

# A Quantile Approach to Asset Pricing Models

Tjeerd de Vries \*

May 20, 2021

## Abstract

This paper develops a new way of analyzing the performance of asset pricing models. I show that many classical asset pricing bounds, such as the [Hansen and Jagannathan \(1991\)](#) (HJ) bound, can be improved upon by looking at derivative contracts. The resulting bound is found to be much tighter than the HJ bound in empirical data. A direct implication is that the SDF process is more volatile than previously assumed and poses new challenges to consumption based asset pricing models. A central ingredient of this new bound is the risk-neutral quantile function. Two additional applications consider the use of this function: (i) As a predictor of Value-at-Risk (ii) As a forward looking measure of crash risk. Both applications underscore the importance of analyzing quantiles of the data, instead of the more prevalent variance and equity premia.

**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](mailto:tjdevrie@ucsd.edu). I would like to thank Alexis Toda, Allan Timmermann, James Hamilton, Xinwei Ma, Yixiao Sun and Rossen Valkanov for useful feedback. All errors are my own.

# 1 Introduction

Nonparametric methods are useful to employ in the macro-finance literature as they can be used to analyze misspecification of asset pricing models (e.g. 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 make inference about 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. Based on this lower bound, they conclude that a basic SDF of the form  $\beta g_c^{-\gamma}$  leads to exorbitant levels of risk-aversion.<sup>1</sup> This result is known as the equity-premium puzzle and the quest to resolve it became one the central interests of financial economists (e.g. Campbell and Cochrane (1999), Bansal and Yaron (2004) and Barro (2006)).

The idea propagated in this paper is that one can often get stronger conclusions about these moment bounds by studying derivatives written on the underlying stock. The intuition is that, by varying the strike value of an underlying derivative (e.g. digital put option), you can obtain additional information on the behavior of the SDF not captured by the individual asset itself. This idea is consistent with the work of Ross (1976); Breeden and Litzenberger (1978), who show that options help complete the market and can be used to identify the SDF.

The focus of this paper is centered around bounds coming from digital put options, as they naturally induce a discrepancy measure between the physical and risk-neutral distribution (see Section 2.2). The benefits of this insight are the following. (i) The bound is valid and performs well even if returns are heavy tailed, this in contrast to the Hansen and Jagannathan (1991) bound (HJ bound hereafter), a point I illustrate in Section 2.3. (ii) The bound can be optimized through varying the strike of the underlying derivative, which leads to sharper conclusions about the SDF volatility. Section 4 details that the SDF is much more volatile than what could be concluded based on the HJ bound. (iii) Many asset pricing models imply that the HJ bound is stronger than the quantile bound (Section 2.4). Hence, the empirical observation that the quantile bound is stronger than the HJ bound provides model-free evidence against such asset pricing models.

A key ingredient in the generalization of these asset pricing bounds is the risk-neutral quantile function. Besides its use to bound moments of the SDF, this function carries important conditional market information. A second contribution of this paper is to show how quantiles from the risk-neutral distribution can be used to predict the latent physical quantile function. I leverage on recent insights from Martin (2017), Martin and Wagner (2019) and Chabi-Yo and Loudis (2020) to derive a market observable proxy that governs the discrepancy between the physical and risk-neutral quantile function. In combination with quantile regression, this leads to promising Value-at-Risk (VaR) forecasts over longer time horizons. Section 6 documents that VaR forecasts using option implied information improve upon the adaptive CAViaR model of Engle and

---

<sup>1</sup>Here,  $\beta$  is a time discount factor,  $g_c$  is consumption growth and  $\gamma$  is the degree of risk-aversion of a representative agent with CRRA utility.

[Manganelli \(2004\)](#). Additionally, I prove that the risk-neutral quantile function is the optimal predictor of the physical quantile function when returns are conditionally lognormal.

A final contribution is the proposal of a new forward looking proxy of the premium on crash risk. This proxy is a natural outflow of the quantile theory developed in this paper. I document that the premium fluctuates significantly over time and spikes during financial crises. Standard consumption based asset pricing models have difficulty to reconcile this phenomenon. Moreover, the obtained premium is completely forward looking and thus avoids the historical sample bias critique of [Goyal and Welch \(2008\)](#).

