Quantile Approach to Asset Pricing Models

Tjeerd de Vries *

October 11, 2021

Abstract

This paper develops a generalization of the Hansen-Jagannathan bound that incorporates information beyond the mean and variance of returns. The resulting bound compares the physical and risk neutral distribution for every τ-quantile, where $\tau \in (0, 1)$. An empirical application with S&P500 return data shows that the new bound is stronger than the Hansen-Jagannathan bound for small values of τ. The long run risk model cannot reconcile this feature of the data, due to the absence of disaster risk. I extend this finding using conditioning information and document that disaster risk is time-varying, using a semiparametric approach. I also propose a new measure of quantile forecastability and show that many stylized facts about the equity premium carry over to the quantile setting.

Keywords: Asset pricing, Stochastic discount factor, 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 and Brendan Beare for useful feedback. All errors are my own.

1 Introduction

Nonparametric methods are useful for analyzing misspecification of asset pricing models in the macro-finance literature (Hansen and Jagannathan (1991), Snow (1991), Stutzer (1995), Bansal and Lehmann (1997), Alvarez and Jermann (2005), Backus et al. (2014) and Liu (2020)). The underlying theme of these methods is to estimate a statistic of the observable asset returns, such as the Sharpe ratio, and use this statistic to bound unobservable moments of the stochastic discount factor (SDF). The seminal paper of Hansen and Jagannathan (1991) shows how the Sharpe ratio leads to a lower bound on the volatility of the SDF. Hansen and Jagannathan (1991) (henceforth HJ) conclude that any SDF needs to be sufficiently volatile to explain historically high Sharpe ratios. This poses a challenge to consumption based asset pricing models, since historical consumption growth in the US is smooth. A typical problem which emerges is that, in order to overcome the HJ bound, one has to impose outrageous levels of risk aversion and this contradicts other stylized facts of the data (Cochrane, 2005).

However, since the HJ bound uses only the mean and variance of excess returns, there might be other information contained in the data which can be exploited to obtain sharper conclusions about the SDF volatility and misspecification of asset pricing models. This paper fills that gap, by developing moment bounds on the SDF which use information from the entire distribution of asset returns. The new bound on the SDF volatility is motivated as a Sharpe ratio on digital put options, that is, a derivative contract paying out one dollar if the stock falls below a specified strike price. By varying the strike, we obtain a continuum of bounds, which compares the physical and risk neutral distribution for every τ-quantile, where $\tau \in (0, 1)$. For this reason, the new bound is referred to as the quantile bound.

The benefits of this approach are threefold. (i) The quantile bound is valid and performs well even if the distribution of returns is heavy tailed (in contrast to the HJ bound); a point I illustrate in Section 2.1. (ii) Different consumption based asset pricing models imply different shapes of the quantile bound. In particular, some models, like the long-run risk model of Bansal and Yaron (2004), predict that the HJ bound is stronger than the quantile bound. I find counter

evidence to this implication in the data (Section 4). I go on to argue that this can be interpreted as a failure of the model to incorporate disaster risk. (iii) The shape of the quantile bound provides a cleaner interpretation of disaster risk. Specifically, adding disaster risk to asset pricing models not only increases the market Sharpe ratio, but it also induces a peak in the quantile bound for small τ. I find support for this hypothesis in the data. This finding lends further support to the idea that a fear of disasters is driving expected returns, in contrast to, for example, changing expectations about economic growth as predicted by the long-run risk model.

The analysis above exploits the difference between the unconditional physical and risk neutral quantile. Complementary to this, we can ask how the quantile wedge evolves using conditional information? In contrast to the conditional equity premium, the literature offers little guidance on the predictors of this quantile wedge. I develop a new measure which, under mild economic constraints, lower bounds the quantile wedge using option data. Quantile regression estimates indicate that the lower bound provides a tight approximation to the latent quantile wedge. Several insights can be gained from this exercise, namely that the quantile wedge fluctuates significantly over time, it spikes during financial crises and it is hard to predict. All these properties are natural analogs to the corresponding characteristics of the conditional equity premium (Goyal and Welch, 2008).

It is, however, not a trivial exercise to extend the equity premium result to the quantile setting. In the former case, a notion like the out-of-sample R^2_{oos} of Campbell and Thompson (2008) provides a natural measure of predictive performance. In addition, the low predictability of the equity premium can be understood as the absence of near arbitrage opportunities, since predictable returns can be leveraged to obtain high conditional Sharpe ratios. A second contribution of this paper is to carry over these results to the quantile setting. Specifically, I propose a clean substitute for R^2_{oos}, called $R^1_{oos}(\tau)$, which is a measure of out-of-sample predictability tailored to quantile regression. I find that that the out-of-sample predictability of return quantiles is in the order of several percentage points. I argue that low predictability of quantiles is unsurprising if a rich option market exists, by showing that certain investment

strategies can exploit quantile predictability. This is anchored to the idea that near arbitrage opportunities are too good to be true (Cochrane and Saa-Requejo, 2000).

The conditional results also allow me to get a real time measure of the forward looking premium on crash risk. The crash risk premium measures the (conditional) expected return on buying a security that pays out 1\$, in case the market return is below some threshold α. For small α, this is a direct measure of the influence of disaster risk on the economy. I document that the crash risk premium fluctuates significantly over time and spikes during financial crises. This is a feature which standard IID disaster risk models cannot reconcile and it provides support for models that incorporate time-varying disaster risk, such as Wachter (2013) or Gabaix (2012). An attractive feature of my approach is that results are robust to misspecification due to the semiparametric nature of measuring the crash risk premium. Moreover, since I use forward looking option data, the approach is not subject to the historical in-sample bias critique of Goyal and Welch (2008).

1.1 Literature review

Hansen and Jagannathan (1991) prove a nonparametric bound on the SDF volatility and use it to establish a duality relation with the maximum Sharpe ratio. Many researchers followed up with higher order bounds (Snow (1991); Almeida and Garcia (2012); Liu (2020)) and entropy bounds (Stutzer (1995); Bansal and Lehmann (1997); Alvarez and Jermann (2005); Backus et al. (2014)).

The idea in this paper to use derivatives to obtain stronger SDF bounds dates back to Ross (1976) and Breeden and Litzenberger (1978). Their work shows that options can complete the market and the risk-neutral PDF can be obtained as the second derivative of the call option price curve. This idea was put to practice by Aït-Sahalia and Lo (1998, 2000) and Jackwerth (2000) to estimate the SDF or risk-aversion in a nonparametric way. Even though the approach of Aït-Sahalia and Lo (1998, 2000) renders an estimate of the entire SDF, it requires estimating a ratio of two density functions, which is particularly hard to estimate in the tails. The results proved in this paper yield a bound on certain moments of the SDF, but we only require an estimate of a CDF and

quantile function, which, in general, can be estimated at a faster rate.

The results in this paper affirm that the highest Sharpe ratio is not attained by the market portfolio. Instead, I find that it is more profitable to engage in (digital) put selling strategies, which is consistent with Coval and Shumway (2001) and Bates (2008). Bates (2008) proposes a structural model to rationalize this finding using crash-averse preferences. I provide an alternative statistical rationale, leveraging on the empirical observation that asset returns are heavy tailed (Danielsson and De Vries, 2000). In particular, imposing a Pareto distribution on asset returns with a sufficiently fat tail implies the existence of put selling strategies that yield higher Sharpe ratios than a direct investment in the market portfolio.

This paper also connects to the burgeoning literature on using options to obtain forward looking estimates of the equity premium (Martin (2017), Chabi-Yo and Loudis (2020)). However, instead of focusing on the conditional expectation of excess returns, I use option data to predict conditional return quantiles. The evaluation and performance of conditional return predictors is well understood in the literature, especially after fundamental contributions of Goyal and Welch (2008) and Campbell and Thompson (2008). To extend this to the quantile setting, I draw on earlier work of Koenker and Machado (1999).

The relation between option data and expected market crashes has a long history and features in the work of Bates (1991, 2000, 2008), Bollerslev and Todorov (2011), Backus et al. (2011) and Ross (2015). Few papers further consider the premium that is associated to crash risk, which is analyzed at the end of this paper. The evidence we document about time variation in the crash premium and spikes during crises is most closely related to Bates (2000), who documents similar results by fitting a two-factor jump-diffusion model. In contrast, the results in this paper do not require distributional assumptions about the market return, since we leverage on the semiparametric approach of Chabi-Yo and Loudis (2020). Our empirical observation also complements theoretical work of Rietz (1988), Barro (2006), Gabaix (2012) and Wachter (2013) about the time varying influence of disasters on asset returns.

The rest of this paper is organized as follows. Section 2 discusses the quantile bound, Section 3 outlines how to estimate the new bound with actual data and

Section 4 describes the main empirical results. Section 5 extends the empirical results using conditioning information. Finally, Section 6 concludes.

2 Quantile bound on the SDF volatility

Let R be the return on any tradable asset and their portfolios. The assumption of no arbitrage and the law of one price imply the existence of stochastic discount factor (SDF), denoted by M, such that the following relation holds

$$\mathbb{E}\left[MR\right] = 1. \tag{2.1}$$

The expectation in (2.1) is taken with respect to some probability measure $\mathbb{P}$, which represents the true probability distribution of the data generating process. Alternatively, the relation in (2.1) can be reformulated in terms of the risk neutral measure. That is, there exists a probability measure $\widetilde{\mathbb{P}}$, equivalent to $\mathbb{P}$, such that[1]

$$\widetilde{\mathbb{E}}\left[R\right] = \mathbb{E}\left[M\right]. \tag{2.2}$$

The expectation on the left in (2.2) is taken with respect to $\widetilde{\mathbb{P}}$. Throughout the paper, I use tilde ($\sim$) to denote quantities which are calculated under the risk-neutral measure $\widetilde{\mathbb{P}}$. If we assume the existence of a risk free asset with return R_f, then $\mathbb{E}\left[M\right] = 1/R_f$ and we uncover the familiar fact in (2.2) that the risk neutral measure absorbs risk premia. Mathematically, the connection between (2.1) and (2.2) follows since $M/\mathbb{E}\left[M\right]$ is the Radon-Nikodym of the measures $\mathbb{P}$ and $\widetilde{\mathbb{P}}$. The SDF can potentially depend on many state variables. To avoid having to specify or estimate these state variables, I work with the projected SDF

$$M = \mathbb{E}\left[\mathfrak{M}|R\right].$$

Here, $\mathfrak{M}$ is the SDF that depends on all the state variables. The projected SDF has the same pricing implications for contingent claims written on return R for which we have data (Cochrane, 2005, pp. 66–67).[2]

[1] Two probability measures $\mathbb{P}, \widetilde{\mathbb{P}}$ are said to be equivalent whenever $\mathbb{P}(A) = 0 \iff \widetilde{\mathbb{P}}(A) = 0$.

[2] Formally, M is a measurable function of R, but I avoid denoting this dependence explicitly to simplify notation.

As explained by Backus et al. (2011), the pricing kernel M is of primary importance to the macro-finance literature, whereas the risk-neutral measure is the focal point of the option-pricing literature. Backus et al. (2011) combine the two literatures and compare implied disaster probabilities. Hansen and Jagannathan (1991) use returns to estimate a lower bound on the volatility of any SDF. I combine the insights of both approaches to obtain a bound on the SDF volatility that exploits option information and elucidates how any difference between the risk-neutral and physical distribution leads to volatile SDF. Before stating the bound, I introduce $\widetilde{Q}_\tau(R)$ to denote the risk neutral τ-quantile of the return R. By definition, the risk neutral quantile function satisfies

$$\widetilde{\mathbb{P}}\left(R \leq \widetilde{Q}_\tau(R)\right) = \tau.$$

Abusing notation, I use the shorthand $\widetilde{\mathbb{P}}\left(\widetilde{Q}_\tau(R)\right) = \tau$ for simplicity. The new SDF volatility bound is summarized in the following Theorem.

Theorem 2.1 (Quantile bound)**.** *For any* $\tau \in (0,1)$*:*

$$\frac{\sigma(M)}{\mathbb{E}\left[M\right]} \geq \frac{\left|\tau - \mathbb{P}(\widetilde{Q}_\tau(R))\right|}{\sqrt{\mathbb{P}(\widetilde{Q}_\tau(R)) \times (1 - \mathbb{P}(\widetilde{Q}_\tau(R)))}}. \tag{2.3}$$

Remark. If we assume that a risk free asset exists, then $\mathbb{E}\left[M\right] = 1/R_f$ and (2.3) simplifies to

$$\sigma(M) \geq \frac{\left|\tau - \mathbb{P}(\widetilde{Q}_\tau(R))\right|}{\sqrt{\mathbb{P}(\widetilde{Q}_\tau(R)) \times (1 - \mathbb{P}(\widetilde{Q}_\tau(R)))R_f}}. \tag{2.4}$$

Proof. See Appendix A.1. ■

I refer to Theorem 2.1 as the quantile bound, since a key ingredient in (2.3) is the risk neutral quantile function. The quantity $\phi(\tau) = \mathbb{P}(\widetilde{Q}_\tau(R))$ can be interpreted as the ordinal dominance curve of the measures $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ (Hsieh et al., 1996). 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 2.1 makes precise the sense in which any discrepancy between the physical and risk neutral distribution induces SDF volatility. Compare this to

the classical HJ bound:

$$\frac{\sigma(M)}{\mathbb{E}[M]} \geq \frac{|\mathbb{E}[R] - 1/\mathbb{E}[M]|}{\sigma(R)}. \tag{2.5}$$

The lower bound in (2.5) tells us that any excess return induces SDF volatility. Essentially, (2.5) 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. One can imagine that other statistics of the data can be exploited as well to learn something about the SDF. This is precisely what (2.3) does, by comparing the physical and risk neutral distribution for every τ-quantile. In this way, the quantile bound picks up other non-linearities in the data such as skewness or other higher order cumulants. Moreover, the quantile bound is well defined regardless of any moment restrictions and thus robust to fat-tails. Both fat-tails and skewness are essential features of financial return data (Martin and Gao, 2021) and hence the quantile bound could offer useful information beyond that contained in the HJ bound, which is confined to the mean-variance paradigm.

The Sharpe ratio on the right hand side of (2.5) summarizes the risk-return trade-off of a mean-variance optimizing agent. Analogously, the quantile bound can be understood as the Sharpe ratio of insurance against crash risk for small quantiles. Consider a security which pays out 1\$, whenever the asset return is below $\widetilde{Q}_\tau(R)$, for some small τ. The price of such a security is given by

$$\mathbb{E}[M]\,\widetilde{\mathbb{E}}\left[\mathbb{1}\left(R \leq \widetilde{Q}_\tau(R)\right)\right] = \mathbb{E}[M]\,\tau. \tag{2.6}$$

Similarly, the (discounted) expected return is

$$\mathbb{E}[M]\,\mathbb{E}\left[\mathbb{1}\left(\widetilde{Q}_\tau(R)\right)\right] = \mathbb{E}[M]\,\mathbb{P}(\widetilde{Q}_\tau(R)). \tag{2.7}$$

And the risk associated to this investment is given by

$$\sigma\left(\mathbb{1}\left(R \leq \widetilde{Q}_\tau(R)\right)\right) = \sqrt{\mathbb{P}(\widetilde{Q}_\tau(R))(1 - \mathbb{P}(\widetilde{Q}_\tau(R)))}. \tag{2.8}$$

Combining (2.6), (2.7) and (2.8) to form the Sharpe ratio we recover (2.3).

The interpretation above sheds light on the economic interpretation of the

two bounds. For example, the HJ bound is used to motivate the equity premium puzzle. That is, standard consumption based asset pricing models have difficulty to overcome the HJ bound for reasonable levels of risk aversion, due to historical high returns on the market portfolio (Cochrane, 2005). The disaster risk model of Rietz (1988) and Barro (2006) argues that this because investors are fearful of extreme (negative) shocks to consumption. Following this reasoning, we would expect high Sharpe ratios for small quantiles in (2.3) due to the crash risk interpretation sketched above. In this sense, (2.3) is a more direct measure of the influence of disasters on the behavior of the SDF. I analyze this further in Section 2.2.

2.1 An illustrative example

To illustrate how the new bound in Theorem 2.1 compares to the HJ bound, I consider a setup with heavy-tailed returns. For ease of exposition, I assume that a risk free asset exists with return R_f. Let $U \sim \mathbf{UNIF}[0,1]$ (Uniform distribution on [0,1]) and consider the following specification

$$M = AU^{\alpha}, \;\; R = BU^{-\beta} \qquad \text{with} \qquad \alpha, \beta, A, B > 0.$$

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

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

Under physical measure $\mathbb{P}$, the distribution of returns is $R \sim \mathbf{PAR}\left(B, \frac{1}{\beta}\right)$, since

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

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

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

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 and variance are respectively given by

$$\mathbb{E}[M] = \frac{A}{\alpha+1} \quad \text{and} \quad \sigma^2(M) = \frac{A^2}{2\alpha+1} - \frac{A^2}{(\alpha+1)^2}.$$

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

$$\frac{AB}{\alpha-\beta+1} = 1. \tag{2.9}$$

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

$$\frac{A}{\alpha+1} = \frac{1}{R_f}. \tag{2.10}$$

In this case, the Sharpe ratio on the asset return is given by

$$\frac{\mathbb{E}[R] - R_f}{\sigma(R)} = \frac{\frac{B}{1-\beta} - \frac{\alpha+1}{A}}{\sqrt{\frac{B^2}{1-2\beta} - \left(\frac{B}{1-\beta}\right)^2}}. \tag{2.11}$$

I show that the quantile bound in this environment is stronger than the HJ bound under two different calibrations. To understand the intuition behind this result, I summarize some key properties of the model.

Lemma 2.2. *In the setup described above, the following properties hold:*

(i) Under $\widetilde{\mathbb{P}}$, $R \sim \boldsymbol{PAR}\left(B, \frac{\alpha+1}{\beta}\right)$.

(ii) The quantile bound depends only on the (left) tail index α *of* M. *In par-*

ticular,

$$\frac{1}{R_f}\frac{\left|\tau-\mathbb{P}(R\leq\widetilde{Q}_\tau)\right|}{\sqrt{\mathbb{P}(R\leq\widetilde{Q}_\tau)(1-\mathbb{P}(R\leq\widetilde{Q}_\tau))}}=\frac{A}{1+\alpha}\frac{\left|\tau-1+(1-\tau)^{\frac{1}{\alpha+1}}\right|}{\sqrt{(1-(1-\tau)^{\frac{1}{\alpha+1}})(1-\tau)^{\frac{1}{\alpha+1}}}}.$$

(iii) If $\beta\uparrow\frac{1}{2}$, then the HJ bound converges to 0.

