates of (2.4). The point estimates are close to the $[0, 1]$ benchmark in the right-tail ($\tau \geq 0.8$), but not in the left-tail ($\tau \leq 0.2$). Additionally, the joint restriction that $[\beta_0(\tau), \beta_1(\tau)] = [0, 1]$ is (borderline) rejected for all $\tau \leq 0.2$, at all horizons. In contrast, the null hypothesis is never rejected for $\tau \geq 0.8$. The fact that the risk-neutral distribution provides a good approximation of the physical distribution in the right-tail is confirmed by the measures of fit, $R^1(\tau)$ and $R_{oos}^1(\tau)$, which are also shown in Table 1. Specifically, both in- and out-of-sample, the risk-neutral quantile fits the physical distribution much better in the right-tail.

Remark. The standard errors for the quantile regression in Table 1 are obtained by the smooth extended tapered block bootstrap (SETBB) of [Gregory et al. \(2018\)](#), which is robust to heteroscedasticity and weak dependence.⁵ This robustness is important in the estimation, since I use overlapping returns which creates time dependence in the error term, akin to the overlapping observation problem in OLS ([Hansen and Hodrick, 1980](#)). SETBB also renders an estimate of the covariance matrix between $\hat{\beta}_0(\tau)$ and $\hat{\beta}_1(\tau)$, which can be used to test joint restrictions on the

³This approach uses a kernel density and adds several correction terms to approximate the risk-neutral density. I follow [Barletta and Santucci de Magistris \(2018\)](#) and use a principal components step to avoid overfitting in the tails.

⁴See http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html#Research

⁵It may seem counterintuitive that the standard errors decrease in the tails, which are generally harder to estimate. However, since the regressor $\tilde{Q}_{t,\tau}$ changes with τ , there is an opposing effect that can cause the standard errors to decrease in the tails. This happens if $\tilde{Q}_{t,\tau}$ is more variable in the tails, akin to the intuition in OLS that more variability in the regressor decreases the standard error. In the data, $\tilde{Q}_{t,\tau}$ is much more variable in the tails compared to the median.

coefficients.⁶

Table 1: **Risk-neutral quantile regression**

Horizon	τ	$\hat{\beta}_0(\tau)$	$\hat{\beta}_1(\tau)$	Wald test (p -value)	$R^1(\tau)[\%]$	$R_{oos}^1(\tau)[\%]$	$\overline{Hit}[\%]$
30 days*	0.05	0.43 (0.220)	0.56 (0.235)	0.01	6.28	6.11	-2.67 (0.699)
	0.1	0.45 (0.244)	0.54 (0.254)	0.03	3.45	1.01	-3.56 (1.162)
	0.2	0.69 (0.375)	0.30 (0.382)	0.10	0.55	0.89	-3.73 (1.695)
	0.5	-0.60 (0.307)	1.61 (0.305)	0.00	1.65	2.24	-8.07 (2.567)
	0.8	-0.09 (0.163)	1.09 (0.158)	0.23	12.44	12.50	-3.24 (2.221)
	0.9	0.03 (0.113)	0.97 (0.108)	0.96	20.41	21.88	-0.04 (1.509)
	* (Obs. 4333)	0.95	0.12 (0.119)	0.89 (0.113)	0.51	27.07	31.31 (1.120)
60 days**	0.05	0.45 (0.321)	0.54 (0.364)	0.00	3.12	13.14	-3.33 (0.831)
	0.1	0.58 (0.329)	0.41 (0.354)	0.01	1.79	3.50	-5.57 (1.312)
	0.2	0.78 (0.418)	0.21 (0.428)	0.04	0.38	-0.03	-6.60 (2.425)
	0.5	-1.02 (0.446)	2.01 (0.440)	0.03	2.81	3.86	-7.88 (4.047)
	0.8	-0.08 (0.209)	1.08 (0.200)	0.12	12.70	12.23	-5.53 (3.012)
	0.9	0.04 (0.150)	0.96 (0.141)	0.57	21.66	22.79	-1.94 (1.816)
	** (Obs. 4312)	0.95	0.04 (0.132)	0.96 (0.123)	0.91	31.07	34.19 (1.102)
90 days***	0.05	0.60 (0.422)	0.37 (0.502)	0.02	2.90	15.63	-2.95 (1.054)
	0.1	0.59 (0.403)	0.40 (0.447)	0.00	3.46	3.84	-6.36 (1.463)
	0.2	0.57 (0.601)	0.43 (0.621)	0.09	0.83	1.93	-7.53 (3.031)
	0.5	-0.99 (0.624)	1.99 (0.613)	0.04	2.58	0.18	-10.85 (4.873)
	0.8	-0.23 (0.254)	1.23 (0.240)	0.28	15.47	16.54	-6.66 (3.447)
	0.9	-0.02 (0.176)	1.02 (0.163)	0.86	23.18	27.92	-1.12 (2.008)
	*** (Obs. 4291)	0.95	0.08 (0.160)	0.93 (0.146)	0.87	32.14	39.88 (1.407)

Note: This table reports the QR estimates of (2.4) over the sample period 2003–2021 at different horizons, using overlapping returns. Standard errors are shown in parentheses and based on SETBB with a block length of 5 times the prediction horizon. Wald test denotes the p -value of the joint restriction $[\beta_0(\tau), \beta_1(\tau)] = [0, 1]$. $R^1(\tau)$ denotes the goodness of fit measure (2.5). $R_{oos}^1(\tau)$ is the out-of-sample goodness of fit (2.6), using a rolling window of size 10 times the prediction horizon. Hit refers to the sample expectation defined in (3.2) and standard errors are reported in parentheses, which are obtained by stationary bootstrap based on 10,000 bootstrap samples.

⁶I use the `QregBB` function from the *R* package `QregBB`, available on the author's Github page: <https://rdrr.io/github/gregorkb/QregBB/man/QregBB.html>. The only user required input for this method is the block length in the bootstrap procedure.

3 Implications for the SDF

I comment on three implications of Table 1 that relate to model misspecification and properties of the SDF that have previously received attention in the literature. I also present a conditional lognormal model and derive closed form expression for the limiting values of the QR estimates, which are shown to be incompatible with Table 1.

3.1 Equity Premium Puzzle

The results in Table 1 show that the physical distribution is close to risk-neutral in the right-tail, but not in the left-tail. Investors in the market portfolio thus get compensated for bearing downside risk, but not upside risk. This result has important repercussions for explanations of the equity premium puzzle. To see this, consider the following decomposition of the equity premium⁷

$$\begin{aligned}\mathbb{E}_t [R_{m,t \rightarrow N}] - R_{f,t \rightarrow N} &= \int_0^1 (Q_{t,\tau} - \tilde{Q}_{t,\tau}) d\tau \\ &= \underbrace{\int_0^{\underline{\tau}} (Q_{t,\tau} - \tilde{Q}_{t,\tau}) d\tau}_{\text{disaster risk}} + \int_{\underline{\tau}}^1 (Q_{t,\tau} - \tilde{Q}_{t,\tau}) d\tau,\end{aligned}\quad (3.1)$$

where $\underline{\tau}$ is a percentile close to zero. The first term on the right-hand side aggregates the local difference between the risk-neutral and physical quantile in the left-tail, which I define as the contribution of disaster risk. The results in Table 1 show that these differences are the primary determinant for the equity premium, as in the right-tail we have $Q_{t,\tau} \approx \tilde{Q}_{t,\tau}$. The latter finding is consistent with the modeling assumption in (time varying) disaster risk models that shocks to the market return are negative conditional on a disaster occurring (see, e.g., the condition $\theta < 0$ in Example 2.1). Hence, an asset pricing model seeking to explain the (conditional) equity premium of the market portfolio must embed a source of disaster risk.

3.2 Pricing Kernel Monotonicity and Stochastic Dominance

Pricing kernel monotonicity refers to the property that $M_{t \rightarrow N}(R_{m,t \rightarrow N}) := \mathbb{E}[M_{t \rightarrow N}|R_{m,t \rightarrow N}]$ is a decreasing function of the market return. Asset pricing models that link the SDF to the marginal rate of substitution imply that the pricing kernel is indeed a decreasing function. Empirically, one finds strong evidence that the pricing kernel is not monotonic, which is puzzling as it contradicts that a representative investor is risk-averse (see Aït-Sahalia and Lo (1998), Jackwerth (2000), Rosenberg and Engle (2002), Bakshi et al. (2010), Beare and Schmidt (2016) and Cuesdeanu and Jackwerth

⁷See Appendix A.1 for a derivation.

(2018)). However, a formal statistical test that can detect violations of monotonicity is challenging as one needs uniform confidence bands for the estimated SDF, which requires tools from empirical process theory (Beare and Schmidt, 2016).

I consider a different approach based on stochastic dominance. Proposition A.1 in the Appendix shows that pricing kernel monotonicity implies that the physical distribution is first-order stochastic dominant (FOSD) over the risk-neutral distribution, i.e. $F_t(x) \leq \tilde{F}_t(x)$ for all x . The latter condition can be rephrased as $F_t(\tilde{Q}_{t,\tau}) \leq \tau$ for all $\tau \in (0, 1)$. A violation of stochastic dominance, and hence pricing kernel monotonicity, is thus implied if there is statistical evidence that $F_t(\tilde{Q}_{t,\tau}) > \tau$ for a single τ . To investigate this possibility, let⁸

$$\begin{aligned} \text{Hit}_{t \rightarrow N} &= \mathbb{1}(R_{m,t \rightarrow N} < \tilde{Q}_{t,\tau}) - \tau, \\ \overline{\text{Hit}} &= \frac{1}{T} \sum_{t=1}^T \text{Hit}_{t \rightarrow N}. \end{aligned} \quad (3.2)$$

Hence, $\overline{\text{Hit}}$ provides an estimate of $\mathbb{E}(F_t(\tilde{Q}_{t,\tau}) - \tau)$ which ought to be negative for all τ under FOSD.⁹ The “ $\overline{\text{Hit}}$ ” column in Table 1 reports the value of (3.2), which is positive for $\tau = 0.95$ at the 30 and 60 day horizon. However, these estimates are not significant at the conventional levels and a violation of FOSD cannot be concluded.

3.3 Belief Recovery

A recent literature asks to what extent Arrow prices can be used to learn about the underlying probability distribution of the data, or the subjective probabilities used by investors. Since Arrow prices are confounded by risk aversion, it is impossible to identify the underlying probabilities from Arrow prices alone, unless one imposes additional restrictions (Ross, 2015; Borovička et al., 2016; Bakshi et al., 2018; Qin et al., 2018; Jackwerth and Menner, 2020). For example, Ross (2015) uses the Perron-Frobenius theorem to recover investors’ beliefs, which agrees with the underlying physical measure under rational expectations. Complementary to this insight, the QR estimates in Table 1 show that the right-tail of the physical distribution can approximately be recovered from the right-tail of the risk-neutral distribution, which aligns with the investor’s belief under rational expectations. In contrast, the left-tail of the physical distribution cannot be recovered even though the risk-neutral quantile serves as a lower bound. In Section 5, I propose a risk-adjustment term to recover the left-tail of the physical distribution as well from option data.

⁸The $\text{Hit}_{t \rightarrow N}$ function was first introduced by Engle and Manganelli (2004) in a different context.

⁹ $\overline{\text{Hit}}$ also yields another measure of the difference between F_t and \tilde{F}_t . Consistent with the quantile regression estimates, the Hit statistic shows that F_t and \tilde{F}_t are similar in the right-tail, but different in the left-tail.

3.4 A Violation of the Conditional Lognormal Assumption

Figure 1b illustrates that a lognormal assumption on the physical and risk-neutral distribution implies that the tails of the corresponding quantile functions are quite similar, which contradicts the results in Table 1. Since the lognormal assumption features prominently in asset pricing, it is worthwhile to analyze in more detail the behavior of the QR estimates under this assumption.

To describe the environment, consider the following discretized version of the [Black and Scholes \(1973\)](#) model. There is a riskless asset that offers a certain return, $R_{f,t \rightarrow N} \equiv R_f = e^{r_f N}$, and a risky asset with return

$$R_{m,t \rightarrow N} = \exp([\mu_t - \frac{1}{2}\sigma_t^2]N + \sigma_t\sqrt{N}Z_{t+N}), \quad (3.3)$$

where μ_t represents the conditional mean return, σ_t is the conditional volatility, and Z_{t+N} is a random shock that follows a standard normal distribution. In this setup, $M_{t \rightarrow N} := \exp(-[r_f + \xi_t^2/2]N - \xi_t\sqrt{N}Z_{t+N})$ is a valid SDF with conditional Sharpe ratio

$$\xi_t = \frac{\mu_t - r_f}{\sigma_t}.$$

Hence, under risk-neutral measure, the conditional distribution of $R_{m,t \rightarrow N}$ is given by

$$\log \tilde{R}_{m,t \rightarrow N} \sim \mathcal{N}\left((r_f - \frac{1}{2}\sigma_t^2)N, \sigma_t^2 N\right). \quad (3.4)$$

Notice that σ_t is implicitly observed from the risk-neutral distribution, but μ_t is unobserved with mean $\mu := \mathbb{E}[\mu_t]$ and variance $\sigma_\mu^2 := \text{Var}(\mu_t) < \infty$. The following result characterizes the limiting behavior of the QR estimates (2.4) in the lognormal model when the variance of the equity premium is small. A convenient way to model this is by means of a drifting sequence $\sigma_\mu^T \rightarrow 0$ as $T \rightarrow \infty$, which captures the intuition that the volatility of the equity premium is much smaller than the return volatility.

Proposition 3.1 (QR in Lognormal Model). *In the lognormal model described above with return observations $\{R_{m,t \rightarrow N}\}_{t=1}^T$ and risk-neutral quantile functions $\{\tilde{Q}_{t,\tau}\}_{t=1}^T$, the following hold.*

- (i) *Suppose that conditional on time t , μ_t follows a normal distribution $\mu_t \sim \mathcal{N}(\mu, \sigma_\mu^2)$, independent of σ_t . Let $Q_{t,\tau}(\sigma_t, \sigma_\mu)$ denote the physical quantile function of $R_{m,t \rightarrow N}$ conditional on σ_t only. Then, for all $\tau \in \mathcal{I} :=$ a closed subset*

of $[\varepsilon, 1 - \varepsilon]$ for $0 < \varepsilon < 1$, the physical quantile function satisfies

$$\begin{aligned} Q_{t,\tau}(\sigma_t, \sigma_\mu) &= \exp \left[\left(\mu - \frac{1}{2} \sigma_t^2 \right) N + \left(\sqrt{\sigma_\mu^2 N^2 + \sigma_t^2 N} \right) \Phi^{-1}(\tau) \right] \\ &= \tilde{Q}_{t,\tau} e^{(\mu - r_f)N} (1 + \mathcal{O}(\sigma_\mu N)), \end{aligned}$$

where $\Phi^{-1}(\tau)$ denotes the quantile function of the standard normal distribution.

- (ii) Consider a drifting sequence for σ_μ , denoted by $\sigma_\mu^T \rightarrow 0$ as $T \rightarrow \infty$. Then, under Assumption A.4 in the Appendix, the estimated parameters in the quantile regression

$$\left[\hat{\beta}_0(\sigma_\mu^T; \tau), \hat{\beta}_1(\sigma_\mu^T; \tau) \right] = \arg \min_{(\beta_0, \beta_1) \in \mathbb{R}^2} \sum_{t=1}^T \rho_\tau(R_{m,t \rightarrow N} - \beta_0 - \beta_1 \tilde{Q}_{t,\tau}),$$

satisfy

$$\left[\hat{\beta}_0(\sigma_\mu^T; \tau), \hat{\beta}_1(\sigma_\mu^T; \tau) \right] = \left[0, e^{(\mu - r_f)N} \right] + o_p(1). \quad (3.5)$$

Furthermore, the quantile forecast based on the QR estimates satisfies

$$\hat{\beta}_0(\sigma_\mu^T; \tau) + \hat{\beta}_1(\sigma_\mu^T; \tau) \tilde{Q}_{t,\tau} = Q_{t,\tau} + o_p(1). \quad (3.6)$$

Proof. See Appendix A.3. ■

Proposition 3.1(i) shows that the risk-neutral quantile function is a good predictor of $Q_{t,\tau}(\sigma_t; \sigma_\mu)$ when σ_μ is small, and the difference between the two functions is governed by the unconditional equity premium $e^{(\mu - r_f)N}$. In this case, Proposition 3.1(ii) suggests that the QR estimates are almost constant across τ and close to $[0, e^{(\mu - r_f)N}]$. This result obtains without assuming that μ_t follows a normal distribution. The wedge between $Q_{t,\tau}(\sigma_t; \sigma_\mu)$ and $\tilde{Q}_{t,\tau}$ not explained by the equity premium can be attributed to uncertainty about μ_t , which increases the variance of the physical distribution. The assumption that σ_μ is small relative to σ_t accords with empirical findings of Martin (2017, Table I), who finds that $2.4\% \leq \sigma_\mu \leq 4.6\%$, whereas σ_t hovers around 20%. Unreported simulations show that the approximation in (3.5) obtains closely when the model is calibrated to match these stylized facts. As a result, the physical quantile forecast based on the QR estimates in (3.6) is also highly accurate.

Table 1 already indicates evidence against the lognormal model since the QR estimates in the left- and right-tail are rather different, in contradiction with (3.5). To further assess the implications of the lognormal model, I analyze the accuracy of the physical quantile forecast in (3.6) out-of-sample. Specifically, I use QR based on

the first t_0 observations to estimate the model

$$Q_{t,\tau}(R_{m,t \rightarrow N}) = \hat{\beta}_{0,t_0}(\tau) + \hat{\beta}_{1,t_0}(\tau)\tilde{Q}_{t,\tau}, \quad (3.7)$$

where the t_0 -subscript in β_{\cdot,t_0} refers to the fact that the coefficients are estimated using observations up to time t_0 . Using an expanding window to estimate β_{\cdot,t_0} , the model produces dynamic quantile forecasts of the form

$$\hat{Q}_{t,\tau}^{\text{logn}} = \hat{\beta}_{0,t}(\tau) + \hat{\beta}_{1,t}(\tau)\tilde{Q}_{t,\tau}. \quad (3.8)$$

In the lognormal case, Proposition 3.1(ii) suggests that $Q_{t,\tau}(R_{m,t \rightarrow N}) \approx \hat{Q}_{t,\tau}^{\text{logn}}$. This approximation can be tested using the joint restriction

$$H_0 : [\beta_0(\tau), \beta_1(\tau)] = [0, 1],$$

in the quantile regression

$$\min_{\beta_0, \beta_1 \in \mathbb{R}} \sum_t \rho_\tau \left(R_{m,t \rightarrow N} - \beta_0 - \beta_1 \hat{Q}_{t,\tau}^{\text{logn}} \right).$$

The results are summarized in Table 2 and show that the point estimates are quite far from the $[0, 1]$ benchmark. The Wald test on the joint restriction tends to reject H_0 far enough in the tail, but for $\tau = 0.2$ the null hypothesis is never rejected due to the large standard errors. Additionally, the $R^1(\tau)$ statistic shows that the explanatory power is low relative to Table 1, even though the sample sizes are different. Hence, the results are incompatible with (3.5) and (3.6) and provide evidence against the conditional lognormal assumption, which is in line with evidence from the literature (see e.g. [Martin \(2017, Result 4\)](#)).

4 Quantile Bound on the SDF Volatility and Model Implications

The previous section showed that the physical and risk-neutral distribution locally differ most in the left-tail. In this section, I show that these local differences imply that the SDF must be highly volatile; an observation that is closely related to the [Hansen and Jagannathan \(1991\)](#) bound.

4.1 Quantile Bound

Since this section deals with unconditional distributions, I omit the t subscript from the notation and write R to denote a generic return. For ease of notation, I also

Table 2: Expanding quantile prediction with risk-neutral quantile

	Horizon (in days)	$\widehat{\beta}_0(\tau)$	$\widehat{\beta}_1(\tau)$	Wald test (<i>p</i> -value)	$R^1(\tau)[\%]$	Obs
$\underline{\tau = 0.05}$	30	0.54 (0.186)	0.42 (0.194)	0.00	4.36	3804
	60	0.55 (0.293)	0.39 (0.314)	0.11	1.84	3753
	90	0.78 (0.335)	0.11 (0.353)	0.01	0.88	3702
$\underline{\tau = 0.1}$	30	0.59 (0.233)	0.39 (0.240)	0.03	2.39	3804
	60	0.80 (0.382)	0.16 (0.397)	0.06	0.22	3753
	90	0.74 (0.416)	0.21 (0.433)	0.09	1.34	3702
$\underline{\tau = 0.2}$	30	0.83 (0.422)	0.15 (0.429)	0.14	0.28	3804
	60	0.87 (0.508)	0.11 (0.517)	0.23	0.25	3753
	90	0.73 (0.528)	0.26 (0.534)	0.37	0.52	3702

Note: This table reports the QR estimates of (3.8) using an expanding window based on an initial 500 observations. The sample period is 2003-2021. *Wald test* denotes the *p*-value of the joint restriction $[\beta_0(\tau), \beta_1(\tau)] = [0, 1]$. Standard errors are reported in parentheses and calculated using the SETBB with a block length of 5 times the prediction horizon. $R^1(\tau)$ denotes the goodness of fit measure (2.5).

define $\phi(\tau) := F(\widetilde{Q}_\tau(R))$, which can be interpreted as the ordinal dominance curve of the measures \mathbb{P} and $\widetilde{\mathbb{P}}$ (Hsieh and Turnbull, 1996). Finally, let $\aleph^+ := \{M : M \geq 0 \text{ and } \mathbb{E}(MR) = 1\}$, the space of all nonnegative SDFs (Hansen and Jagannathan, 1991). The volatility bound on the SDF can now be stated as follows.

Theorem 4.1 (Quantile bound). *Assume no-arbitrage, then for any $M \in \aleph^+$, we have*

$$\frac{\sigma(M)}{\mathbb{E}[M]} \geq \frac{|\tau - \phi(\tau)|}{\sqrt{\phi(\tau)(1 - \phi(\tau))}} \quad \forall \tau \in (0, 1). \quad (4.1)$$

If a risk-free asset exists, then $\mathbb{E}[M] = 1/R_f$ and (4.1) simplifies to

$$\sigma(M) \geq \frac{|\tau - \phi(\tau)|}{\sqrt{\phi(\tau)(1 - \phi(\tau))} R_f} \quad \forall \tau \in (0, 1). \quad (4.2)$$

Proof. See Appendix A.4. ■

If $\mathbb{P} = \widetilde{\mathbb{P}}$, agents are risk-neutral and the dominance curve evaluates to $\phi(\tau) = \tau$. In that case the quantile bound degenerates to zero. Theorem 4.1 makes precise the sense in which any local difference between the physical and risk-neutral distribution

leads to a volatile SDF. Compare this to the classical HJ bound:

$$\sigma(M) \geq \frac{|\mathbb{E}(R) - R_f|}{\sigma(R)R_f}. \quad (4.3)$$

The lower bound in (4.3) shows that any excess return leads to a volatile SDF. Essentially, (4.3) uses three sources of information: (i) The mean of the physical distribution (ii) The mean of the risk-neutral distribution (iii) The variance of the physical distribution. The lower bound in (4.3) is also a global measure of distance between \mathbb{P} and $\tilde{\mathbb{P}}$, since the mean and volatility are averages across the whole distribution.

In contrast, the bound in (4.1) compares the physical and risk-neutral distribution for every τ -quantile, which is a *local* measure of distance between \mathbb{P} and $\tilde{\mathbb{P}}$. To clarify this local interpretation, I use a classical result of Hoeffding (see Appendix A.5):

$$\text{COV}[R, M] = - \int_0^\infty \text{COV}[\mathbb{1}(R \leq x), M] dx. \quad (4.4)$$

Equation (4.4) shows that $\text{COV}[\mathbb{1}(R \leq x), M]$ locally measures the dependence between the SDF and return. The average over all local measures is equal to the global measure of dependence given by $\text{COV}[R, M]$. HJ apply the Cauchy-Schwarz inequality to $\text{COV}[R, M]$, in order to derive their bound on the SDF volatility. In contrast, I apply Cauchy-Schwarz to the local measure of dependence, $\text{COV}[\mathbb{1}(R \leq \tilde{Q}_\tau), M]$. This local measure is expected to yield sharper bounds on the SDF volatility if, for example, there is high tail dependence between the SDF and asset return.¹⁰

If there is no priced jump or stochastic volatility risk, $\phi(\tau)$ is determined by the equity premium and the HJ bound in (4.3) captures all relevant information. On the other hand, if there are risk premia for jumps or stochastic volatility, the measures \mathbb{P} and $\tilde{\mathbb{P}}$ differ both in shape and location (Broadie et al., 2009), which implies that $\phi(\tau)$ contains information that is not captured by the HJ bound. Moreover, the quantile bound is robust to fat-tails, since it is well defined regardless of any moment restrictions on the returns. Fat-tails and other higher order shape restrictions such as negative skewness are essential features of financial return data (see, e.g. Gao and Martin (2021)).