The rest of this paper is organized as follows. Section 2 shows how quantiles from the risk-neutral distribution can be used to bound the volatility of the SDF, as an alternative to the HJ bound. Section 3 outlines how to estimate the new bound with actual data and details the construction of confidence intervals. Thereafter, Section 4 documents that the new bound is significantly stronger than the HJ bound using a combination of historical return and option data. This provides new evidence that the volatility of the SDF is much higher than previously documented. I also show how this can be interpreted as model free evidence against popular asset pricing models, such as CAPM. Section 5 derives a lower bound on the physical and risk-neutral quantile function, which can be estimated with option data. The results are used in Section 6 to propose long term VaR estimates. I also prove that my proposed estimator is optimal if return data are lognormal. Section 7 uses the result from Section 5 to analyze the premium associated with crash-risk. Finally, Section 8 concludes.

## 2 A new bound on the SDF volatility

### 2.1 Notation and assumptions

Let  $S_t$  be the price of a generic stock at time  $t$  and the associated (gross) return is denoted by  $R_{t+1} := S_{t+1}/S_t$ . The notation  $R_{m,t+1}$  is reserved for the market return. The (gross) risk-free rate, when it exists, is denoted by  $R_{f,t+1}$  and known at the start of period  $t$ . Throughout most of the paper, I assume that the market is *arbitrage-free*. No arbitrage guarantees the existence of a stochastic discount factor (SDF)  $M_{t+1} > 0$  ([Harrison and Kreps, 1979](#)), which by definition satisfies

$$\mathbb{E}_t [M_{t+1} R_{t+1}] = 1. \quad (2.1)$$

The  $t$ -subscript is used to signify the expectation conditional on time  $t$ . If the market is incomplete, then the SDF is not unique, i.e. there are multiple stochastic processes  $M_{t+1}$  that satisfy (2.1). The SDF process serves as the Radon-Nikodym derivative when passing from physical to risk-neutral measure, a fact I often exploit in this paper.<sup>2</sup> Quantities calculated under risk-neutral measure are denoted by a tilde, e.g. the  $\tau$ -quantile of the random variable  $R_{t+1}$ , and denoted by  $\tilde{Q}_\tau(R_{t+1})$ , satisfies

$$\tilde{\mathbb{P}}_t [R_{t+1} \leq \tilde{Q}_\tau(R_{t+1})] = \tau.$$

---

<sup>2</sup>Market incompleteness implies that there are multiple risk-neutral measures.

The dependence of the quantile on time  $t$  is omitted as it should be clear from the context. Sometimes, I suppress the dependence of the quantile on the random variable and write  $Q_\tau$  or  $Q(\tau)$  instead of  $Q_\tau(R_{t+1})$  for the same reason. 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

$$M_{t+1} = \mathbb{E}_t [\mathfrak{M}_{t+1} | R_{t+1}].$$

Here,  $\mathfrak{M}_{t+1}$  is the SDF that depends on all the state variables. The projected SDF has the same pricing implications for contingent claims written on asset  $S_t$  for which we have data (Cochrane, 2005). Formally,  $M_{t+1}$  is a measurable function of  $R_{t+1}$ , but I avoid denoting this dependence explicitly to avoid notational clutter.

## 2.2 A quantile version of the HJ bound

In this Section I make the additional assumption that the market is incomplete. Multiple SDF processes exist in incomplete markets and the HJ bound offers a simple nonparametric lower bound on the volatility of any SDF process. In particular, if the risk-free rate exists, the conditional HJ bound states that

$$\frac{|\mathbb{E}_t [R_{t+1} - R_{f,t+1}]|}{\sigma_t(R_{t+1})R_{f,t+1}} \leq \sigma_t(M_{t+1}).$$

Here  $\sigma_t(\cdot)$  denotes conditional volatility. The first result of this paper shows how quantiles of the risk-neutral distribution can be used to give an alternative bound on the SDF volatility. For this reason, the new bound is referred to as the quantile bound throughout the paper. The quantile bound is naturally understood as the Sharpe ratio on a digital put option, that is, a derivative contract that pays out 1\$ whenever the stock  $S_{t+1}$  is less than some strike price  $K$ .