Proof. (i) Since R_fM is the Radon-Nikodym derivative that induces a change of measure from $\mathbb{P}$ to $\widetilde{\mathbb{P}}$, it follows that

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

The last line follows from (2.10).

(ii) It is easy to show that the quantiles of a $\mathbf{PAR}\left(C,\zeta\right)$ distribution are given by

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

It therefore follows that the risk-neutral quantiles in this example are given by

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

As a result

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

Hence, the quantile bound evaluates to

$$\frac{\left|\tau - \mathbb{P}(R \leq \widetilde{Q}_\tau)\right|}{R_f \sigma\left(\mathbb{1}\left(R \leq \widetilde{Q}_\tau\right)\right)} = \frac{A}{1+\alpha} \frac{\left|\tau - 1 + (1-\tau)^{\frac{1}{\alpha+1}}\right|}{\sqrt{(1-(1-\tau)^{\frac{1}{\alpha+1}})(1-\tau)^{\frac{1}{\alpha+1}}}}. \tag{2.12}$$

(iii). The HJ bound, as given by the Sharpe ratio in (2.11), goes to 0 as $\beta \uparrow 1/2$ since $\sigma(R) \uparrow \infty$. ■

Properties (ii) and (iii) 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 still render a good bound. Moreover, we do not need to impose any restrictions on the parameter space to calculate the quantile bound, whereas the HJ bound requires $\beta < 1/2$. However, the 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).

I now calibrate the model in two different ways to illustrate the difference between the quantile and HJ bound. The first calibration is targeted to match some of the salient features of the US market return. To match the typical Pareto exponent for the US market return, I set $\beta = 1/3$. For simplicity, assume $R_f = 1$ and to match the observed equity premium of 8%, I pick $B = 1.08 \times (1-\beta) = 0.72$. The implied return volatility is far above typical estimates (16% in Cochrane (2005)), but since this example is provided to gain intuition the discrepancy is ignored. Equations (2.9) and (2.10) are used to solve for the parameters A, α from the SDF distribution. In the alternative calibration, I set $\beta = 1/2.2$ (heavier tails, but still finite variance) and pick B again to match the equity premium. Once more, (2.9) and (2.10) are used to solve for A, α. Table 1 summarizes the resulting parameter values for reference, together with the corresponding Sharpe ratio and SDF volatility.

Figure 1 contrasts the HJ bound to the quantile bound. One can see that the supremum of the quantile bound (displayed in red) exceeds the HJ bound in both cases. The display on the right shows that the performance of the quantile

Table 1: Model calibration

	A	α	B	β	$\sigma(R)$	Sharpe ratio	$\sigma(M)$
Calibration 1	1.19	0.19	0.72	0.33	0.62	0.13	0.16
Calibration 2	1.11	0.11	0.59	0.45	1.63	0.05	0.10

Note: *Calibration of SDF model with Pareto returns. Both calibrations impose an equity premium of 8% and (gross) risk-free rate* $R_{f,t+1} = 1$. $\sigma(R)$ *denotes the return volatility and* $\sigma(M)$ *the SDF volatility.*

bound is better whenever the distribution of asset returns is more heavy tailed. In the left display of Figure 1, the supremum of the quantile bound is slightly stronger than the HJ bound, but only marginally so. In contrast, the display on the right of Figure 1 highlights that there is a range of values for which the quantile bound is stronger than the HJ bound, owing to the fatter tails of the return distribution. Figure 2 compares the quantile bound to the HJ bound when β varies. The parameters A, α are solved for implicitly, as in the calibration of Table 1. The graph shows that the quantile bound is better when $\beta \notin [0.19, 0.28]$. Hence, there is a range of values for which the HJ bound improves upon the quantile bound. However, within that range the differences are rather small. In general, the quantile bound seems to track the behavior of the SDF volatility much closer and still delivers reasonable approximations if returns are very heavy tailed ($\beta > 0.28$) or more light tailed ($\beta < 0.19$).

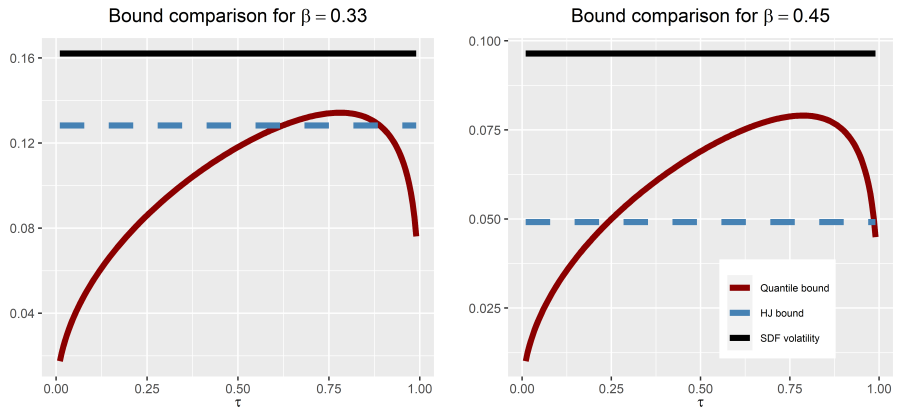

Figure 1: Plots of the quantile bound (in red), HJ bound (blue), true SDF volatility (black) and estimated quantile bound (green) for different values of β.

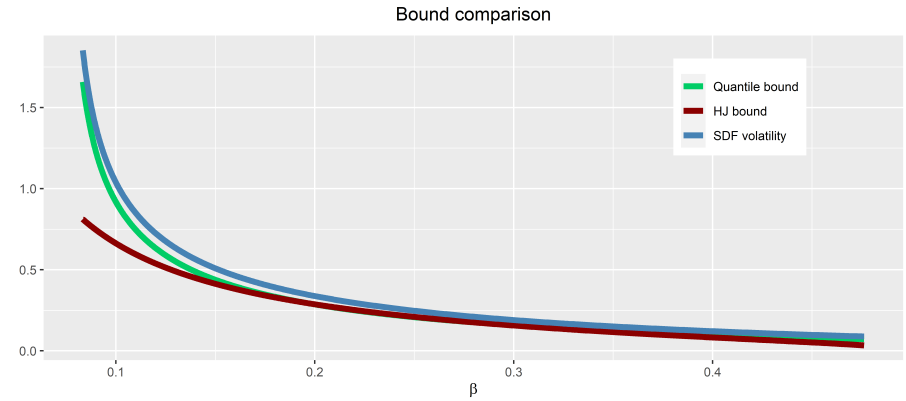

Figure 2: Plot of the Quantile bound (green) and HJ bound (red) when β varies. The true SDF volatility is denoted by the blue line.

2.2 Quantile and HJ bound for common asset pricing models

In this section I compare the tightness of the quantile bound in Theorem 2.1 to the HJ bound using common asset pricing models. This is of interest, since some asset pricing models imply that the HJ bound is always tighter than the quantile bound. For other models this is true under common parameter calibration. Since real data in Section 4 show that the quantile bound is significantly stronger than the HJ bound, this can be taken as evidence against such models. Appendix E contains similar results in this direction using other well known asset pricing bounds.

Example 2.1 (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. Since the HJ bound is derived by applying the Cauchy-Schwarz inequality to $\mathbb{COV}(R_m, M)$, the inequality binds if M is a linear combination of R_m. Hence, under CAPM, the HJ bound is always stronger than the quantile bound regardless of the distribution of R_m.

For the following two examples I need Stein's Lemma (Cochrane, 2005, p. 163):

Lemma 2.3 (Stein's Lemma)**.** *If X_1, X_2 are bivariate normal, $g : \mathbb{R} \to \mathbb{R}$ is differentiable and $\mathbb{E}\,|g'(X_1)| < \infty$, then*

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

Example 2.2 (Joint normality)**.** Suppose that M and R are jointly normally distributed. This obviously violates no-arbitrage but could be defended as an approximation over short time horizons, when the variance is small (see Example 2.3). The proof of the quantile bound in Theorem 2.1 gives the following identity

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

By an approximation argument, Stein's lemma still applies with $g(x) = \mathbb{1}\left(x \leq k\right)$ and $g'(x) = \delta_k(x)$ (Dirac delta function). Therefore,

$$\left|\mathbb{COV}\left(\mathbb{1}\left(R \leq \widetilde{Q}_\tau\right), M\right)\right| = f(\widetilde{Q}_\tau)\left|\mathbb{COV}(R, M)\right|. \tag{2.13}$$

Here, $f(\cdot)$ is the marginal density of R. Standard SDF properties also yield the well known identity

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

To evaluate the strengths of the quantile and HJ bound, consider the relative efficiency

$$\begin{aligned}\frac{\text{HJ bound}}{\text{Quantile bound}} &= \frac{\frac{\left|\mathbb{E}[R] - R_f\right|}{\sigma(R) R_f}}{\frac{\left|\tau - \mathbb{P}(R \leq \widetilde{Q}_\tau)\right|}{\sqrt{\mathbb{P}(R \leq \widetilde{Q}_\tau)(1 - \mathbb{P}(R \leq \widetilde{Q}_\tau))}\, R_f}} \\ &\overset{(2.13)}{=} \frac{\sqrt{\mathbb{P}(R \leq \widetilde{Q}_\tau)(1 - \mathbb{P}(R \leq \widetilde{Q}_\tau))}}{\sigma(R) f(\widetilde{Q}_\tau)}.\end{aligned} \tag{2.14}$$

To see that the HJ bound is always stronger than the quantile bound, minimize (2.14) with respect to τ. Temporarily write $x = \widetilde{Q}_\tau$ and $F(x) = \mathbb{P}(R \leq x)$ and consider

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

Minimizing $\Gamma(x)$ is equivalent to minimizing (2.14) and first order conditions

yield

$$[f(x) - 2F(x)f(x)]f(x)^2 - 2f(x)f'(x)[F(x)(1 - F(x))] = 0. \tag{2.15}$$

Since f, F are the respective PDF and CDF of the normal random variable R, it follows that $f'(\mu_R) = 0$ and $F(\mu_R) = 1/2$, where μ_R is the mean of R. As a result, (2.15) holds when $\widetilde{Q}_\tau = x = \mu_R$. For this choice, $\mathbb{P}(R \leq \widetilde{Q}_\tau) = 1/2$ and $f(\widetilde{Q}_\tau) = 1/\sqrt{2\pi\sigma(R)^2}$. Therefore, (2.14) obeys the bound

$$\frac{\sqrt{\mathbb{P}(R \leq \widetilde{Q}_\tau)(1 - \mathbb{P}(R \leq \widetilde{Q}_\tau))}}{\sigma(R)f(\widetilde{Q}_\tau)} \geq \frac{\sqrt{2\pi}}{2} \approx 1.25.$$

Hence, the HJ bound is always better in a model where the SDF and wealth portfolio are assumed to be jointly normal.

Example 2.3 (Joint lognormality)**.** Let Z_R and Z_M be standard normal random variables with correlation ρ and consider the specification

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

Here, 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.2, where I prove that

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

This expression is independent of μ_R. An application of l'Hôspital's rule reveals that the relative efficiency converges to $\sqrt{2\pi}/2$ if $\lambda \to 0^+$. This is the same relative efficiency in Example 2.2, which is unsurprising as the linearization becomes exact in the limit as $\lambda \to 0^+$. This ratio is less than 1 if $\sigma \geq 0.91$ and $\lambda = 1$. In practice, annualized market return volatility is about 16%, which means that the HJ bound is stronger than the quantile bound under any

reasonable parameterization if the SDF and asset return are lognormal.

Example 2.4 (Disaster risk)**.** The disaster risk model of Rietz (1988) and Barro (2006) posits that risk-premia are driven by extreme events that affect consumption growth. I follow the specification in Backus et al. (2011), who assume that the representative agent has power utility and the log pricing kernel is given by

$$\log M = \log(\beta) - \gamma \Delta c.$$

Innovations in consumption growth are driven by two independent shocks

$$\Delta c = \varepsilon + \eta. \tag{2.16}$$

Here, $\varepsilon \sim N(\mu, \sigma^2)$ and

$$\eta | (J = j) \sim N(j\theta, j\nu^2), \quad J \sim \textbf{Poisson}(\kappa).$$

The interpretation of η is that of a jump component (disaster) which induces negative shocks to consumption growth. κ governs the jump intensity for the Poisson distribution. I use the same calibration as Backus et al. (2011). In line with their paper, the market portfolio is considered as a claim on levered consumption, i.e. an asset that pays dividends C^λ. I convert the model implied volatility bounds to monthly units, to facilitate the comparison with the long-run risk model and the empirical bounds obtained in Section 4.

The quantile bound, HJ bound and SDF volatility are obtained using simulation and illustrated in the left upper panel of Figure 3. We see that the quantile bound has a sharp peak at $\tau = 0.035$, 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 is in line with the empirical evidence in Section 4.

The reason for this result can be understood from the upper right panel of Figure 3, which shows the physical and risk neutral distribution of return on equity.[3] 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

[3] Appendix D.1 shows how to calculate the physical and risk neutral CDF using cumulant generating functions.

put options must have high Sharpe ratios as their prices are high (insurance against disaster risk), but the actual probability of disaster is so low that the risk associated to selling such insurance is limited. The fact that there is a range of τ values for which the quantile bound is stronger than the HJ bound depends on the risk-aversion parameter γ. Backus et al. (2011) calibrate the model with $\gamma = 5.19$, which is high enough to overcome the HJ bound, however, unreported simulations show that $\gamma = 4$ is no longer enough to overcome the HJ bound. Hence, the representative agent needs to be risk-averse enough to induce the pattern in the left top panel of Figure 3.

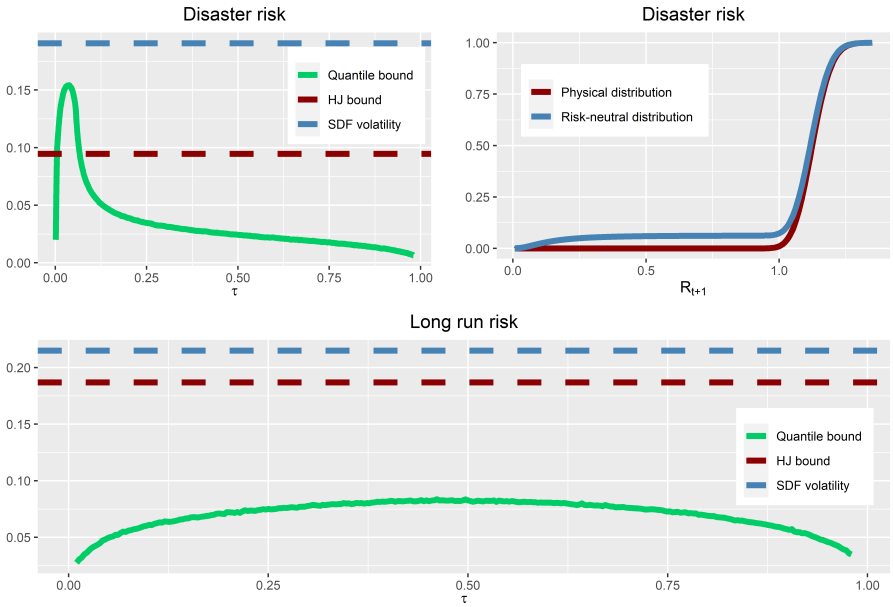

Figure 3: The upper left panel compares the HJ bound and quantile bound in disaster risk model and the right upper panel shows the physical and risk neutral distribution. The bottom panel shows the HJ and quantile bound for the long run risk model. The bounds are in monthly units.

The following example uses time t conditional information. I denote the SDF from time t to $t+1$, by M_{t+1}. The subscript refers to a random variable that is realized at time $t+1$, conditioned on time t.

Example 2.5 (Long-run risk)**.** The long-run risk (LRR) model of Bansal and Yaron (2004) posits that consumption growth is driven by a small and persistent component that captures long run risk. Moreover, the existence of a representative agent with Epstein and Zin (1989) recursive preferences is assumed. After

calibration, this model is successful in matching many of the salient features of the US market return data. I consider the extended model of Bansal et al. (2012), which allows for correlation between consumption growth shocks and dividend growth. In particular, the following dynamics are assumed:

$$\begin{aligned} x_{t+1} &= \rho x_t + \varphi_e \sigma_t e_{t+1} \\ \sigma^2_{t+1} &= \bar{\sigma}^2 + \nu(\sigma^2_t - \bar{\sigma}^2) + \sigma_w w_{t+1} \\ \Delta c_{t+1} &= \mu_c + x_t + \sigma_t \eta_{t+1} \\ \Delta d_{t+1} &= \mu_d + \phi x_t + \pi \sigma_t \eta_{t+1} + \varphi \sigma_t u_{d,t+1}. \end{aligned}$$

Here, Δc_{t+1} and Δd_{t+1} denote log consumption and dividend growth, while σ_t is conditional volatility of log consumption growth. The parameter ρ governs the persistence of long-term risk. The log SDF dynamics follow from the Euler equation and the Epstein and Zin (1989) preferences

$$\log M_{t+1} = \theta \log \beta - \frac{\theta}{\psi} \Delta c_{t+1} + (\theta - 1) r_{c,t+1},$$

where $r_{c,t+1}$ is the continuous return on the consumption asset. I omit further details on the parameter interpretation and calibration approach, as this is extensively discussed in Bansal et al. (2012). To compare the HJ bound to the quantile bound, I use the same calibration of parameters as Bansal et al. (2012). To approximate the quantile and HJ bound, I simulate the model for 110,000 months and drop the first 10,000 observations as burn-in sample. The results are summarized in the bottom panel of Figure 3. Unlike the disaster risk model, the quantile bound is always weaker than the HJ bound. Moreover, the quantile bound is almost symmetric around $\tau = 0.5$, at which the maximum is attained. Hence, it is not profitable to sell insurance against disaster risk. This contradicts the empirical estimates from Section 4, which suggests that the quantile bound is stronger than the HJ bound and implies it is most profitable to sell insurance against disaster risk. Unreported simulations show that a coefficient of risk aversion needed for the quantile bound to overcome the HJ bound is 90, which is far beyond reasonable levels of risk aversion.

3 Estimation of the unconditional quantile bound

In this section I discuss the estimation of the (unconditional) quantile bound. I assume the existence of a risk free asset, so we can use the quantile bound in (2.4). The bound is comprised of three unknowns: the risk-neutral quantile function $\widetilde{Q}_\tau(R)$, the physical probability measure $\mathbb{P}(R \leq x)$ and the risk-free rate R_f. I follow Liu (2020) and fix R_f at a pre specified level. This is because the US risk-free rate can be inferred with high precision and its influence is minor compared to the estimation of the first two functions. I sketch the intuition for the function estimates below; a more detailed description is provided in Appendix B.2.