¹⁰See McNeil et al. (2015, Chapter 7.2.4) for a formal definition of tail dependence.

4.2 Quantile Bound and HJ Bound in Asset Pricing Models

In this section, I discuss the quantile and HJ bound in (common) asset pricing models. These models make different predictions regarding the shape and strength of the quantile bound relative to the HJ bound. As such, the quantile bound can be used to diagnose model misspecification.

Example 4.1 (CAPM). The Capital Asset Pricing Model (CAPM) specifies the SDF as

$$M = \alpha - \beta R_m.$$

Here, R_m denotes the return on the market portfolio. In this case $M \notin \mathbb{N}^+$, since the SDF can become negative. However, this probability is very small over short time horizons or we can think of M as an approximation to $M^* := \max(0, M) \in \mathbb{N}^+$. Since the HJ bound is derived by applying the Cauchy-Schwarz inequality to $\text{COV}(R_m, M)$, the inequality binds if M is a linear combination of R_m . Hence, under CAPM, the HJ bound is strictly stronger than the quantile bound regardless of the distribution of R_m .

Example 4.2 (Joint normality). Suppose that M and R are jointly normally distributed and denote the mean and variance of R by μ_R and σ_R^2 respectively. The normality assumption violates no-arbitrage since M can be negative, but could be defended as an approximation over short time horizons when the variance is small (see Example 4.3). In Appendix A.5, I prove that

$$\left| \text{COV} \left(\mathbb{1} \left(R \leq \tilde{Q}_\tau \right), M \right) \right| = f_R(\tilde{Q}_\tau) |\text{COV}(R, M)|, \quad (4.5)$$

where $f_R(\cdot)$ is the marginal density of R .¹¹ This identity gives an explicit expression for the weighting factor in Hoeffding's identity (4.4). In Appendix A.5, I also derive an explicit expression for the relative efficiency between the local and HJ bound

$$\frac{\text{HJ bound}}{\text{quantile bound}} = \frac{\sqrt{\phi(\tau)(1 - \phi(\tau))}}{\sigma_R f_R(\tilde{Q}_\tau)}. \quad (4.6)$$

To see that the HJ bound is always stronger than the quantile bound, minimize (4.6) with respect to τ . Appendix A.6 shows that the minimizer τ^* satisfies $\tilde{Q}_{\tau^*} = \mu_R$. For this choice, $\phi(\tau^*) = \mathbb{P}(R \leq \tilde{Q}_{\tau^*}) = 1/2$ and $f_R(\tilde{Q}_{\tau^*}) = 1/\sqrt{2\pi\sigma_R^2}$. Therefore, (4.6) obeys the bound

$$\frac{\sqrt{\phi(\tau)(1 - \phi(\tau))}}{\sigma_R f(\tilde{Q}_\tau)} \geq \frac{\sqrt{2\pi}}{2} \approx 1.25.$$

Hence, the HJ bound is always stronger in a model where the SDF and return are

¹¹Notice that this is the marginal density under physical measure \mathbb{P} .

jointly normal.

Example 4.3 (Joint lognormality). Let Z_R and Z_M be standard normal random variables with correlation ρ and consider the specification

$$R = e^{(\mu_R - \frac{\sigma_R^2}{2})\lambda + \sigma_R \sqrt{\lambda} Z_R}$$

$$M = e^{-(r_f + \frac{\sigma_M^2}{2})\lambda + \sigma_M \sqrt{\lambda} Z_M},$$

where λ governs the time scale. Simple algebra shows that the no-arbitrage condition, $\mathbb{E}[RM] = 1$, is satisfied when $\mu_R - r_f = -\rho\sigma_R\sigma_M$. It is hard to find an analytical solution for the relative efficiency between the HJ and quantile bound in this case, but linearization leads to a closed form expression which is quite accurate in simulations. The details are described in Appendix A.7, where I show that

$$\min_{\tau \in (0,1)} \frac{\text{HJ bound}}{\text{quantile bound}} \approx \frac{1}{2} \sqrt{\frac{2\pi\sigma_R^2\lambda}{\exp(\sigma_R^2\lambda) - 1}}. \quad (4.7)$$

This expression is independent of μ_R . An application of l'Hôpital's rule reveals that the relative efficiency converges to $\sqrt{2\pi}/2$ if $\lambda \rightarrow 0^+$.¹² The ratio in (4.7) is less than 1 if $\sigma \geq 0.92$ and $\lambda = 1$. If $R = R_m$ (the market return), and the annualized market return volatility is about 16%, then the HJ bound is stronger than the quantile bound under any reasonable parameterization if the SDF and market return are lognormal.

Example 4.4 (Pareto distribution). I now give an example of a model where the quantile bound can be stronger than the HJ bound, due to the fat tails of the return distribution. Let $U \sim \mathbf{Unif}[0, 1]$ (Uniform distribution on $[0,1]$) and consider the following specification:

$$M = AU^\alpha, \quad R = BU^{-\beta} \quad \text{with} \quad \alpha, \beta, A, B > 0. \quad (4.8)$$

A random variable $X \sim \mathbf{Par}(C, \zeta)$ follows a Pareto distribution with scale parameter $C > 0$ and shape parameter $\zeta > 0$ if the CDF is given by

$$\mathbb{P}(X \leq x) = \begin{cases} 1 - (x/C)^{-\zeta} & x \geq C \\ 0 & x < C. \end{cases}$$

The assumption (4.8) implies that returns follow a Pareto distribution, both under the physical and risk-neutral measure. This fact allows me to obtain an explicit expression for the quantile bound. I summarize these properties in the Proposition below.

¹²This is the same relative efficiency in Example 4.2, which is unsurprising as the linearization becomes exact in the limit as $\lambda \rightarrow 0^+$.

Proposition 4.2. Let the SDF and return be given by (4.8). Then,

- (i) Under \mathbb{P} , the distribution of returns is Pareto: $R \sim \text{Par}\left(B, \frac{1}{\beta}\right)$.
- (ii) Under $\tilde{\mathbb{P}}$, the distribution of returns is Pareto: $R \sim \text{Par}\left(B, \frac{\alpha+1}{\beta}\right)$.

(iii) The Sharpe ratio on the asset return is given by

$$\frac{\mathbb{E}[R] - R_f}{\sigma(R)} = \frac{\frac{B}{1-\beta} - \frac{\alpha+1}{A}}{\sqrt{\frac{B^2}{1-2\beta} - \left(\frac{B}{1-\beta}\right)^2}}. \quad (4.9)$$

(iv) The quantile bound is given by

$$\frac{1}{R_f} \frac{|\tau - \phi(\tau)|}{\sqrt{\phi(\tau)(1 - \phi(\tau))}} = \frac{A}{1 + \alpha} \frac{\left| \tau - 1 + (1 - \tau)^{\frac{1}{\alpha+1}} \right|}{\sqrt{(1 - (1 - \tau)^{\frac{1}{\alpha+1}})(1 - \tau)^{\frac{1}{\alpha+1}}}}.$$

(v) If $\beta \nearrow \frac{1}{2}$, the HJ bound converges to 0. ■

Proof. See Appendix A.8.

Proposition 4.2(iv) shows that the quantile bound is independent of the Pareto tail index β . Properties (iv) and (v) provide some intuition when the quantile bound is stronger than the HJ bound. Namely, heavier tails of the distribution of R (as measured by β) lead to a lower Sharpe ratio. However, the quantile bound is unaffected by β since it only depends on the tail index α . Therefore, when β gets close to $1/2$, the HJ bound is rather uninformative whereas the quantile bound may fare better. Moreover, no additional restrictions on the parameter space are necessary to calculate the quantile bound, while the HJ bound requires $\beta < 1/2$.¹³

Figure 2 shows two instances of the quantile and HJ bound using different parameter calibrations. Both calibrations are targeted to match an equity premium of 8% and risk-free rate of 0%, but in Panel 2b, the distribution of returns has a fatter tail compared to Panel 2a. In both calibrations, the quantile bound has a range of values for which it is stronger than the HJ bound. In line with Proposition 4.2, we see that the range is larger in Panel (b), since the HJ bound is less informative owing to the heavier tails of R . However, the quantile bound attains its maximum in the right-tail since that is the region where the physical and risk-neutral measure differ most. This result is inconsistent with the empirical results from Table 1, which indicate that the physical and risk-neutral measure are nearly identical in the right-tail.

¹³The latter restriction is not unreasonable for asset returns, since typical tail index estimates suggest $\beta \in [1/4, 1/3]$ (Danielsson and de Vries, 2000).

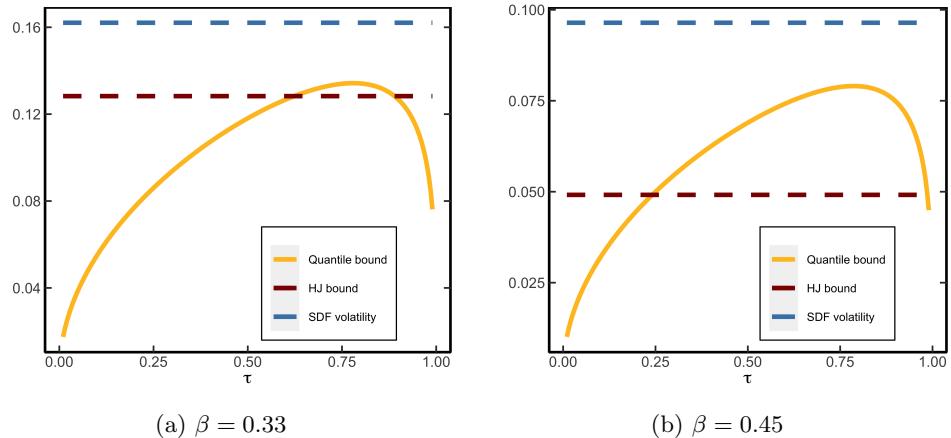


Figure 2: HJ and Quantile bound for heavy tailed returns. Both panels plot the Quantile bound, HJ bound and true SDF volatility for the Pareto model (4.8). In Panel (b), the distribution of returns has a fatter tail compared to Panel (a). Panel (a) uses the parameters $[A, \alpha, B, \beta] = [1.19, 0.19, 0.72, 0.33]$. Panel (b) uses the parameters $[A, \alpha, B, \beta] = [1.11, 0.11, 0.59, 0.45]$. Both calibrations imply an equity premium of 8% and (net) risk-free rate of 0%.

Example 2.1 (Continued). The disaster risk model discussed above is calibrated according to the results in [Backus et al. \(2011, Table II\)](#). The market return in this model is considered as a levered claim on consumption growth, i.e. an asset that pays dividends proportional to G_{t+1}^λ . Here λ governs the variability of the claim to equity. I convert the model implied volatility bounds to monthly units, to facilitate the comparison with the empirical bounds obtained in Section 4.4.

The quantile bound, HJ bound and SDF volatility are illustrated in Figure 3a (without jumps) and Figure 3b (with jumps). The quantile bound with jumps has a sharp peak at $\tau = 0.037$, after which it decreases monotonically. Interestingly, there is a range of τ values for which the quantile bound is sharper than the HJ bound. This result can be understood from the physical and risk-neutral distribution in Figure 1a. The risk-neutral distribution displays a heavy left-tail, owing to the implied disaster risk embedded in the SDF. As a result, it is extremely profitable to sell digital put options which pay out in case of a disaster. These put options must have high Sharpe ratios as their prices are high (insurance against disaster risk), but the actual probability of disaster is low enough that the risk associated to selling such insurance is limited.

4.3 Estimating the Quantile Bound Empirically

To empirically estimate the quantile bound (4.2), I use the same 30-day S&P500 returns as discussed in Section 2.3. However, in this case, I use non-overlapping re-

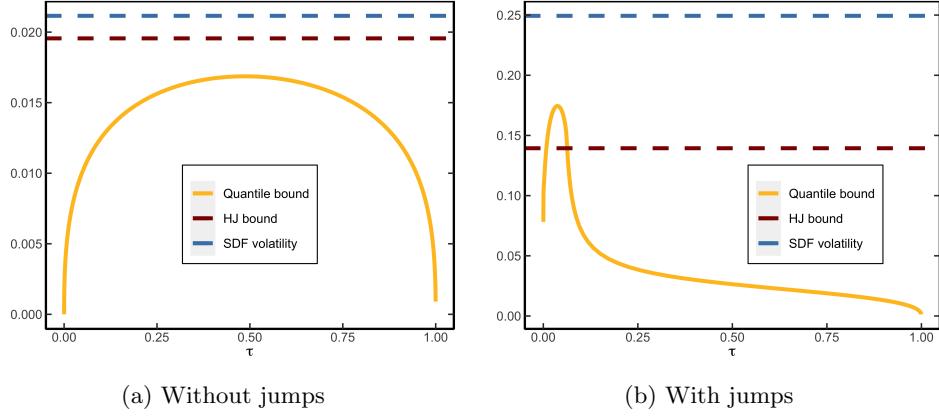


Figure 3: HJ and Quantile bound in disaster risk model without and with jumps. Panels (a) and (b) show the HJ and Quantile bound for the disaster risk model (Example 2.1) without and with jumps respectively. The bounds and true SDF volatility are reported in monthly units.

turns that cover the period 1996–2021.¹⁴ These returns are sampled at the middle of each month, resulting in a total of 312 observations. Over this period, the Sharpe ratio is 13%, and the HJ bound therefore implies that the monthly SDF is quite volatile.

The quantile bound consists of three unknowns that need to be estimated: (i) the physical distribution F ; (ii) the risk-neutral quantile function \tilde{Q}_τ , and; (iii) the risk-free rate R_f . To estimate the unconditional risk-free rate, denoted by \hat{R}_f , I rely on the historical average of monthly interest rates. Next, to obtain an estimate of the physical distribution, I employ a kernel (CDF) estimator, given by:

$$\hat{F}(x) := \frac{1}{T} \sum_{t=1}^T \Phi \left(\frac{x - R_{m,t \rightarrow N}}{h} \right), \quad (4.10)$$

where $\Phi(\cdot)$ is the Epanechnikov kernel and h is the bandwidth determined by cross-validation. This choice of estimator ensures that the quantile bound is a smooth function of τ , which reduces the impact of outliers relative to the discontinuous empirical CDF.

Moving on, I apply the procedure outlined in Section 2.3 to estimate \tilde{F}_t (the conditional risk-neutral CDF). Subsequently, I average the conditional distributions

¹⁴I use non-overlapping returns in this section to facilitate testing and to make the results comparable to other nonparametric bounds, which are typically estimated based on non-overlapping returns (see e.g. Liu (2021)).

to estimate the unconditional CDF:

$$\hat{\tilde{F}}(x) := \frac{1}{T} \sum_{t=1}^T \tilde{F}_t(x).$$

Under appropriate assumptions about the distribution of returns, $\hat{\tilde{F}}$ converges to \tilde{F} as $T \rightarrow \infty$. An estimate of the unconditional risk-neutral quantile function is then obtained from

$$\hat{\tilde{Q}}(\tau) := \inf \left\{ x \in \mathbb{R} : \tau \leq \hat{\tilde{F}}(x) \right\}. \quad (4.11)$$

Finally, based on the physical CDF (4.10) and risk-neutral quantile function (4.11), I estimate the quantile bound by

$$\hat{\theta}(\tau) := \frac{|\tau - \hat{\phi}(\tau)|}{\sqrt{\hat{\phi}(\tau)(1 - \hat{\phi}(\tau))\hat{R}_f}}, \quad \tau \in [\varepsilon, 1 - \varepsilon] \subseteq (0, 1), \quad (4.12)$$

where $\hat{\phi}(\tau) := \hat{F}(\hat{\tilde{Q}}(\tau))$ is the estimated ordinal dominance curve (ODC) and ε is a small positive number.

4.4 Quantile Bound for S&P500 Returns

Figure 4a illustrates that the estimated physical and risk-neutral measures differ most in the left-tail. The quantile bound shows that this difference leads to a volatile SDF, which is shown in Figure 4b. The lower bound on the SDF volatility implied by the quantile bound is much stronger than the HJ bound in the left-tail. This finding aligns with empirical evidence, which documents that high Sharpe ratios can be attained by selling out-of-the money put options (see Broadie et al. (2009) and the references therein). The supremum of the quantile bound occurs around the 5th percentile, implying that the monthly SDF volatility must exceed 31%. This value is more than twice the level indicated by the sample HJ bound.¹⁵ Moreover, the shape of the quantile bound is quite similar to the quantile bound implied by the disaster risk model in Figure 3b.

The graphical evidence suggests that the quantile bound renders a stronger bound on the SDF volatility than the HJ bound. To test this hypothesis more formally, I fix a priori the probability level at 0.037 ($\tau = 0.037$), which renders the sharpest bound on the SDF volatility in the disaster risk model (Example 2.1). At this probability level, the quantile bound is 26% in the data, which is roughly double the level implied

¹⁵The non-monotonicity in the right-tail of the quantile bound occurs because $\tilde{F}(x) > F(x)$, for x large enough. That is, the physical distribution does not first-order stochastically dominate the risk-neutral distribution. This result is consistent with the negative $\overline{H_t}$ estimates in Table 1.

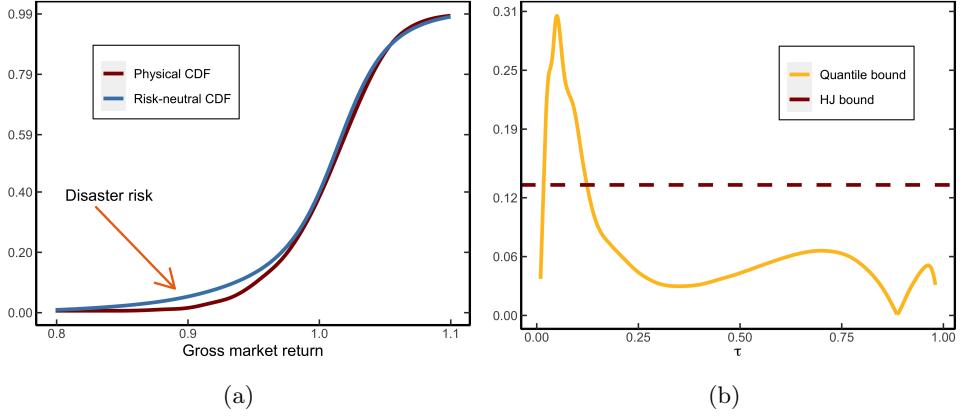


Figure 4: **Physical/risk-neutral CDF and quantile bound for monthly S&P500 returns.** Panel (a) shows the unconditional physical and risk-neutral CDF for monthly S&P500 returns, over the period 1996-2021. Panel (b) shows the quantile bound as function of τ , together with the HJ bound.

by the HJ bound.

To see whether this difference is statistically significant, I consider the following test statistic

$$\mathcal{T} := \widehat{\theta}(0.037) - \frac{|\bar{R}_m - \widehat{R}_f|}{\widehat{\sigma} \widehat{R}_f}. \quad (4.13)$$

The first term on the right denotes the estimated quantile bound (4.12) evaluated at the 3.7th percentile, using the entire time series of returns $\{R_{m,t \rightarrow N}\}$. The second term denotes the estimated HJ bound, using \bar{R}_m and $\widehat{\sigma}$ as the respective sample mean and standard deviation of $\{R_{m,t \rightarrow N}\}$. A value of $\mathcal{T} > 0$ indicates that the quantile bound is stronger than the HJ bound. To test this restriction, consider the null and alternative hypothesis:

$$H_0 : \mathcal{T} \leq 0 \quad (4.14)$$

$$H_1 : \mathcal{T} > 0.$$

Since the distribution of (4.13) is hard to characterize, I use stationary bootstrap to approximate the p -value under the null hypothesis. The stationary bootstrap is used to generate time indices from which we recreate (with replacement) bootstrapped returns $\{R_{m,t \rightarrow N}^*\}$ (Politis and Romano, 1994). The same bootstrapped time indices are used to re-estimate the physical CDF and risk-neutral quantile function. I repeat the bootstrap exercise 100,000 times and for each bootstrap sample, I calculate the test statistic \mathcal{T}^* . Finally, the empirical p -value is obtained as the fraction of times $\mathcal{T}^* \leq 0$. The last column in Table 3 shows that the p -value is 7.5%, which provides preliminary evidence that the quantile bound significantly exceeds the HJ bound.

Table 3: **Sample bounds and bootstrap result**

Sample size	HJ bound	quantile bound	<i>p</i> -value
312	0.133	0.260	0.075

Note: This table reports the HJ and quantile bound for monthly S&P500 returns over the period 1996–2021. The quantile bound is evaluated at $\tau = 0.0374$. The final column denotes the *p*-value of the null hypothesis in (4.14). The *p*-value is obtained from 100,000 bootstrap samples and counts the fraction of times that $\mathcal{T}^* \leq 0$.

When the HJ bound is stronger than the quantile bound, many of the bootstrap samples may not include disaster shocks. Over the entire sample period, there are only two instances where returns were less than -20%: in September 2008 and February 2020. When considering bootstrap samples that include both of these months, the *p*-value is only 3.6%. In contrast, the *p*-value increases to 22% for bootstrap samples that exclude these months. These findings underscore the sensitivity of the test to the presence of disaster shocks. Overall, the results suggest that, unconditionally, the SDF needs to be highly volatile to be consistent with local differences between the physical and risk-neutral measure in the left-tail. This finding is consistent with the disaster risk model that incorporates jumps in Figure 3b.

5 Lower Bound on Physical Quantile and Dark Matter

The findings in Section 2 and 4 show that the risk-neutral quantile function is not a good approximation to the physical quantile function in the left-tail. In this section, I derive a risk-adjustment term that provides a more accurate approximation of the conditional quantile function, $Q_{t,\tau}$, for small values of τ . The latent quantile function in the left-tail is of particular interest since it serves as a direct measure of *conditional disaster risk*.¹⁶

5.1 A Lower Bound on the Quantile Difference

To analyze the difference between $Q_{t,\tau}$ and $\tilde{Q}_{t,\tau}$, I use some elementary tools from functional analysis. The quantile function can be regarded as a map φ between normed spaces, taking as input a distribution function and returning the quantile function: $\varphi(F_t) = F_t^{-1} = Q_{t,\tau}$. Expanding φ around the observed risk-neutral CDF yields

$$Q_{t,\tau} - \tilde{Q}_{t,\tau} = \varphi(F_t) - \varphi(\tilde{F}_t) = \varphi'_{\tilde{F}_t}(F_t - \tilde{F}_t) + o\left(\|F_t - \tilde{F}_t\|\right), \quad (5.1)$$

¹⁶In risk management, $Q_{t,\tau}$ is known as the Value-at-Risk.

where $\|\cdot\|$ is a norm on a suitable linear space¹⁷ and $\varphi'_{\tilde{F}_t}(F_t - \tilde{F}_t)$ is the Gâteaux derivative of φ at \tilde{F}_t in the direction of F_t :

$$\begin{aligned}\varphi'_{\tilde{F}_t}(F_t - \tilde{F}_t) &:= \lim_{\lambda \downarrow 0} \frac{\varphi[(1-\lambda)\tilde{F}_t + \lambda F_t]}{\lambda} \\ &= \left. \frac{\partial}{\partial \lambda} \varphi((1-\lambda)\tilde{F}_t + \lambda F_t) \right|_{\lambda=0}.\end{aligned}\quad (5.2)$$

Heuristically, the Gâteaux derivative can be thought of as measuring the change in the quantile function when we move the risk-neutral distribution in the direction of the physical distribution. Appendix A.9 shows that the Gâteaux derivative is given by

$$\varphi'_{\tilde{F}_t}(F_t - \tilde{F}_t) = \frac{\tau - F_t(\tilde{Q}_{t,\tau})}{\tilde{f}_t(\tilde{Q}_{t,\tau})} = \frac{\tau - \phi_t(\tau)}{\tilde{f}_t(\tilde{Q}_{t,\tau})}, \quad (5.3)$$

where $\phi_t(\tau) = F_t(\tilde{Q}_{t,\tau})$ denotes the conditional ODC. I proceed under the working hypothesis that the remainder term in (5.1) is “small” in the sup-norm, $\|g\|_\infty = \sup_x |g(x)|$.

Assumption 5.1. *The remainder term in (5.1) can be neglected.*

Remark. The assumption implies that the first order approximation in (5.1) is accurate. The condition that $\|F_t - \tilde{F}_t\|_\infty$ is small can be understood as excluding near-arbitrage opportunities, since the quantile bound in Theorem 4.1 shows that substantial pointwise differences between $F_t(\cdot)$ and $\tilde{F}_t(\cdot)$ lead to a very volatile SDF. Appendix E.3 illustrates the approximation in the Black and Scholes (1973) model.

I combine (5.1) and (5.3) in conjunction with Assumption 5.1 to obtain the approximation

$$Q_{t,\tau} \approx \tilde{Q}_{t,\tau} + \underbrace{\frac{\tau - F_t(\tilde{Q}_{t,\tau})}{\tilde{f}_t(\tilde{Q}_{t,\tau})}}_{\text{risk-adjustment}}. \quad (5.4)$$

The second term on the right can be thought of as a risk-adjustment term to capture the unobserved wedge between $Q_{t,\tau}$ and $\tilde{Q}_{t,\tau}$. The approximation in (5.4) contains the terms $\tilde{Q}_{t,\tau}$ and $\tilde{f}_t(\tilde{Q}_{t,\tau})$, which are directly observed at time t using the Breeden and Litzenberger (1978) formula in (2.7). However, $F_t(\cdot)$ is unknown and hence (5.4) cannot be used directly to approximate $Q_{t,\tau}$.

To make further progress, I show that the numerator term, $\tau - F_t(\tilde{Q}_{t,\tau})$, can be bounded with option data under certain economic constraints. Chabi-Yo and Loudis

¹⁷Formally, the space can be defined as $\{\Delta : \Delta = c(F - G), F, G \in \mathbb{D}, c \in \mathbb{R}\}$ and \mathbb{D} is the space of distribution functions (Serfling, 2009). See van der Vaart (2000, Section 20.1) and Serfling (2009, p. 217) for further details about the approximation.

(2020) show that in representative agent models, the SDF is a function of the market return:

$$\frac{\mathbb{E}_t[M_{t \rightarrow N}]}{M_{t \rightarrow N}} = \frac{\frac{u'(W_t x_0)}{u'(W_t x)}}{\tilde{\mathbb{E}}_t \left[\frac{u'(W_t x_0)}{u'(W_t x)} \right]} \quad \text{with } x = R_{m,t \rightarrow N} \text{ and } x_0 = R_{f,t \rightarrow N}, \quad (5.5)$$

where W_t is the agent's wealth at time t and $u(x)$ represents the agent's utility function. Define

$$\zeta(x) := \frac{u'(W_t R_{f,t \rightarrow N})}{u'(W_t x)} \quad \text{and} \quad \theta_k = \frac{1}{k!} \left(\frac{\partial^k \zeta(x)}{\partial x^k} \right)_{x=R_{f,t \rightarrow N}}. \quad (5.6)$$

Notice that $\zeta(\cdot)$ is simply the inverse of the marginal rate of substitution (MRS) and θ_k are the coefficients of its Taylor expansion around $R_{f,t \rightarrow N}$. I make the following assumptions about the market return and the MRS of the representative agent.

Assumption 5.2. *In the representative agent model, it holds that (i) $\tilde{\mathbb{E}}_t[R_{m,t \rightarrow N}^3] < \infty$; and (ii) $\zeta^{(4)}(x) \leq 0$.*

Part (i) allows for fat tails in the risk-neutral distribution as long as the third moment exists. This assumption relaxes the implicit assumption made by Chabi-Yo and Loudis (2020) that infinitely many moments exist. Figure F7 in the Appendix illustrates that the risk-neutral distribution frequently exhibits a finite number of moments, some of which may not exceed 4, particularly in turbulent market conditions. Chabi-Yo and Loudis (2020) present sufficient conditions for part (ii) to hold, which relate to the sign of the fifth derivative of the utility function of the representative agent. Specifically, for common utility functions such as CRRA or HARA utility, parameter restrictions are needed to ensure that (ii) holds.¹⁸

Under Assumption 5.2, the difference between the physical and risk-neutral distribution in the left-tail can be bounded as follows.

Theorem 5.3 (Infeasible Lower Bound). *Let Assumption 5.2 hold and assume that the risk-neutral CDF is absolutely continuous with respect to Lebesgue measure. Define τ^* so that $G(\tilde{Q}_{t,\tau^*}) = \tilde{\mathbb{E}}_t(G(R_{m,t \rightarrow N}))$, where*

$$G(R_{m,t \rightarrow N}) := \int_{R_{f,t \rightarrow N}}^{R_{m,t \rightarrow N}} \zeta^{(4)}(t)(R_{m,t \rightarrow N} - t)^3 dt.$$

¹⁸For example, for CRRA utility, the risk-aversion coefficient cannot be too large. See Appendix C for a detailed discussion.

Then for all $\tau \leq \tau^*$,

$$\tau - F_t(\tilde{Q}_{t,\tau}) \geq \frac{\sum_{k=1}^3 \theta_k (\tau \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)} - \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)}[\tilde{Q}_{t,\tau}])}{1 + \sum_{k=1}^3 \theta_k \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)}}, \quad (5.7)$$

where

$$\begin{aligned} \tilde{\mathbb{M}}_{t \rightarrow N}^{(n)} &:= \tilde{\mathbb{E}}_t [(R_{m,t \rightarrow N} - R_{f,t \rightarrow N})^n] \\ \tilde{\mathbb{M}}_{t \rightarrow N}^{(n)}[k_0] &:= \tilde{\mathbb{E}}_t [\mathbb{1}(R_{m,t \rightarrow N} \leq k_0) (R_{m,t \rightarrow N} - R_{f,t \rightarrow N})^n]. \end{aligned} \quad (5.8)$$

Proof. See Appendix A.10. ■

Remark. The condition that $\tau \leq \tau^*$ is sufficient but not necessary, as the proof of Theorem 5.3 shows. In practice τ^* is unknown since $G(\cdot)$ depends on the unknown utility function of the representative agent. Appendix C.5 shows that $\tau^* \approx 0.5$ in the data for CRRA utility and different levels of risk-aversion. In light of this result, it seems that $\tau \in \{0.05, 0.1, 0.2\}$ is sufficiently conservative for the lower bound to hold, and I use these values in the empirical application in Section 5.3.

The (un)truncated moments in (5.8) can be computed from option prices (see Appendix A.11), but the bound in (5.7) is still infeasible since θ_k is unknown.¹⁹ However, Chabi-Yo and Loudis (2020) show that $\{\theta_k\}_{k=1}^3$ relate to the coefficient of relative risk-aversion, relative prudence and relative temperance of the representative agent. Based on this observation and using results from the expected utility literature (Eckhoudt and Schlesinger, 2006), the authors propose an additional restriction on θ_k that allows me to obtain a feasible lower bound.

Corollary 5.4 (Feasible lower bound). *Suppose the assumptions of Theorem 5.3 hold. In addition, suppose that (i) $\theta_k \geq \frac{1}{R_{f,t \rightarrow N}^k}$ for $k = 1, 3$ and $\theta_2 \leq \frac{-1}{R_{f,t \rightarrow N}^2}$; (ii) $\tilde{\mathbb{M}}_{t \rightarrow N}^{(3)} \leq 0$. Then,*

$$\tau - F_t(\tilde{Q}_{t,\tau}) \geq \frac{\sum_{k=1}^3 \frac{(-1)^{k+1}}{R_{f,t \rightarrow N}^k} (\tau \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)} - \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)}[\tilde{Q}_{t,\tau}])}{1 + \sum_{k=1}^3 \frac{(-1)^{k+1}}{R_{f,t \rightarrow N}^k} \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)}} =: \text{LB}_{t,\tau}, \quad (5.9)$$

for all τ such that $\tilde{Q}_{t,\tau} \leq \min \left(R_{f,t \rightarrow N} - \sqrt{\widetilde{\text{VAR}}_t(R_{m,t \rightarrow N})}, \tilde{Q}_{t,\tau^*} \right)$, where \tilde{Q}_{t,τ^*} is defined in Theorem 5.3.

Proof. See Appendix A.10. ■

Assumption (i) is inspired by the expected utility framework (Chabi-Yo and Loudis, 2020, Result 5). Chabi-Yo and Loudis (2020, Table 6) provide empirical

¹⁹In Appendix D, I use comparative statics for common utility functions to analyze the tail difference between the physical and risk-neutral distribution.

evidence that (i) is tight when estimating the conditional equity premium. Part (ii) is a very mild restriction on risk-neutral skewness, which is almost always negative at every date and time horizon. This empirical fact is well known.²⁰ Furthermore, the bound only holds for quantiles far enough in the left-tail. Compared to Theorem 5.3, the additional condition needed for the bound to hold is that $\tilde{Q}_{t,\tau} \leq R_{f,t \rightarrow N} - \sqrt{\widetilde{\text{VAR}}_t(R_{m,t \rightarrow N})}$, which covers a wide range of quantiles in the left-tail, since in the data $\sqrt{\widetilde{\text{VAR}}_t(R_{m,t \rightarrow N})}$ is in the order of 10^{-3} for 90-day returns, whereas the risk-free rate is typically around 1.²¹

The bound in Corollary 5.4 does not require any parameter estimation and can be calculated solely based on time t information. Corollary 5.4 therefore complements the recent literature on the recovery of beliefs. Ross (2015) shows that one can recover F_t , if the pricing kernel is transition independent. Subsequent work (Borovička et al., 2016; Qin et al., 2018; Jackwerth and Menner, 2020) casts doubt on the transition independence assumption and shows that recovery is generally impossible. Complimentary to these results, Corollary 5.4 shows that one can still establish a lower bound on the left-tail of the physical distribution, using a different set of (mild) economic constraints. Moreover, Section 2.3 showed that the right-tail of F_t can be recovered from the risk-neutral distribution due to the near absence of risk-adjustment. Both results combined show that approximate recovery of F_t using option prices might still possible.

Based on the “Hit” column in Table 1 and stochastic dominance, I concluded that the risk-neutral quantile underestimates the physical quantile in the left-tail. The following Corollary shows how the lower bound on the physical distribution can be used to bound the wedge between the physical and risk-neutral quantile function.

Corollary 5.5 (Feasible quantile bound). *Under Assumption 5.1 and the same assumptions of Corollary 5.4, it follows that*

$$Q_{t,\tau} - \tilde{Q}_{t,\tau} \geq \underbrace{\frac{\text{LB}_{t,\tau}}{f_t(\tilde{Q}_{t,\tau})}}_{\text{risk-adjustment}} =: \text{RA}_{t,\tau}. \quad (5.10)$$

Proof. By Assumption 5.1, the approximation in (5.4) holds, which in combination

²⁰Chabi-Yo and Loudis (2020) argue that all odd risk-neutral moments should be negative, since they expose the investor to unfavorable market conditions.

²¹At the 30- and 60-day horizon, the risk-neutral standard deviation is even smaller.

with Corollary 5.4 renders

$$\begin{aligned} Q_{t,\tau} - \tilde{Q}_{t,\tau} &\stackrel{(5.4)}{\approx} \frac{\tau - F_t(\tilde{Q}_{t,\tau})}{\tilde{f}_t(\tilde{Q}_{t,\tau})} \\ &\geq \frac{1}{\tilde{f}_t(\tilde{Q}_{t,\tau})} \left(\frac{\sum_{k=1}^3 \frac{(-1)^{k+1}}{R_{f,t \rightarrow N}^k} (\tau \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)} - \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)}[\tilde{Q}_{t,\tau}])}{1 + \sum_{k=1}^3 \frac{(-1)^{k+1}}{R_{f,t \rightarrow N}^k} \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)}} \right). \end{aligned} \quad (5.11)$$

■

Corollary 5.5 establishes a lower bound on the difference between the physical and risk-neutral quantile function, which can be thought of as a risk-adjustment term. If the lower bound in (5.10) happens to be tight, it gives us a direct measure of conditional disaster risk. In addition, the term on the right of (5.10) is not subject to the historical sample bias critique of [Welch and Goyal \(2008\)](#) and it is available at a daily frequency. The motivation for introducing disaster risk is to explain the historically high equity premium ([Rietz, 1988](#); [Barro, 2006](#)). However, a high (conditional) equity premium does not necessarily arise due to disaster risk and, moreover, calibrating these models is difficult due to the lack of disasters in the data ([Martin, 2013](#)). For these reasons, it is of interest to approximate $Q_{t,\tau}$ in the left-tail using a model-free approach. The next Sections show that the assumption of a tight lower bound in (5.10) cannot be rejected, which allows me to obtain a model-free measure of disaster risk.

5.2 Computation of the Lower Bound

Before assessing how tight the lower bound is in Corollary 5.5, I first outline the procedure to calculate $\text{LB}_{t,\tau}$ and $\tilde{f}_t(\tilde{Q}_{t,\tau})$. Since $\frac{d}{d\tau} \tilde{Q}_t(\tau) = 1/f_t(\tilde{Q}_{t,\tau})$, I approximate the denominator term in (5.10) by

$$\frac{1}{\tilde{f}_t(\tilde{Q}_{t,\tau})} \approx \frac{\tilde{Q}_t(\tau + h) - \tilde{Q}_t(\tau - h)}{2h},$$

where h is the bandwidth of the τ -grid.²² Second, to calculate $\text{LB}_{t,\tau}$ from (5.9), I use the estimated quantile curve $\tilde{Q}_{t,\tau}$ in combination with the formula for high order risk-neutral moments in Appendix A.11.

Since virtually all macro-finance models assume that the market return is negatively correlated with the SDF, one expects the market return to have a higher probability of a crash under risk-neutral measure than under physical measure. This

²²I slightly abuse notation to emphasize that the derivative is taken w.r.t. τ , so that $\tilde{Q}_t(\tau + h)$ denotes $\tilde{Q}_{t,\tau+h}$.

means that Corollary 5.5 has non trivial content in the data if $\text{LB}_{t,\tau} \geq 0$. I confirm that $\text{LB}_{t,\tau} \geq 0$ for all dates considered, using the same τ 's from Table 5 below. Appendix Table E1 contains summary statistics of $\text{RA}_{t,\tau}$, which show that the risk-adjustment term is right-skewed, more pronounced in the right-tail and economically meaningful in magnitude, with outliers that can spike up to 29%.

5.3 Tightness of the Lower Bound: In-sample Evidence

To test whether the lower bound in Corollary 5.5 is tight, I form *excess quantile returns*: $R_{m,t \rightarrow N} - \tilde{Q}_{t,\tau}$. Since $\tilde{Q}_{t,\tau}$ is observed at time t , we have $Q_{t,\tau}(R_{m,t \rightarrow N} - \tilde{Q}_{t,\tau}) = Q_{t,\tau}(R_{m,t \rightarrow N}) - \tilde{Q}_{t,\tau}$. Subsequently, I use QR to estimate the model

$$\begin{aligned} Q_{t,\tau}(R_{m,t \rightarrow N}) - \tilde{Q}_{t,\tau}(R_{m,t \rightarrow N}) &= \beta_0(\tau) + \beta_1(\tau)\text{RA}_{t,\tau}, \\ [\hat{\beta}_0(\tau), \hat{\beta}_1(\tau)] &= \arg \min_{(\beta_0, \beta_1) \in \mathbb{R}^2} \sum_{t=1}^T \rho_\tau(R_{m,t \rightarrow N} - \tilde{Q}_{t,\tau} - \beta_0 - \beta_1\text{RA}_{t,\tau}). \end{aligned} \quad (5.12)$$

Regression (5.12) is similar to the excess return regressions of [Welch and Goyal \(2008\)](#). Under the null hypothesis that the lower bound is tight, we have

$$H_0 : [\beta_0(\tau), \beta_1(\tau)] = [0, 1]. \quad (5.13)$$

Less restrictive, one can test whether $\beta_0(\tau) = 0$ and $\beta_1(\tau) > 0$, which implies that the statistical “factor” $\text{RA}_{t,\tau}$ explains the conditional quantile wedge.²³

Table 4 shows the result of regression (5.12). The null hypothesis of a tight lower bound cannot be rejected for $\tau = 0.2$, but is rejected for $\tau \in \{0.05, 0.1\}$ at all horizons. In case the null hypothesis is rejected, the $\beta_1(\tau)$ -coefficient is larger than 1, which is consistent with the theory that $\text{RA}_{t,\tau}$ represents a lower bound on the difference between the physical and risk-neutral distribution. Table 4 also documents p -values of the joint hypothesis in (5.13), which again is not rejected for $\tau = 0.2$ only. In all cases, the lower bound is economically meaningful since $\beta_1(\tau)$ is significantly different from 0, while $\beta_0(\tau) = 0$ can never be rejected. To shed more light on the explanatory power of the lower bound, I use the $R^1(\tau)$ measure-of-fit

$$R^1(\tau) = 1 - \frac{\min_{b_0, b_1} \sum \rho_\tau(R_{m,t \rightarrow N} - b_0 - b_1\text{RA}_{t,\tau})}{\min_{b_0} \sum \rho_\tau(R_{m,t \rightarrow N} - b_0)}. \quad (5.14)$$

Table 4 shows that the explanatory power of $\text{RA}_{t,\tau}$ is modest, generally around a few

²³For example, if we start with a quantile factor model $Q_{t,\tau} = \tilde{Q}_{t,\tau} + \beta(\tau)\text{RA}_{t,\tau}$, the model has one testable implication for the data: the intercept in a quantile regression of $R_{m,t \rightarrow N} - \tilde{Q}_{t,\tau}$ on $\text{RA}_{t,\tau}$ should be zero. Quantile factor models have recently been proposed by [Chen et al. \(2021\)](#).

percent. Low explanatory power is typical in the equity premium literature, since high R^2 -values lead to high Sharpe ratios, as market timing strategies can be designed to exploit return predictability (Ross, 2005; Campbell and Thompson, 2008). The same logic applies in the quantile case, since the quantile bound in Theorem 4.1 shows that large pointwise differences between the physical and risk-neutral distribution lead to near arbitrage opportunities. If one could predict this difference, one could again design a trading strategy that exploits the predictability, this time using options rather than a direct timing investment in the market portfolio.

Table 4: **Quantile regression lower bound**

	Horizon (in days)	$\hat{\beta}_0(\tau)$	$\hat{\beta}_1(\tau)$	Wald test (<i>p</i> -value)	$R^1(\tau)[\%]$	Obs
$\underline{\tau = 0.05}$	30	-0.01 (0.006)	4.43 (0.350)	0.00	6.03	4333
	60	-0.01 (0.018)	5.53 (0.717)	0.00	3.60	4312
	90	-0.02 (0.040)	6.37 (1.385)	0.00	4.91	4291
$\underline{\tau = 0.1}$	30	-0.01 (0.006)	2.17 (0.420)	0.02	3.18	4333
	60	-0.02 (0.014)	3.25 (0.602)	0.00	2.23	4312
	90	-0.02 (0.024)	3.05 (0.703)	0.00	4.43	4291
$\underline{\tau = 0.2}$	30	-0.01 (0.006)	1.33 (0.418)	0.03	0.41	4333
	60	-0.02 (0.013)	1.50 (0.506)	0.49	0.48	4312
	90	-0.02 (0.022)	1.36 (0.694)	0.76	1.46	4291

Note: This table reports the QR estimates of (5.12) over the sample period 2003–2021 at different horizons, using overlapping returns. Standard errors are shown in parentheses and calculated using SETBB with a block length of maturity times 5. Wald test denotes the *p*-value of the joint restriction $[\beta_0(\tau), \beta_1(\tau)] = [0, 1]$. $R^1(\tau)$ denotes the goodness-of-fit measure (5.14).

Another way to test whether the lower bound in Corollary 5.5 represents a meaningful risk-adjustment term is to test directly whether the following is a good proxy for the latent conditional quantile function

$$\widehat{Q}_{t,\tau} = \widetilde{Q}_{t,\tau} + \text{RA}_{t,\tau}. \quad (5.15)$$

I use QR to estimate the model

$$Q_{t,\tau}(R_{m,t \rightarrow N}) = \beta_0(\tau) + \beta_1(\tau)\widehat{Q}_{t,\tau}. \quad (5.16)$$

If (5.15) is a good predictor of the conditional quantile function, we have under the null hypothesis

$$H_0 : \beta_0(\tau) = 0, \quad \beta_1(\tau) = 1. \quad (5.17)$$

Table 5 summarizes the estimates of (5.16) for several quantiles. The results uniformly improve upon the risk-neutral estimates in Table 1. First, the point estimates for $[\beta_0(\tau), \beta_1(\tau)]$ are closer to the $[0, 1]$ benchmark. Second, the Wald test on the joint restriction in (5.17) is never rejected and third, the in-sample explanatory power is higher. The same conclusion applies when comparing the predictive results to the expanding quantile regression from Table 2. These results suggest that $\hat{Q}_{t,\tau}$ can be regarded as a good proxy for the latent conditional quantile function $Q_{t,\tau}$ in the left-tail.

Two additional remarks are in order. First, since we estimate $\hat{Q}_{t,\tau}$ from option prices, there might be a concern for attenuation bias due to measurement error (Angrist et al., 2006). Appendix E.4 provides simulation evidence which shows that attenuation bias is negligible in a setup that mimics the empirical application. Second, there is a possibility of quantile crossing, which means that the predicted quantiles, $\hat{\beta}_0(\tau) + \hat{\beta}_1(\tau)\text{RA}_{t,\tau}$, are not monotone with respect to τ . This problem frequently arises in dynamic quantile models (Gouriéroux and Jasiak, 2008). It appears however, not a concern in our case, since crossing occurs only 0.04% of the time for the 30 day horizon. For the other horizons, crossing happens about 0.1% of the time.²⁴

5.4 Tightness of the Lower Bound: Out-of-sample Evidence

Since the in-sample results for the physical quantile approximation suggest $\beta_0(\tau) = 0$ and $\beta_1(\tau) = 1$, it is natural to test how well this works out-of-sample by means of predicting $Q_{t,\tau}(R_{m,t \rightarrow N})$ directly with $\hat{Q}_{t,\tau}$, which does not require any parameter estimation.

In order to interpret $\hat{Q}_{t,\tau}$ as a valid approximation to the latent quantile function, we need to ensure that $\hat{Q}_{t,\tau}$ is not subject to the quantile crossing problem. I verify that crossing does not occur for any prediction horizon and quantile level. To assess the out-of-sample performance further, I use the following out-of-sample metric

$$R_{oos}^1(\tau) = 1 - \frac{\sum \rho_\tau(R_{m,t \rightarrow N} - \hat{Q}_{t,\tau})}{\sum \rho_\tau(R_{m,t \rightarrow N} - \bar{Q}_{t,\tau})}. \quad (5.18)$$

The out-of-sample $R_{oos}^1(\tau)$ is also displayed in Table 5. The predictor variable $\hat{Q}_{t,\tau}$

²⁴Recall that the explanatory variable in the estimation, $\text{RA}_{t,\tau}$, changes with the quantile level τ . Hence, based on the estimated coefficients $[\hat{\beta}_0(\tau), \hat{\beta}_1(\tau)]$, we cannot tell whether the estimated quantile is monotone in τ .