Since our estimation strategy involves returns sampled over time we have to incorporate conditioning information. I use R_{t+1} to denote the asset return realized at time $t+1$, conditioned on time t information. I use the same notation for the risk free rate $R_{f,t+1}$, which is assumed to be known at time t.

To estimate the risk neutral quantile function, I first estimate $\widetilde{\mathbb{P}}(R \leq x)$ and then use inversion to get an estimate of $\widetilde{Q}_\tau(R)$. To make this idea operational, we rely on the following result of Breeden and Litzenberger (1978)

$$\widetilde{\mathbb{P}}_t(R_{t+1} \leq K/S_t) = R_{f,t+1}\frac{\partial}{\partial K}\text{Put}_t(K),$$

where $\text{Put}_t(K)$ is the time t price of a European put option with strike K, expiring at time $t+1$ and S_t is the time t stock price of the underlying asset. Hence, with enough put option prices, we can identify the conditional risk neutral CDF. To obtain an estimate of the unconditional CDF, we simply average the conditional CDFs over time

$$\widetilde{\mathbb{P}}_T\left(R \leq x\right) \coloneqq \frac{1}{T}\sum_{t=1}^{T}\widetilde{\mathbb{P}}_t\left(R_{t+1} \leq x\right).^{4} \tag{3.1}$$

Under suitable restrictions on the distribution of returns, we expect this to converge to the unconditional return distribution. An estimate for the uncon-

[4] I follow the empirical process literature and use subscript T to denote a functional estimate.

ditional risk-neutral quantile curve is then obtained from

$$\widetilde{Q}_T(\tau) \coloneqq \inf\left\{x \in \mathbb{R} : \tau \leq \widetilde{\mathbb{P}}_T(x)\right\}.$$

It is a non-trivial exercise to obtain solid estimates via this procedure, due to the lack of a continuum of option prices, interpolation issues and missing data for option prices far in– and out-of-the money. A detailed description is given in Appendix B.2, using a modification of the procedure proposed by Figlewski (2008).

Secondly, we need to estimate $\mathbb{P}(R \leq x)$. A first thought might be to use the empirical CDF for observed return data, however this estimator is not suitable due to discontinuities around each observation. This results in an estimator for the quantile bound which is too volatile and jagged. Instead, a kernel CDF estimator is used to avoid the discontinuity issue. Specifically, the distribution function is estimated by

$$\mathbb{P}_{T,\text{smooth}}(R_{t+1} \leq x) = \frac{1}{T}\sum_{t=1}^{T} \Phi\left(\frac{x - R_{t+1}}{h}\right),$$

where Φ is the integral of the Epanechnikov kernel and h is the bandwidth. The optimal bandwidth is determined via cross-validation. Combining the pieces, I obtain the following estimator for the quantile bound[5]

$$\widehat{\theta}_{\text{smooth}}(\tau) \coloneqq \frac{\left|\tau - \mathbb{P}_{T,\text{smooth}}(\widetilde{Q}_T(\tau))\right|}{\sqrt{\mathbb{P}_{T,\text{smooth}}(\widetilde{Q}_T(\tau))(1 - \mathbb{P}_{T,\text{smooth}}(\widetilde{Q}_T(\tau)))R_f}} \qquad \varepsilon \leq \tau \leq 1 - \varepsilon. \tag{3.2}$$

4 Empirical application

This Section presents estimates of the quantile bound using a combination of forward looking option data and historical market returns. Based on calibration of the disaster risk model, we show that the quantile bound renders a stronger bound on the SDF volatility than the HJ bound.

[5] I use a slight abuse of notation and write $\mathbb{P}_{T,\text{smooth}}(\widetilde{Q}_T(\tau))$, instead of $\mathbb{P}_{T,\text{smooth}}(R_{m,t+1} \leq \widetilde{Q}_T(\tau))$

4.1 Data and estimation of the unconditional risk neutral quantile curve

I use daily option data on the S&P500 from OptionMetrics covering the period 01-01-1996 until 12-31-2019 to estimate the unconditional risk neutral quantile curve. Before estimating the risk neutral quantile curve, I use several data cleaning procedures, which are detailed in Appendix B.1. Subsequently, the unconditional risk neutral quantile curve is estimated following the steps in Appendix B.2. Figure 4 shows the estimated unconditional quantile curves for several tenors. It is apparent that longer time horizons carry a higher (risk-neutral) probability of up and downswings. This makes intuitive sense, but is not apparent a priori, since the quantile curves document profit and losses over a fixed horizon and not during any time within the horizon.

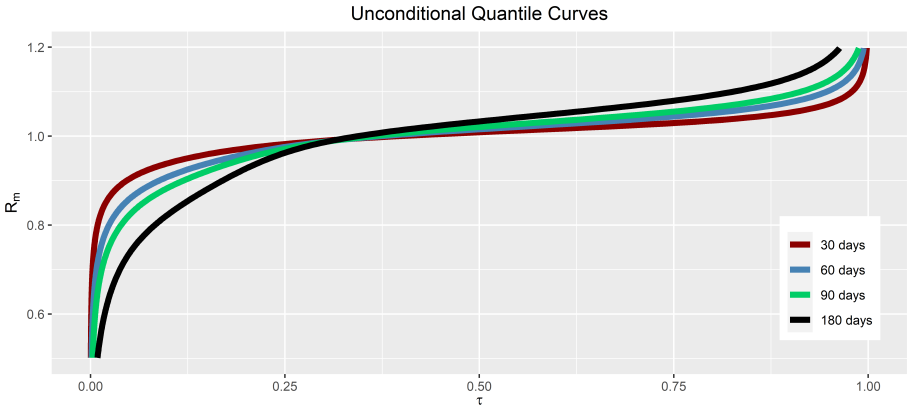

Figure 4: Plots of the unconditional (risk-neutral) quantile function for various maturities.

4.2 Quantile bound for 30-day returns

I now turn to the estimation of the unconditional quantile bound for 30-day returns, using the estimator in (3.2). The unconditional physical CDF estimate is based on non overlapping historical 30-day returns on the S&P500 index over the period 1996-2019, calculated at the middle of the month. This consists of a total of $T = 288$ return observations. The unconditional risk neutral quantile function is estimated following the procedure in Appendix B.2, using only the dates at which the historical market returns $R_{m,t+1}$ are calculated, i.e. I average

over dates t corresponding to the start of the return period of $R_{m,t+1}$. Figure 5 shows the estimated quantile bound (in red), as well as the Sharpe ratio (in blue). Notice that the estimated quantile bound is quite similar to the quantile bound predicted by the disaster risk model (Figure 3). This substantiates the belief that very high Sharpe ratios can be obtained for selling insurance against crash risk, since the probability of a disaster is low, but the price is high since agents seek insurance against extreme crash risk.

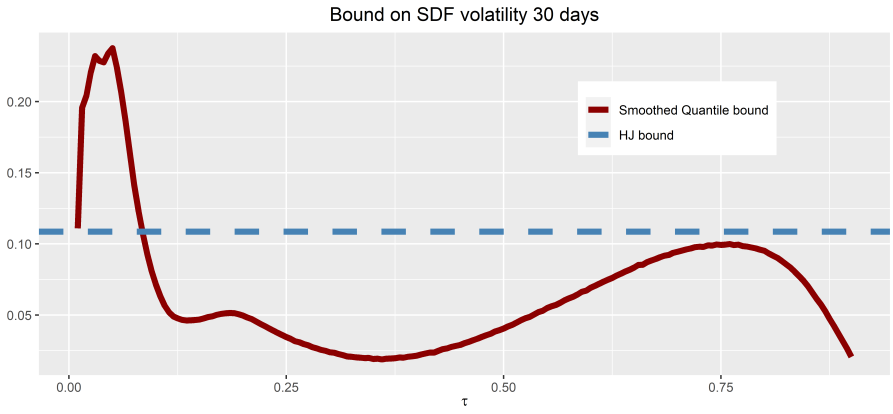

Figure 5: Plot of the quantile bound as function of τ. The solid red line is the estimated quantile bound. The dashed blue line depicts the HJ bound.

4.2.1 Testing the quantile and HJ bound

To show more formally whether the quantile bound improves upon the HJ bound, I use a bootstrap approach. To make this operational, I fix, a priori, the τ-quantile at probability level $\tau = 0.035$, which renders the sharpest bound on the SDF volatility in the disaster risk model (Example 2.4). Then, I consider the test statistic

$$\mathcal{T}(R_{m,t+1}) := \widehat{\theta}_{\text{smooth}}(0.035) - \frac{\left|\bar{R}_m - R_f\right|}{\widehat{\sigma}(R_m)R_f}. \tag{4.1}$$

The first term on the right hand side denotes the estimated quantile bound (3.2) evaluated at the 0.035-quantile, using the entire time series in our sample. The second term on the right hand side denotes the estimated HJ-bound, with $\bar{R}_m$ and $\widehat{\sigma}(R_m)$ denoting the sample mean and standard deviation of $R_{m,t+1}$ respectively. A value of $\mathcal{T}(R_{m,t+1}) \geq 0$ indicates that the quantile bound is

Table 2: Bootstrap result

R_f	Sample size	τ	Bootstrap samples	HJ bound	Quantile bound	p-value
1	288	0.035	100,000	0.1087	0.2287	0.0761

stronger than the HJ bound.

Since the distribution of (4.1) is hard to characterize, I use block bootstrap to approximate the p-value of the null hypothesis

$$H_0 : \mathcal{T}(R_{m,t+1}) \leq 0.$$

The block bootstrap is used to generate time indices from which we recreate (with replacement) bootstrapped returns $R^\star_{m,t+1}$. The same bootstrapped time indices are used to estimate the risk neutral quantile curves and leads to the quantile bound $\widehat{\theta}^\star_{\text{smooth}}(0.035)$, as well as the HJ bound

$$\frac{\left|\bar{R}^\star_m - R_f\right|}{\widehat{\sigma}(R^\star_m)R_f}.$$

I repeat the bootstrap exercise 100,000 and calculate the bootstrapped test statistic

$$\mathcal{T}(R^\star_{m,t+1}) \coloneqq \widehat{\theta}^\star_{\text{smooth}}(0.035) - \frac{\left|\bar{R}^\star_m - R_f\right|}{\widehat{\sigma}(R^\star_m)R_f}. \tag{4.2}$$

Finally, the p-value is obtained as the fraction of times $\mathcal{T}(R^\star_{m,t+1}) \leq 0$. Table 2 shows that the estimated p-value is 0.0761, which suggests that the quantile bound renders a higher Sharpe ratio compared to a direct investment in the market portfolio.

4.2.2 Implications

What are the economic implications of this finding? First, in Section 2.2, I argued that under CAPM, the HJ bound is always stronger than the quantile bound. Therefore, the empirical finding that the quantile bound is stronger than the HJ bound is evidence against CAPM. The LRR model of Bansal et al. (2012) can neither reconcile this feature of the data under common parameter calibration. The evidence from Figure 5 and Table 2 supports the view that the misspecification of the LRR model results from the inability to incorporate

disaster risk. For the LRR model, I establish by simulation that a risk aversion coefficient of 90 is needed for the quantile bound to overcome the HJ bound. Moreover, in the LRR model, the quantile bound is almost symmetric around $\tau = 0.5$ and decreases steeply when τ is close to $\{0, 1\}$, which again contradicts empirical evidence that high Sharpe ratios can be achieved for selling insurance against disaster risk. In conclusion, the empirical results indicate that the disaster risk model compares favorably to the LRR model and the misspecification of the LRR model is caused by the failure to account for disaster risk.

The observation that put options can be leveraged to yield higher Sharpe ratios than a direct investment in the market portfolio has been noted before in the literature (Bates, 2008). However, what seems to have gone unnoticed is the connection to model misspecification, in the sense that the empirical regularity puts a tight constraint on a model's risk-aversion coefficient to replicate this finding. In particular, this is true for the LRR model or statistical models that impose joint lognormality between the SDF and market return (Example 2.3). In the most extreme case, for CAPM or models that impose joint normality (Example 2.2), we cannot even find parameters that replicate the empirical regularity. Broadly speaking, our results underscore the importance of incorporating information beyond the mean and variance to analyze model misspecification.

5 Conditional relation between physical and risk-neutral quantile

Section 4 shows that the unconditional quantile bound delivers high Sharpe ratios for small τ. This section complements that finding by looking at conditional information. I show how the numerator term in the quantile bound governs the difference between the physical and risk neutral quantile, up to a first order correction. I interpret this first order correction as a risk adjustment term, which, under some conditions, can be approximated by option data available at time t. I develop a new measure of out-of-sample quantile predictability and argue that quantiles cannot be too predictable if a liquid option market exists.

5.1 Lower bound on physical quantile

I start by deriving a market observable lower bound on the physical quantile function. Write $F_t(x) \coloneqq \mathbb{P}_t(R_{t+1} \leq x)$ for the conditional physical distribution function of R_{t+1}, $f_t(\cdot)$ for the PDF and $Q_{t,\tau}$ for the corresponding τ-quantile. As before, a tilde superscript denotes the conditional CDF, PDF or τ-quantile under risk neutral measure. The idea is to use von Mises (1947) calculus for statistical functionals to find a first order approximation to the physical quantile function. A statistical functional $\phi : \mathbb{D} \to \mathbb{E}$ is typically a map between two normed spaces $\mathbb{D}$ and $\mathbb{E}$. In our application, we take $\phi(G) = G^{-1}(\tau)$, where G is a CDF and $G^{-1}(\tau)$ is the quantile function. Following Van der Vaart (2000, Section 20.1), we hope that the following is a good approximation

$$Q_{t,\tau} - \widetilde{Q}_{t,\tau} = \phi(F_t) - \phi(\widetilde{F}_t) \approx \phi'_{\widetilde{F}_t}(F_t - \widetilde{F}_t), \tag{5.1}$$

where $\phi'_{\widetilde{F}_t}(F_t - \widetilde{F}_t)$ is the Gâteaux derivative of ϕ at $\widetilde{F}_t$ in the direction of F_t

$$\begin{aligned}\phi'_{\widetilde{F}_t}(F_t - \widetilde{F}_t) &\coloneqq \lim_{\lambda \downarrow 0} \frac{\phi\left[(1-\lambda)\widetilde{F}_t + \lambda F_t\right]}{\lambda} \\ &= \frac{\partial}{\partial \lambda} \phi\left((1-\lambda)\widetilde{F}_t + \lambda F\right)\Big|_{\lambda=0}.\end{aligned} \tag{5.2}$$

Heuristically, the Gâteaux derivative in this context can be thought of as measuring the change in the quantile function when we move the distribution of returns from risk neutral in the direction of the physical distribution. Appendix A.6 shows that, when $\phi(\cdot)$ is the quantile map, the Gâteaux derivative is given by

$$\phi'_{\widetilde{F}_t}(F_t - \widetilde{F}_t) = \frac{\tau - F_t(\widetilde{Q}_{t,\tau})}{\widetilde{f}_t(\widetilde{Q}_{t,\tau})}. \tag{5.3}$$

Combining (5.1) and (5.3) and solving for the physical quantile function renders the following approximation

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

Observe that the numerator in the risk adjustment term equals the numerator term in the conditional version of the quantile bound. We do not bother about precise conditions that assess the quality of the approximation, but instead use it as a guiding concept rather than a formal mathematical theorem.[6] This is often how von Mises calculus is used in statistical analysis (Van der Vaart, 2000; Serfling, 2009).

The approximation in (5.4) contains the terms $\widetilde{Q}_{t,\tau}$ and $\widetilde{f}_t(\widetilde{Q}_{t,\tau})$, which are directly observable at time t, as they can be computed using (a variation of) the Breeden and Litzenberger (1978) formula in (3.1). However, the physical CDF $F_t(\cdot)$ is unknown and hence (5.4) cannot directly be used to predict $Q_{t,\tau}$.

Under additional assumptions, the numerator term $\tau - F_t(\widetilde{Q}_{t,\tau})$ can be estimated with market data. To show this, I link $F_t(\cdot)$ to a risk-neutral covariance term, similar to Chabi-Yo and Loudis (2020). Use the reciprocal of the SDF to pass from physical to risk-neutral measure

$$\begin{aligned} F_t(\widetilde{Q}_{t,\tau}) = \mathbb{E}_t\left[\mathbb{1}\left(R_{t+1} \leq \widetilde{Q}_{t,\tau}\right)\right] &= \widetilde{\mathbb{E}}_t\left[R_{t+1} \leq \widetilde{Q}_{t,\tau}\right]\frac{\mathbb{E}_t\left[M_{t+1}\right]}{M_{t+1}} \\ &= \widetilde{\mathbb{COV}}_t\left[\mathbb{1}\left(R_{t+1} \leq \widetilde{Q}_{t,\tau}\right), \frac{\mathbb{E}_t\left[M_{t+1}\right]}{M_{t+1}}\right] + \tau. \end{aligned} \tag{5.5}$$

To proceed, assume that $R_{t+1} = R_{m,t+1}$ (the market return). This allows me to get a more explicit expression of the SDF as follows. Chabi-Yo and Loudis (2020) show that in a one-period model with a representative agent who has utility function $u(\cdot)$, and derives utility over final wealth

$$\frac{\mathbb{E}_t\left[M_{t+1}\right]}{M_{t+1}} = \frac{\frac{u'(W_t x_0)}{u'(W_t x)}}{\widetilde{\mathbb{E}}_t\left[\frac{u'(W_t x_0)}{u'(W_t x)}\right]} =: f(x) \qquad \text{with } x = R_{m,t+1} \text{ and } x_0 = R_{f,t+1}.$$

Here W_t is the agent's initial wealth at time t. Taylor expansion around $x = x_0$ yields

$$f(x) = 1 + \sum_{k=1}^{\infty} \theta_k (x - x_0)^k \qquad \text{with } \theta_k = \frac{1}{k!}\left(\frac{\partial^k f(x)}{\partial x^k}\right)_{x=x_0}.$$

The θ_k-coefficients depend on the specific utility representation employed, but

[6] Such conditions might impose the existence of $\widetilde{f}_t$ or give bounds on the remainder term.

are conditionally non-random. I substitute the above in (5.5) and obtain

$$\begin{aligned}&\widetilde{\mathbb{COV}}_t\left[\mathbb{1}\left(R_{m,t+1}\leq\widetilde{Q}_{t,\tau}\right),\frac{\widetilde{\mathbb{E}}_t(M_{t+1})}{M_{t+1}}\right]\\&=\frac{\sum_{k=1}^{\infty}\theta_k\left(\widetilde{\mathbb{E}}_t\left[\mathbb{1}\left(R_{m,t+1}\leq\widetilde{Q}_{t,\tau}\right)(R_{m,t+1}-R_{f,t+1})^k\right]-\tau\widetilde{\mathbb{E}}_t\left[(R_{m,t+1}-R_{f,t+1})^k\right]\right)}{1+\sum_{k=1}^{\infty}\theta_k\widetilde{\mathbb{E}}_t\left[(R_{m,t+1}-R_{f,t+1})^k\right]},\end{aligned}\tag{5.6}$$

since $\widetilde{\mathbb{E}}_t(\mathbb{1}(R_{m,t+1}\leq\widetilde{Q}_\tau))=\tau$. Results in Appendix A.3 show how to compute higher order moments of the (un)truncated excess market return under risk neutral measure as the integral of the estimated quantile function. To enhance notation, I follow Chabi-Yo and Loudis (2020) and write

$$\begin{aligned}\widetilde{\mathbb{M}}_{t+1}^{(n)}&:=\widetilde{\mathbb{E}}_t\left[(R_{m,t+1}-R_{f,t+1})^n\right]\\\widetilde{\mathbb{M}}_{t+1}^{(n)}[k_0]&:=\widetilde{\mathbb{E}}_t\left[\mathbb{1}\left(R_{m,t+1}\leq k_0\right)(R_{m,t+1}-R_{f,t+1})^n\right].\end{aligned}$$

This means (5.6) can be rewritten to

$$\widetilde{\mathbb{COV}}_t\left[\mathbb{1}\left(R_{m,t+1}\leq\widetilde{Q}_{t,\tau}\right),\frac{1}{M_{t+1}R_{f,t+1}}\right]=\frac{\sum_{k=1}^{\infty}\theta_k\left(\widetilde{\mathbb{M}}_{t+1}^{(k)}[\widetilde{Q}_{t,\tau}]-\tau\widetilde{\mathbb{M}}_{t+1}^{(k)}\right)}{1+\sum_{k=1}^{\infty}\theta_k\widetilde{\mathbb{M}}_{t+1}^{(k)}}.\tag{5.7}$$

Combining Equation (5.7) and (5.5) in (5.4) leads to the first order approximation

$$Q_{t,\tau}\approx\widetilde{Q}_{t,\tau}+\frac{1}{\widetilde{f}_t(\widetilde{Q}_{t,\tau})}\left(\frac{\sum_{k=1}^{\infty}\theta_k\left(\tau\widetilde{\mathbb{M}}_{t+1}^{(k)}-\widetilde{\mathbb{M}}_{t+1}^{(k)}[\widetilde{Q}_{t,\tau}]\right)}{1+\sum_{k=1}^{\infty}\theta_k\widetilde{\mathbb{M}}_{t+1}^{(k)}}\right).\tag{5.8}$$

The right hand side of (5.8) depends on known quantities that can be calculated at time t with option data, except for the unknown parameters θ_k. However, Chabi-Yo and Loudis (2020) show that we can make assumption about θ_k that lead to a lower bound . I adopt the following assumptions from their paper:

Assumption 5.1. $\widetilde{\mathbb{M}}_{t+1}^{(k)} \leq 0$ *if* k *is odd. Furthermore, for* $k_0 \leq R_{f,t+1}$

$$\widetilde{\mathbb{M}}_{t+1}^{(1)}[k_0] \leq 0, \quad \widetilde{\mathbb{M}}_{t+1}^{(2)}[k_0] \geq 0$$
$$\widetilde{\mathbb{M}}_{t+1}^{(3)}[k_0] \leq 0, \quad \widetilde{\mathbb{M}}_{t+1}^{(4)}[k_0] \geq 0.$$

Assumption 5.2. *Preference parameters* θ_k *satisfy the following inequalities for* $k \geq 1$

$$\theta_k \leq 0 \text{ if } k \text{ is even and } \theta_k \geq 0 \text{ if } k \text{ is odd}$$

Assumption 5.2 needs to be strengthened as follows to obtain the completely nonparametric bound in Corollary 5.5:

Assumption 5.3. *The first three preference parameters can be expressed as*

$$\theta_k = \frac{(-1)^{k+1}}{R_{f,t+1}^k} \quad \text{for } k \in \{1, 2, 3\}.$$

Remark. Chabi-Yo and Loudis (2020) discuss the economic relevance of these assumptions. Assumption 5.1 concerns odd moments of excess market returns, which are typically negative, since they relate to unfavorable market conditions. Assumption 5.2 is natural given Assumption 5.1, since investors require compensation for exposure to risk-neutral moments. Assumption 5.3 strengthens Assumption 5.2 and is needed to obtain a completely nonparametric bound in Corollary 5.5. One can test the validity of Assumption 5.3 in the data. Chabi-Yo and Loudis (2020) do so and find that Assumption 5.3 cannot be rejected.

Under these assumptions, we can bound the discrepancy between the conditional physical and risk-neutral distribution.

Theorem 5.4 (Lower bound)**.** *Let assumptions 5.1 and 5.2 hold. Assume that the risk-neutral CDF is absolutely continuous w.r.t. Lebesgue measure and* $\sup_k \|R_{m,t+1}\|_k := \sup_k \widetilde{\mathbb{E}}(|R_{m,t+1}|^k)^{1/k} < \infty$*. Finally, define* τ^* *so that*

$$\widetilde{Q}_{t,\tau^*} = R_{f,t+1} - \sup_k \|R_{m,t+1} - R_{f,t+1}\|_k .$$

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

$$\tau - \mathbb{P}_t\left(R_{m,t+1} \leq \widetilde{Q}_{t,\tau}\right) \geq \left(\frac{\sum_{k=1}^{3} \theta_k \left(\tau \widetilde{\mathbb{M}}_{t+1}^{(k)} - \widetilde{\mathbb{M}}_{t+1}^{(k)}[\widetilde{Q}_{t,\tau}]\right)}{1 + \sum_{k=1}^{3} \theta_k \widetilde{\mathbb{M}}_{t+1}^{(k)}}\right).$$

Proof. See Appendix A.4. ■

Corollary 5.5. *If, additionally, Assumption 5.3 holds, then for all* $\tau \leq \tau^*$

$$\tau - \mathbb{P}_t\left(R_{m,t+1} \leq \widetilde{Q}_{t,\tau}\right) \geq \left(\frac{\sum_{k=1}^{3} \frac{(-1)^{k+1}}{R_{f,t+1}^{k}} \left(\tau \widetilde{\mathbb{M}}_{t+1}^{(k)} - \widetilde{\mathbb{M}}_{t+1}^{(k)}[\widetilde{Q}_{t,\tau}]\right)}{1 + \sum_{k=1}^{3} \frac{(-1)^{k+1}}{R_{f,t+1}^{k}} \widetilde{\mathbb{M}}_{t+1}^{(k)}}\right) =: LRB_t(\tau). \tag{5.9}$$

Proof. See Appendix A.4. ■

Under the working hypothesis that the remainder term in the quantile approximation in (5.4) is negligible, the following Corollary is immediate

Corollary 5.6. *Suppose that the remainder term in* (5.4) *is negligible, and Assumptions 5.1–5.3 hold, then for all* $\tau \leq \tau^*$

$$Q_{t,\tau} - \widetilde{Q}_{t,\tau} \geq \frac{LRB_t(\tau)}{\widetilde{f}_t(\widetilde{Q}_{t,\tau})}. \tag{5.10}$$

Notice that Theorem 5.4 establishes a lower bound on how far the risk-neutral distribution can diverge from the physical distribution conditional on time t. The lower bound that results depends on unknown preference parameters θ_k, which, in principle, can be estimated from return and option data (Chabi-Yo and Loudis, 2020, Section 3.3). The bound in Corollary 5.5 is completely nonparametric and can be calculated solely based on time t information. Corollary 5.5 complements the recent literature on the recovery of beliefs. Ross (2015) shows that one can recover $F_t(\cdot)$, 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 this, Corollary 5.5 shows that one can still establish a lower bound under a different set of (mild) economic constraints. Corollary 5.6 establishes an interesting relation between the physical and risk neutral quantile, showing that the difference between the

two can be bounded using option data available at time t. The hypothesis that the remainder term in (5.4) is small is confirmed with simulation results in Appendix F, using the Black and Scholes (1973) model. If the lower bound in (5.10) happens to be tight, one can use it as a predictor variable to forecast Value-at-Risk. Moreover, this predictor variable is not subject to the historical sample bias critique of Goyal and Welch (2008). I now show, using various tests, that the lower bounds in Corollary 5.5 and 5.6 are indeed tight.

5.2 Testing (5.9) in the data

One expects the market return to have a higher probability of a crash under risk-neutral measure than using actual probabilities. This means that Corollary 5.5 has non trivial content if $LRB_t(\tau) \geq 0$ in the data. I confirm that this is the case for all dates considered, using the parameters in Table 3.

Since the lower bound in (5.9) is non trivial, I can test whether bound is tight, by running an OLS regression of the form

$$\tau - \mathbb{1}\left(R_{m,t+1} \leq \widetilde{Q}_{t,\tau}\right) = \beta_0(\tau) + \beta_1(\tau) LRB_t(\tau). \tag{5.11}$$

If Corollary 5.5 holds in the data we expect $\beta_1(\tau) \geq 1$. Moreover, if the lower bound is tight, we expect $\beta_1(\tau) = 0$ and $\beta_1(\tau) = 1$. Table 3 contains the regression results for tenors 30 and 60 days. With the exception of the 60 day horizon and $\tau = 0.01$, the results confirm that $\beta_1(\tau) \geq 1$. In addition, some of the point estimates of $\beta_1(\tau)$ are close to 1 and in all cases we cannot reject the null hypothesis that the lower bound is tight.

Following the literature on the conditional equity premium Campbell and Thompson (2008); Martin (2017), I also consider the out-of-sample performance of the predictor variable $LRB_t(\tau)$. That is, I impose the restriction $\beta_0(\tau) = 0, \beta_1(\tau) = 1$ and predict the dependent variable in (5.11) directly by $LRB_t(\tau)$. The out-of-sample R^2 gives a measure of forecast ability using $LRB_t(\tau)$ as a predictor variable

$$R^2_{oos} = 1 - \frac{\sum \nu_t^2}{\sum e_t^2}, \tag{5.12}$$

where ν_t is the forecast error using $LRB_t(\tau)$ as a predictor and e_t is the forecast error using a historical (rolling) mean as a predictor. Table 3 shows that R^2_{oos} is

Table 3: OLS estimates of (5.11)

Maturity:	30 days					60 days				
	$\widehat{\beta}_0(\tau)$	$\widehat{\beta}_1(\tau)$	χ^2_2	$R^2(\%)$	$R^2_{oos}(\%)$	$\widehat{\beta}_0(\tau)$	$\widehat{\beta}_1(\tau)$	χ^2_2	$R^2(\%)$	$R^2_{oos}(\%)$
$\tau = 0.1$	0.02 (0.0226)	1.35 (1.7757)	0.82	0.08	2.35	0.05 (0.0212)	0.09 (1.0951)	0.01	0	4.17
$\tau = 0.2$	0.01 (0.0352)	1.50 (2.1784)	1	0.1	3.92	0.02 (0.0301)	1.38 (0.9256)	1	0.21	8.47
$\tau = 0.25$	0.01 (0.0372)	1.48 (2.0292)	1	0.09	4.76	0.02 (0.0391)	1.07 (1.1599)	0.94	0.11	9.13
$\tau = 0.3$	0.02 (0.0369)	1.04 (1.7433)	0.87	0.04	5.21	0.01 (0.0408)	1.07 (1.1215)	0.97	0.11	10.05

Note: *Standard errors are shown in parentheses and calculated using the kernel-based approach of Andrews (1991) with automatic bandwidth selection.* χ^2_2 *gives the p-value of the Wald test on the joint restriction:* $\widehat{\beta}_0(\tau) = 0, \widehat{\beta}_1(\tau) = 1$. R^2_{oos} *denotes the out-of-sample* R^2 *using a 500 day rolling window.*

positive for all regression specifications, indicating that $LRB_t(\tau)$ also has merit as an out-of-sample predictor.

5.3 Testing (5.10) in the data

The OLS regressions in Table 3 suggest that the lower bound in Corollary 5.5 is tight. One might therefore expect that the lower bound on the physical quantile function in (5.10) is tight as well. I test this following a procedure similar to Section 5.2. In particular, at the start of each time period t, the conditional quantile $Q_{t,\tau}$ is estimated by

$$\widehat{Q}_{t,\tau} = \widetilde{Q}_{t,\tau} + \frac{LRB_t(\tau)}{\widetilde{f}_t(\widetilde{Q}_{t,\tau})}. \tag{5.13}$$

Subsequently, quantile regression (Koenker and Bassett, 1978) is used to estimate the model

$$Q_{t,\tau} = \beta_0(\tau) + \beta_1(\tau)\widehat{Q}_{t,\tau}. \tag{5.14}$$

If (5.13) is a good predictor of the conditional quantile function, we expect

$$H_0: \qquad \beta_0(\tau) = 0, \quad \beta_1(\tau) = 1. \tag{5.15}$$

To test the null hypothesis in (5.15) requires the asymptotic covariance matrix $\Sigma(\tau)$ in the CLT

$$\sqrt{T}\left(\begin{bmatrix}\widehat{\beta}_0(\tau)\\ \widehat{\beta}_1(\tau)\end{bmatrix} - \begin{bmatrix}0\\1\end{bmatrix}\right) \rightsquigarrow N\left(\mathbf{0}, \Sigma(\tau)\right).$$

It is not trivial to obtain a correct estimate of $\Sigma(\tau)$, since we deal with overlapping return data and the covariance matrix typically depends on the unknown density of the data. The overlapping data problem has been addressed by Hansen and Hodrick (1980); Hodrick (1992) and many others in the context of OLS regression, but the extension to quantile regression remained elusive. Quite recently, Gregory et al. (2018) proposed the smooth extended tapered block bootstrap (SETBB) to estimate $\Sigma(\tau)$, if the time series data are weakly dependent (Assumptions C.1–C.7 in their paper). Moreover, SETBB avoids the thorny issue of estimating the density function of the data. Further details about this method can be found in Gregory et al. (2018). I use SETBB to estimate $\widehat{\Sigma}(\tau)$, from which standard errors can be calculated, as well as a Wald statistic to test the null hypothesis (5.15).[7] Let $\widehat{\theta}(\tau) := [\widehat{\beta}_0(\tau), \widehat{\beta}_1(\tau)]^\top - [0,1]^\top$, then under H_0 we obtain the distribution of the Wald statistic

$$T\left[\widehat{\theta}(\tau)\widehat{\Sigma}(\tau)^{-1}\widehat{\theta}(\tau)\right] \rightsquigarrow \chi^2_2, \tag{5.16}$$

where χ^2_2 is the Chi-squared distribution with 2 degrees of freedom. I use the prediction horizon as the block length in the bootstrap procedure, which is the only user required input for SETBB. The results in Table 4 are quite striking, as the point estimates of $[\beta_0(\tau), \beta_1(\tau)]^\top$ are close to the benchmark $[0,1]^\top$. Moreover, the joint restriction in (5.15) is not rejected for all days and quantiles.

Since the in-sample results suggest $\beta_0(\tau) = 0$ and $\beta_1(\tau) = 1$, it is natural to test how well this works out-of-sample by evaluating the predictive model

$$Q_{t,\tau}(R_{m,t+1}) = \widehat{Q}_{t,\tau}(R_{m,t+1}).$$

It is worth emphasizing that the left hand side is unknown at time t, whereas the right hand side is known at time t. Moreover, this predictive model does not require any estimation and parallels the discussion in Martin (2017), who used the same idea to measure the equity premium.

An initial concern about the out-of-sample predictor (5.13) is the possibility of crossing, which means that the predicted quantiles $\widehat{Q}_{t,\tau}$ are not monotone

[7] To calculate $\widehat{\Sigma}(\tau)$ with SETBB, I 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.

with respect to τ. This problem frequently arises in dynamic quantile models (Gouriéroux and Jasiak, 2008). It appears, however, that this is not a concern for the nonparametric predictor in (5.13). Using the quantiles from Table 4, I confirm that crossing does not occur for the 30 day horizon, and it only happens for $5/3961 = 0.12\%$ of the data when the prediction horizon is 60 days.

To assess the out-of-sample performance further, we cannot use out-of-sample R^2 (5.12) from Campbell and Thompson (2008), since this is appropriate for OLS regression only. However, a natural substitute for quantile regression is available. Koenker and Machado (1999) proposed a goodness-of fit criterion for quantile regression, which resembles the traditional OLS R^2

$$R^1(\tau) = 1 - \frac{\min_{b_0,b_1} \sum \rho_\tau(R_{m,t+1} - b_0 - \widehat{Q}_{t,\tau} b_1)}{\min_{b_0} \sum \rho_\tau(R_{m,t+1} - b_0)}, \tag{5.17}$$

where $\rho_\tau(x) = x(\tau - \mathbb{1}(x < 0))$ is the loss function from quantile regression (Koenker and Bassett, 1978) and $\widehat{Q}_{t,\tau}$ is the prediction variable from (5.13). It is well known that b_0 in the denominator of (5.17) equals the in-sample τ-quantile. The natural out-of-sample analogue of (5.12) is then

$$R^1_{oos}(\tau) \coloneqq 1 - \frac{\sum \rho_\tau(R_{m,t+1} - \widehat{Q}_{t,\tau})}{\sum \rho_\tau(R_{m,t+1} - \overline{Q}_{t,\tau})}, \tag{5.18}$$

where $\overline{Q}_{t,\tau}$ is the historical rolling quantile of the market return until time t. The out-of-sample $R^1_{oos}(\tau)$ is also displayed in Table 4. In all but one case the predictor variable $\widehat{Q}_{t,\tau}$ improves upon the historical rolling quantile out-of-sample. The outperformance is in the order of several percentages, which is also typical for expected market return predictors (Martin, 2017, Table 2). Ross (2005) and Campbell and Thompson (2008) point out this is typical for expected return predictors, for otherwise one could design a market strategy that exploits the predictability of returns and this would result in extremely high Sharpe ratios. The same intuitive argument applies in the case of quantile prediction. If $Q_{t,\tau}(R_{m,t+1})$ is highly predictable at time t by, say $\widehat{Q}_{t,\tau}$, one could form a market timing strategy using digital put options that exploits the

Table 4: Quantile regression estimates of (5.14)

Maturity:	30 days					60 days				
	$\widehat{\beta}_0(\tau)$	$\widehat{\beta}_1(\tau)$	Wald test	$R^1(\tau)$[%]	$R^1_{oos}(\tau)$[%]	$\widehat{\beta}_0(\tau)$	$\widehat{\beta}_1(\tau)$	Wald test	$R^1(\tau)$[%]	$R^1_{oos}(\tau)$[%]
$\tau = 0.1$	0.27 (0.2631)	0.73 (0.2740)	0.14	5.15	-0.83	0.14 (0.3603)	0.87 (0.3818)	0.75	4.97	7.69
$\tau = 0.2$	0.23 (0.3534)	0.76 (0.3595)	0.42	1.77	0.03	0.11 (0.4262)	0.89 (0.4352)	0.88	2.09	2.84
$\tau = 0.25$	0.26 (0.4051)	0.74 (0.4092)	0.45	0.98	0.51	0.00 (0.4687)	1.01 (0.4730)	1	1.36	2.57
$\tau = 0.3$	0.16 (0.4166)	0.84 (0.4183)	0.74	0.43	1.05	0.14 (0.3938)	0.87 (0.3945)	0.79	0.5	2.15

Note: *Standard errors are shown in parentheses and calculated using the SETBB method of Gregory et al. (2018) with block length equal to the maturity length and 1,000 Monte Carlo bootstrap samples.* Wald test *gives the p-value of the Wald test on the joint restriction:* $\widehat{\beta}_0(\tau) = 0, \widehat{\beta}_1(\tau) = 1$. $R^1(\tau)$ *denotes the in-sample goodness-of fit criterion (5.17).* $R^1_{oos}(\tau)$ *is the out-of-sample goodness-of fit (5.18), using a 1000 day rolling window.*

difference

$$\underbrace{\mathbb{E}_t\left[\mathbb{1}\left(R_{m,t+1} \leq \widehat{Q}_{t,\tau}\right)\right]}_{\approx\tau} - \underbrace{\widetilde{\mathbb{E}}_t\left(\mathbb{1}\left(R_{m,t+1} \leq \widehat{Q}_{t,\tau}\right)\right)}_{\text{digital put price at time } t}. \tag{5.19}$$

Again, high Sharpe ratios can result from this market strategy. For this to work, a liquid market in derivatives needs to exist which allows one to approximate the payoff in (5.19) using more primitive options. This seems to be sensible for the S&P500 returns which I use in this application. The next Section shows that one can obtain perfect predictions of $Q_{t,\tau}(R_{m,t+1})$ if returns are conditionally lognormal.