Table 5: **Risk-adjusted quantile regression**

	Horizon (in days)	$\hat{\beta}_0(\tau)$	$\hat{\beta}_1(\tau)$	Wald test (<i>p</i> -value)	$R^1(\tau)[\%]$	$R_{oos}^1(\tau)[\%]$	Obs
$\underline{\tau = 0.05}$	30	0.29 (0.283)	0.70 (0.301)	0.13	6.28	9.94	4333
	60	0.30 (0.434)	0.71 (0.484)	0.06	3.40	17.81	4312
	90	0.36 (0.594)	0.64 (0.687)	0.10	4.26	21.98	4291
$\underline{\tau = 0.1}$	30	0.28 (0.310)	0.72 (0.322)	0.27	3.57	4.02	4333
	60	0.38 (0.444)	0.61 (0.472)	0.22	2.35	9.22	4312
	90	0.31 (0.613)	0.70 (0.664)	0.13	4.19	13.22	4291
$\underline{\tau = 0.2}$	30	0.57 (0.499)	0.43 (0.507)	0.47	0.58	2.53	4333
	60	0.44 (0.617)	0.56 (0.630)	0.40	0.57	4.28	4312
	90	0.23 (0.760)	0.78 (0.774)	0.56	0.70	5.99	4291

Note: This table reports the QR estimates of (5.16) over the sample period 2003-2021. Standard errors are shown in parentheses and calculated using the SETBB, with block length equal to 5 times the maturity and 1,000 Monte Carlo bootstrap samples. Wald test gives the *p*-value of the Wald test on the joint restriction: $\hat{\beta}_0(\tau) = 0, \hat{\beta}_1(\tau) = 1$. $R^1(\tau)$ denotes the in-sample goodness-of-fit criterion (2.5). $R_{oos}^1(\tau)$ is the out-of-sample goodness-of fit (5.18), using a rolling window size of $10 \times$ maturity.

improves upon the historical rolling quantile out-of-sample in all cases. In particular, this outperformance is most pronounced at the 5th percentile, which is expected since option data are known to provide useful information about extreme downfalls in the stock market (Bates, 2008; Bollerslev and Todorov, 2011). In Appendix E.2, I run a battery of robustness tests which show that, out-of-sample, $RA_{t,\tau}$ better predicts the quantile than other benchmarks such as the risk-neutral quantile or the VIX index. The latter result is particularly encouraging since the VIX predictor uses in-sample information.

5.5 Time Varying Disaster Risk and Dark Matter

The in- and out-of-sample results show that $\hat{Q}_{t,\tau}$ is a good proxy for the latent conditional quantile function. Time fluctuation in $\hat{Q}_{t,\tau}$ for small τ can therefore be interpreted as time varying disaster risk. The left panels of Figure 5 show the evolution of $\hat{Q}_{t,\tau}$ over time for the 30 and 60 day horizon, with $\tau = 0.05$. The time fluctuation in both series is evident from the graph and provides empirical support for the thesis that disaster risk is time-varying, as in the models of Gabaix (2012), Wachter (2013), Isoré and Szczerbowicz (2017), Farhi and Gourio (2018) and Seo

and Wachter (2019). According to Martin (2013), estimating disaster risk models is challenging, particularly over short time horizons. Thus, the appeal of using $\hat{Q}_{t,\tau}$ as a measure of disaster risk lies in its model-free nature.

Ross (2015) refers to the impact that changes in perceived disaster probabilities can have on asset prices as *dark matter*: “It is unseen and not directly observable but it exerts a force that can change over time and that can profoundly influence markets”. We can illuminate this dark matter somewhat by studying the lowest quantile forecasts in Figure 5, which are produced over the 4th quarter of 2008 and the Covid crisis. During these periods, the quantile forecasts drop below 72%, which suggests that a loss of -28% or more had an expected probability of 5%. To put things in perspective, it happened only once since 1926 that the S&P500 index recorded a monthly loss of -28% or more.²⁵ Hence, based on historical estimates, a back of the envelope calculation puts the probability of a loss of -28% or more at 1/1139, which is 57 times lower than 5%.

The right panels of Figure 5 give another view on this dark matter, as they show the evolution of risk-adjustment over time. The largest spikes occur once more at the height of the financial crisis and the difference between the physical and risk-neutral measure can be as large as 25% (30 day horizon) or 16% (60 day horizon). This difference suggests that the risk-neutral quantile decreases disproportionately more than the physical quantile during crises.

It is well known that risk-neutral quantiles are smaller than historical quantiles in the left-tail of the distribution. However, this fact alone does not tell us about the market’s forecast of a decline, since historical probabilities can be different from the market forecast or the risk-neutral distribution differs significantly from the physical distribution, due to a risk premium on disaster insurance. The discussion above shows that both effects are at play, but the right panels of Figure 5 suggest that the insurance effect is more dominant during a crisis, since the physical and risk-neutral quantile are further apart. The equity decomposition in (3.1) shows that this leads to an increase in the conditional equity premium, and the quantile bound implies that the conditional SDF must be highly volatile in such periods.

²⁵I use historical monthly SP500 return from WRDS that are available from January 1926 and renders a total of 1139 observations.

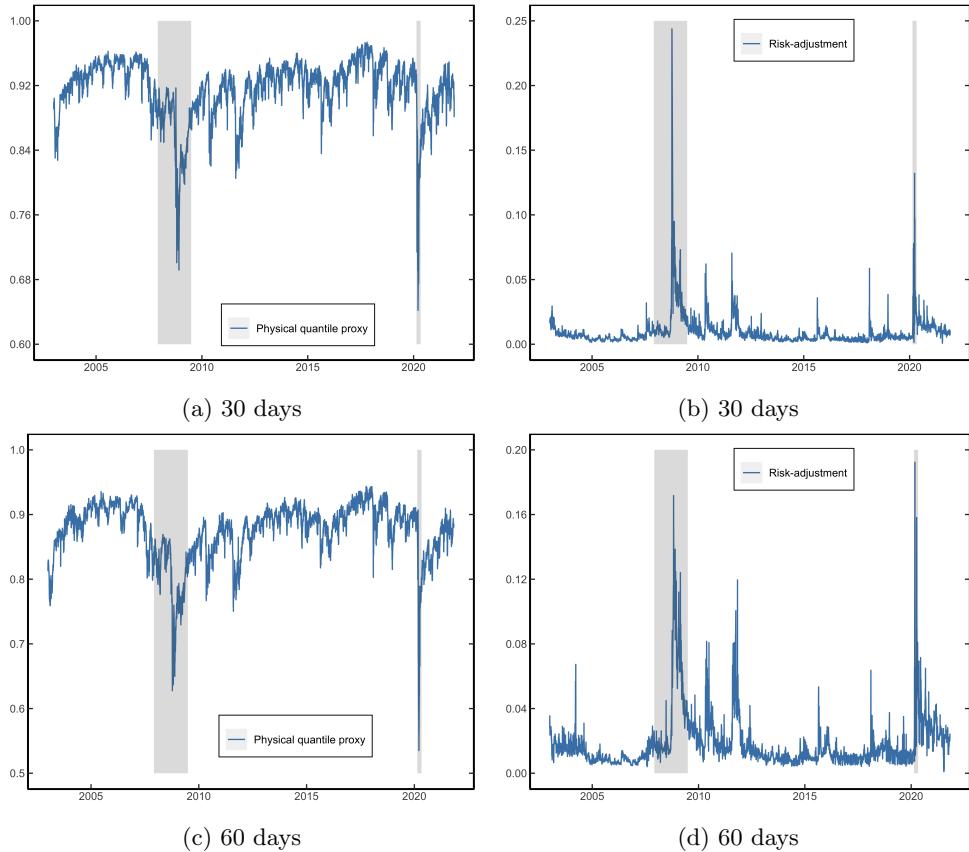


Figure 5: Time variation in risk-adjustment at the 5th percentile. The left panels show the real time quantile predictor $\hat{Q}_{t,\tau}$ from (5.15) for $\tau = 0.05$. The right panels show the risk-adjustment term $RA_{t,\tau}$ from (5.10). The two shaded bars signify the Great Recession period (Dec 2007 – June 2009) and COVID-19 crisis (Feb 2020 – April 2020).

6 Conclusion

I use return and option data on the S&P500 in combination with quantile regression to obtain a local measure of distance between the conditional physical and risk-neutral distribution. In the data, the difference between both distributions is most pronounced in the left-tail, whereas the right tails are nearly indistinguishable.

To build on this finding, I introduce a quantile bound on the SDF volatility, which demonstrates that the volatility of the SDF is largely driven by local differences in the left-tail of the physical and risk-neutral measure. Compared to the closely related HJ bound, the quantile bound is significantly stronger in the data. An asset pricing model seeking to explain these stylized facts must therefore satisfy two requirements: (i) produce physical and risk-neutral distributions that are similar in the right-tail but different in the left-tail, and (ii) generate a quantile bound that is stronger than

the HJ bound in the left-tail. These requirements pose a challenge to asset pricing models that do not embed a jump component in the physical and risk-neutral measure.

Finally, I suggest using a risk-adjustment term observed from option prices to measure the real time difference between the physical and risk-neutral quantile in the left-tail. This risk-adjustment term serves as a good predictor of the quantile wedge, exhibiting spikes during crises and significant fluctuations over time, providing model-free evidence of time-varying disaster risk.

References

- Yacine Aït-Sahalia and Andrew W. Lo. Nonparametric estimation of state-price densities implicit in financial asset prices. *Journal of Finance*, 53(2):499–547, 1998. doi:[10.1111/0022-1082.215228](https://doi.org/10.1111/0022-1082.215228).
- Yacine Aït-Sahalia and Andrew W. Lo. Nonparametric risk management and implied risk aversion. *Journal of Econometrics*, 94(1):9–51, 2000. doi:[10.1016/S0304-4076\(99\)00016-0](https://doi.org/10.1016/S0304-4076(99)00016-0).
- Caio Almeida and René Garcia. Assessing misspecified asset pricing models with empirical likelihood estimators. *Journal of Econometrics*, 170(2):519–537, 2012. doi:[10.1016/j.jeconom.2012.05.020](https://doi.org/10.1016/j.jeconom.2012.05.020).
- 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. doi:[10.1111/j.1468-0262.2005.00643.x](https://doi.org/10.1111/j.1468-0262.2005.00643.x).
- Joshua Angrist, Victor Chernozhukov, and Iván Fernández-Val. Quantile regression under misspecification, with an application to the U.S. wage structure. *Econometrica*, 74(2):539–563, 2006. doi:[10.1111/j.1468-0262.2006.00671.x](https://doi.org/10.1111/j.1468-0262.2006.00671.x).
- David Backus, Mikhail Chernov, and Ian Martin. Disasters implied by equity index options. *Journal of Finance*, 66(6):1969–2012, 2011. doi:[10.1111/j.1540-6261.2011.01697.x](https://doi.org/10.1111/j.1540-6261.2011.01697.x).
- David Backus, Mikhail Chernov, and Stanley Zin. Sources of entropy in representative agent models. *Journal of Finance*, 69(1):51–99, 2014. doi:[10.1111/jofi.12090](https://doi.org/10.1111/jofi.12090).
- Gurdip Bakshi, Dilip Madan, and George Panayotov. Returns of claims on the upside and the viability of U-shaped pricing kernels. *Journal of Financial Economics*, 97(1):130–154, 2010. doi:[10.1016/j.jfineco.2010.03.009](https://doi.org/10.1016/j.jfineco.2010.03.009).

- Gurdip Bakshi, Fousseni Chabi-Yo, and Xiaohui Gao. A recovery that we can trust? Deducing and testing the restrictions of the recovery theorem. *Review of Financial Studies*, 31(2):532–555, 2018. doi:[10.1093/rfs/hhx108](https://doi.org/10.1093/rfs/hhx108).
- Ravi Bansal and Bruce N. Lehmann. Growth-optimal portfolio restrictions on asset pricing models. *Macroeconomic Dynamics*, 1(2):333–354, 1997. doi:[10.1017/S1365100597003039](https://doi.org/10.1017/S1365100597003039).
- Andrea Barletta and Paolo Santucci de Magistris. Analyzing the risks embedded in option prices with *rndfittool*. *Risks*, 6(2):1–15, 2018. doi:[10.3390/risks6020028](https://doi.org/10.3390/risks6020028).
- Robert J. Barro. Rare disasters and asset markets in the twentieth century. *Quarterly Journal of Economics*, 121(3):823–866, 2006. doi:[10.1162/qjec.121.3.823](https://doi.org/10.1162/qjec.121.3.823).
- David S. Bates. The crash of '87: was it expected? The evidence from options markets. *Journal of Finance*, 46(3):1009–1044, 1991. doi:[10.1111/j.1540-6261.1991.tb03775.x](https://doi.org/10.1111/j.1540-6261.1991.tb03775.x).
- David S. Bates. Post-'87 crash fears in the S&P500 futures option market. *Journal of Econometrics*, 94(1):181–238, 2000. doi:[10.1016/S0304-4076\(99\)00021-4](https://doi.org/10.1016/S0304-4076(99)00021-4).
- David S. Bates. The market for crash risk. *Journal of Economic Dynamics and Control*, 32(7):2291–2321, 2008. doi:[10.1016/j.jedc.2007.09.020](https://doi.org/10.1016/j.jedc.2007.09.020).
- Brendan K. Beare and Lawrence D. W. Schmidt. An empirical test of pricing kernel monotonicity. *Journal of Applied Econometrics*, 31(2):338–356, 2016. doi:[10.1002/jae.2422](https://doi.org/10.1002/jae.2422).
- Geert Bekaert and Marie Hoerova. The VIX, the variance premium and stock market volatility. *Journal of Econometrics*, 183(2):181–192, 2014. doi:[10.1016/j.jeconom.2014.05.008](https://doi.org/10.1016/j.jeconom.2014.05.008).
- Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973. doi:[10.1086/260062](https://doi.org/10.1086/260062).
- Tim Bollerslev and Viktor Todorov. Tails, fears, and risk premia. *Journal of Finance*, 66(6):2165–2211, 2011. doi:[10.1111/j.1540-6261.2011.01695.x](https://doi.org/10.1111/j.1540-6261.2011.01695.x).
- Jaroslav Borovička, Lars Peter Hansen, and José A. Scheinkman. Misspecified recovery. *Journal of Finance*, 71(6):2493–2544, 2016. doi:[10.1111/jofi.12404](https://doi.org/10.1111/jofi.12404).
- Douglas T. Breeden and Robert H. Litzenberger. Prices of state-contingent claims implicit in option prices. *Journal of Business*, 51(4):621–651, 1978. doi:[10.1086/296025](https://doi.org/10.1086/296025).

Mark Broadie, Mikhail Chernov, and Michael Johannes. Understanding index option returns. *Review of Financial Studies*, 22(11):4493–4529, 2009. doi:[10.1093/rfs/hhp032](https://doi.org/10.1093/rfs/hhp032).

John Y. Campbell and Samuel B. Thompson. Predicting excess stock returns out of sample: Can anything beat the historical average? *Review of Financial Studies*, 21(4):1509–1531, 2008. doi:[10.1093/rfs/hhm055](https://doi.org/10.1093/rfs/hhm055).

George Casella and Roger L. Berger. *Statistical Inference*. Duxbury/Thomson Learning. Belmont, CA, 2002.

Fousseni Chabi-Yo and Johnathan Loudis. The conditional expected market return. *Journal of Financial Economics*, 137(3):752–786, 2020. doi:[10.1016/j.jfineco.2020.03.009](https://doi.org/10.1016/j.jfineco.2020.03.009).

Liang Chen, Juan J. Dolado, and Jesús Gonzalo. Quantile factor models. *Econometrica*, 89(2):875–910, 2021. doi:[10.3982/ECTA15746](https://doi.org/10.3982/ECTA15746).

Victor Chernozhukov, Iván Fernández-Val, and Siyi Luo. Distribution regression with sample selection, with an application to wage decompositions in the UK, 2018. URL <https://arxiv.org/abs/1811.11603>.

John H. Cochrane. *Asset Pricing*. Princeton University Press, 2005.

Joshua D. Coval and Tyler Shumway. Expected option returns. *Journal of Finance*, 56(3):983–1009, 2001. doi:[10.1111/0022-1082.00352](https://doi.org/10.1111/0022-1082.00352).

Horatio Cuesdeanu and Jens Carsten Jackwerth. The pricing kernel puzzle in forward looking data. *Review of Derivatives Research*, 21(3):253–276, 2018. doi:[10.1007/s11147-017-9140-8](https://doi.org/10.1007/s11147-017-9140-8).

Jon Danielsson and Casper G. de Vries. Value-at-risk and extreme returns. *Annales d'Économie et de Statistique*, (60):239–270, 2000. doi:[10.2307/20076262](https://doi.org/10.2307/20076262).

Louis Eeckhoudt and Harris Schlesinger. Putting risk in its proper place. *American Economic Review*, 96(1):280–289, 2006. doi:[10.1257/000282806776157777](https://doi.org/10.1257/000282806776157777).

Robert F. Engle and Simone Manganelli. CAViaR: Conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics*, 22(4):367–381, 2004. doi:[10.1198/073500104000000370](https://doi.org/10.1198/073500104000000370).

Emmanuel Farhi and François Gourio. Accounting for macro-finance trends: Market power, intangibles, and risk premia. *Brookings Papers on Economic Activity*, (2):147–250, 2018. doi:[10.1353/eca.2018.0024](https://doi.org/10.1353/eca.2018.0024).

- Stephen Figlewski. Estimating the implied risk-neutral density for the US market portfolio. In *Volatility and Time Series Econometrics: Essays in Honor of Robert Engle*, chapter 15, pages 323–353. Oxford University Press, 03 2010. doi:[10.1093/acprof:oso/9780199549498.003.0015](https://doi.org/10.1093/acprof:oso/9780199549498.003.0015).
- Damir Filipović, Eberhard Mayerhofer, and Paul Schneider. Density approximations for multivariate affine jump-diffusion processes. *Journal of Econometrics*, 176(2):93–111, 2013. doi:[10.1016/j.jeconom.2012.12.003](https://doi.org/10.1016/j.jeconom.2012.12.003).
- Xavier Gabaix. Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance. *Quarterly Journal of Economics*, 127(2):645–700, 2012. doi:[10.1093/qje/qjs001](https://doi.org/10.1093/qje/qjs001).
- Can Gao and Ian Martin. Volatility, valuation ratios, and bubbles: An empirical measure of market sentiment. *Journal of Finance*, 76(6):3211–3254, 2021. doi:[10.1111/jofi.13068](https://doi.org/10.1111/jofi.13068).
- Christian Gouriéroux and Joann Jasiak. Dynamic quantile models. *Journal of Econometrics*, 147(1):198–205, 2008. doi:[10.1016/j.jeconom.2008.09.028](https://doi.org/10.1016/j.jeconom.2008.09.028).
- Karl B. Gregory, Soumendra N. Lahiri, and Daniel J. Nordman. A smooth block bootstrap for quantile regression with time series. *Annals of Statistics*, 46(3):1138–1166, 2018. doi:[10.1214/17-AOS1580](https://doi.org/10.1214/17-AOS1580).
- Lars Peter Hansen and Robert J. Hodrick. Forward exchange rates as optimal predictors of future spot rates: An econometric analysis. *Journal of Political Economy*, 88(5):829–853, 1980. doi:[10.1086/260910](https://doi.org/10.1086/260910).
- 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. doi:[10.1086/261749](https://doi.org/10.1086/261749).
- Fushing Hsieh and Bruce W. Turnbull. Nonparametric and semiparametric estimation of the receiver operating characteristic curve. *Annals of Statistics*, 24(1):25–40, 1996. doi:[10.1214/aos/1033066197](https://doi.org/10.1214/aos/1033066197).
- Elton P. Hsu and S. R. Srinivasa Varadhan. *Probability Theory and Applications*. American Mathematical Society, 1999.
- Marlène Isoré and Urszula Szczerbowicz. Disaster risk and preference shifts in a New Keynesian model. *Journal of Economic Dynamics and Control*, 79:97–125, 2017. doi:[10.1016/j.jedc.2017.04.001](https://doi.org/10.1016/j.jedc.2017.04.001).
- Jens Carsten Jackwerth. Recovering risk aversion from option prices and realized returns. *Review of Financial Studies*, 13(2):433–451, 2000. doi:[10.1093/rfs/13.2.433](https://doi.org/10.1093/rfs/13.2.433).

- Jens Carsten Jackwerth and Marco Menner. Does the Ross recovery theorem work empirically? *Journal of Financial Economics*, 137(3):723–739, 2020. doi:[10.1016/j.jfineco.2020.03.006](https://doi.org/10.1016/j.jfineco.2020.03.006).
- Roger Koenker and Gilbert Bassett. Regression quantiles. *Econometrica*, 46(1): 33–50, 1978. doi:[10.2307/1913643](https://doi.org/10.2307/1913643).
- Roger Koenker and José A. F. Machado. Goodness of fit and related inference processes for quantile regression. *Journal of the American Statistical Association*, 94(448):1296–1310, 1999. doi:[10.1080/01621459.1999.10473882](https://doi.org/10.1080/01621459.1999.10473882).
- Roger W. Lee. The moment formula for implied volatility at extreme strikes. *Mathematical Finance*, 14(3):469–480, 2004. doi:[10.1111/j.0960-1627.2004.00200.x](https://doi.org/10.1111/j.0960-1627.2004.00200.x).
- Erich L. Lehmann. Some concepts of dependence. *Annals of Mathematical Statistics*, 37(5):1137–1153, 1966. doi:[10.1214/aoms/1177699260](https://doi.org/10.1214/aoms/1177699260).
- Matthew Linn, Sophie Shive, and Tyler Shumway. Pricing kernel monotonicity and conditional information. *Review of Financial Studies*, 31(2):493–531, 2018. doi:[10.1093/rfs/hhx095](https://doi.org/10.1093/rfs/hhx095).
- Yan Liu. Index option returns and generalized entropy bounds. *Journal of Financial Economics*, 139(3):1015–1036, 2021. doi:[10.1016/j.jfineco.2020.08.011](https://doi.org/10.1016/j.jfineco.2020.08.011).
- Ian Martin. Consumption-based asset pricing with higher cumulants. *Review of Economic Studies*, 80(2):745–773, 2013. doi:[10.1093/restud/rds029](https://doi.org/10.1093/restud/rds029).
- Ian Martin. What is the expected return on the market? *Quarterly Journal of Economics*, 132(1):367–433, 2017. doi:[10.1093/qje/qjw034](https://doi.org/10.1093/qje/qjw034).
- Ian Martin and Christian Wagner. What is the expected return on a stock? *Journal of Finance*, 74(4):1887–1929, 2019. doi:[10.1111/jofi.12778](https://doi.org/10.1111/jofi.12778).
- Alexander J. McNeil, Rüdiger Frey, and Paul Embrechts. *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press, 2015.
- Dimitris N. Politis and Joseph P. Romano. The stationary bootstrap. *Journal of the American Statistical Association*, 89(428):1303–1313, 1994. doi:[10.1080/01621459.1994.10476870](https://doi.org/10.1080/01621459.1994.10476870).
- Likuan Qin and Vadim Linetsky. Long-term risk: A martingale approach. *Econometrica*, 85(1):299–312, 2017. doi:[10.3982/ECTA13438](https://doi.org/10.3982/ECTA13438).
- Likuan Qin, Vadim Linetsky, and Yutian Nie. Long forward probabilities, recovery, and the term structure of bond risk premiums. *Review of Financial Studies*, 31(12):4863–4883, 2018. doi:[10.1093/rfs/hhy042](https://doi.org/10.1093/rfs/hhy042).

- Thomas A. Rietz. The equity risk premium a solution. *Journal of Monetary Economics*, 22(1):117–131, 1988. doi:[10.1016/0304-3932\(88\)90172-9](https://doi.org/10.1016/0304-3932(88)90172-9).
- Joshua V. Rosenberg and Robert F. Engle. Empirical pricing kernels. *Journal of Financial Economics*, 64(3):341–372, 2002. doi:[10.1016/S0304-405X\(02\)00128-9](https://doi.org/10.1016/S0304-405X(02)00128-9).
- Stephen A. Ross. *Neoclassical Finance*. Princeton University Press, 2005.
- Stephen A. Ross. The recovery theorem. *Journal of Finance*, 70(2):615–648, 2015. doi:[10.1111/jofi.12092](https://doi.org/10.1111/jofi.12092).
- Paul Schneider and Fabio Trojani. (Almost) model-free recovery. *Journal of Finance*, 74(1):323–370, 2019. doi:[10.1111/jofi.12737](https://doi.org/10.1111/jofi.12737).
- Sang Byung Seo and Jessica A. Wachter. Option prices in a model with stochastic disaster risk. *Management Science*, 65(8):3449–3469, 2019. doi:[10.1287/mnsc.2017.2978](https://doi.org/10.1287/mnsc.2017.2978).
- Robert J. Serfling. *Approximation Theorems of Mathematical Statistics*. John Wiley & Sons, 2009.
- Karl N. Snow. Diagnosing asset pricing models using the distribution of asset returns. *Journal of Finance*, 46(3):955–983, 1991. doi:[10.1111/j.1540-6261.1991.tb03773.x](https://doi.org/10.1111/j.1540-6261.1991.tb03773.x).
- Michael Stutzer. A Bayesian approach to diagnosis of asset pricing models. *Journal of Econometrics*, 68(2):367–397, 1995. doi:[10.1016/0304-4076\(94\)01656-K](https://doi.org/10.1016/0304-4076(94)01656-K).
- Engin A. Sungur. Dependence information in parameterized copulas. *Communications in Statistics - Simulation and Computation*, 19(4):1339–1360, 1990. doi:[10.1080/03610919008812920](https://doi.org/10.1080/03610919008812920).
- Aad W. van der Vaart. *Asymptotic Statistics*. Cambridge University Press, 2000.
- Jessica A. Wachter. Can time-varying risk of rare disasters explain aggregate stock market volatility? *Journal of Finance*, 68(3):987–1035, 2013. doi:[10.1111/jofi.12018](https://doi.org/10.1111/jofi.12018).
- Ivo Welch and Amit Goyal. A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies*, 21(4):1455–1508, 2008. doi:[10.1093/rfs/hhm014](https://doi.org/10.1093/rfs/hhm014).

A Proofs

This section contains proofs and detailed calculations of results used in the main paper.

A.1 Decomposing the Equity Premium

For any atomless integrable random variable X with CDF $F(\cdot)$ and quantile function $Q = F^{-1}$, we have

$$\mathbb{E}(X) = \int_{\mathbb{R}} x dF(x) = \int_0^1 Q(\tau) d\tau.$$

The second identity holds by the change of variables formula for the Lebesgue-Stieltjes integral. In case F has a density, the formula follows from a simple substitution $x \rightarrow Q(\tau)$. Hence,

$$\mathbb{E}_t [R_{m,t \rightarrow N}] - R_{f,t \rightarrow N} = \mathbb{E}_t [R_{m,t \rightarrow N}] - \tilde{\mathbb{E}}_t (R_{m,t \rightarrow N}) = \int_0^1 (Q_{t,\tau} - \tilde{Q}_{t,\tau}) d\tau.$$

A.2 Stochastic Dominance and Pricing Kernel Monotonicity

In this section I provide more details on the relation between stochastic dominance and pricing kernel monotonicity. To begin with, recall that the physical distribution is first-order stochastic dominant (FOSD) over the risk-neutral distribution if and only if $F_t(x) \leq \tilde{F}_t(x)$. The latter definition is equivalent to $F_t(\tilde{Q}_{t,\tau}) \leq \tau$ for all $\tau \in (0, 1)$, which follows from the substitution $x \rightarrow \tilde{Q}_{t,\tau}$.

To see the connection with pricing kernel monotonicity, we recall from [Beare and Schmidt \(2016\)](#) that pricing kernel monotonicity is equivalent to $\phi_t(\tau) := F_t(\tilde{Q}_{t,\tau})$ being a convex function for all τ .²⁶ Figure A1 shows two different ODCs; the blue line corresponds to a situation where FOSD holds and the pricing kernel is monotonic (hence convex), whereas the yellow line shows a scenario where FOSD does not hold and convexity automatically fails. The geometric argument for why non-monotonicity is implied by a failure of FOSD is conveyed by the figure: if FOSD fails, the yellow line must cross the 45-degree line for some $\tau \in (0, 1)$, which automatically implies that the ODC is non-convex since the ODC has to satisfy $\phi_t(1) = 1$, because the physical and risk-neutral measure are equivalent. The proposition below thus follows.

Proposition A.1. *If the pricing kernel is a monotonically decreasing function of the market return, the physical measure first-order stochastically dominates the risk-*

²⁶[Beare and Schmidt \(2016\)](#) actually consider the reverse function $\phi_t(\tau) = \tilde{F}_t(Q_{t,\tau})$, so that pricing kernel monotonicity is equivalent to $\phi_t(\cdot)$ being concave.

neutral measure. Conversely, a violation of FOSD implies a violation of pricing kernel monotonicity.

A violation of FOSD is puzzling from the viewpoint of expected utility maximization. In this framework, the SDF is given by $u'(R_{m,t \rightarrow N})/\mathbb{E}_t(u'(R_{m,t \rightarrow N}))$, where $u(\cdot)$ is a utility function and the initial endowment is normalized to one for simplicity. The following proposition shows that a sufficient (but not necessary) condition for FOSD to hold is that $u'(\cdot)$ is non-increasing; a rather ubiquitous assumption in asset pricing models.

Proposition A.2. *In the expected utility framework, a sufficient condition for the physical measure to first-order stochastically dominate the risk-neutral measure is that $u'(\cdot)$ is non-increasing.*

Proof. Using the SDF to change from physical to risk-neutral measure, it follows that FOSD is equivalent to

$$\begin{aligned} F_t(x) &\leq \tilde{F}_t(x) \\ \iff \mathbb{E}_t [\mathbb{1}(R_{m,t \rightarrow N} \leq x)] &\leq \mathbb{E}_t \left[\frac{u'(R_{m,t \rightarrow N})}{\mathbb{E}_t[u'(R_{m,t \rightarrow N})]} \mathbb{1}(R_{m,t \rightarrow N} \leq x) \right] \\ \iff 0 &\leq \text{COV}_t(\mathbb{1}(R_{m,t \rightarrow N} \leq x), u'(R_{m,t \rightarrow N})). \end{aligned}$$

By Lemma D.1, the covariance above is nonnegative if $u'(\cdot)$ is non-increasing. ■

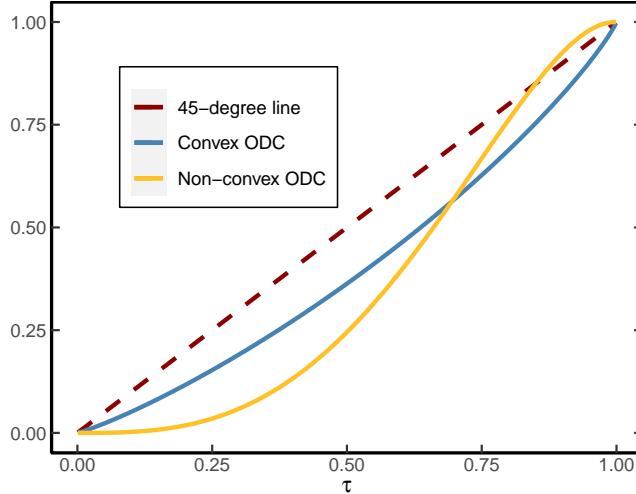


Figure A1: Ordinal dominance curve with and without first-order stochastic dominance. This figure shows two different ordinal dominance curves. The blue ODC corresponds to a situation where the physical measure FOSD the risk-neutral measure, whereas the yellow line shows a situation where FOSD fails.

A.3 Proof of Proposition 3.1

We separately show (i) and (ii) of Proposition 3.1. To prove these results, I use the following lemma.

Lemma A.3. *In the lognormal model, the physical and risk-neutral quantile functions conditional on μ_t, σ_t are given by, respectively*

$$Q_{t,\tau} = \exp \left[\left(\mu_t - \frac{1}{2} \sigma_t^2 \right) N + \sigma_t \sqrt{N} \Phi^{-1}(\tau) \right] \quad (\text{A.1})$$

$$\tilde{Q}_{t,\tau} = \exp \left[\left(r_f - \frac{1}{2} \sigma_t^2 \right) N + \sigma_t \sqrt{N} \Phi^{-1}(\tau) \right], \quad (\text{A.2})$$

where $\Phi^{-1}(\cdot)$ denotes the quantile function of the standard normal distribution. If $\mu_t \sim \mathcal{N}(\mu, \sigma_\mu^2)$ and independent from σ_t , the physical quantile function conditional on σ_t , but not μ_t , equals

$$Q_{t,\tau}(\sigma_t, \sigma_\mu) = \exp \left[\left(\mu - \frac{1}{2} \sigma_t^2 \right) N + \left(\sqrt{\sigma_\mu^2 N^2 + \sigma_t^2 N} \right) \Phi^{-1}(\tau) \right]. \quad (\text{A.3})$$

Proof. The quantile function of a random variable X such that $\log X \sim \mathcal{N}(a, b^2)$, is given by $\exp(a + b\Phi^{-1}(\tau))$. Therefore, the quantile functions conditional on μ_t, σ_t in (A.1) and (A.2) follow immediately from the conditional lognormal assumption. In (A.3), the function is conditioned on σ_t , but not μ_t . Since μ_t is assumed to be normally distributed and independent from σ_t , it follows that

$$\left(\mu_t - \frac{1}{2} \sigma_t^2 \right) N + \sigma_t \sqrt{N} Z_{t+N} | \sigma_t \sim \mathcal{N} \left(\left(\mu - \frac{1}{2} \sigma_t^2 \right) N, \sigma_\mu^2 N^2 + \sigma_t^2 N \right).$$

The expression in (A.3) can now be obtained again using the general formula of the lognormal quantile function. ■

Proof of Proposition 3.1(i). Recall that $\sqrt{a^2 + b^2} \leq \sqrt{a^2 + b^2 + 2ab} = a+b$, provided $a, b \geq 0$. This inequality shows that

$$\begin{aligned} & \exp \left[\left(\sqrt{\sigma_\mu^2 N^2 + \sigma_t^2 N} - \sigma_t \sqrt{N} \right) \Phi^{-1}(\tau) \right] \\ & \leq \exp \left[\left(\sqrt{\sigma_\mu^2 N^2 + \sigma_t^2 N} - \sigma_t \sqrt{N} \right) |\Phi^{-1}(\tau)| \right] \\ & \leq \exp \left(\sigma_\mu N |\Phi^{-1}(\tau)| \right) \\ & = 1 + \mathcal{O}(\sigma_\mu N), \end{aligned}$$

uniformly in $\tau \in \mathcal{I}$ and the support of σ_t . In combination with Lemma A.3, it follows

that

$$\begin{aligned} Q_{t,\tau}(\sigma_t, \sigma_\mu) &= \tilde{Q}_{t,\tau} e^{(\mu-r_f)N} \exp \left[\left(\sqrt{\sigma_\mu^2 N^2 + \sigma_t^2 N} - \sigma_t \sqrt{N} \right) \Phi^{-1}(\tau) \right] \\ &= \tilde{Q}_{t,\tau} e^{(\mu-r_f)N} (1 + \mathcal{O}(\sigma_\mu N)). \end{aligned}$$

■

In order to prove Proposition 3.1(ii), I need additional regularity conditions stated in Assumption A.4 below. The following notation for the quantile empirical process will be used:

$$\begin{aligned} L_{T,\tau}(\beta, \sigma_\mu) &:= \frac{1}{T} \sum_{t=1}^T \rho_\tau(R_{m,t \rightarrow N} - \beta_0 - \beta_1 \tilde{Q}_{t,\tau}) \\ L_\tau(\beta, \sigma_\mu) &:= \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \mathbb{E} \left[\rho_\tau(R_{m,t \rightarrow N} - \beta_0 - \beta_1 \tilde{Q}_{t,\tau}) \right]. \end{aligned}$$

Assumption A.4. *In the lognormal model, assume additionally that*

- (i) $\mathbb{E}[R_{m,t \rightarrow N}]$ and $\mathbb{E}[\tilde{Q}_{t,\tau}]$ are finite,
- (ii) $L_\tau(\beta, 0)$ has an identifiably unique minimum β^* , i.e. for all $\varepsilon > 0$

$$\inf_{\|\beta - \beta^*\| > \varepsilon} L_\tau(\beta, 0) - L_\tau(\beta^*, 0) > 0.$$

- (iii) as $T \rightarrow \infty$, for any compact set \mathcal{B} and sequence $b_T \searrow 0$,

$$\sup_{\beta \in \mathcal{B}} \|L_\tau(\beta, \sigma_\mu^T) - L_\tau(\beta, 0)\| = o(1) \quad (\text{Uniform continuity}). \quad (\text{A.4a})$$

$$\sup_{\sigma_\mu \leq b_T} \sup_{\beta \in \mathcal{B}} \|L_{T,\tau}(\beta, \sigma_\mu) - L_\tau(\beta, \sigma_\mu)\| = o_p(1) \quad (\text{Uniform LLN}). \quad (\text{A.4b})$$

Proof of Proposition 3.1(ii). Consider the population minimization problem of quantile regression at $\sigma_\mu = 0$

$$[\beta_0^*(0; \tau), \beta_1^*(0; \tau)] := \arg \min_{(\beta_0, \beta_1) \in \mathbb{R}^2} L_\tau(\beta, 0). \quad (\text{A.5})$$

Assumptions A.4(i,ii) ensure that the objective function is well defined and the solution in (A.5) is unique for all $\tau \in \mathcal{I}$. At $\sigma_\mu = 0$, $Q_{t,\tau} = e^{(\mu-r)N}$, so that $[\beta_0^*(0; \tau), \beta_1^*(0; \tau)] = [0, e^{(\mu-r)N}]$. To ease notation in the following derivation, I write $\hat{\beta}(\sigma_\mu^T) := \arg \min_\beta L_{T,\tau}(\beta, \sigma_\mu^T)$ and $\beta^*(0) = \arg \min_\beta L_\tau(\beta, 0)$. It then follows

that for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$\begin{aligned}
& \mathrm{P} \left(\left\| \widehat{\beta}(\sigma_\mu^T) - \beta^*(0) \right\| > \varepsilon \right) \\
& \leq \mathrm{P} \left(L_\tau(\widehat{\beta}(\sigma_\mu^T), 0) - L_\tau(\beta^*(0), 0) > \delta \right) \\
& = \mathrm{P} \left(L_\tau(\widehat{\beta}(\sigma_\mu^T), 0) - L_{T,\tau}(\widehat{\beta}(\sigma_\mu^T), \sigma_\mu^T) + L_{T,\tau}(\widehat{\beta}(\sigma_\mu^T), \sigma_\mu^T) - L_\tau(\beta^*(0), 0) > \delta \right) \\
& \leq \mathrm{P} \left(L_\tau(\widehat{\beta}(\sigma_\mu^T), 0) - L_{T,\tau}(\widehat{\beta}(\sigma_\mu^T), \sigma_\mu^T) + L_{T,\tau}(\beta^*(0), \sigma_\mu^T) - L_\tau(\beta^*(0), 0) > \delta \right) \\
& \leq \mathrm{P} \left(2 \sup_{\beta \in \mathcal{B}} \|L_\tau(\beta, 0) - L_{T,\tau}(\beta, \sigma_\mu^T)\| > \delta \right).
\end{aligned}$$

The second line follows from identification and the second to last line from the minimization property of $\widehat{\beta}(\sigma_\mu^T)$. Therefore, it suffices to show that

$$\sup_{\beta \in \mathcal{B}} \|L_\tau(\beta, 0) - L_{T,\tau}(\beta, \sigma_\mu^T)\| = o_p(1).$$

This claim follows from

$$\begin{aligned}
& \sup_{\beta \in \mathcal{B}} \|L_\tau(\beta, 0) - L_{T,\tau}(\beta, \sigma_\mu^T)\| \\
& \leq \sup_{\beta \in \mathcal{B}} \|L_\tau(\beta, 0) - L_\tau(\beta, \sigma_\mu^T)\| + \|L_\tau(\beta, \sigma_\mu^T) - L_{T,\tau}(\beta, \sigma_\mu^T)\| \\
& \leq \sup_{\beta \in \mathcal{B}} \|L_\tau(\beta, 0) - L_\tau(\beta, \sigma_\mu^T)\| + \sup_{\sigma_\mu \leq b_T} \sup_{\beta \in \mathcal{B}} \|L_\tau(\beta, \sigma_\mu) - L_{T,\tau}(\beta, \sigma_\mu)\|.
\end{aligned}$$

The first term is $o(1)$ by (A.4a) and the second term is $o_p(1)$ by (A.4b), which completes the proof. The claim in (3.6) easily follows from (3.5). ■

A.4 Proof of Theorem 4.1

Proof. I suppress the dependence of the τ -quantile on R and write $\tilde{Q}_\tau := \tilde{Q}_\tau(R)$. Starting from the definition of a risk-neutral quantile, it follows that

$$\begin{aligned}
\tau &= \tilde{\mathbb{P}} \left[R \leq \tilde{Q}_\tau \right] = \tilde{\mathbb{E}} \left[\mathbb{1} \left(R \leq \tilde{Q}_\tau \right) \right] = \frac{1}{\mathbb{E}(M)} \mathbb{E} \left[M \mathbb{1} \left(R \leq \tilde{Q}_\tau \right) \right] \\
&= \frac{1}{\mathbb{E}(M)} \left[\mathbb{C}\mathbb{O}\mathbb{V} \left(M, \mathbb{1} \left(R \leq \tilde{Q}_\tau \right) \right) + \mathbb{E}[M] \mathbb{E} \left[\mathbb{1} \left(R \leq \tilde{Q}_\tau \right) \right] \right] \\
&= \frac{1}{\mathbb{E}(M)} \mathbb{C}\mathbb{O}\mathbb{V} \left(M, \mathbb{1} \left(R \leq \tilde{Q}_\tau \right) \right) + \underbrace{\mathbb{E} \left[\mathbb{1} \left(R \leq \tilde{Q}_\tau \right) \right]}_{=\phi(\tau)}. \tag{A.6}
\end{aligned}$$

Rearranging then yields

$$\frac{1}{\mathbb{E}(M)} \mathbb{C}\mathbb{O}\mathbb{V} \left(M, \mathbb{1} \left(R \leq \tilde{Q}_\tau \right) \right) = \tau - \phi(\tau).$$

Using Cauchy-Schwarz renders the inequality

$$\begin{aligned} \frac{1}{\mathbb{E}(M)} \sigma(M) \sigma \left(\mathbb{1} \left(R \leq \tilde{Q}_\tau \right) \right) &\geq |\tau - \phi(\tau)| \\ \frac{\sigma(M)}{\mathbb{E}(M)} &\geq \frac{|\tau - \phi(\tau)|}{\sigma \left(\mathbb{1} \left(R \leq \tilde{Q}_\tau \right) \right)}. \end{aligned} \quad (\text{A.7})$$

Finally, since $\mathbb{1} \left(R \leq \tilde{Q}_\tau \right)$ is a Bernoulli random variable, it follows that

$$\sigma \left(\mathbb{1} \left(R \leq \tilde{Q}_\tau \right) \right) = \sqrt{\phi(\tau)(1 - \phi(\tau))}. \quad (\text{A.8})$$

Theorem 4.1 now follows after substituting (A.8) into (A.7). \blacksquare

A.5 Quantile Bound when SDF and Return are Jointly Normal

In this Section I derive (4.5) and (4.6), when M and R are jointly normal. First consider (4.6). The proof of the quantile bound in Theorem 4.1 gives the following identity

$$\frac{|\tau - \phi(\tau)|}{R_f} = |\mathbb{COV} \left(\mathbb{1} \left(R \leq \tilde{Q}_\tau \right), M \right)|.$$

Standard SDF properties also yield the well known result

$$\frac{|\mathbb{E}(R) - R_f|}{R_f} = |\mathbb{COV} (R, M)|.$$

These results, combined with (4.5) prove (4.6), since

$$\begin{aligned} \frac{\text{HJ bound}}{\text{quantile bound}} &= \frac{\frac{|\mathbb{E}[R] - R_f|}{\sigma_R R_f}}{\frac{|\tau - \phi(\tau)|}{\sqrt{\phi(\tau)(1 - \phi(\tau))} R_f}} \\ &\stackrel{(4.5)}{=} \frac{\sqrt{\phi(\tau)(1 - \phi(\tau))}}{\sigma_R f_R(\tilde{Q}_\tau)}, \end{aligned}$$

where $f_R(\tilde{Q}_\tau)$ is the marginal density of R .

Finally, I make use of the following covariance identities to prove (4.5).

Lemma A.5 (Hoeffding). *For any square integrable random variable X and Z with*

marginal CDFs F_X, F_Z and joint CDF $F_{X,Z}$, it holds that

$$\text{COV}[\mathbb{1}(Z \leq z), X] = - \int_{-\infty}^{\infty} [F_{X,Z}(x, z) - F_X(x)F_Z(z)] dx \quad (\text{A.9})$$

$$\text{COV}[Z, X] = - \int_{-\infty}^{\infty} \text{COV}[\mathbb{1}(Z \leq z), X] dz. \quad (\text{A.10})$$

Proof. See [Lehmann \(1966\)](#). ■

I also need a relation for the bivariate normal distribution. Suppose that X, Z are jointly normal with correlation ρ , mean μ_X, μ_Z and variance σ_X^2, σ_Z^2 , then

$$\frac{\partial \Phi_2(x, z; \rho, \mu_X, \mu_Z, \sigma_X^2, \sigma_Z^2)}{\partial \rho} = \sigma_X \sigma_Z \phi_2(x, z; \rho, \mu_X, \mu_Z, \sigma_X^2, \sigma_Z^2), \quad (\text{A.11})$$

where $\Phi_2(\cdot)$ denotes the bivariate normal CDF and $\phi_2(\cdot)$ denotes the bivariate normal PDF ([Sungur, 1990](#)). We can now prove a covariance identity for jointly normal random variables.

Proposition A.6. *Suppose R and M are jointly normal with correlation ρ , then*

$$-\text{COV}[\mathbb{1}(R \leq x), M] = \phi_R(x) \cdot \text{COV}[R, M], \quad (\text{A.12})$$

where $\phi_R(\cdot)$ is the marginal density of R .

Proof. To lighten notation, I suppress the dependence on $\mu_R, \mu_M, \sigma_R^2, \sigma_M^2$ in the joint CDF and PDF. We then have

$$\begin{aligned} -\text{COV}[\mathbb{1}(R \leq x), M] &= \int_{-\infty}^{\infty} \Phi_2(x, m; \rho) - \Phi_2(x, m; 0) dm \\ &= \int_{-\infty}^{\infty} \int_0^{\rho} \sigma_R \sigma_M \phi_2(x, m; y) dy dm \\ &= \sigma_R \sigma_M \rho \phi_R(x) = \text{COV}[M, R] \phi_R(x), \end{aligned}$$

where, in the first line, I use (A.9) together with $F_R(r)F_M(m) = \Phi_2(r, m; 0)$, the second line follows from (A.11) and the third line follows from Fubini's theorem to swap the order of integration and $\int_{-\infty}^{\infty} \phi_2(x, m; y) dm = \phi_R(x)$. ■

Remark. The second covariance identity in (A.10) shows that $\text{COV}[\mathbb{1}(R \leq x), M]$ is a measure of local dependence. In case of joint normality (A.12), the weight is given by the marginal PDF. For other distributions, the weighting factor is more complicated, but sometimes can be given an explicit form using a local Gaussian representation (see [Chernozhukov et al. \(2018\)](#)).

A.6 Minimum Percentile of Quantile Bound with Normal SDF

This section shows that the relative efficiency between the HJ bound and quantile bound is minimized when $\tilde{Q}_\tau = \mu_R$. To see this, write $x = \tilde{Q}_\tau$, and use $F(\cdot)$ to denote the physical CDF of R . I also drop the R subscript for f to avoid notational clutter. Consider

$$\Gamma(x) = \frac{F(x)(1 - F(x))}{f(x)^2}.$$

Minimizing $\Gamma(x)$ is equivalent to minimizing (4.6) and first order conditions imply that the optimal x^* satisfies

$$[f(x^*) - 2F(x^*)f(x^*)]f(x^*)^2 - 2f(x^*)f'(x^*)[F(x^*)(1 - F(x^*))] = 0. \quad (\text{A.13})$$

Since f, F are the respective PDF and CDF of the normal random variable R , it follows that $f'(\mu_R) = 0$ and $F(\mu_R) = 1/2$. As a result, (A.13) holds when $\tilde{Q}_{\tau^*} = x^* = \mu_R$.