5.4 Prediction conditional lognormal environment

The previous Section shows that a predictable quantile of the return distribution can be leveraged to obtain high Sharpe ratios. This Section shows that quantile regression, using the risk neutral quantile as a regressor, gives an optimal forecast of the physical quantile when returns are conditionally lognormal. I then show, using various different angles, that the conditional lognormal assumption is implausible.

Consider the following discretized version of the Black-Scholes model, with a riskless asset that offers a certain return $R_{f,t+1} = e^{r_f \lambda}$ and a risky asset with return $R_{t+1} = \exp([\mu - \frac{1}{2}\sigma_t^2]\lambda + \sigma_t\sqrt{\lambda}Z_{t+1})$, where Z_{t+1} is standard normal, σ_t is the conditional ($\mathcal{F}_t$-measurable) volatility of returns and λ denotes the time difference in years between period t and $t+1$. In this setup, $M_{t+1} :=$

$\exp(-[r_f + \xi_t^2/2]\lambda - \xi_t\sqrt{\lambda}Z_{t+1})$, is a valid SDF with conditional Sharpe ratio

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

The implied dynamics under risk-neutral measure are given by

$$R_{t+1} = \exp\left((r_f - \frac{1}{2}\sigma_t^2)\lambda + \sigma_t\sqrt{\lambda}Z_{t+1}\right). \tag{5.20}$$

The next Theorem shows that quantile regression using the risk-neutral quantile as a regressor renders an optimal forecast.

Theorem 5.7. *Consider the lognormal model described above with return observations* $\{R_{t+1}\}_{t=1}^T$ *stacked in the* $T \times 1$ *vector* R. *Let* $\widetilde{X}_t(\tau) \coloneqq [1\ \widetilde{Q}_{t,\tau}(R_{t+1})]^\top$ *and denote the* $T \times 2$ *matrix of stacked* $\widetilde{X}_t(\tau)$ *by* $\widetilde{X}(\tau)$. *Define the regression quantile* $\widehat{\beta}(\tau; R, \widetilde{X}(\tau))$ *as the solution to the quantile regression with the risk-neutral quantile as a covariate*

$$\widehat{\beta}(\tau; R, \widetilde{X}(\tau)) \in \underset{\beta\in\mathbb{R}^2}{\arg\min} \sum_{t=1}^{T} \rho_\tau\left(R_{t+1} - \widetilde{X}_t(\tau)^\top\beta\right),$$

where $\rho_\tau(u) = (\tau - \mathbb{1}(u \le 0))u$.

Similarly, let $X_t(\tau) \coloneqq [1\ Q_{t,\tau}(R_{t+1})]^\top$, $X(\tau)$ *the* $T \times 2$ *matrix of stacked* $X_t(\tau)$ *and define* $\widehat{\beta}(\tau; R, X(\tau))$ *as the solution to the quantile regression using the physical quantile as a covariate*

$$\widehat{\beta}(\tau; R, X(\tau)) \in \underset{\beta\in\mathbb{R}^2}{\arg\min} \sum_{t=1}^{T} \rho_\tau\left(R_{t+1} - X_t(\tau)^\top\beta\right). \tag{5.21}$$

Then

$$\widetilde{X}_{T+1}(\tau)^\top\widehat{\beta}(\tau; R, \widetilde{X}(\tau)) = X_{T+1}(\tau)^\top\widehat{\beta}(\tau; R, X(\tau)). \tag{5.22}$$

That is, quantile prediction based on the risk-neutral quantile is numerically identical to quantile prediction based on the physical quantile.

Proof. See Appendix A.5. ■

The importance of Theorem 5.7 comes from the fact that $\widetilde{Q}_{t,\tau}(R_{t+1})$ can be inferred from market data at time t and hence no information is lost by

using this as a predictor of the physical quantile, as opposed to using the actual physical quantile as a predictor variable. If the conditional physical quantile function is known, a quantile regression of R_{t+1} on

$$\beta_0(\tau) + \beta_1(\tau) Q_{t,\tau}(R_{t+1}),$$

renders the limiting values $\widehat{\beta}_0(\tau) \to 0, \widehat{\beta}_1(\tau) \to 1$. To predict the next quantile, one could use

$$\widehat{\beta}_0(\tau) + \widehat{\beta}_1(\tau) Q_{T+1,\tau}(R_{T+2}) = Q_{T+1,\tau}(R_{T+2}) + o_p(1). \tag{5.23}$$

Theorem 5.7 details that the predicted value on the left hand side of (5.23) is the same as the prediction obtained using quantile regression with observable $\widetilde{Q}_\tau(R_{t+1})$ as regressor. A similar situation occurs in Principal Component Analysis, where for OLS regression, it is enough to identify the principal component up to some rotation matrix (Bai, 2003).

The assumption underlying the result is that the only source of variation in the distribution of returns is changes in conditional volatility. This is in essence the same idea underlying the popular GARCH models. However, I have abstracted away from specifying what actually drives the volatility process, as opposed to GARCH type models. Hence, the result of Theorem 5.7 is valid for any conditional volatility specification. The result comes at the cost of modeling the returns as conditionally lognormal. There is ample evidence that returns are not conditionally lognormal (Martin, 2017), but given the popularity of the lognormal assumption in financial models it is still an interesting benchmark to consider.

Theorem 5.7 can actually be used as model free evidence against the conditional lognormal assumption. To see this, we can use quantile regression to estimate coefficients $\widehat{\beta}_{0,t}(\tau) + \widehat{\beta}_{1,t}(\tau)$, using $\widetilde{Q}_{t,\tau}$ as the only covariate to explain the quantiles of $R_{m,t+1}$. The t-subscript in $\beta_{\cdot,t}$ refers to the fact that the coefficients are estimated using information up to time t. To allow for time variation in the parameters I use a rolling window of size 1,000. The estimated coefficients

are used to produce dynamic quantile forecasts of the form

$$\widehat{Q}_{t,\tau} = \widehat{\beta}_{0,t} + \widehat{\beta}_{1,t}\widetilde{Q}_{t,\tau}.$$

If the return data follow the lognormal dynamics in (5.20), we expect by Theorem 5.7, that

$$\widehat{Q}_{t,\tau} \approx Q_{t,\tau}.$$

This hypothesis can be formalized by testing for the joint restriction

$$H_0 : [\beta_0(\tau), \beta_1(\tau)] = [0, 1]^\top. \tag{5.24}$$

in the quantile regression

$$\min_{\beta_0(\tau),\beta_1(\tau)} \sum_t \rho_\tau(R_{m,t+1} - \beta_0(\tau) - \beta_1(\tau)\widehat{Q}_{t,\tau}).$$

The same Wald restriction test is used as in (5.16), using SETBB to compute the covariance matrix. The results are summarized in Table 5 for several prediction horizons. Interestingly, $\widehat{Q}_{t,\tau}$ tends to underestimate the true quantiles, since the number of violations is higher than τ in all cases. The Wald restriction test rejects the null hypothesis (5.24) at any conventional significance level. This formalizes the evidence against conditional lognormal models. Additionally, if returns are conditionally lognormal, Theorem 5.7 assures that quantiles are highly predicable using $\widetilde{Q}_{t,\tau}$ as a covariate in a quantile regression. Section 5.3 argues this leads to near arbitrage opportunities, which, once more, seems to be at odds with the actual data.

The inconsistencies that arise from the conditional lognormal assumption has been documented before by Martin (2017, Result 4), using a completely different approach. The conclusion from that paper applies more generally though, since it allows for time variation in the mean and risk-free rate.

5.5 Crash risk premium

Evidence from Table 3 suggests that $LRB_t(\tau)$ is a good approximation of $\tau - \mathbb{P}_t(R_{m,t+1} \leq \widetilde{Q}_{t,\tau})$. With some rewriting, we can link $LRB_t(\tau)$ to the premium

Table 5: Risk neutral quantile prediction performance

Horizon	30 days			60 days			90 days		
100τ	1	2.5	5	1	2.5	5	1	2.5	5
% Viol.	1.57	3.44	5.75	1.24	3.17	7.19	2.29	3.43	5.73
Wald test	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Note: *Quantile prediction evaluation using a rolling window of length 1,000.* % Viol. *denotes the percentage of violations of the predicted quantile at level* τ. *Wald test contains p-values corresponding to the joint restriction* $[\beta_0(\tau), \beta_1(\tau)]^\top = [0, 1]^\top$.

on crash risk, which I formally define as

$$CRP_{t+1} \coloneqq \frac{1}{T}\left(\mathbb{E}_t\left[\mathbb{1}\left(R_{m,t+1} \leq \alpha\right)\right] - \widetilde{\mathbb{E}}_t\left[\mathbb{1}\left(R_{m,t+1} \leq \alpha\right)\right]\right). \tag{5.25}$$

Here, T, is a scaling factor which converts the premium to annual units. For example, if the time difference between $t+1$ and t is 30 days, T would be 30/365. To illustrate the crash risk interpretation, suppose that $\alpha = 0.8$, then (5.25) corresponds to the (annualized) return on buying an asset that pays out 1\$ whenever the market return, one period from now, incurs a loss of 20% or more. Using $\widetilde{Q}_\tau = \alpha$, so that $\tau = \widetilde{\mathbb{P}}_t(\alpha)$, we get

$$CRP_{t+1} = -\frac{1}{T} LRB_t\left(\widetilde{\mathbb{P}}_t(\alpha)\right).$$

By construction, this proxy is completely forward looking and can be calculated without using historical data, thereby avoiding the historical bias critique of Goyal and Welch (2008). Also, the physical probability of a crash in this setup concurs with the one from Martin (2017) if the representative agent has log preferences (see Appendix D.2).

Figure 6 shows the crash risk premium for $\alpha \in \{0.95, 0.9, 0.85\}$ for 30 and 60 day horizons. The premium is negative throughout the sample period, which is expected given that the lower bound is positive in the data for the parameters reported in Table 3. The graph documents that the risk premium was sharply decreasing during the period associated with the 2008 financial crisis. The lowest dip occurs on September 28, 2008, which, until the stock market crash of 2020, was the largest point drop in history.[8] Hence, the forward look-

[8] See https://www.thebalance.com/stock-market-crash-of-2008-3305535#:~:text=The%20stock%20market%20crash%20of%202008%20occurred%20on%20Sept.,largest%

ing crash premium measure aligns well with periods that are associated with more uncertainty.[9] This is consistent with the evidence of Bates (2000), who documents that implicit jump risk was highest right after considerable market drops. Bates (2000) derives this conclusion by fitting a 2-factor stochastic volatility/jump model, whereas the results in Figure 6 are essentially semiparametric, leveraging on the existence of a representative agent with well defined utility whose wealth is invested completely in the market. This makes our results more robust to model misspecification.

It is evident from Figure 6 that the crash risk premium fluctuates significantly over time. This raises the question of what the drivers are of the crash risk premium? Bates (2008) develops a model with "crash-averse" preferences, which sheds light on the drivers of the crash risk premium and can explain several empirical anomalies. However, since the model of Bates (2008) imposes i.i.d. evolution of returns, it cannot replicate the time variation in Figure 6 and neither can the disaster risk model (Example 2.4) for the same reason. To get this time variation, one has to resort to more exotic processes, such as the jump-diffusion model with time varying jump intensity proposed by Wachter (2013).

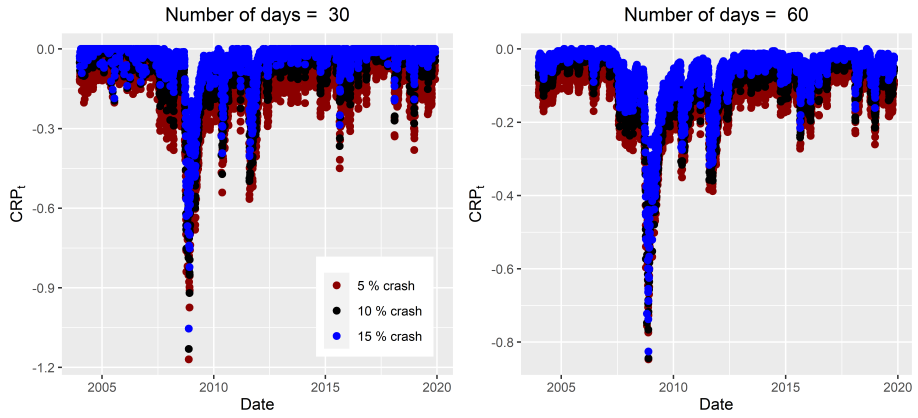

Figure 6: Plots of the crash risk premium CRP_{t+1}, where $\alpha \in \{0.95, 0.9, 85\}$.

20point%20drop%20in%20history. Recall also that the COVID pandemic is not in our sample.

[9] These estimates are conservative, since from the proof of Corollary, we have: 5.5 $\tau - \mathbb{P}_t(\widetilde{Q}_\tau) \geq LRB_t(\tau)$.

6 Conclusion

This paper proposes a new bound (quantile bound) on the volatility of the SDF, which, in contrast to the HJ bound, compares the physical distribution to the risk-neutral distribution at any τ-quantile level. I show that the quantile bound compares favorably to the HJ bound in scenarios where return data are heavy tailed or models that incorporate disaster risk. Among others, the quantile bound uses information of the data beyond the mean and variance, which are the focal statistics for the HJ bound. In the data, I find suggestive evidence of the presence of disaster risk and show that the quantile bound is stronger than the HJ bound. I argue that this points to misspecification of asset pricing models, such as CAPM or LRR, since they fail to incorporate disaster risk.

Subsequently, I analyze the conditional difference between physical and risk neutral quantiles. A von Mises expansion is used to analyze this difference, which, under mild economic constraints, can be approximated by moments of the risk neutral distribution. Quantile regression estimates confirm that the lower bound predicts well in the data. This complements the vast literature in finance which is concerned with conditional mean forecasts (Goyal and Welch, 2008; Campbell and Thompson, 2008). I take the analogy further and develop a measure of out-of-sample performance, which shows that the lower bound outperforms the historical quantile in predicting the conditional return quantiles. The outperformance is marginal, which, as I argue, is to be expected if a liquid option market exists. A final application of this result considers the forward looking premium on crash risk, which is shown to fluctuate significantly over time and spikes downward during periods associated with market distress. This imposes yet another challenge to asset pricing models, as conditional fluctuations in the crash risk premium seem to be a pervasive characteristic of the data, which the vanilla disaster risk model cannot reconcile.