A.7 Quantile Bound when SDF and Return are Log-normal

This section provides a closed form approximation for the relative efficiency between the HJ and quantile bound under joint lognormality. The result depends on Stein's Lemma ([Casella and Berger, 2002](#), Lemma 3.6.5).²⁷

Lemma A.7 (Stein's Lemma). *If X_1, X_2 are bivariate normal, $g : \mathbb{R} \rightarrow \mathbb{R}$ is differentiable and $\mathbb{E}|g'(X_1)| < \infty$, then*

$$\mathbb{COV}(g(X_1), X_2) = \mathbb{E}[g'(X_1)]\mathbb{COV}(X_1, X_2).$$

To prove the approximation, recall the distribution assumption

$$\begin{aligned} R &= e^{(\mu_R - \frac{\sigma_R^2}{2})\lambda + \sigma_R \sqrt{\lambda} Z_r} \\ M &= e^{-(r_f + \frac{\sigma_M^2}{2})\lambda + \sigma_M \sqrt{\lambda} Z_M}. \end{aligned}$$

Both Z_R and Z_M are standard normal random variables with correlation ρ . First, approximate M by a first order Taylor expansion, which gives

$$\widehat{M} = e^{-(r_f + \frac{\sigma_M^2}{2})\lambda} + Z_M \sigma_M \sqrt{\lambda} e^{-(r_f + \frac{\sigma_M^2}{2})\lambda}.$$

²⁷I use the form of Stein's Lemma reported in [Cochrane \(2005, p. 163\)](#), which follows from Stein's lemma as reported in [Casella and Berger \(2002\)](#).

Notice that $\widehat{M} = M + o_p(\sqrt{\lambda})$. Consequently, by Stein's Lemma

$$\begin{aligned}\text{COV}(R, M) &\approx \text{COV}(R, \widehat{M}) = \sigma_M \sqrt{\lambda} e^{-(r_f + \frac{\sigma_M^2}{2})\lambda} \text{COV}(R, Z_M) \\ &= \sigma_M \sqrt{\lambda} e^{-(r_f + \frac{\sigma_M^2}{2})\lambda} \mathbb{E} \left[\sigma_R \sqrt{\lambda} \exp \left(\left[\mu_R - \frac{\sigma_R^2}{2} \right] \lambda + \sigma_R \sqrt{\lambda} Z_R \right) \right] \text{COV}(Z_R, Z_M) \\ &= \sigma_M \sigma_R \lambda e^{-(r_f + \frac{\sigma_M^2}{2})\lambda} e^{\mu_R \lambda} \text{COV}(Z_R, Z_M).\end{aligned}$$

By Proposition A.6,

$$\begin{aligned}\text{COV}(\mathbb{1}(\log R \leq x), M) &\approx \text{COV}(\mathbb{1}(\log R \leq x), \widehat{M}) \\ &= \sigma_M \sqrt{\lambda} e^{-(r_f + \frac{\sigma_M^2}{2})\lambda} \text{COV}(\mathbb{1}(\log R \leq x), Z_M) \\ &= \sigma_M \sqrt{\lambda} e^{-(r_f + \frac{\sigma_M^2}{2})\lambda} \text{COV}(\mathbb{1}((\mu_R - \sigma_R^2/2)\lambda + \sigma_R \sqrt{\lambda} Z_R \leq x), Z_M) \\ &= -\sigma_M \sqrt{\lambda} e^{-(r_f + \frac{\sigma_M^2}{2})\lambda} f(x) \text{COV}(Z_R, Z_M).\end{aligned}$$

Here, f is the density of a normal random variable with mean $(\mu_R - \sigma_R^2/2)\lambda$ and variance $\lambda\sigma_R^2$. As a result,

$$\left| \frac{\mathbb{E}[R] - e^{\lambda r_f}}{\tau - \phi(\tau)} \right| \approx \frac{\sigma_R \sqrt{\lambda} e^{\mu_R \lambda}}{f(x)}. \quad (\text{A.14})$$

The same reasoning in Example 4.2 implies that the relative efficiency between the HJ and quantile bound can be approximated by

$$\frac{\text{HJ bound}}{\text{quantile bound}} = \frac{\frac{|\mathbb{E}[R] - R_f|}{\sigma(R) R_f}}{\frac{|\tau - \phi(\tau)|}{\sqrt{\phi(\tau)(1-\phi(\tau))} R_f}} \quad (\text{A.15})$$

$$\stackrel{(\text{A.14})}{\approx} \frac{\sqrt{\mathbb{P}(r \leq x) \cdot (1 - \mathbb{P}(r \leq x))}}{\sigma(R)} \times \frac{\sigma_R \sqrt{\lambda} e^{\mu_R \lambda}}{f(x)}, \quad (\text{A.16})$$

where $r = \log R$ and $x = \log \tilde{Q}_\tau$. Using the same reasoning as in Example 4.2, the expression on the right hand side of (A.15) is minimized by choosing $x = \log \tilde{Q}_\tau^*$ s.t. $\mathbb{P}(R \leq \tilde{Q}_\tau^*) = 1/2$. In that case the relative efficiency equals

$$\frac{\sqrt{2\pi\sigma_R^2} \sqrt{\lambda} e^{\mu_R \lambda}}{2\sqrt{[\exp(\sigma_R^2 \lambda) - 1] \exp(2\mu_R \lambda)}} = \frac{1}{2} \sqrt{\frac{2\pi\sigma_R^2 \lambda}{\exp(\sigma_R^2 \lambda) - 1}}.$$

A.8 Quantile Bound Pareto Distribution

Proof of Proposition 4.2. (i) The distribution of returns is Pareto, since

$$\begin{aligned}\mathbb{P}(R \leq x) &= \mathbb{P}(U^{-\beta} \leq x/B) \\ &= \mathbb{P}\left(U \geq (x/B)^{-\frac{1}{\beta}}\right) = 1 - \left(\frac{x}{B}\right)^{-\frac{1}{\beta}}, \quad x \geq B.\end{aligned}$$

(ii) Since $R_f M$ is the Radon-Nikodym derivative that induces a change of measure from \mathbb{P} to $\tilde{\mathbb{P}}$, it follows that

$$\begin{aligned}\tilde{\mathbb{P}}(R \leq x) &= R_f \mathbb{E}[M \mathbb{1}(R \leq x)] \\ &= R_f \int_0^1 A u^\alpha \mathbb{1}(B u^{-\beta} \leq x) du \\ &= R_f A \int_0^1 u^\alpha \mathbb{1}\left(u \geq \left(\frac{x}{B}\right)^{-\frac{1}{\beta}}\right) du \\ &= \frac{R_f A}{\alpha + 1} \left(1 - \left(\frac{x}{B}\right)^{-\frac{\alpha+1}{\beta}}\right) \\ &= 1 - \left(\frac{x}{B}\right)^{-\frac{\alpha+1}{\beta}}.\end{aligned}$$

The last line follows from (A.19) below.

(iii) Routine calculations show that the mean and variance of R are given by (provided $\beta < 1/2$)

$$\mathbb{E}[R] = \frac{B}{1-\beta} \quad \sigma^2(R) = \frac{B^2}{1-2\beta} - \left(\frac{B}{1-\beta}\right)^2. \quad (\text{A.17})$$

Likewise, the distribution of the SDF follows from

$$\mathbb{P}(M \leq x) = \mathbb{P}(AU^\alpha \leq x) = \left(\frac{x}{A}\right)^{\frac{1}{\alpha}}, \quad 0 \leq x \leq A.$$

In this case, M is said to have a Pareto lower tail. The expectation is given by

$$\mathbb{E}[M] = \frac{A}{\alpha + 1}.$$

The constraint $\mathbb{E}[MR] = 1$ forces

$$\frac{AB}{\alpha - \beta + 1} = 1. \quad (\text{A.18})$$

In addition from $\mathbb{E}[M] = \frac{1}{R_f}$ it follows

$$\frac{A}{\alpha+1} = \frac{1}{R_f}. \quad (\text{A.19})$$

The Sharpe ratio can now be computed from (A.17) and (A.19).

- (iv) It is straightforward to show that the quantiles of a $\mathbf{Par}(C, \zeta)$ distribution are given by

$$Q_\tau = C \times (1 - \tau)^{-1/\zeta}.$$

It therefore follows that the risk-neutral quantile function is equal to

$$\tilde{Q}_\tau = B(1 - \tau)^{-\frac{\beta}{\alpha+1}}.$$

As a result

$$\begin{aligned} \mathbb{P}(R \leq \tilde{Q}_\tau) &= \mathbb{P}\left(R \leq B(1 - \tau)^{-\frac{\beta}{\alpha+1}}\right) \\ &= 1 - \left(\frac{B}{B(1 - \tau)^{\frac{-\beta}{\alpha+1}}}\right)^{\frac{1}{\beta}} \\ &= 1 - (1 - \tau)^{\frac{1}{\alpha+1}}. \end{aligned}$$

Hence, the quantile bound evaluates to

$$\frac{1}{R_f} \frac{|\tau - \phi(\tau)|}{\sqrt{\phi(\tau)(1 - \phi(\tau))}} = \frac{A}{1 + \alpha} \frac{|\tau - 1 + (1 - \tau)^{\frac{1}{\alpha+1}}|}{\sqrt{(1 - (1 - \tau)^{\frac{1}{\alpha+1}})(1 - \tau)^{\frac{1}{\alpha+1}}}}.$$

- (v) The HJ bound, as given by the Sharpe ratio in (4.9), goes to 0 as $\beta \nearrow 1/2$ since $\sigma(R) \nearrow \infty$. ■

A.9 Derivation of Gâteaux Derivative

In this Section I derive (5.3). For ease of exposition, I drop the time subscripts. For $\lambda \in [0, 1]$, define $\tilde{F}_\lambda := (1 - \lambda)\tilde{F} + \lambda F$. The following (trivial) identity will prove helpful²⁸

$$\tau = \tilde{F}_\lambda \tilde{F}_\lambda^{-1}. \quad (\text{A.20})$$

To further simplify notation, write $q(\lambda) := \tilde{F}_\lambda^{-1}$. Then (A.20) becomes

$$\tau = (1 - \lambda)\tilde{F}(q(\lambda)) + \lambda F(q(\lambda)).$$

²⁸This “equality” may actually only be an inequality for some τ , but this is immaterial to the argument.

Applying the implicit function theorem, we obtain

$$q'(\lambda) = -\frac{-\tilde{F}(q(\lambda)) + F(q(\lambda))}{(1-\lambda)\tilde{f}(q(\lambda)) + \lambda f(q(\lambda))}.$$

Plug in $\lambda = 0$ to get

$$q'(0) = -\frac{-\tilde{F}(q(0)) + F(q(0))}{\tilde{f}(q(0))}. \quad (\text{A.21})$$

Notice that

$$\tilde{F}_\lambda|_{\lambda=0} = \tilde{F} \implies q(\lambda)|_{\lambda=0} = q(0) = \tilde{F}^{-1}. \quad (\text{A.22})$$

Substitute (A.22) into (A.21) to obtain

$$q'(0) = -\frac{-\tilde{F}(\tilde{F}^{-1}) + F(\tilde{F}^{-1})}{\tilde{f}(\tilde{F}^{-1})} = \frac{\tau - F(\tilde{F}^{-1})}{\tilde{f}(\tilde{F}^{-1})}. \quad (\text{A.23})$$

Notice that $q'(0)$ is exactly equal to the Gâteaux derivative from the definition in (5.2), since

$$\frac{\partial}{\partial \lambda} \varphi \left[(1-\lambda)\tilde{F} + \lambda F \right] \Big|_{\lambda=0} = \frac{\partial}{\partial \lambda} q(\lambda) \Big|_{\lambda=0} = q'(0).$$

A.10 Proof of Theorem 5.3 and Corollary 5.4

In the proofs that follow, I repeatedly use Taylor's theorem with integral remainder, which is stated for completeness.

Lemma A.8 (Taylor's theorem). *Let $\zeta^{(3)}(\cdot)$ be absolutely continuous on the closed interval between a and x , then*

$$\zeta(x) = \sum_{k=0}^3 \frac{\zeta^{(k)}(a)}{k!} (x-a)^k + \int_a^x \frac{\zeta^{(4)}(t)}{3!} (x-t)^3 dt.$$

Before I prove Theorem 5.3 and Corollary 5.4, I collect several results about the SDF in representative agent models.

Lemma A.9. *Assume a representative agent model with SDF given by (5.5), then*

$$\tau - F_t(\tilde{Q}_{t,\tau}) = -\frac{\widetilde{\mathbb{COV}}_t \left[\mathbb{1} \left(R_{m,t \rightarrow N} \leq \tilde{Q}_{t,\tau} \right), \zeta(R_{m,t \rightarrow N}) \right]}{\widetilde{\mathbb{E}}_t [\zeta(R_{m,t \rightarrow N})]}, \quad (\text{A.24})$$

where $\zeta(\cdot)$ is defined in (5.6).

Proof. Use the reciprocal of the SDF to pass from physical to risk-neutral measure

$$\begin{aligned} F_t(\tilde{Q}_{t,\tau}) &= \mathbb{E}_t \left[\mathbb{1} \left(R_{m,t \rightarrow N} \leq \tilde{Q}_{t,\tau} \right) \right] = \tilde{\mathbb{E}}_t \left[\mathbb{1} \left(R_{m,t \rightarrow N} \leq \tilde{Q}_{t,\tau} \right) \frac{\mathbb{E}_t [M_{t \rightarrow N}]}{M_{t \rightarrow N}} \right] \\ &= \widetilde{\mathbb{COV}}_t \left[\mathbb{1} \left(R_{m,t \rightarrow N} \leq \tilde{Q}_{t,\tau} \right), \frac{\mathbb{E}_t [M_{t \rightarrow N}]}{M_{t \rightarrow N}} \right] + \tau. \end{aligned} \quad (\text{A.25})$$

Rearranging the above and using the definition of $\zeta(\cdot)$ in (5.6), as well as (5.5), we obtain (A.24). \blacksquare

Lemma A.10. *Under Assumption 5.2,*

$$\tilde{\mathbb{E}}_t [\zeta(R_{m,t \rightarrow N})] \leq \sum_{k=0}^3 \theta_k \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)} = 1 + \sum_{k=1}^3 \theta_k \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)},$$

where $\zeta(x)$ is the marginal rate of substitution defined in (5.6).

Proof. In the integral of Lemma A.8, substitute $s = (t - a)/(x - a)$ to get

$$\begin{aligned} \zeta(x) &= \sum_{k=0}^3 \frac{\zeta^{(k)}(a)}{k!} (x - a)^k + (x - a)^4 \int_0^1 \frac{\zeta^{(4)}(a + s(x - a))}{3!} (1 - s)^3 ds \\ &\leq \sum_{k=0}^3 \frac{\zeta^{(k)}(a)}{k!} (x - a)^k, \end{aligned}$$

since $\zeta^{(4)}(x) < 0$ by Assumption 5.2(ii). Using this result with $a = R_{f,t \rightarrow N}$ and taking expectations, we obtain

$$\tilde{\mathbb{E}}_t [\zeta(R_{m,t \rightarrow N})] \leq \sum_{k=0}^3 \theta_k \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)}. \quad \blacksquare$$

Proof of Theorem 5.3. By Taylor's theorem,

$$\begin{aligned} -\widetilde{\mathbb{COV}}_t \left[\mathbb{1} \left(R_{m,t \rightarrow N} \leq \tilde{Q}_{t,\tau} \right), \zeta(R_{m,t \rightarrow N}) \right] &= \sum_{k=1}^3 \theta_k \left(\tau \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)} - \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)}[\tilde{Q}_{t,\tau}] \right) \\ &- \widetilde{\mathbb{COV}}_t \left[\mathbb{1} \left(R_{m,t \rightarrow N} \leq \tilde{Q}_{t,\tau} \right), \frac{1}{3!} \int_{R_{f,t \rightarrow N}}^{R_{m,t \rightarrow N}} \zeta^{(4)}(t)(R_{m,t \rightarrow N} - t)^3 dt \right] \\ &\geq \sum_{k=1}^3 \theta_k \left(\tau \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)} - \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)}[\tilde{Q}_{t,\tau}] \right). \end{aligned} \quad (\text{A.26})$$

The last line follows from Lemma A.11 below. Hence,

$$\begin{aligned}\tau - F_t(\tilde{Q}_{t,\tau}) &= -\frac{\widetilde{\mathbb{COV}}_t \left[\mathbb{1} \left(R_{m,t \rightarrow N} \leq \tilde{Q}_{t,\tau} \right), \zeta(R_{m,t \rightarrow N}) \right]}{\widetilde{\mathbb{E}}_t [\zeta(R_{m,t \rightarrow N})]} \\ &\geq \frac{\sum_{k=1}^3 \theta_k \left(\tau \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)} - \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)}[\tilde{Q}_{t,\tau}] \right)}{1 + \sum_{k=1}^3 \theta_k \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)}},\end{aligned}$$

where the first identity follows from Lemma A.9 and the inequality follows from (A.26) and Lemma A.10. \blacksquare

Lemma A.11. *Suppose that Assumption 5.2 holds. In addition, define τ^* so that $G(\tilde{Q}_{t,\tau^*}) = \widetilde{\mathbb{E}}_t(G(R_{m,t \rightarrow N}))$, where*

$$G(R_{m,t \rightarrow N}) := \int_{R_{f,t \rightarrow N}}^{R_{m,t \rightarrow N}} \zeta^{(4)}(t)(R_{m,t \rightarrow N} - t)^3 dt.$$

Then for all $\tau \leq \tau^*$,

$$\widetilde{\mathbb{COV}}_t \left[\mathbb{1} \left(R_{m,t \rightarrow N} \leq \tilde{Q}_{t,\tau} \right), \int_{R_{f,t \rightarrow N}}^{R_{m,t \rightarrow N}} \zeta^{(4)}(t)(R_{m,t \rightarrow N} - t)^3 dt \right] \leq 0. \quad (\text{A.27})$$

Proof. If $\zeta^{(4)} \equiv 0$, then (A.27) trivially holds. Hence, assume that $\zeta^{(4)}$ is not identically equal to zero. First we show that $G(R_{m,t \rightarrow N})$ is increasing on $(0, R_{f,t \rightarrow N})$, since by Leibniz' rule

$$G'(R_{m,t \rightarrow N}) = -3 \int_{R_{m,t \rightarrow N}}^{R_{f,t \rightarrow N}} \zeta^{(4)}(t)(R_{m,t \rightarrow N} - t)^2 dt \geq 0.$$

The inequality follows since $\zeta^{(4)}(t) < 0$ by Assumption 5.2(ii). Temporarily write $K = \tilde{Q}_{t,\tau}$ to ease notation and consider

$$\Gamma(K) = \widetilde{\mathbb{COV}}_t \left[\mathbb{1} \left(R_{m,t \rightarrow N} \leq K \right), \int_{R_{f,t \rightarrow N}}^{R_{m,t \rightarrow N}} \zeta^{(4)}(t)(R_{m,t \rightarrow N} - t)^3 dt \right].$$

By Leibniz' rule again, we get

$$\Gamma'(K) = \tilde{f}_t(K) \left(G(K) - \widetilde{\mathbb{E}}_t(G(R_{m,t \rightarrow N})) \right).$$

Since $G(R_{f,t \rightarrow N}) = 0$, $G(R_{m,t \rightarrow N}) \leq 0$ and $G(R_{m,t \rightarrow N})$ is increasing on $(0, R_{f,t \rightarrow N})$, we know that $\Gamma'(K) \leq 0$ for all $K \leq K^* < R_{f,t \rightarrow N}$, where K^* is defined such that $G(K^*) = \widetilde{\mathbb{E}}_t(G(R_{m,t \rightarrow N}))$. To complete the proof, define τ^* so that it satisfies $\tilde{Q}_{\tau^*} = K^*$. \blacksquare

Proof of Corollary 5.4. Using Assumption (i) and (ii), we get $\theta_2 \tilde{\mathbb{M}}_{t \rightarrow N}^{(2)} \leq -1/R_{f,t \rightarrow N}^2 \tilde{\mathbb{M}}_{t \rightarrow N}^{(2)}$

and $\theta_3 \tilde{\mathbb{M}}_{t \rightarrow N}^{(3)} \leq 1/R_{f,t \rightarrow N}^3 \tilde{\mathbb{M}}_{t \rightarrow N}^{(3)}$, from which it follows that

$$1 + \sum_{k=1}^3 \theta_k \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)} \leq 1 - \frac{1}{R_{f,t \rightarrow N}^2} \tilde{\mathbb{M}}_{t \rightarrow N}^{(2)} + \frac{1}{R_{f,t \rightarrow N}^3} \tilde{\mathbb{M}}_{t \rightarrow N}^{(3)}. \quad (\text{A.28})$$

Second, recall that for $K > 0$

$$\tilde{F}_t(K) \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)} - \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)}[K] = -\widetilde{\mathbb{COV}}_t [\mathbb{1}(R_{m,t \rightarrow N} \leq K), (R_{m,t \rightarrow N} - R_{f,t \rightarrow N})^k].$$

If $k = 1, 3$, then Chebyshev's sum inequality D.1 implies that

$$\Gamma(K) := \widetilde{\mathbb{COV}}_t [\mathbb{1}(R_{m,t \rightarrow N} \leq K), (R_{m,t \rightarrow N} - R_{f,t \rightarrow N})^k] \leq 0.$$

Hence under Assumption (i),

$$\theta_k \left(\tilde{F}_t(K) \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)} - \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)}[K] \right) \geq \frac{1}{R_{f,t \rightarrow N}^k} \left(\tilde{F}_t(K) \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)} - \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)}[K] \right) \quad \text{for } k = 1, 3. \quad (\text{A.29})$$

If $k = 2$, we obtain from Leibniz' rule

$$\Gamma'(K) = \tilde{f}_t(K) \left[(K - R_{f,t \rightarrow N})^2 - \widetilde{\mathbb{VAR}}_t(R_{m,t \rightarrow N}) \right]. \quad (\text{A.30})$$

It follows that (A.30) is positive if $K \leq R_{f,t \rightarrow N} - \sqrt{\widetilde{\mathbb{VAR}}_t(R_{m,t \rightarrow N})} =: K^{**}$. Combining (A.29) and (A.30), we get for $K \leq K^{**}$

$$\theta_k \left(\tilde{F}_t(K) \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)} - \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)}[K] \right) \geq \frac{(-1)^{k+1}}{R_{f,t \rightarrow N}^k} \left(\tilde{F}_t(K) \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)} - \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)}[K] \right). \quad (\text{A.31})$$

Collecting the results from (A.28) and (A.31) and using the general upper bound (5.7) from Theorem 5.3, it follows that

$$\begin{aligned} \tau - F_t(\tilde{Q}_{t,\tau}) &\stackrel{(5.7)}{\geq} \frac{\sum_{k=1}^3 \theta_k (\tau \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)} - \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)}[\tilde{Q}_{t,\tau}])}{1 + \sum_{k=1}^3 \theta_k \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)}} \\ &\geq \frac{\sum_{k=1}^3 \frac{(-1)^{k+1}}{R_{f,t \rightarrow N}^k} (\tau \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)} - \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)}[\tilde{Q}_{t,\tau}])}{1 + \sum_{k=1}^3 \frac{(-1)^{k+1}}{R_{f,t \rightarrow N}^k} \tilde{\mathbb{M}}_{t \rightarrow N}^{(k)}}, \end{aligned}$$

for all τ such that $\tilde{Q}_{t,\tau} \leq \min(K^*, K^{**})$, where K^* is defined in Theorem 5.3. ■

A.11 Formulas for market moments

This Section presents formulas for the (un)truncated risk-neutral moments of the excess market return. I use a slight abuse of notation and write $\tilde{Q}(\tau) := \tilde{Q}_\tau(R_{m,t \rightarrow N})$,

to emphasize that the integrals below are taken with respect to τ .

Proposition A.12. *Any risk-neutral moment can be computed from the risk-neutral quantile function, since*

$$\tilde{\mathbb{E}}_t [(R_{m,t \rightarrow N} - R_{f,t \rightarrow N})^n] = \int_0^1 [\tilde{Q}_\tau (R_{m,t \rightarrow N} - R_{f,t \rightarrow N})]^n d\tau = \int_0^1 [\tilde{Q}(\tau) - R_{f,t \rightarrow N}]^n d\tau. \quad (\text{A.32})$$

Moreover, any truncated risk-neutral moment can be calculated by

$$\tilde{\mathbb{E}}_t [(R_{m,t \rightarrow N} - R_{f,t \rightarrow N})^n \mathbb{1}(R_{m,t \rightarrow N} \leq k_0)] = \int_0^{\tilde{F}_t(k_0)} [\tilde{Q}(\tau) - R_{f,t \rightarrow N}]^n d\tau.$$

Proof. For any random variable X and integer n such that the n -th moment exists, we have

$$\mathbb{E}[X^n] = \int_0^1 [Q_X(\tau)]^n d\tau.$$

This follows straightforward from the substitution $x = Q(\tau)$. Now use that for any constant $a \in \mathbb{R}$, $Q_{X-a}(\tau) = Q_X(\tau) - a$ to derive (A.32). The truncated formula follows similarly. ■

Remark. Frequently I use $k_0 = \tilde{Q}_\tau$, in which case the truncated moment formula reduces to

$$\tilde{\mathbb{E}}_t \left[(R_{m,t \rightarrow N} - R_{f,t \rightarrow N})^n \mathbb{1}(R_{m,t \rightarrow N} \leq \tilde{Q}_\tau) \right] = \int_0^\tau [\tilde{Q}(p) - R_{f,t \rightarrow N}]^n dp.$$

B Estimating the Risk-neutral Quantile Function

B.1 Data Description

To estimate the risk-neutral quantile curve for each point in time, I use daily option prices from OptionMetrics covering the period 01-01-1996 until 12-31-2021. The data consist of European Put and Call options on the S&P 500 index. The option contract further contains information on the highest closing bid and lowest closing ask price and price of the forward contract on the underlying security. I use the midpoint of the bid and ask price to proxy for the unobserved option price. In addition, I obtain data on the daily risk-free rate from Kenneth French' website.²⁹ Finally, I obtain stock price data on the closing price of the S&P 500 from WRDS.

I use an additional cleaning procedure for the option data, prior to estimating the martingale measure. All observations are dropped for which the highest closing bid

²⁹See http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html#Research

price equals zero, as well as all option prices that violate no-arbitrage bounds. Subsequently, I drop all option prices with maturity less than 7 days or greater than 500 days. After the cleaning procedure, I'm left with 23,264,113 option-day observations.

I discard all observations prior to 2003 for the quantile regression application, since there are many days in the period 1996-2003 that have insufficient option data to estimate $\tilde{Q}_{t,\tau}$ at the 30, 60 and 90-day horizon. Occasionally it happens that I cannot estimate the risk-neutral quantile on a specific day in the post 2003 period and I discard these days as well.³⁰

B.2 Estimation Procedure

There is a substantial literature on how to extract the martingale measure from option prices. I use the **RND Fitting Tool** application on MATLAB, which is developed by [Barletta and Santucci de Magistris \(2018\)](#).³¹ The tool is based on the orthogonal polynomial expansion of [Filipović et al. \(2013\)](#). In short, the idea is to approximate the conditional risk-neutral density function by an expansion of the form

$$\tilde{f}_t(x) \approx \phi(x) \left[1 + \sum_{k=1}^K \sum_{i=0}^k c_k w_{i,k} x^k \right],$$

where $\phi(x)$ is an arbitrary density and the polynomial term serves to tilt the density function towards the risk-neutral distribution. Further details about the estimation of the coefficients $w_{i,k}$ and c_k can be found in [Filipović et al. \(2013\)](#).

For my purpose, I need to choose the kernel function $\phi(\cdot)$, the estimation method for c_k and the degree of the expansion K . I follow the recommendation of [Barletta and Santucci de Magistris \(2018\)](#) and use the double beta distribution for the kernel and principal component analysis to estimate c_k . This is the most robust method for S&P500 options. To avoid overfitting, I use $K = 3$ if the number of option data is less than 70, $K = 6$ if the number is less than 100 and $K = 8$ otherwise. This choice renders a good approximation for most time periods.

I interpolate the estimated risk-neutral densities for a given time period. Occasionally, there are no two interpolation points. In such cases, I drop the observations to avoid negative density estimates due to extrapolation. Since the RND Fitting Tool is designed for an equal number of put and call options, I use Put-Call parity to convert in-the-money call prices to put prices and vice versa. Subsequently, I use

³⁰The number of days I cannot estimate the risk-neutral PDF is very small, about 2% in total. Most of these days occur at the beginning of the sample period.

³¹The application can be downloaded from the author's GITHUB page: <https://github.com/abarletta/rndfittool>

Black-Scholes implied volatilities to interpolate the Call-Put option price curve near the forward price. This transformation ensures that the risk-neutral density does not have a discontinuity for strike prices that are close to being at-the-money (Figlewski, 2010). Finally, I integrate the density function and take the inverse to obtain the risk-neutral quantile curve

$$\tilde{Q}_{t,\tau} := \inf \left\{ x \in \mathbb{R} : \tau \leq \tilde{F}_t(x) \right\}, \quad \text{where } \tilde{F}_t(x) = \int_0^x \tilde{f}_t(y) dy.$$

C Verifying Assumption 5.2(ii) in Representative Agent Models

The proof of Theorem 5.3 relies on Assumption 5.2(ii). This section derives parameter restrictions for common utility functions that are needed so that Assumption 5.2(ii) is satisfied. Most of these restrictions closely resemble those of Chabi-Yo and Loudis (2020). I also illustrate the lower bound with actual data assuming CRRA utility.

C.1 Log utility

In this case $u(x) = \log x$. It follows that $\zeta(x) = x/R_{f,t \rightarrow N}$. Clearly $\zeta^{(4)}(x) = 0$ and Assumption 5.2 holds.

C.2 CRRA utility

More generally, consider $u(x) = \frac{x^{1-\gamma}}{1-\gamma}$ for $\gamma \geq 0$. It follows that $\zeta(x) = (\frac{x}{R_{f,t \rightarrow N}})^\gamma$ and hence

$$\zeta^{(4)}(x) = \frac{1}{R_{f,t \rightarrow N}^\gamma} \gamma(\gamma-1)(\gamma-2)(\gamma-3)x^{\gamma-4}.$$

Part (ii) of Assumption 5.2 holds if $\gamma \in [0, 1]$, but also if $\gamma \in [2, 3]$. Notice that the additional restrictions in the feasible lower bound in Corollary 5.4 cannot be accommodated by this model. To see this, observe that $\theta_2 \leq -1/R_{f,t \rightarrow N}^2$ implies that $\gamma(\gamma-1)/2 \leq -1/R_{f,t \rightarrow N}^2$, which cannot hold for any reasonable interest rate. This failure illustrates that a representative agent model with CRRA utility is misspecified in that it cannot produce a sizable risk-premium on skewness.³²

C.3 CARA utility

In this case, $u(x) = 1 - e^{-\gamma x}$ and $\zeta(x) = e^{\gamma^*(x-R_{f,t \rightarrow N})}$, where $\gamma^* = W_t \gamma$. Since $\zeta^{(4)} > 0$, Assumption 5.2 does not hold.

³²See in particular Chabi-Yo and Loudis (2020, Equation (A.5)), which shows that θ_2 is related to the risk-premium on market skewness.

C.4 HARA utility

The utility function is given by $u(x) = \frac{1-\gamma}{\gamma} \left(\frac{ax}{1-\gamma} + b \right)^\gamma$, where $a > 0$ and $\frac{ax}{1-\gamma} + b > 0$. Successive differentiation renders

$$\zeta^{(4)}(x) = \frac{-\gamma(\gamma+1)(\gamma+2)(aW_t)^4 \left(\frac{aW_t x}{1-\gamma} + b \right)^{-\gamma-3} \left(\frac{aW_t R_{f,t \rightarrow N}}{1-\gamma} + b \right)^{\gamma-1}}{(1-\gamma)^3}.$$

We see that $\gamma \in [0, 1)$ is a sufficient condition for $\zeta^{(4)}(x) \leq 0$.

C.5 Lower Bound in the Data for CRRA utility

Figure C2 illustrates the infeasible lower bound as well as the quantile approximation for CRRA utility with different levels of risk aversion. The risk-neutral distribution is obtained from option data over a 90-day horizon on October 28, 2015. Panels C2a and C2c show the infeasible lower bound from Theorem 5.3 when risk aversion is 2.2 and 2.9 respectively. Consistent with the theorem, the infeasible lower bound is below $\tau - F_t(\tilde{Q}_{t,\tau})$ in the left-tail, and seems to hold for a large range of τ 's, in particular for all $\tau \leq 0.5$. The right panels show the quantile approximation (5.10) based on the infeasible lower bound. We see that the risk-adjusted quantile approximation comes much closer to the physical quantile relative to the risk-neutral quantile function.

D Crash Probability in Representative Agent Models

In this Section, I derive several results about conditional tail probabilities in representative agent models. Specifically, I show how these probabilities can be calculated using common utility functions and how they are affected by a change in the underlying parameter (comparative statics). The results do not assume a specific distribution of the market return and generalize some known results in the literature which assume log-normality.

D.1 Crash Probability with Log Utility

[Chabi-Yo and Loudis \(2020\)](#), Remark 1) show that their bounds on the equity premium equal the bounds of [Martin \(2017\)](#) when the representative agent has log preferences. Here, I derive the analogous result for the subjective crash probability of a log investor reported by [Martin \(2017\)](#), Result 2). In our notation, [Martin \(2017\)](#)

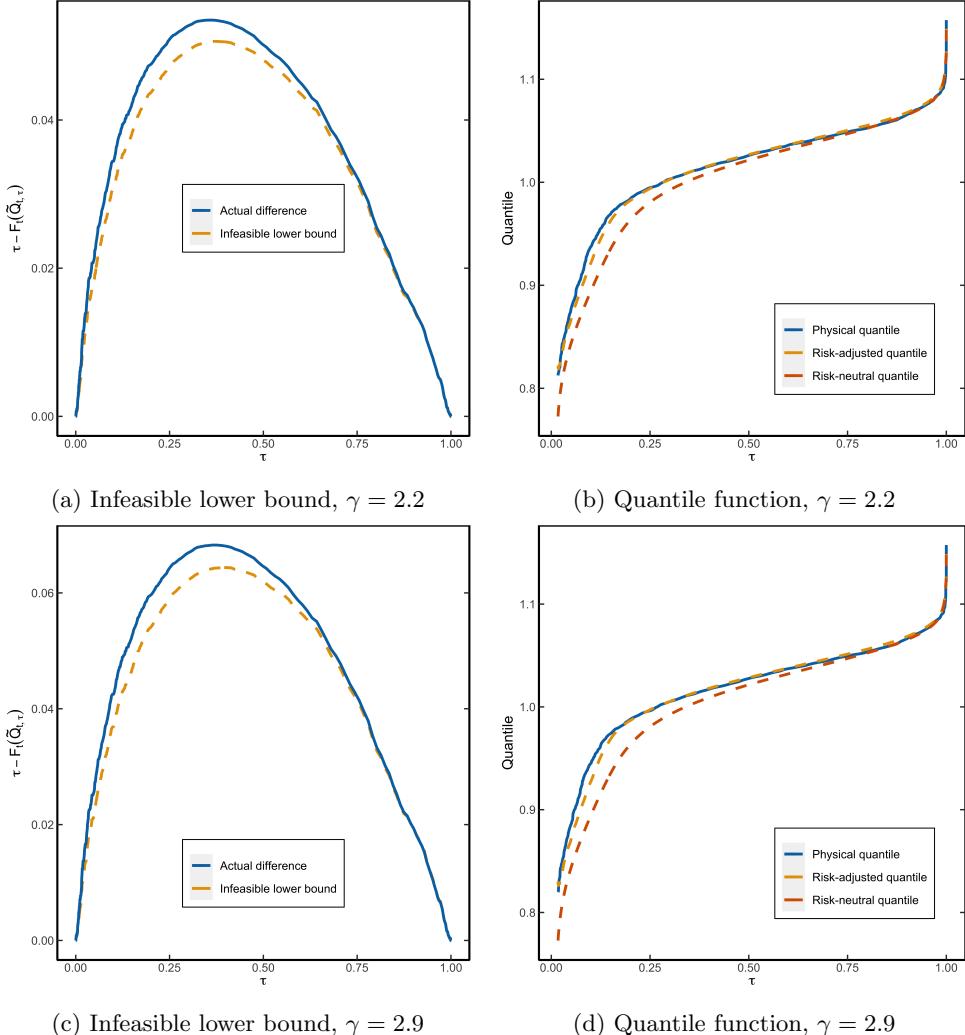


Figure C2: Lower bound with CRRA utility for 90-day returns. This figure shows the lower bound on $\tau - F_t(\tilde{Q}_{t,\tau})$ as well as the quantile approximation $Q_{t,\tau} \approx \tilde{Q}_{t,\tau} + RA_{t,\tau}$ in a representative agent model with CRRA utility function, $u(x) = x^{1-\gamma}/(1-\gamma)$, for $\gamma \in \{2.2, 2.9\}$. The left panels show the infeasible lower bound $LB_{t,\tau}$, and the true risk-adjustment, $\tau - F_t(\tilde{Q}_{t,\tau})$. The right panels show the physical, risk-neutral and risk-adjusted quantile functions. The risk-adjusted quantile function uses the infeasible lower bound. The risk-neutral distribution is coming from option data on the S&P500 on October 28, 2015 with a maturity of 90 days.

shows that

$$\mathbb{P}_t(R_{m,t \rightarrow N} < \alpha) = \alpha \left[\text{Put}'_t(\alpha S_t) - \frac{\text{Put}_t(\alpha S_t)}{\alpha S_t} \right], \quad (\text{D.1})$$

where Put'_t is the derivative of the put option price curve seen as a function of the strike. Under log preferences and using (A.25), it follows that

$$\begin{aligned}\mathbb{P}_t(R_{m,t \rightarrow N} < \tilde{Q}_{t,\tau}) &= \tau + \frac{1}{R_{f,t \rightarrow N}} \widetilde{\text{COV}}_t \left[\mathbb{1}(R_{m,t \rightarrow N} \leq \tilde{Q}_{t,\tau}), R_{m,t \rightarrow N} \right] \\ &= \tau + \frac{1}{R_{f,t \rightarrow N}} \left(\widetilde{\mathbb{E}}_t \left[\mathbb{1}(R_{m,t \rightarrow N} \leq \tilde{Q}_{t,\tau}) R_{m,t \rightarrow N} \right] - \widetilde{\mathbb{E}}_t(R_{m,t \rightarrow N}) \widetilde{\mathbb{E}}_t \left(\mathbb{1}(R_{m,t \rightarrow N} \leq \tilde{Q}_{t,\tau}) \right) \right) \\ &= \frac{1}{R_{f,t \rightarrow N}} \widetilde{\mathbb{E}}_t \left[\mathbb{1}(R_{m,t \rightarrow N} \leq \tilde{Q}_{t,\tau}) R_{m,t \rightarrow N} \right].\end{aligned}\quad (\text{D.2})$$

The result now follows upon substituting $\tilde{Q}_\tau = \alpha$, since Martin (2017) shows that (D.2) equals the right hand side of (D.1).

D.2 Crash Probability with CRRA utility

I now consider the case in which the representative agent has constant relative risk aversion (CRRA) utility, $u(x) = x^{1-\gamma}/(1-\gamma)$, where γ is the relative risk aversion parameter. First, I show that the excess market return is non-decreasing in γ *regardless* of the distribution of the market return.³³ The proof uses the following lemma, which is a special case of the FKG inequality (Hsu and Varadhan, 1999, Theorem 1.3).

Lemma D.1 (Chebyshev sum inequality). *Let X be a random variable and let g, h both be non-increasing or non-decreasing. Then,*

$$\mathbb{E}(g(X)h(X)) \geq \mathbb{E}(g(X))\mathbb{E}(h(X)).$$

The inequality is reversed if one is non-increasing and the other is non-decreasing.

Proof. Let X_1, X_2 be IID copies of X and assume that g, h are non-decreasing. It follows that

$$(g(X_1) - g(X_2))(h(X_1) - h(X_2)) \geq 0. \quad (\text{D.3})$$

Taking expectations on both sides completes the proof. The same proof goes through if g, h are non-increasing. If one is non-increasing and the other is non-decreasing, the inequality in (D.3) is reversed. ■

Proposition D.2. *Assume that a representative investor has CRRA utility, with $\gamma \geq 0$ and $\mathbb{E}_t \left[R_{m,t \rightarrow N}^{\gamma+1} \log R_{m,t \rightarrow N} \right] < \infty$. Then, $\mathbb{E}_t[R_{m,t \rightarrow N}] - R_{f,t \rightarrow N}$, is non-decreasing in γ .*

Remark. I suppress the dependence of the physical expectation on γ in the notation for readability.

³³Cochrane (2005) derives this result when the distribution is lognormal.

Proof. According to Chabi-Yo and Loudis (2020, Equation (53)), we have

$$\mathbb{E}_t [R_{m,t \rightarrow N}] - R_{f,t \rightarrow N} = \frac{\tilde{\mathbb{E}}_t [R_{m,t \rightarrow N}^{\gamma+1}]}{\tilde{\mathbb{E}}_t [R_{m,t \rightarrow N}^\gamma]} - R_{f,t \rightarrow N} =: g(\gamma).$$

It is enough to show that $g'(\gamma) \geq 0$ for $\gamma \geq 0$. Taking first order conditions, we need to show that

$$\tilde{\mathbb{E}}_t [R_{m,t \rightarrow N}^{\gamma+1} \log R_{m,t \rightarrow N}] \tilde{\mathbb{E}}_t [R_{m,t \rightarrow N}^\gamma] \geq \tilde{\mathbb{E}}_t [R_{m,t \rightarrow N}^{\gamma+1}] \tilde{\mathbb{E}}_t [R_{m,t \rightarrow N}^\gamma \log R_{m,t \rightarrow N}]. \quad (\text{D.4})$$

Introduce another probability measure \mathbb{P}^* , defined by

$$\mathbb{E}_t^* [Z] := \frac{\tilde{\mathbb{E}}_t [ZR_{m,t \rightarrow N}^\gamma]}{\tilde{\mathbb{E}}_t [R_{m,t \rightarrow N}^\gamma]}. \quad (\text{D.5})$$

We can rewrite (D.4) into

$$\mathbb{E}_t^* [R_{m,t \rightarrow N}^\gamma \log R_{m,t \rightarrow N}] \geq \mathbb{E}_t^* [R_{m,t \rightarrow N}^\gamma] \mathbb{E}_t^* [\log R_{m,t \rightarrow N}]. \quad (\text{D.6})$$

Inequality (D.6) now follows from Lemma D.1. \blacksquare

I mimic the steps above to show that the physical distribution differs more from the risk-neutral distribution at every point in the support, whenever risk aversion is increasing. As before, the dependence of the physical measure on γ is omitted.

Proposition D.3. *Assume that a representative investor has CRRA utility, with $\gamma \geq 0$ and $\mathbb{E}_t [R_{m,t \rightarrow N}^\gamma \log R_{m,t \rightarrow N}] < \infty$, then $F_t(x)$ is non-increasing in γ . In particular, $\tau - F_t(\tilde{Q}_{t,\tau})$ is non-decreasing in γ .*

Proof. I start from the relation

$$F_t(x) = \tilde{\mathbb{E}}_t \left[\frac{R_{m,t \rightarrow N}^\gamma}{\tilde{\mathbb{E}}_t [R_{m,t \rightarrow N}^\gamma]} \mathbb{1}(R_{m,t \rightarrow N} \leq x) \right].$$

From first order conditions, we need to show that

$$\begin{aligned} \tilde{\mathbb{E}}_t [\log(R_{m,t \rightarrow N}) \mathbb{1}(R_{m,t \rightarrow N} \leq x) R_{m,t \rightarrow N}^\gamma] \tilde{\mathbb{E}}_t [R_{m,t \rightarrow N}^\gamma] &\leq \\ \tilde{\mathbb{E}}_t [R_{m,t \rightarrow N}^\gamma \mathbb{1}(R_{m,t \rightarrow N} \leq x)] \tilde{\mathbb{E}}_t [\log(R_{m,t \rightarrow N}) R_{m,t \rightarrow N}^\gamma] &. \end{aligned}$$

Using the same change of measure as in (D.5), we obtain the equivalent statement

$$\mathbb{E}_t^* [\log(R_{m,t \rightarrow N}) \mathbb{1}(R_{m,t \rightarrow N} \leq x)] \leq \mathbb{E}_t^* [\mathbb{1}(R_{m,t \rightarrow N} \leq x)] \mathbb{E}_t^* [\log R_{m,t \rightarrow N}].$$

This inequality holds, since $\log(y)$ and $\mathbb{1}(y \leq x)$ are respectively increasing and non-increasing in y , hence the result follows from Lemma D.1. Using the substitution $x \rightarrow \tilde{Q}_{t,\tau}$, it follows that $\tau - F_t(\tilde{Q}_{t,\tau})$, is non-decreasing in γ . ■

D.3 Exponential utility

Here, I assume that the representative agent has exponential utility, $u(x) = 1 - e^{-\gamma^* x}$, where γ^* is the absolute risk aversion. According to Chabi-Yo and Loudis (2020, Equation (55)), the following expression for the equity premium obtains

$$\mathbb{E}_t [R_{m,t \rightarrow N}] - R_{f,t \rightarrow N} = \frac{\tilde{\mathbb{E}}_t [R_{m,t \rightarrow N} e^{\gamma R_{m,t \rightarrow N}}]}{\tilde{\mathbb{E}}_t [e^{\gamma R_{m,t \rightarrow N}}]} - R_{f,t \rightarrow N},$$

where $\gamma = \gamma^* W_t$ is relative risk aversion and W_t represents the agent's wealth at time t . Since there is a one-to-one relation between γ and γ^* , it follows from the results in Section D.2 that the equity premium is increasing in γ^* and so is the distance between the physical and risk-neutral distribution, as measured by: $\tau - F_t(\tilde{Q}_{t,\tau})$.

E Risk-adjustment in the Data and Robustness of the Lower Bound

E.1 Risk-adjustment term in the data

In the empirical application, I compute the risk-adjustment term, $RA_{t,\tau} = LB_{t,\tau}/\tilde{f}_t(\tilde{Q}_{t,\tau})$, for the 30, 60 and 90 day horizon. Table E1 contains summary statistics of $RA_{t,\tau}$. The risk-adjustment term is right-skewed and is most significant for the 5th and 10th percentile. Moreover, over the 30 day horizon, it can spike up to 25% and averages to about 1% in the far left-tail. In annual units, this average implies that the risk-neutral and physical quantile differ by 11%, when $\tau = 0.05$.

E.2 Robustness of the risk-adjustment term and risk-neutral quantile

The risk-adjustment term, $RA_{t,\tau}$, tries to capture the difference between the physical and risk-neutral quantile in the left-tail. What are some other measures that are available at a daily frequency and contain information about the quantile wedge? One candidate is the VIX index, which is defined as

$$VIX_t^2 = \frac{2R_{f,t \rightarrow N}}{N} \left[\int_0^{F_t} \frac{1}{K^2} Put_t(K) dK + \int_{F_t}^{\infty} \frac{1}{K^2} Call_t(K) dK \right],$$

Table E1: **Summary statistics risk-adjustment term**

	Horizon (in days)	Mean	Median	Std. dev.	Min	Max
$\tau = 0.05$	30	0.92	0.63	1.07	0.08	24.38
	60	1.81	1.31	1.67	0.10	19.23
	90	2.65	2.02	2.02	0.02	18.63
$\tau = 0.1$	30	0.70	0.45	0.87	0.06	12.22
	60	1.71	1.19	1.66	0.25	19.89
	90	2.86	2.12	2.32	0.04	24.47
$\tau = 0.2$	30	0.47	0.25	0.74	0.04	10.93
	60	1.14	0.69	1.50	0.12	23.57
	90	1.97	1.22	2.33	0.26	28.92

Note: This table reports summary statistics of the risk-adjustment term, $\text{RA}_{t,\tau} = \text{LB}_{t,\tau}/\tilde{f}_t(\tilde{Q}_{t,\tau})$, in (5.15) at different time horizons and different quantile levels over the sample period 2003-2021. All statistics are in percentage point.

where N is the time to expiration, F_t is the forward price on the S&P500, and $\text{Put}_t(K)$ (resp. $\text{Call}_t(K)$) is the put (resp. call) option price on the S&P 500 with strike K . [Martin \(2017\)](#) shows that VIX measures risk-neutral entropy

$$\text{VIX}_t^2 = \frac{2}{N} \tilde{L}_t \left(\frac{R_{m,t \rightarrow N}}{R_{f,t \rightarrow N}} \right),$$

where entropy is defined as $\tilde{L}_t(X) := \log \tilde{\mathbb{E}}_t[X] - \tilde{\mathbb{E}}_t[\log X]$. Entropy, just like variance, is a measure of spread in the distribution. However, entropy places more weight on left-tail events than variance, since entropy places more weight on out-of-the money puts. As such, VIX is a good candidate to explain potential differences between $Q_{t,\tau}$ and $\tilde{Q}_{t,\tau}$. Second, the Chicago Board Options Exchange provides daily data on VIX for the 30 day horizon.