A Proofs and computations

A.1 Proof of Theorem 2.1 (quantile bound)

Proof. I suppress the dependence of the τ-quantile on R and write $\widetilde{Q}_\tau \coloneqq \widetilde{Q}_\tau(R)$. Start from the definition of a risk-neutral quantile

$$\begin{aligned}\tau &= \widetilde{\mathbb{P}}\left[R \leq \widetilde{Q}_\tau\right] = \widetilde{\mathbb{E}}\left[\mathbb{1}\left(R \leq \widetilde{Q}_\tau\right)\right] = R_f \mathbb{E}_t\left[M\mathbb{1}\left(R \leq \widetilde{Q}_\tau\right)\right] \\ &= R_f\left[\mathbb{COV}\left(M, \mathbb{1}\left(R \leq \widetilde{Q}_\tau\right)\right) + \mathbb{E}\left[M\right]\mathbb{E}\left[\mathbb{1}\left(R \leq \widetilde{Q}_\tau\right)\right]\right] \\ &= R_f\,\mathbb{COV}\left(M, \mathbb{1}\left(R \leq \widetilde{Q}_\tau\right)\right) + \underbrace{\mathbb{E}\left[\mathbb{1}\left(R \leq \widetilde{Q}_\tau\right)\right]}_{=\mathbb{P}(R\leq\widetilde{Q}_\tau)}. \end{aligned} \tag{A.1}$$

Rearranging then yields

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

Using Cauchy-Schwarz renders the inequality

$$\begin{aligned}\frac{\left|\tau - \mathbb{P}(R \leq \widetilde{Q}_\tau)\right|}{R_f} &\leq \sigma(M)\sigma\left(\mathbb{1}\left(R \leq \widetilde{Q}_\tau\right)\right) \\ \frac{\left|\tau - \mathbb{P}(R \leq \widetilde{Q}_\tau)\right|}{\sigma\left(\mathbb{1}\left(R \leq \widetilde{Q}_\tau\right)\right) R_f} &\leq \sigma(M). \end{aligned} \tag{A.2}$$

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

$$\sigma\left(\mathbb{1}\left(R \leq \widetilde{Q}_\tau\right)\right) = \sqrt{\mathbb{P}(\widetilde{Q}_\tau(R)) \times (1 - \mathbb{P}(\widetilde{Q}_\tau(R)))}. \tag{A.3}$$

Theorem 2.1 now follows after substituting (A.3) into (A.2). ∎

A.2 Lognormal return and SDF

This Section provides a closed form approximation for the relative efficiency between the HJ and quantile bound under joint lognormality. Let

$$R_{t+1} = e^{(\mu_R - \frac{\sigma_R^2}{2})\lambda + \sigma_R\sqrt{\lambda}Z_r}$$
$$M_{t+1} = e^{-(r_f + \frac{\sigma_M^2}{2})\lambda + \sigma_M\sqrt{\lambda}Z_M}.$$

Both Z_R and Z_M are standard normal random variables with correlation ρ. First, approximate M_{t+1} by a first order Taylor expansion, which gives

$$\widehat{M_{t+1}} = e^{-(r_f + \frac{\sigma_M^2}{2})\lambda} + Z_M\sigma_M\sqrt{\lambda}e^{-(r_f + \frac{\sigma_M^2}{2})\lambda}.$$

Notice that $\widehat{M_{t+1}} = M_{t+1} + o_p(\sqrt{\lambda})$. Consequently, by Stein's Lemma

$$\begin{aligned}
&\mathbb{COV}(R_{t+1}, M_{t+1}) \approx \mathbb{COV}(R_{t+1}, \widehat{M_{t+1}}) = \sigma_M\sqrt{\lambda}e^{-(r_f + \frac{\sigma_M^2}{2})\lambda}\,\mathbb{COV}(R_{t+1}, Z_M) \\
&= \sigma_M\sqrt{\lambda}e^{-(r_f + \frac{\sigma_M^2}{2})\lambda}\mathbb{E}\left[\sigma_R\sqrt{\lambda}\exp\left(\left[\mu_R - \frac{\sigma_R^2}{2}\right]\lambda + \sigma_R\sqrt{\lambda}Z_R\right)\right]\mathbb{COV}(Z_R, Z_M) \\
&= \sigma_M\sigma_R\lambda e^{-(r_f + \frac{\sigma_M^2}{2})\lambda}e^{\mu_R\lambda}\,\mathbb{COV}(Z_R, Z_M).
\end{aligned}$$

Again by Stein's Lemma

$$\begin{aligned}
&\mathbb{COV}(\mathbb{1}\left(\log R_{t+1} \le x\right), M_{t+1}) \approx \mathbb{COV}\left(\mathbb{1}\left(\log R_{t+1} \le x\right), \widehat{M_{t+1}}\right) \\
&= \sigma_M\sqrt{\lambda}e^{-(r_f + \frac{\sigma_M^2}{2})\lambda}\,\mathbb{COV}\left(\mathbb{1}\left(\log R_{t+1} \le x\right), Z_M\right) \\
&= \sigma_M\sqrt{\lambda}e^{-(r_f + \frac{\sigma_M^2}{2})\lambda}\,\mathbb{COV}\left(\mathbb{1}\left((\mu_R - \sigma_R^2/2)\lambda + \sigma_R\sqrt{\lambda}Z_R \le x\right), Z_M\right) \\
&= \sigma_M\sqrt{\lambda}e^{-(r_f + \frac{\sigma_M^2}{2})\lambda}f\left(x\right)\mathbb{COV}\left(Z_R, Z_M\right).
\end{aligned}$$

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,

$$\left|\frac{\mathbb{E}\left[R_{t+1}\right] - e^{\lambda r_f}}{\tau - \mathbb{P}(R_{t+1} \le \widetilde{Q}_\tau)}\right| \approx \frac{\sigma_R\sqrt{\lambda}e^{\mu_R\lambda}}{f(x)}. \tag{A.4}$$

The same reasoning in Example 2.2 implies that the relative efficiency between the HJ and quantile bound can be approximated by

$$\frac{\text{HJ bound}}{\text{Quantile bound}} = \frac{\frac{|\mathbb{E}[R_{t+1}]-R_{f,t+1}|}{\sigma(R_{t+1})R_{f,t+1}}}{\frac{|\tau-\mathbb{P}(R_{t+1}\leq\widetilde{Q}_\tau)|}{\sqrt{\mathbb{P}(R_{t+1}\leq\widetilde{Q}_\tau)(1-\mathbb{P}(R_{t+1}\leq\widetilde{Q}_\tau))}R_{f,t+1}}} \overset{(A.4)}{\approx} \frac{\sqrt{\mathbb{P}(r_{t+1}\leq x)(1-\mathbb{P}(r_{t+1}\leq x)}}{\sigma(R_{t+1})} \times \frac{\sigma_R\sqrt{\lambda}e^{\mu_R\lambda}}{f(x)}. \quad \text{(A.5)}$$

Here, $r_{t+1} = \log R_{t+1}$ and $x = \log \widetilde{Q}_\tau$. Using the same reasoning as in Example 2.2, the expression on the RHS of (A.5) is minimized by choosing $x = \log \widetilde{Q}^*_\tau$ s.t. $\mathbb{P}(R_{t+1} \leq \widetilde{Q}^*_\tau) = 1/2$. 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.3 Formulas for market moments

This Section presents formulas for the (un)truncated risk-neutral moments of excess market return. An alternative way to calculate these is provided in Chabi-Yo and Loudis (2020, Appendix B). I use a slight abuse of notation and write $\widetilde{Q}_{R_{t+1}}(\tau) \coloneqq \widetilde{Q}_{t,\tau}(R_{t+1})$, to obviate that the integrals below are taken with respect to τ.

Proposition A.1. *Higher order risk-neutral moments can be computed directly from the risk-neutral quantile function*

$$\widetilde{\mathbb{E}}_t\left[(R_{t+1}-R_{f,t+1})^n\right] = \int_0^1 [\widetilde{Q}_{R_{t+1}-R_{f,t+1}}(\tau)]^n d\tau = \int_0^1 [\widetilde{Q}_{R_{t+1}}(\tau)-R_{f,t+1}]^n d\tau. \quad \text{(A.6)}$$

And the truncated higher order risk-neutral moments also follow from

$$\widetilde{\mathbb{E}}_t\left[(R_{t+1}-R_{f,t+1})^n \mathbb{1}\left(R_{t+1}\leq k_0\right)\right] = \int_0^{\widetilde{F}_t(k_0)} [\widetilde{Q}_{R_{t+1}}(\tau)-R_{f,t+1}]^n d\tau.$$

Where $\widetilde{F}_t(x) \coloneqq \widetilde{\mathbb{P}}_t(R_{t+1}\leq x)$ is the risk-neutral CDF. Frequently I use $k_0 = \widetilde{Q}_{t,\tau}(R_{t+1})$, in which case the truncated moment formula reduces to

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

Proof. For any random variable X and integer n s.t. the n-th moment exist

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

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.6). The truncated formula follows similarly. ■

A.4 Proof of Theorem 5.4 and Corollary 5.5

Proof of Theorem 5.4 and Corollary 5.5. I split the proof in three parts.

Part 1: Showing that $\widetilde{\mathbb{COV}}_t\left[\mathbb{1}\left(R_{m,t+1} \leq \widetilde{Q}_{t,\tau}\right), (R_{m,t+1} - R_{f,t+1})^k\right] \leq 0$ *for* k *odd.*

Temporarily write $X = R_{m,t+1}$. To prove the claim above I distinguish 3 cases. Take two i.i.d. copies X_1, X_2 with the same law as X under risk-neutral measure and consider

$$\Lambda \coloneqq \underbrace{\left(\mathbb{1}\left(X_1 \leq \widetilde{Q}_{t,\tau}\right) - \mathbb{1}\left(X_2 \leq \widetilde{Q}_{t,\tau}\right)\right)}_{=I} \underbrace{\left((X_1 - R_{f,t+1})^k - (X_2 - R_{f,t+1})^k\right)}_{=II}. \tag{A.7}$$

Case 1: $(I = -1)$. This implies $X_2 < X_1$. Since k is odd I get $II > 0$ so that $\Lambda < 0$.

Case 2: $(I = 0)$. This implies $\Lambda = 0$.

Case 3: $(I = 1)$. This implies $X_1 \leq X_2$ and hence $II < 0$. Therefore $\Lambda < 0$.

Combining all three cases I get that $\Lambda \leq 0$ almost surely. Take conditional (risk-neutral) expectations on both sides of (A.7), using the non-positivity of Λ and the independence of X_1, X_2 proves that the covariance term is negative. Since by assumption $\theta_k \geq 0$ when k is odd I obtain

$$\begin{aligned}
&\widetilde{\mathbb{E}}_t\left[\mathbb{1}\left(R_{m,t+1} \leq \widetilde{Q}_{t,\tau}\right)(R_{m,t+1} - R_{f,t+1})^k\right] - \tau\widetilde{\mathbb{E}}_t\left[(R_{m,t+1} - R_{f,t+1})^k\right] \leq 0 \\
&\implies \tau\widetilde{\mathbb{M}}_{t+1}^{(k)} - \widetilde{\mathbb{M}}_{t+1}^{(k)}[\widetilde{Q}_{t,\tau}] \geq 0 \implies \theta_k\left(\tau\widetilde{\mathbb{M}}_{t+1}^{(k)} - \widetilde{\mathbb{M}}_{t+1}^{(k)}[\widetilde{Q}_{t,\tau}]\right) \geq 0.
\end{aligned}$$

Part 2: Showing that $\widetilde{\mathbb{COV}}_t\left[\mathbb{1}\left(R_{m,t+1} \leq \widetilde{Q}_{t,\tau}\right), (R_{m,t+1} - R_{f,t+1})^k\right] \geq 0$ *for* k *even and* τ *small enough.*

This requires more delicate reasoning. First note that the covariance term goes to zero as $\tau \to 0$ as a consequence of the Cauchy-Schwarz inequality and the continuity of probability measures. Hence, to show that for τ small enough the covariance term is positive, it suffices that the covariance term, seen as a function of τ, has positive slope for τ small enough. To show this, write the covariance as

$$\widetilde{\mathbb{E}}_t\left[(\mathbb{1}\left(R_{m,t+1} \leq \widetilde{Q}_{t,\tau}\right) - \tau)(R_{m,t+1} - R_{f,t+1})^k\right]. \tag{A.8}$$

Consider the associated function

$$\begin{aligned}\Gamma(\tau) &\coloneqq \widetilde{\mathbb{E}}_t\left[(\mathbb{1}\left(R_{m,t+1} \leq \widetilde{Q}_{t,\tau}\right) - \tau)(R_{m,t+1} - R_{f,t+1})^k\right] \\ &= \int_{-\infty}^{\widetilde{Q}_{t,\tau}} (R - R_{f,t+1})^k \widetilde{f}_{R_{m,t+1}}(R)dR - \tau\int_{-\infty}^{\infty}(R - R_{f,t+1})^k \widetilde{f}_{R_{m,t+1}}(R)dR.\end{aligned}$$

Here $\widetilde{f}_{R_{m,t+1}}(\cdot)$ is the (risk-neutral) PDF of the market return. From Leibniz' rule

$$\begin{aligned}\frac{\partial}{\partial\tau}\Gamma(\tau) &= (\widetilde{Q}_{t,\tau} - R_{f,t+1})^k \widetilde{f}_{R_{m,t+1}}(\widetilde{Q}_{t,\tau})\frac{\partial\widetilde{Q}_{t,\tau}}{\partial\tau} - \widetilde{\mathbb{E}}_t\left[(R_{m,t+1} - R_{f,t+1})^k\right] \\ &= (\widetilde{Q}_{t,\tau} - R_{f,t+1})^k - \widetilde{\mathbb{E}}_t\left[(R_{m,t+1} - R_{f,t+1})^k\right],\end{aligned} \tag{A.9}$$

since, by the rules for derivatives of inverses

$$\frac{\partial\widetilde{Q}_{t,\tau}}{\partial\tau} = \frac{1}{\widetilde{f}_{R_{m,t+1}}(\widetilde{Q}_{t,\tau})}.$$

Because I assume that $\sup_k \|R_{m,t+1}\|_k < \infty$, it follows that (A.9) is positive for all $\tau \in [0, \tau^*]$, where τ^* solves

$$\widetilde{Q}_{t,\tau^*} = R_{f,t+1} - \sup_k \|R_{m,t+1} - R_{f,t+1}\|_k.$$

In conclusion, I have shown that the covariance (A.8) vanishes when $\tau \to 0^+$ and the slope of (A.8) is positive for all $\tau \leq \tau^*$. This means that (A.8) is

positive for all $\tau \leq \tau^*$. Thus for all such $\tau \in (0, \tau^*]$

$$\tau \widetilde{\mathbb{M}}_{t+1}^{(k)} - \widetilde{\mathbb{M}}_{t+1}^{(k)}[\widetilde{Q}_{t,\tau}] \leq 0.$$

Hence, since $\theta_k \leq 0$ for k even

$$\theta_k \left(\tau \widetilde{\mathbb{M}}_{t+1}^{(k)} - \widetilde{\mathbb{M}}_{t+1}^{(k)}[\widetilde{Q}_{t,\tau}] \right) \geq 0.$$

Part 3: Combining both cases

I have now established $\theta_k(\tau \widetilde{\mathbb{M}}_{t+1}^{(k)} - \widetilde{\mathbb{M}}_{t+1}^{(k)}[\widetilde{Q}_{t,\tau}]) \geq 0$ for all k and $\tau \leq \tau^*$. Therefore

$$\begin{aligned} Q_{t,\tau} - \widetilde{Q}_{t,\tau} &\approx \frac{1}{f_t(\widetilde{Q}_{t,\tau})} \left(\frac{\sum_{k=1}^{\infty} \theta_k \left(\tau \widetilde{\mathbb{M}}_{t+1}^{(k)} - \widetilde{\mathbb{M}}_{t+1}^{(k)}[\widetilde{Q}_{t,\tau}] \right)}{1 + \sum_{k=1}^{\infty} \theta_k \widetilde{\mathbb{M}}_{t+1}^{(k)}} \right). \\ &\geq \frac{1}{f_t(\widetilde{Q}_{t,\tau})} \left(\frac{\sum_{k=1}^{3} \theta_k \left(\tau \widetilde{\mathbb{M}}_{t+1}^{(k)} - \widetilde{\mathbb{M}}_{t+1}^{(k)}[\widetilde{Q}_{t,\tau}] \right)}{1 + \sum_{k=1}^{3} \theta_k \widetilde{\mathbb{M}}_{t+1}^{(k)}} \right). \end{aligned} \tag{A.10}$$

If additionally Assumption 5.3 holds, then

$$\theta_1 = \frac{1}{R_{f,t+1}}, \theta_2 = -\frac{1}{R_{f,t+1}^2}, \text{ and } \theta_3 = \frac{1}{R_{f,t+1}^3}.$$

Using this in (A.10) gives

$$Q_{t,\tau} - \widetilde{Q}_{t,\tau} \geq \frac{1}{f_t(\widetilde{Q}_{t,\tau})} \left(\frac{\sum_{k=1}^{3} \frac{(-1)^{k+1}}{R_{f,t+1}^k} \left(\tau \widetilde{\mathbb{M}}_{t+1}^{(k)} - \widetilde{\mathbb{M}}_{t+1}^{(k)}[\widetilde{Q}_{t,\tau}] \right)}{1 + \sum_{k=1}^{3} \frac{(-1)^{k+1}}{R_{f,t+1}^k} \widetilde{\mathbb{M}}_{t+1}^{(k)}} \right).$$

■

A.5 Proof of Theorem 5.7

Proof. By definition

$$\tau = \mathbb{P}_t(R_{t+1} \leq Q_{t,\tau}) = \mathbb{P}_t \left(\exp\left(-\frac{1}{2}\sigma_t^2 \lambda + \sigma_t \sqrt{\lambda} Z_{t+1} \right) \leq \exp(-\mu\lambda) Q_{t,\tau} \right).$$

Similarly

$$\tau = \widetilde{\mathbb{P}}_t(R_{t+1} \le \widetilde{Q}_{t,\tau}) = \widetilde{\mathbb{P}}_t\left(\exp\left(-\frac{1}{2}\sigma_t^2\lambda + \sigma_t\sqrt{\lambda}Z_{t+1}\right) \le \exp(-r_f\lambda)\widetilde{Q}_{t,\tau}\right).$$

As a result

$$e^{(\mu - r_f)\lambda}\widetilde{Q}_{t,\tau} = Q_{t,\tau}. \tag{A.11}$$

Recall that the quantile regression estimator is equivariant to reparametrization of design: for any 2×2 nonsingular matrix A, we have

$$\widehat{\beta}(\tau; R, XA) = A^{-1}\widehat{\beta}(\tau; R, X).$$

By Equation (A.11)

$$X(\tau) = \widetilde{X}(\tau) \times \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & e^{(\mu - r_f)\lambda} \end{bmatrix}}_{:=A}.$$

Therefore

$$\widehat{\beta}(\tau; R, X(\tau)) = \widehat{\beta}(\tau; R, \widetilde{X}(\tau)A) = A^{-1}\widehat{\beta}(\tau; R, \widetilde{X}(\tau)).$$

Hence, the predicted quantile using the physical quantile regression (5.21) equals

$$\begin{aligned} [1 \quad Q_{T+1,\tau}(R_{T+2})]\ \widehat{\beta}(\tau; R, X(\tau)) &= [1 \quad Q_{T+1,\tau}(R_{T+2})]A^{-1}\ \widehat{\beta}(\tau; R, \widetilde{X}(\tau)) \\ &= [1 \quad \widetilde{Q}_{T+1,\tau}(R_{T+2})]\ \widehat{\beta}(\tau; R, \widetilde{X}(\tau)). \end{aligned}$$

This is exactly (5.22). ■

A.6 Computation of Gâteaux derivative (5.3)

In this section we prove (5.3), which states that

$$\phi'_{\widetilde{F}_t}(F_t - \widetilde{F}_t) = \frac{\tau - F_t(\widetilde{Q}_{t,\tau})}{\widetilde{f}_t(\widetilde{Q}_{t,\tau})}.$$

For ease of exposition, we drop the time subscripts. For $\lambda \in [0,1]$, define $\widetilde{F}_\lambda \coloneqq (1-\lambda)\widetilde{F} + \lambda F$. The following (trivial) identity will prove helpful[10]

$$\tau = \widetilde{F}_\lambda \widetilde{F}_\lambda^{-1}. \tag{A.12}$$

To further simplify notation, write $q(\lambda) \coloneqq \widetilde{F}_\lambda^{-1}$. Then (A.12) becomes

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

Applying the implicit function theorem, we obtain

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

Plug in $\lambda = 0$ to get

$$q'(0) = -\frac{-\widetilde{F}(q(0)) + F(q(0))}{\widetilde{f}(q(0))}. \tag{A.13}$$

Notice that

$$\widetilde{F}_\lambda\big|_{\lambda=0} = \widetilde{F} \implies q(\lambda)\big|_{\lambda=0} = q(0) = \widetilde{F}^{-1}. \tag{A.14}$$

Substitute (A.14) into (A.13) to obtain

$$q'(0) = -\frac{-\widetilde{F}(\widetilde{F}^{-1}) + F(\widetilde{F}^{-1})}{\widetilde{f}(\widetilde{F}^{-1})} = \frac{\tau - F(\widetilde{F}^{-1})}{\widetilde{f}(\widetilde{F}^{-1})}. \tag{A.15}$$

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

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

[10] This "equality" may actually only be an inequality for some τ, but this is immaterial to the argument.

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 data from OptionMetrics covering the period 01-01-1996 until 12-31-2019. This consists of European Put and Call option data with time to expiration less than 500 days on the S&P 500 index. The option contract further contains data on the highest closing bid and lowest closing ask price and price of the forward contract on the underlying security. In addition, I obtain data on the daily risk-free rate from Kenneth French' website.[11] Finally, stock price data on the closing price of the S&P 500 are obtained via WRDS.

Prior to estimating the martingale measure, I use an additional data cleaning procedure for the option data. All observations are dropped for which the highest closing bid price equals zero, as well as all option prices that violate no-arbitrage bounds. This is similar to the cleaning procedure of Martin (2017) and leaves a total of 16,624,104 option-day observations.

B.2 Estimating the risk-neutral quantile function

There is a substantial literature on how to extract the martingale measure from option prices. I follow Figlewski (2008), with some minor modifications emphasized below to estimate the conditional and unconditional risk-neutral quantile curve for tenors $30, 60, 90, 180$ and 360 days.

(i) Construct option prices: Use the midpoint of the highest closing bid and lowest closing ask price to obtain the option price.

(ii) Convert option prices to Black-Scholes implied volatilities (IVs): Use out-of-the money put and call option prices to construct IVs, since these tend to be more liquid than in-the-money options. I regard a put-option out-of-the-money if the strike prices is less than the forward price on the underlying security. This is consistent with Martin (2017). Forward prices are provided by OptionMetrics.

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

(iii) Interpolate the IVs using a smoothing cubic spline: I use a smoothing cubic spline (Wahba, 1990) with 7 knots to interpolate the IVs at a dense set of equidistant strike prices

$$K \in \{K_{\min}, K_{\min} + \Delta, K_{\min} + 2\Delta, \ldots, K_{\max}\} \quad \Delta = \frac{K_{\max} - K_{\min}}{500},$$

where $K_{\min}, K_{\max}$ are the respective minimum and maximum strike prices observed in the sample. This differs from Figlewski (2008), who recommends a 4th degree smoothing spline and a single knot. The 4th degree spline in Figlewski (2008) comes from the necessity to obtain a smooth density function, which corresponds to the second derivative of the put-call-option price curve. Since I only need to estimate a CDF, a 3^{rd} degree polynomial suffices and prevents overfitting. The use of 7 knots is arbitrary, as is the single knot in Figlewski (2008), but renders acceptable estimates of the martingale measure in most cases.

(iv) Smooth the IV curve for at-the-money options: There tends to be a discontinuity in the smoothed IV curve for at-the-money options, since puts and calls at the same strike trade on slightly different IVs. Let F_{t+1} be the price of the forward contract, I consider all strike prices $K \in [K_{\text{low}}, K_{\text{high}}]$, where K_{low} is the lowest traded strike such that $F_{t+1}(1-0.02) \leq K_{\text{low}}$ and K_{high} the highest traded strike which satisfies $K_{\text{high}} \leq (1+0.02)F_{t+1}$. Following Figlewski (2008), I use a weighted average of $IV_{\text{put}}(K), IV_{\text{call}}(K)$ to estimate

$$IV(K) = \alpha IV_{\text{put}} + (1-\alpha) IV_{\text{call}} \qquad K \in [K_{\text{low}}, K_{\text{high}}],$$

where

$$\alpha = \frac{K_{\text{high}} - K}{K_{\text{high}} - X_{\text{low}}}.$$

(v) Obtain the central part of the risk-neutral CDF: The smoothed Black-Scholes IVs are converted back to option prices. By the (nonparametric) result of Breeden and Litzenberger (1978)

$$1 + R_{f,t+1} \frac{\partial}{\partial K} \text{Call}_{t+1}(K) = \widetilde{\mathbb{P}}_t(S_{t+1} \leq K) = \widetilde{\mathbb{P}}_t\left(R_{m,t+1} \leq \frac{K}{S_t}\right). \quad \text{(B.1)}$$

I convert put prices to call option prices using put-call parity. Let K_{n-1}, K_n, K_{n+1} be consecutive strike prices on the strike grid, then the partial derivative in (B.1) is approximated by

$$1 + R_{f,t+1}\left[\frac{C_{n+1} - C_{n-1}}{K_{n+1} - K_{n-1}}\right] \approx \widetilde{\mathbb{P}}_t\left(R_{m,t+1} \leq \frac{K}{S_t}\right).$$

(vi) Obtain the risk-neutral CDF in the tails: It remains to estimate the risk-neutral CDF for $K \leq K_2$ and $K \geq K_{499}$, which concerns the left and right tail of the distribution. I follow Figlewski (2008) and fit a generalized extreme value (GEV) distribution in the left and right tail. The GEV distribution function is given by (see De Haan and Ferreira (2007, p. 6)):

$$F_{\text{GEV}}(x) = \exp(-(1+\gamma x)^{-1/\gamma}).$$

A location parameter μ and scale parameter σ can be introduced via the transformation

$$x = \frac{S_{t+1} - \mu}{\sigma}.$$

The function $F_{\text{GEV}}(S_{t+1}; \mu, \sigma, \gamma)$ effectively contains three parameters which will be calibrated to the implied risk-neutral CDF. For the left tail, I use $\tau = 0.02, \tau_2 = 0.03$ and $\tau_3 = 0.05$ and find the associated strikes from the empirical risk-neutral CDF, denoted by $K(\tau_1), K(\tau_2), K(\tau_3)$ respectively. Thereafter, I solve the following non-linear system of equations to match the quantiles of F_{GEV} with the empirical quantiles:

$$\begin{aligned} F_{\text{GEV}}(K(\tau_1)) &= \tau_1 \\ F_{\text{GEV}}(K(\tau_2)) &= \tau_2 \\ F_{\text{GEV}}(K(\tau_3)) &= \tau_3. \end{aligned}$$

I follow the same procedure to extend the risk-neutral CDF to the right tail, by matching the GEV function at the empirical quantile $\tau_1 = 0.95, \tau_2 = 0.97, \tau_3 = 0.98$. I discard tenors for which the 0.02 or 0.98-quantiles are not available, to prevent extrapolation error in the tail. I depart from Figlewski (2008) by matching only quantiles of the CDF, as opposed to

matching the shape of the PDF.

(vii) Interpolate the risk-neutral CDFs for specific tenor: Typically, options with a specfic tenor we want to estimate (e.g. 30 days) are not traded. To overcome this issue, I linearly interpolate the conditional risk-neutral curves with tenors closest to the tenor of interest. For example, suppose on a given day only options with tenor 20 and 45 days are traded and we are interested to obtain the 30-day risk-neutral CDF. In that case, the 20 and 45-day risk-neutrals CDFs are linearly interpolated so as to obtain the 30-day risk-neutral CDF. This is similar to Martin (2017, Appendix A). The conditional risk-neutral quantile curve is constructed as the left continuous inverse of the estimated risk-neutral CDF

$$\widetilde{Q}_{t,\tau} = \inf\left\{x \in \mathbb{R} : \tau \leq \widetilde{\mathbb{P}}_t(R_{m,t+1} \leq x)\right\}.$$

(viii) Estimate the unconditional risk-neutral CDF: Once the conditional risk-neutral measure have been estimated using the previous steps, I estimate the unconditional CDF via

$$\widetilde{\mathbb{P}}_T\left(R_m \leq x\right) \coloneqq \frac{1}{T}\sum_{t=1}^{T}\widetilde{\mathbb{P}}_t\left(R_{m,t+1} \leq x\right).$$

(ix) Estimate the unconditional risk-neutral quantile function: Given the unconditional risk-neutral CDF estimate in Step (viii), I estimate the unconditional risk-neutral quantile function with

$$\widetilde{Q}_T(\tau) \coloneqq \inf\left\{x \in \mathbb{R} : \tau \leq \widetilde{\mathbb{P}}_T(x)\right\}. \tag{B.2}$$

C Bootstrap results

This section presents simulation results of the estimation of the quantile bound and the performance of the bootstrap approach to construct confidence intervals.

C.1 Simulation results

Here I present simulation evidence that the bootstrap procedure outlined in Section 4.2 delivers acceptable confidence intervals in case returns are conditionally lognormal. Assume that the DGP is given by

$$R_{t+1} = \exp\left((\mu - \frac{1}{2}\sigma_t^2)\lambda + \sigma_t\sqrt{\lambda}Z_{t+1}\right), \qquad Z_{t+1} \overset{\text{IID}}{\sim} N(0,1). \tag{C.1}$$

I generate σ_t independent of Z_{t+1} according to

$$\sigma_t \overset{\text{IID}}{\sim} \textbf{UNIF}\left[0.06, 0.26\right].$$

It is assumed that σ_t is realized at the start of period t, so that the conditinoal distribution of R_{t+1} is given by

$$\mathbb{P}_t(R_{t+1} \leq x) = \Phi\left(\frac{\log(x) - (\mu - \frac{1}{2}\sigma_t^2)\lambda}{\sigma_t\sqrt{\lambda}}\right).$$

Here, $\Phi(\cdot)$ is the CDF of the standard normal distribution. The unconditional CDF is obtained by integrating out σ_t

$$\mathbb{P}(R_{t+1} \leq x) = \mathbb{E}\left[\mathbb{P}_t(R_{t+1} \leq x)\right] = \mathbb{E}\left[\Phi\left(\frac{\log(x) - (\mu - \frac{1}{2}\sigma_t^2)\lambda}{\sigma_t\sqrt{\lambda}}\right)\right].$$

The conditional SDF is given by

$$M_{t+1} = \exp\left(-(r + \frac{1}{2}\xi_t^2)\lambda - \xi_t\sqrt{\lambda}Z_{t+1}\right),$$

where Z_{t+1} is the same as in (C.1) and ξ_t is the conditional Sharpe ratio

$$\xi_t = \frac{\mu - r}{\sigma_t}.$$

To mimic the empirical application, I use $\lambda = 30/365$ and pick a total of $T = 300$ returns. In addition, I assume that prices of call and put options conditional on

time t are given by the Black-Scholes formula

$$\begin{aligned} \text{Call}_t(K) &= \Phi(d_1)S_t - \Phi(d_2)Ke^{-r\lambda} \\ \text{Put}_t(K) &= \Phi(-d_2)Ke^{-r\lambda} - \Phi(-d_1)S_t \\ d_1 &= \frac{1}{\sigma_t\sqrt{\lambda}}\left[\ln\left(\frac{S_t}{K}\right) + (r - \frac{\sigma_t^2}{2})\lambda\right] \\ d_2 &= d_1 - \sigma_t\sqrt{\lambda}. \end{aligned} \tag{C.2}$$

As in the empirical application, it is assumed that call and put option prices with maturity exactly equal to 30 days are not traded. Instead, I linearly interpolate the risk-neutral CDFs corresponding to maturities 25 and 37 days respectively (see Step (vii) in Appendix B.2). These CDFs are obtained from the Breeden and Litzenberger (1978) formula, assuming 1,000 strike values per maturity (as in Step (v) of Appendix B.2). This is consistent with the latter part of my empirical sample in Section 4.2. Confidence intervals are created following the bootstrap procedure described in Section 4.2 with 10,000 bootstrap samples.[12] For convenience the main parameter settings are summarized in Table 6.

To assess the accuracy of the sup quantile bound estimator ($\sup_\tau \widehat{\theta}_{\text{smooth}}(\tau)$) and the resulting confidence intervals (described in Section 4.2), I repeat the estimation procedure 1,000 times. The results are summarized in Table 7. The sup quantile bound estimator is upward biased. Moreover, the coverage properties of the 95% confidence interval are below 0.95, as they contain the true quantile bound only 78% of the times. I repeat the same exercise with 1,000 return observations instead of 300. The results are in the bottom row of Table 7. The bias decreases significantly in this case, with an average sup quantile bound of 0.1261. Also the coverage properties are much better, as the confidence bands cover the true supremum 86% of the times.

[12] I use the build in *R* function `tsbootstrap` for this calculation.

Table 6: Parameter settings

r	μ	λ	$\bar{\sigma}$	T	S_0	Number of strike values
0	0.07	30/365	0.16	300	3046	2000

Note: *Parameter values for simulation.* r *is the risk-free rate,* μ *is the growth rate,* λ *is time to maturity,* $\bar{\sigma}$ *is the average volatility with* $\sigma_t \sim \boldsymbol{UNIF}[\bar{\sigma} - 0.1, \bar{\sigma} + 0.1]$*,* T *is the time series length for returns* R_t*,* S_0 *is the starting value of the stock (fixed during simulation) and the last column denotes the number of strike values observed for put and call options with maturity 25 and 37 days.*

Table 7: Simulation results

Time length	Coverage	Mean	Median	True quantile bound	SDF vol
300	0.7770	0.1438	0.1425	0.1177	0.1626
1,000	0.8620	0.1261	0.1261	0.1177	0.1626

Note: *Simulation results quantile bound estimation and confidence intervals using 1,000 independent samples.* Time length *is the number of time series observations.* Coverage *denotes the fraction of times the 95% confidence intervals contain the true supremum.* Mean *is the average value of the quantile estimator and* Median *denotes the median value.* True quantile bound *is the quantile bound estimand.* SDF vol *is the true SDF volatility.*

D Detailed derivations representative agent models

In this Section I show two results about representative agent models which are used in the paper. The first Section describes how to obtain the risk-neutral and physical CDF in the disaster risk model. Section D.2 shows that the subjective crash risk probability derived by Martin (2017) under log preferences is identical to the crash probability I obtain building on the work of Chabi-Yo and Loudis (2020).

D.1 Disaster risk probabilities

Consumption growth is i.i.d. by assumption (see Equation (2.16)). It turns out to be convenient to work with cumulant generating functions (CGF) to find the physical and risk-neutral probabilities of equity (Backus et al., 2011). Let Δc be log consumption growth and define

$$k(s; \Delta c) \coloneqq \log \mathbb{E}\left[e^{s\Delta c}\right].$$

k is the CGF of the random variable Δc. Due to the Poisson mixture assumption, the CGF obtains the explicit form

$$k(s;\Delta c) = \mu s + \frac{\sigma^2 s^2}{2} + \kappa\left(e^{\theta s + \frac{\nu^2 s^2}{2}} - 1\right). \tag{D.1}$$

Since return on equity is a claim on levered consumption growth, the associated CGF is

$$k(s;\lambda\Delta c) = k(\lambda s;\Delta c).$$

The equality follows from (D.1). The SDF is given by $M = \beta e^{-\gamma\Delta c}$ and so $q^1 \coloneqq 1/R_{f,t+1} = \beta k(-\gamma)$. Since by definition the risk-neutral probabilities satisfy $\widetilde{p}(\Delta c) = p(\Delta c)m(\Delta c)/q^1$ it follows

$$\widetilde{k}(s;\Delta c) = k(s-\gamma;\Delta c) - k(-\gamma;\Delta c).$$

As derived in Backus et al. (2011), the CGF of the risk-neutral equity return is given by

$$\widetilde{k}(s;\lambda\Delta c) = \widetilde{k}(\lambda s;\Delta c) = k(\lambda s-\gamma;\Delta c) - k(-\gamma;\Delta c).$$

The characteristic function of return on equity under physical and risk-neutral measure are respectively defined by

$$\varphi(s) \coloneqq \exp(k(is)), \quad \widetilde{\varphi}(s) \coloneqq \exp(\widetilde{k}(is)).$$

Here, i is the imaginary unit. Finally, I obtain numerical approximations of the physical and risk neutral probabilities from the Gil-Pelaez (1951) theorem, which states that

$$\mathbb{P}(\lambda\Delta c \leq x) = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \frac{\Im(e^{-isx}\varphi(s))}{s} ds. \tag{D.2}$$

$\widetilde{\mathbb{P}}(\lambda\Delta c \leq x)$ can be obtained in the same fashion, replacing $\varphi(\cdot)$ with $\widetilde{\varphi}(\cdot)$ in (D.2). This renders the quantile bound for logarithmic returns, which is the same for gross returns, as the two are related via a monotonic transformation.

D.2 Crash probability with known utility

Chabi-Yo and Loudis (2020) 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) shows that

$$\mathbb{P}_t\left(R_{m,t+1} < \alpha\right) = \alpha\left[\mathrm{Put}_t'(\alpha S_t) - \frac{\mathrm{Put}_t(\alpha S_t)}{\alpha S_t}\right], \tag{D.3}$$

where Put_t' is the derivative of the put option price curve seen as a function of the strike. Under log preferences and using (5.5), it follows that

$$\begin{aligned}
&\mathbb{P}_t(R_{m,t+1} < \widetilde{Q}_{t,\tau}) = \tau + \frac{1}{R_{f,t+1}}\widetilde{\mathbb{COV}}_t\left[\mathbb{1}\left(R_{m,t+1} \leq \widetilde{Q}_{t,\tau}\right), R_{m,t+1}\right] \\
&= \tau + \frac{1}{R_{f,t+1}}\left(\widetilde{\mathbb{E}}_t\left[\mathbb{1}\left(R_{m,t+1} \leq \widetilde{Q}_{t,\tau}\right)R_{m,t+1}\right] - \widetilde{\mathbb{E}}_t(R_{m,t+1})\widetilde{\mathbb{E}}_t\left(\mathbb{1}\left(R_{m,t+1} \leq \widetilde{Q}_{t,\tau}\right)\right)\right) \\
&= \frac{1}{R_{f,t+1}}\widetilde{\mathbb{E}}_t\left[\mathbb{1}\left(R_{m,t+1} \leq \widetilde{Q}_{t,\tau}\right)R_{m,t+1}\right].
\end{aligned} \tag{D.4}$$

The result now follows upon substituting $\widetilde{Q}_\tau = \alpha$, since Martin (2017) shows that (D.4) equals the right hand side of (D.3).