Table E2 shows the result of the quantile regression

$$Q_{t,\tau}(R_{m,t \rightarrow N}) - \tilde{Q}_{t,\tau}(R_{m,t \rightarrow N}) = \beta_0(\tau) + \beta_1(\tau)\text{RA}_{t,\tau} + \beta_{\text{VIX}}(\tau)\text{VIX}_t. \quad (\text{E.1})$$

We see that β_{VIX} is marginally significant in the left-tail. In contrast, $\beta_1(\tau)$ is even more significant compared to Table 4. Furthermore, the explanatory power of the model that only includes VIX is lower compared to the model that only includes $\text{RA}_{t,\tau}$ (Table 4).

As a second robustness check, I consider how well the direct quantile forecast,

Table E2: **Quantile regression lower bound and VIX**

	$\hat{\beta}_0(\tau)$	$\hat{\beta}_1(\tau)$	$\hat{\beta}_{\text{VIX}}(\tau)$	$R^1(\tau)[\%]$	$R^1(\tau)[\%]$ (VIX only)
$\tau = 0.05$	-0.20 (2.021)	10.09 (0.345)	-0.30 (0.143)	6.34	5.51
$\tau = 0.1$	-0.35 (1.515)	5.06 (0.310)	-0.22 (0.100)	3.41	2.84
$\tau = 0.2$	-0.28 (1.112)	3.62 (0.265)	-0.25 (0.079)	0.61	0.18

Note: This table reports the QR estimates of (E.1) over the 30-day horizon. The sample period is 2003-2021, standard errors are shown in parentheses and calculated using SETBB with a block length of 5 times the forecast horizon. $R^1(\tau)$ denotes the goodness-of-fit measure (2.5). The last column denotes the goodness-of-fit in the model that only uses VIX as covariate. The standard error and point estimate of β_0 is multiplied by 100 for readability.

$\hat{Q}_{t,\tau} = \tilde{Q}_{t,\tau} + \text{RA}_{t,\tau}$, compares to the VIX forecast. Since $\hat{Q}_{t,\tau}$ does not require any parameter estimation, this exercise is a measure of out-of-sample performance. However, VIX does not directly measure $Q_{t,\tau}$ and hence I use an expanding window to obtain the VIX benchmark: $\hat{Q}_{t,\tau}^{\text{VIX}} := \hat{\beta}_0(\tau) + \hat{\beta}_1(\tau)\text{VIX}_t$. Finally, I use the following out-of-sample metric to compare both forecasts

$$R_{oos}^1(\tau) = 1 - \sum_{t=500}^T \rho_\tau(R_{m,t \rightarrow N} - \hat{Q}_{t,\tau}) / \sum_{t=500}^T \rho_\tau(R_{m,t \rightarrow N} - \hat{Q}_{t,\tau}^{\text{VIX}}).$$

Notice that $R_{oos}^1(\tau) > 0$, if $\hat{Q}_{t,\tau}$ attains a lower error than $\hat{Q}_{t,\tau}^{\text{VIX}}$. This exercise is more ambitious, since $\hat{Q}_{t,\tau}^{\text{VIX}}$ makes use of in-sample information. Nonetheless, Figure E3 shows that $\hat{Q}_{t,\tau}$ outperforms the VIX predictor at all quantile levels.

Figure E4 performs a similar exercise in the right-tail, but using $\tilde{Q}_{t,\tau}$ instead of $\hat{Q}_{t,\tau}$, since Table 1 shows that the risk-neutral quantile is a good approximation to $Q_{t,\tau}$ in the right-tail. We see that $\tilde{Q}_{t,\tau}$ outperforms $\hat{Q}_{t,\tau}^{\text{VIX}}$ at all quantile levels. Hence, the risk-neutral approximation in the right-tail is more accurate than using the in-sample VIX measure.

E.3 Lower bound in Black-Scholes model

This section illustrates the accuracy of the quantile approximation in (5.10) in a discretized version of the Black-Scholes model with time varying parameters. Specif-

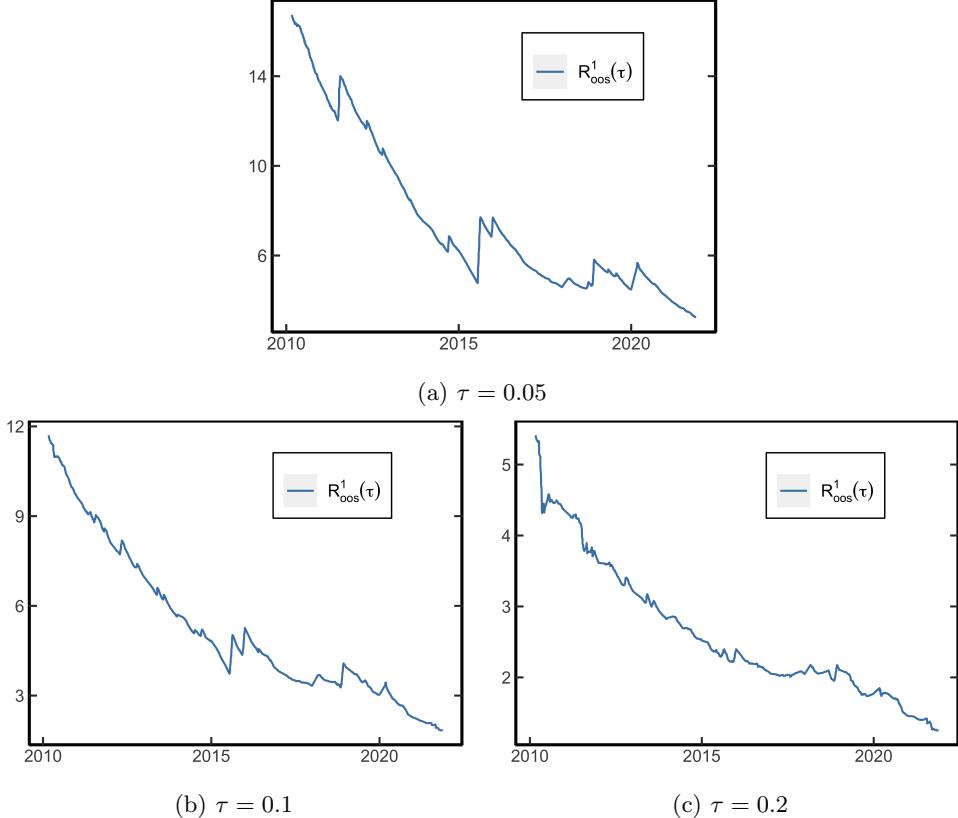


Figure E3: Out-of-sample forecast using risk-adjusted quantile with VIX benchmark. This figure shows the cumulative out-of sample $R^1(\tau)$, defined as $R^1_{oos}(\tau) = 1 - \sum_{t=500}^T \rho_\tau(R_{m,t \rightarrow N} - \hat{Q}_{t,\tau}) / \sum_{t=500}^T \rho_\tau(R_{m,t \rightarrow N} - \hat{Q}_{t,\tau}^{\text{VIX}})$, where $\hat{Q}_{t,\tau} = \tilde{Q}_{t,\tau} + \text{RA}_{t,\tau}$, $\tilde{Q}_{t,\tau}^{\text{VIX}} = \hat{\beta}_0(\tau) + \hat{\beta}_1(\tau) \cdot \text{VIX}_t$, and $\hat{\beta}_0(\tau), \hat{\beta}_1(\tau)$ are the regression estimates from a quantile regression of $R_{m,t \rightarrow N}$ on VIX_t , using data only up to time t . The horizon is 30 days and the QR estimates are dynamically updated using an expanding window over the period 2003–2021. The initial sample uses 500 observations.

ically, I assume the following DGP

$$R_{m,t \rightarrow N} = \exp \left((\mu_t - \frac{1}{2} \sigma_t^2) N + \sigma_t \sqrt{N} Z_{t+N} \right), \quad Z_{t+N} \sim \mathcal{N}(0, 1) \quad (\text{E.2})$$

$$\sigma_t \sim \mathbf{Unif}[0.05, 0.35]$$

$$\mu_t \sim \mathbf{Unif}[-0.02, 0.2].$$

The return distribution under risk-neutral dynamics is given by

$$\tilde{R}_{m,t \rightarrow N} = \exp \left((r_t - \frac{1}{2} \sigma_t^2) N + \sigma_t \sqrt{N} Z_{t+N} \right) \quad (\text{E.3})$$

$$r_t \sim \mathbf{Unif}[0, 0.03]. \quad (\text{E.4})$$

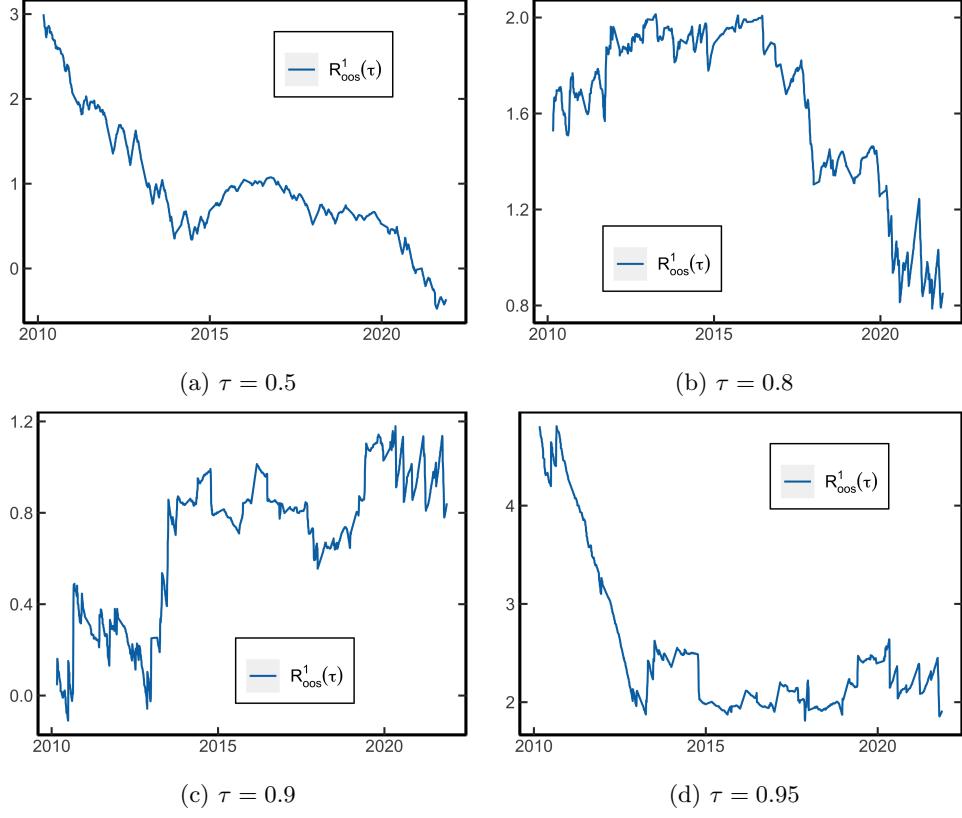


Figure E4: Out-of-sample forecast using risk-neutral quantile with VIX benchmark. This figure shows the cumulative out-of sample $R^1(\tau)$, defined as $R^1_{oos}(\tau) = 1 - \sum_{t=500}^T \rho_\tau(R_{m,t \rightarrow N} - \tilde{Q}_{t,\tau}) / \sum_{t=500}^T \rho_\tau(R_{m,t \rightarrow N} - \tilde{Q}_{t,\tau}^{VIX})$, where $\tilde{Q}_{t,\tau}^{VIX} = \hat{\beta}_0(\tau) + \hat{\beta}_1(\tau) \cdot VIX_t$, and $\hat{\beta}_0(\tau), \hat{\beta}_1(\tau)$ are the regression estimates from a quantile regression of $R_{m,t \rightarrow N}$ on VIX_t , using data only up to time t . The horizon is 30 days and the QR estimates are dynamically updated using an expanding window over the period 2003–2021. The initial sample uses 500 observations.

Finally, assume that all parameters are IID over time and that options are priced according to the Black-Scholes formula, conditional on time t . In this setup, it is fruitless to use historical data to predict future quantiles, since parameters change unpredictably over time. We use $N = 30$ to mimic the monthly application in this paper. It is assumed that the risk-neutral quantile function is known at the start of period t , as it is in the real world, by the result of [Breeden and Litzenberger \(1978\)](#). I use the risk-neutral quantile function to calculate $LB_{t,\tau}$ at time t . Then, following the approximation in (5.10), the physical quantile function is estimated by

$$\hat{Q}_{t,\tau} = \tilde{Q}_{t,\tau} + RA_{t,\tau}. \quad (\text{E.5})$$

We take 3,000 return observations that are generated according to (E.2). This exercise is repeated 1,000 times. To assess the accuracy of the approximation in (E.5), I

use several metrics. For every sample, I estimate a quantile regression of the form

$$Q_\tau(R_{t+1}) = \beta_0(\tau) + \beta_1(\tau)\hat{Q}_{t,\tau},$$

where $\hat{Q}_{t,\tau}$ comes from (E.5). The first two columns in Table E3 report the average values of the QR estimates across the 1,000 simulations. The means are rather close to 0 and 1 respectively for all quantiles. If (5.4) is a good approximation, one expects $Q_{t,\tau} > \hat{Q}_{t,\tau}$, since $\text{LB}_{t,\tau} \leq \tau - F_t(\tilde{Q}_{t,\tau})$. The third column in Table E3 shows this happens for the majority of samples. The fourth column shows the correlation between $Q_{t,\tau}$ and $\hat{Q}_{t,\tau}$, which is very close to one, and corroborates the view that the approximation is quite accurate. Columns four and five document the percentage of non rejection of H_0 , which is indeed quite high. The last column considers non rejection of the joint null hypothesis, which is also high except for the 10th percentile. Overall, Table E3 suggests that (E.5) is a highly accurate predictor of the physical quantile function.

Table E3: Simulation results

	$\mathbb{E}\hat{\beta}_0(\tau)$	$\mathbb{E}\hat{\beta}_1(\tau)$	$Q > \hat{Q}$	$\rho(Q, \hat{Q})$	$\hat{\beta}_0(\tau) = 0$	$\hat{\beta}_1(\tau) = 1$	$[\hat{\beta}_0(\tau), \hat{\beta}_1(\tau)] = [0, 1]$
$\tau = 0.01$	0.01	0.99	0.85	1	0.94	0.96	0.80
$\tau = 0.05$	-0.03	1.04	0.69	0.99	0.90	0.89	0.66
$\tau = 0.1$	-0.06	1.07	0.64	0.99	0.78	0.76	0.47

Note: $\mathbb{E}\hat{\beta}_0(\tau)$ denotes the average QR estimate of $\hat{\beta}_0(\tau)$ and likewise $\mathbb{E}\hat{\beta}_1(\tau)$ shows it for $\hat{\beta}_1(\tau)$. $Q > \hat{Q}$ shows the fraction of times the true physical quantile is larger than our predicted quantile. Columns $H_0 : \hat{\beta}_0(\tau) = 0$ and $H_0 : \hat{\beta}_1(\tau) = 1$ report the fraction of times the individual null hypotheses $\beta_0(\tau) = 0, \beta_1(\tau) = 1$ are not rejected. The last column reports the fraction of times the joint null hypothesis is not rejected.

Example E.1. I illustrate the approximation (5.4) in the Black-Scholes model with fixed parameters: $N = 365$ (one year), $\mu = 0.08, r = 0.02, \sigma = 0.2$.³⁴ We can explicitly calculate F^{-1}, \tilde{F}^{-1} and \tilde{f} owing to the lognormal assumption. Figure E5 shows the risk-neutral quantile function, the approximation (5.4) and the true physical quantile function. Observe that the approximation (5.4) is very accurate in this case.

E.4 Bias in quantile regression

In the empirical application, we have to estimate $\tilde{Q}_{t,\tau}, \tilde{f}(\cdot)$ and $\text{LB}_{t,\tau}$ from market data. Therefore, the estimated coefficients in the quantile regression are biased due to measurement error in the covariate. I present simulation evidence which shows that the bias is small in finite samples.

³⁴For illustrative purposes, I use $N = 1$, instead of $N = 1/12$, otherwise the physical quantile function and its approximation are indistinguishable.

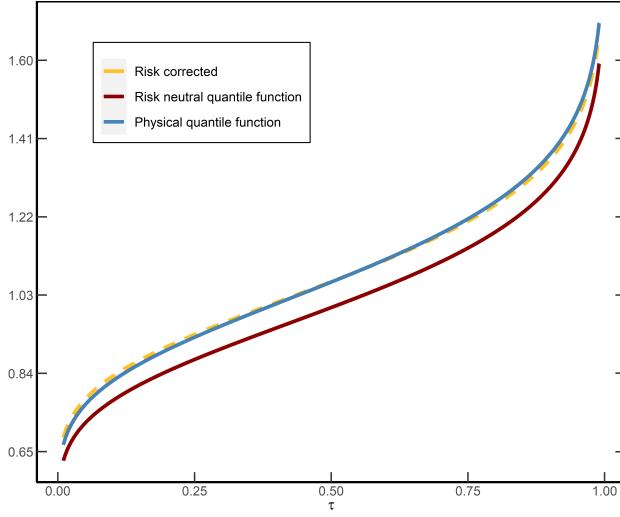


Figure E5: **Quantile approximation in the Black-Scholes model.** This figure shows that quantile approximation (E.5) in the Black-Scholes model with parameters: $\mu = 0.08, r = 0.02, \sigma = 0.2$ over a one-year time horizon.

The setup is as follows. I simulate returns according to model (E.2) and assume that options are priced according to the [Black and Scholes \(1973\)](#) formula at the start of period t . We want to calculate the risk-adjustment term for a maturity of 90 days. As in the empirical application, I assume that options with an exact 90 day maturity are not available, but instead we observe options with maturity 85 and 97 days. I generate a total of 1,000 options every time period with maturities randomly sampled from 85 and 97 days.³⁵ These numbers are roughly consistent with the latter part of the empirical sample. The procedure is repeated for a total of 1,000 time periods. For the entire sample, I compare the estimated and analytical risk-adjustment term, which are given by respectively

$$\begin{aligned} \text{RA}_{t,\tau}^e &:= \widehat{\tilde{Q}}_{t,\tau} + \frac{\widehat{\text{LB}}_{t,\tau}}{\widehat{f}_t(\widehat{\tilde{Q}}_{t,\tau})} \\ \text{RA}_{t,\tau}^a &:= \widetilde{Q}_{t,\tau} + \frac{\text{LB}_{t,\tau}}{\widetilde{f}_t(\widetilde{Q}_{t,\tau})}. \end{aligned}$$

The hats in $\text{RA}_{t,\tau}^e$ signify that the risk-neutral quantile, PDF and lower bound are estimated from the available options at time t , using the procedure in Appendix B.2. The terms in $\text{RA}_{t,\tau}^a$ are obtained from the known analytical expression of the

³⁵So on average there will 500 options with maturity 85 days and 500 with maturity 97 days.

risk-neutral distribution (recall (E.3))). I then use QR to estimate the models

$$Q(R_{t+1}) = \widehat{\beta}_0(\tau) + \widehat{\beta}_{1,e}(\tau)RA_{t,\tau}^e$$

$$Q(R_{t+1}) = \widehat{\beta}_0(\tau) + \widehat{\beta}_{1,a}(\tau)RA_{t,\tau}^a.$$

I use the ratio $\widehat{\beta}_{1,e}/\widehat{\beta}_{1,a}$ to measure the relative bias in the sample. This experiment is repeated 500 times to get a distribution of the relative bias. Figure E6 shows boxplots of the bias for several quantiles. We see that the relative bias is very small and centered around 1. Hence, the error in measurement problem resulting from estimating the risk-adjustment term is limited in this case.

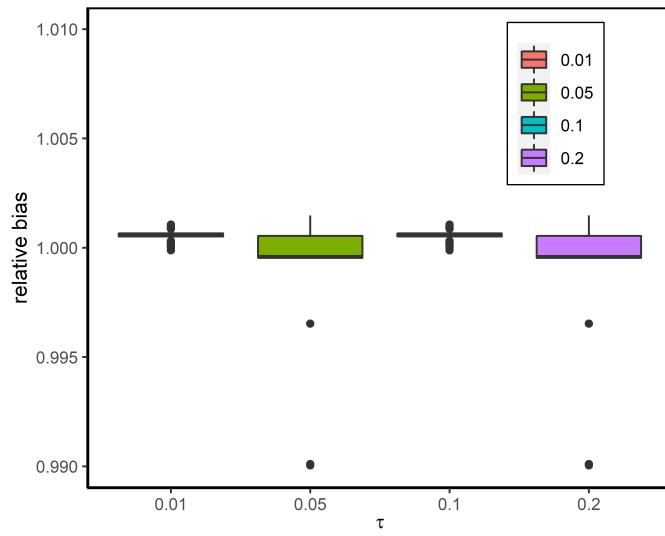


Figure E6: **Bias in QR resulting from measurement error.** This boxplot shows the relative bias in the quantile regression estimate as a result of measurement error.

F Additional Figures

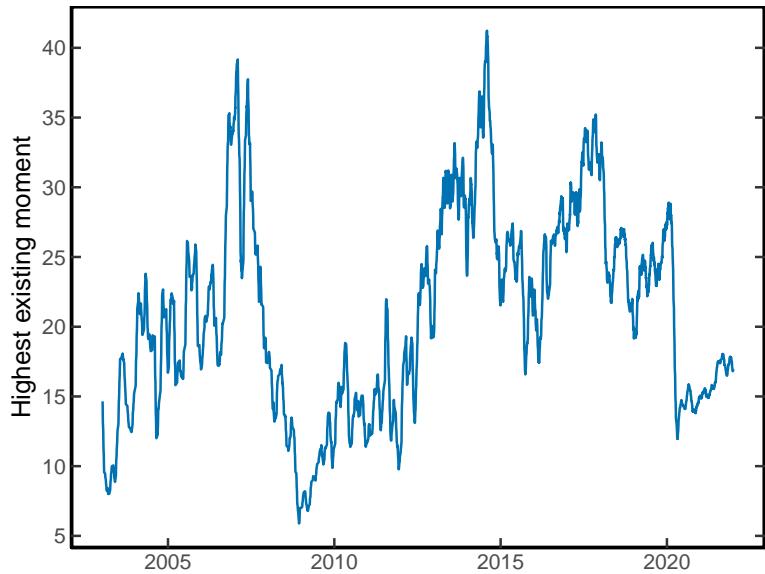


Figure F7: Highest existing risk-neutral moment for 30-day returns. This figure shows $p_t^* := \sup\{p : \tilde{\mathbb{E}}_t(R_{m,t \rightarrow N}^p) < \infty\}$ over time, where $R_{m,t \rightarrow N}$ represents the 30-day return. p_t^* is calculated from the moment formula of Lee (2004), $p_t^* = \frac{1}{2\beta_R} + \frac{\beta_R}{8} + \frac{1}{2}$, where $\beta_R = \limsup_{x \rightarrow \infty} \frac{\sigma_{IV}^2(x)}{|x|/N}$, $\sigma_{IV}(x)$ is the implied volatility at log-moneyness $x = \log(K/(e^{rN}S_0))$, and $N = 30/365$ is the time horizon. β_R is estimated from the call option with highest available strike price. The figure is smoothed using a 30-day moving average.