E Other SDF bounds

The principal method to use quantiles to derive bounds on the volatility of the SDF can be applied to other well known bounds in the literature. In this Section I revisit some of these SDF bounds and show how the quantile relation can be used to obtain results akin to the quantile version of the HJ bound in Theorem **??**. For all the results to follow it is well known under which conditions the bounds are tight. For example, the log bound in Section E.1 is known to bind for the growth-optimal portfolio. Under some conditions the growth-optimal portfolio is equal to the market portfolio. Using the quantile relation to bound the log of SDF could therefore refute the presumption that the market portfolio is growth optimal, if the quantile bound is significantly stronger. For convenience, recall the relation derived in the proof of Theorem

??, which is used repeatedly in this Section to analyze other SDF bounds

$$\tau = R_{f,t+1}\mathbb{E}_t\left[M_{t+1}\mathbb{1}\left(R_{t+1} \leq \widetilde{Q}_{t,\tau}\right)\right]. \tag{E.1}$$

E.1 Bound of Bansal and Lehmann (1997)

Here I consider a bound on the logarithm of the SDF. By an application of Jensen's inequality, we get

$$\begin{aligned} 0 = \log(1) = \log \mathbb{E}_t\left[M_{t+1}R_{t+1}\right] \geq \mathbb{E}_t\left[\log M_{t+1}\right] + \mathbb{E}_t\left[\log R_{t+1}\right] \\ \implies -\mathbb{E}_t\left[\log M_{t+1}\right] \geq \mathbb{E}_t\left[\log R_{t+1}\right]. \end{aligned}$$

This bound, together with its asset pricing implications, is analyzed in detail by Bansal and Lehmann (1997). It is known to bind for the market portfolio in a representative agent model with log utility. Applying a log transformation to (E.1), we obtain for any $\tau \in (0,1)$

$$\log(\tau) = \log(R_{f,t+1}) + \log\left(\mathbb{E}_t\left[M_{t+1}\mathbb{1}\left(R_{t+1} \leq \widetilde{Q}_{t,\tau}\right)\right]\right).$$

Use Jensen's inequality in a similar vein as above and rearranging gives

$$-\mathbb{E}_t\left[\log\left(M_{t+1}\right)\right] \geq \log(R_{f,t+1}) + \mathbb{E}_t\left[\log\left(\mathbb{1}\left(R_{t+1} \leq \widetilde{Q}_{t,\tau}\right)\right)\right] - \log(\tau).$$

Taking expectations on both sides also renders an unconditional version.

E.2 Bound of Snow (1991)

Snow (1991) derives a continuum of bounds of higher order moments on the SDF. In somewhat simplified form, the idea is to use Hölder's inequality to the defining SDF equation

$$1 = \mathbb{E}_t\left[M_{t+1}R_{t+1}\right] \leq \mathbb{E}_t\left[M_{t+1}^p\right]^{\frac{1}{p}}\mathbb{E}_t\left[R_{t+1}^q\right]^{\frac{1}{q}},$$

for Hölder exponents $\frac{1}{p}+\frac{1}{q}=1$ and $p>1$. Rearranging gives the restriction on the p-th norm of the SDF

$$\mathbb{E}_t\left[M_{t+1}^p\right]^{\frac{1}{p}} \geq \mathbb{E}_t\left[R_{t+1}^q\right]^{-\frac{1}{q}}.$$

The quantile relation (E.1) can similarly be exploited by applying Hölder's inequality on the right hand side. This gives

$$\mathbb{E}_t\left[M_{t+1}^p\right]^{\frac{1}{p}} \geq \left(\frac{\tau}{R_{f,t+1}}\right)\mathbb{E}_t\left[\mathbb{1}\left(R_{t+1}\leq \widetilde{Q}_{t,\tau}\right)\right]^{-\frac{1}{q}}.$$

E.3 Bound of Liu (2020)

Liu (2020) develops a continuum of bounds which are based on different moments of the SDF. In particular

$$\mathbb{E}_t\left[M_{t+1}^s\right]\begin{cases}\leq \mathbb{E}_t\left[R_{t+1}^{-\frac{s}{1-s}}\right]^{1-s}, & \text{if } s\in(0,1).\\ \geq \mathbb{E}_t\left[R_{t+1}^{-\frac{s}{1-s}}\right]^{1-s}, & \text{if } s\in(-\infty,0).\end{cases} \tag{E.2}$$

The proof, as in Liu (2020), follows from an application of the *reverse Hölder inequality.*[13] Equality occurs for the return which satisfies

$$\log M_{t+1} = -\frac{1}{1-s}\log R_{t+1} + \text{Constant}.$$

The quantile relation can only be used to obtain the upper bound part in (E.2), since the reverse Hölder inequality requires almost sure positivity of $\mathbb{1}(R_{t+1}\leq \widetilde{Q}_{t,\tau})$ to prove the lower bound. For $p\in(1,\infty)$, apply the reverse Hölder inequality to the relation (E.1) to obtain

$$\tau = R_{f,t+1}\mathbb{E}_t\left[M_{t+1}\mathbb{1}\left(R_{t+1}\leq \widetilde{Q}_\tau\right)\right] \geq R_{f,t+1}\mathbb{E}\left[M_{t+1}^{\frac{-1}{p-1}}\right]^{1-p}\mathbb{E}_t\left[\mathbb{1}\left(R_{t+1}\leq \widetilde{Q}_{t,\tau}\right)^{\frac{1}{p}}\right]^p$$

[13] The reverse Hölder inequality states that for any $p\in(1,\infty)$ and measure space (S,Σ,μ) that satisfies $\mu(S)>0$. Then for all measurable real- or complex-valued functions f and g on S such that $g(s)\neq 0$ for μ-almost all $s\in S$, $\|fg\|_1 \geq \|f\|_{\frac{1}{p}}\|g\|_{\frac{-1}{p-1}}$.

Rearranging and using $s \coloneqq -\frac{1}{p-1} \in (-\infty, 0)$ yields

$$\mathbb{E}_t\left[M_{t+1}^s\right] \geq \left(\frac{\tau}{R_{f,t+1}}\right)^s \mathbb{E}_t\left[\mathbb{1}\left(R_{t+1} \leq \widetilde{Q}_{t,\tau}\right)\right]^{1-s}.$$

F Von Mises approximation: evidence from Black-Scholes

This section illustrates the accuracy of the quantile approximation in (5.10) in a discretized version of the Black-Scholes model with changing parameters. Specifically, I assume the following model

$$\begin{aligned} R_{t+1} &= \exp\left((\mu_t - \frac{1}{2}\sigma_t^2)\lambda + \sigma_t \lambda Z_{t+1}\right) \\ Z_{t+1} &\sim N(0,1) \\ \sigma_t &\sim \textbf{UNIF}\left[0.05, 0.35\right] \\ \mu_t &\sim \textbf{UNIF}\left[-0.02, 0.2\right]. \end{aligned} \tag{F.1}$$

The returns under risk neutral dynamics are given by

$$\begin{aligned} R_{t+1} &= \exp\left((r_t - \frac{1}{2}\sigma_t^2)\lambda + \sigma_t \lambda Z_{t+1}\right) \\ r_t &\sim \textbf{UNIF}\left[0, 0.03\right]. \end{aligned}$$

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 $\lambda = 1/12$ 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, using the result of Breeden and Litzenberger (1978). I use the risk-neutral quantile function to calculate $LRB_t(\tau)$ at time t. Then, following the approximation in (5.10), the physical quantile function is estimated by

$$\widehat{Q}_{t,\tau} = \widetilde{Q}_{t,\tau} + \frac{LRB_t(\tau)}{\widetilde{f}_t(\widetilde{Q}_{t,\tau})}. \tag{F.2}$$

Table 8: Simulation results

	$\mathbb{E}\widehat{\beta}_0(\tau)$	$\mathbb{E}\widehat{\beta}_1(\tau)$	$Q > \widehat{Q}$	$\rho(Q, \widehat{Q})$	$H_0 : \widehat{\beta}_0(\tau) = 0$	$H_0 : \widehat{\beta}_1(\tau) = 1$	$H_0 : [\widehat{\beta}_0(\tau), \widehat{\beta}_1(\tau)] = [0, 1]$
$\tau = 0.01$	0.01	0.99	0.85	1	0.94	0.96	0.8
$\tau = 0.05$	-0.03	1.04	0.69	0.99	0.9	0.89	0.66
$\tau = 0.1$	-0.06	1.07	0.64	0.99	0.78	0.76	0.47

Note: $\mathbb{E}\widehat{\beta}_0(\tau)$ *denotes the average quantile regression estimate of* $\widehat{\beta}_0(\tau)$ *and likewise* $\mathbb{E}\widehat{\beta}_0(\tau)$ *shows it for* $\widehat{\beta}_1(\tau)$. $Q > \widehat{Q}$ *shows the fraction of times the true physical quantile is larger than our predicted quantile. Columns* $H_0 : \widehat{\beta}_0(\tau) = 0$ *and* $H_0 : \widehat{\beta}_1(\tau) = 1$ *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.*

We take 3,000 return observations, which are generated according to (F.1). This exercise is repeated 1,000 times. To assess the accuracy of the approximation in (F.2), I use several metrics. For every sample, I estimate a quantile regression of the form

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

where $\widehat{Q}_{t,\tau}$ comes from (F.2). The first two columns in Table 8 report the average values of the quantile regression 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 $Q_{t,\tau} > \widehat{Q}_{t,\tau}$, since $LRB_t(\tau) \leq \tau - \mathbb{P}_t(\widetilde{Q}_{t,\tau})$. The third column in Table 8 shows this happens for the majority of samples. The fourth column shows the correlation between $Q_{t,\tau}$ and $\widehat{Q}_{t,\tau}$, which is very close to one, which 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 $\tau = 0.1$ quantile. Overall, Table 8 suggests that (F.2) is a highly accurate predictor of the physical quantile function.

Example F.1. Let us illustrate the von Mises approximation (5.4) in the Black-Scholes model with fixed parameters: $\lambda = 1$ (one year), $\mu = 0.08, r = 0.02, \sigma = 0.2$.[14] We can explicitly calculate $F^{-1}, \widetilde{F}^{-1}$ and $\widetilde{f}$ owing to the lognormal assumption. Figure 7 shows the risk neutral quantile function (in green), von Mises approximation (5.4) (in blue) and the true physical quantile function (in red). Observe that the approximation (5.4) is very accurate in this case.

[14] For illustrative purposes, I use $\lambda = 1$, instead of $\lambda = 1/12$, otherwise the physical quantile function and von Mises approximation are indistinguishable.

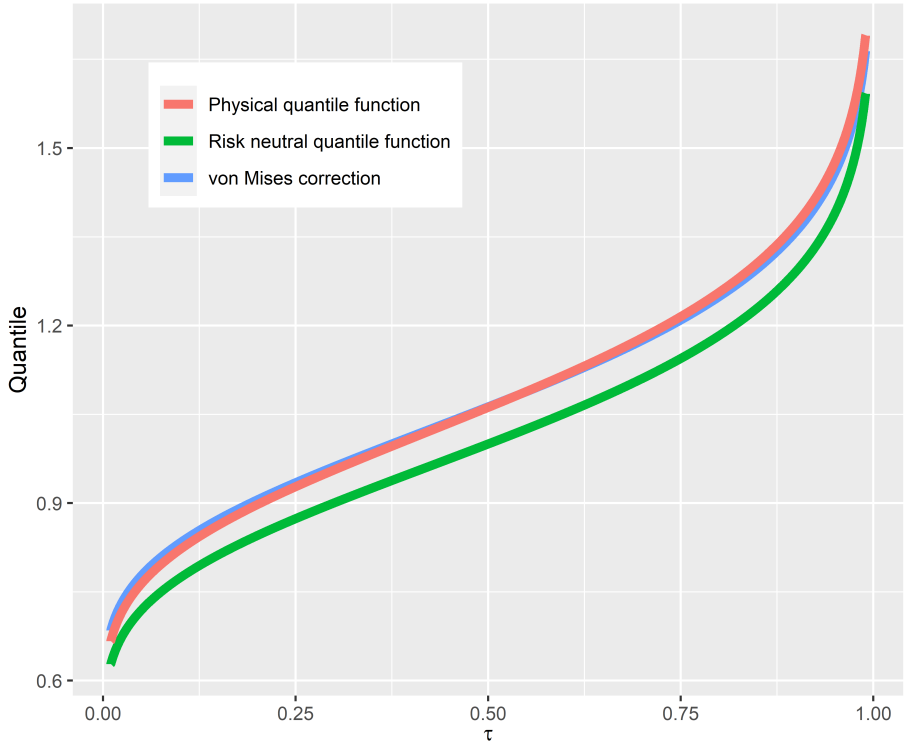

Figure 7: Illustration of quantile approximation (F.2) in the Black-Scholes model.

References

Aït-Sahalia, Y. and Lo, A. W. (1998). Nonparametric estimation of state-price densities implicit in financial asset prices. *The journal of finance*, 53(2):499–547.

Aït-Sahalia, Y. and Lo, A. W. (2000). Nonparametric risk management and implied risk aversion. *Journal of econometrics*, 94(1-2):9–51.

Almeida, C. and Garcia, R. (2012). Assessing misspecified asset pricing models with empirical likelihood estimators. *Journal of Econometrics*, 170(2):519–537.

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

Andrews, D. W. (1991). Heteroskedasticity and autocorrelation consistent co-

variance matrix estimation. *Econometrica: Journal of the Econometric Society*, pages 817–858.

Backus, D., Chernov, M., and Martin, I. (2011). Disasters implied by equity index options. *The journal of finance*, 66(6):1969–2012.

Backus, D., Chernov, M., and Zin, S. (2014). Sources of entropy in representative agent models. *The Journal of Finance*, 69(1):51–99.

Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica*, 71(1):135–171.

Bansal, R., Kiku, D., and Yaron, A. (2012). An empirical evaluation of the long-run risks model for asset prices. *Critical Finance Review*, (1):183–221.

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

Bansal, R. and Yaron, A. (2004). Risks for the long run: A potential resolution of asset pricing puzzles. *The journal of Finance*, 59(4):1481–1509.

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

Bates, D. S. (1991). The crash of 87: was it expected? The evidence from options markets. *The journal of finance*, 46(3):1009–1044.

Bates, D. S. (2000). Post-'87 crash fears in the s&p 500 futures option market. *Journal of econometrics*, 94(1-2):181–238.

Bates, D. S. (2008). The market for crash risk. *Journal of Economic Dynamics and Control*, 32(7):2291–2321.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654.

Bollerslev, T. and Todorov, V. (2011). Tails, fears, and risk premia. *The Journal of Finance*, 66(6):2165–2211.

Borovička, J., Hansen, L. P., and Scheinkman, J. A. (2016). Misspecified recovery. *The Journal of Finance*, 71(6):2493–2544.

Breeden, D. T. and Litzenberger, R. H. (1978). Prices of state-contingent claims implicit in option prices. *Journal of business*, pages 621–651.

Campbell, J. Y. and Thompson, S. B. (2008). Predicting excess stock returns out of sample: Can anything beat the historical average? *The Review of Financial Studies*, 21(4):1509–1531.

Chabi-Yo, F. and Loudis, J. (2020). The conditional expected market return. *Journal of Financial Economics*.

Cochrane, J. H. (2005). *Asset pricing: Revised edition*. Princeton university press.

Cochrane, J. H. and Saa-Requejo, J. (2000). Beyond arbitrage: Good-deal asset price bounds in incomplete markets. *Journal of political economy*, 108(1):79–119.

Coval, J. D. and Shumway, T. (2001). Expected option returns. *The journal of Finance*, 56(3):983–1009.

Danielsson, J. and De Vries, C. G. (2000). Value-at-risk and extreme returns. *Annales d'Economie et de Statistique*, pages 239–270.

De Haan, L. and Ferreira, A. (2007). *Extreme value theory: an introduction*. Springer Science & Business Media.

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

Figlewski, S. (2008). Estimating the implied risk neutral density.

Gabaix, X. (2012). Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance. *The Quarterly journal of economics*, 127(2):645–700.

Gil-Pelaez, J. (1951). Note on the inversion theorem. *Biometrika*, 38(3-4):481–482.

Gouriéroux, C. and Jasiak, J. (2008). Dynamic quantile models. *Journal of Econometrics*, 147(1):198–205.

Goyal, A. and Welch, I. (2008). A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies*, 21(4):1455–1508.

Gregory, K. B., Lahiri, S. N., and Nordman, D. J. (2018). A smooth block bootstrap for quantile regression with time series. *The Annals of Statistics*, 46(3):1138–1166.

Hansen, L. P. and Hodrick, R. J. (1980). Forward exchange rates as optimal predictors of future spot rates: An econometric analysis. *Journal of political economy*, 88(5):829–853.

Hansen, L. P. and Jagannathan, R. (1991). Implications of security market data for models of dynamic economies. *Journal of political economy*, 99(2):225–262.

Hodrick, R. J. (1992). Dividend yields and expected stock returns: Alternative procedures for inference and measurement. *The Review of Financial Studies*, 5(3):357–386.

Hsieh, F., Turnbull, B. W., et al. (1996). Nonparametric and semiparametric estimation of the receiver operating characteristic curve. *Annals of statistics*, 24(1):25–40.

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

Jackwerth, J. C. and Menner, M. (2020). Does the ross recovery theorem work empirically? *Journal of Financial Economics*, 137(3):723–739.

Koenker, R. and Bassett, G. (1978). Regression quantiles. *Econometrica: journal of the Econometric Society*, pages 33–50.

Koenker, R. and Machado, J. A. (1999). Goodness of fit and related inference processes for quantile regression. *Journal of the american statistical association*, 94(448):1296–1310.

Liu, Y. (2020). Index option returns and generalized entropy bounds. *Journal of Financial Economics*.

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

Martin, I. and Gao, C. (2021). Volatility, valuation ratios, and bubbles: An empirical measure of market sentiment. *Journal of Finance, forthcoming*.

Qin, L., Linetsky, V., and Nie, Y. (2018). Long forward probabilities, recovery, and the term structure of bond risk premiums. *The Review of Financial Studies*, 31(12):4863–4883.

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

Ross, S. A. (1976). Options and efficiency. *The Quarterly Journal of Economics*, 90(1):75–89.

Ross, S. A. (2005). *Neoclassical finance*. Princeton University Press.

Ross, S. A. (2015). The recovery theorem. *The Journal of Finance*, 70(2):615–648.

Serfling, R. J. (2009). *Approximation theorems of mathematical statistics*, volume 162. John Wiley & Sons.

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

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

Van der Vaart, A. W. (2000). *Asymptotic statistics*, volume 3. Cambridge university press.

von Mises, R. (1947). On the asymptotic distribution of differentiable statistical functions. *The annals of mathematical statistics*, 18(3):309–348.

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

Wahba, G. (1990). *Spline models for observational data*. SIAM.