**Theorem 2.1** (Quantile bound). *Suppose there is a risk-free asset in the market, then*

$$\frac{\left| \tau - \mathbb{P}_t(R_{t+1} \leq \tilde{Q}_\tau(R_{t+1})) \right|}{\sigma_t \left( \mathbb{1} \left( R_{t+1} \leq \tilde{Q}_\tau(R_{t+1}) \right) \right) R_{f,t+1}} \leq \sigma_t(M_{t+1}). \quad (2.2)$$

*The ratio on the LHS is the Sharpe ratio of the investment in a digital put option, that is, a derivative contract with payoff*

$$\begin{cases} 1 & \text{if } S_{t+1} \leq K \\ 0 & \text{otherwise.} \end{cases}$$

*Proof.* See Appendix B.1. ■

**Corollary 2.2.** *More generally, if there is no risk-free rate in the market*

$$\frac{\sigma_t(M_{t+1})}{\mathbb{E}_t [M_{t+1}]} \geq \frac{\left| \tau - \mathbb{P}_t(R_{t+1} \leq \tilde{Q}_\tau(R_{t+1})) \right|}{\sigma_t \left( \mathbb{1} \left( R_{t+1} \leq \tilde{Q}_\tau(R_{t+1}) \right) \right)}.$$

In this case  $\tilde{Q}_\tau$  is the  $\tau$ -quantile of the distribution induced by the change of measure

$$\tilde{\mathbb{P}}_t(A) := \mathbb{E}_t \left[ \frac{M_{t+1}}{\mathbb{E}_t[M_{t+1}]} \mathbb{1}(A) \right].$$

*Proof.* The proof is analogous to Theorem 2.1, replacing all instances of  $R_{f,t+1}$  by  $\frac{1}{\mathbb{E}_t[M_{t+1}]}$ .  $\blacksquare$

Theorem 2.1 puts a new bound on the volatility of the stochastic discount factor. Notice that  $\tilde{Q}_\tau$  can be obtained at time  $t$  from the [Breedon and Litzenberger \(1978\)](#) formula. To see the link with the Sharpe ratio, observe that price of a digital put option with strike  $K$  equals

$$P_{t+1}(K) = \frac{1}{R_{f,t+1}} \tilde{\mathbb{E}}_t[\mathbb{1}(S_{t+1} \leq K)].$$

Since

$$\tau = \tilde{\mathbb{E}}_t[\mathbb{1}(R_{t+1} \leq \tilde{Q}_\tau)] = \tilde{\mathbb{E}}_t[\mathbb{1}(S_{t+1} \leq \tilde{Q}_\tau S_t)]. \quad (2.3)$$

Hence,

$$\frac{\tau}{R_{f,t+1}} = \frac{1}{R_{f,t+1}} \tilde{\mathbb{E}}_t[\mathbb{1}(S_{t+1} \leq \tilde{Q}_\tau S_t)] = P_{t+1} \left( \underbrace{\tilde{Q}_\tau S_t}_{=: K} \right).$$

Let  $K = \tilde{Q}_\tau S_t$ , then the volatility bound on the LHS of Equation 2.2 can be rewritten to

$$\frac{\left| \tilde{\mathbb{E}}_t[\mathbb{1}(S_{t+1} \leq K)] - \mathbb{E}_t[\mathbb{1}(S_{t+1} \leq K)] \right|}{\sigma_t(\mathbb{1}(S_{t+1} \leq K)) R_{f,t+1}}.$$

This is the Sharpe ratio for a digital put option with strike  $\tilde{Q}_\tau S_t$ . Because the bound on the LHS of Equation (2.2) is valid for any  $\tau \in [a, b] \subseteq [0, 1]$ , I can optimize over  $\tau$  to obtain the sharpest possible bound. I refer to the sharpest possible bound as the *Sup quantile bound*. The quantile bound is also independent of units, meaning that the bound is invariant under monotonic transformations of the returns. This in contrast to the HJ bound. For example, it can happen that a model is able to generate high Sharpe ratios for log returns, but not for gross returns.

*Remark.* Theorem 2.1 can be repeated step by step using unconditional expressions, which gives

$$\frac{\left| \tau - \mathbb{P}(R_{t+1} \leq \tilde{Q}_\tau(R_{t+1})) \right|}{\sigma \left( \mathbb{1}(R_{t+1} \leq \tilde{Q}_\tau(R_{t+1})) \right) R_{f,t+1}} \leq \sigma(M_{t+1}).$$

This will be used in Section 4 to contrast the quantile bound to the unconditional HJ bound.

*Remark.* The quantity  $\phi(\tau) := \mathbb{P}_t[R_{t+1} \leq \tilde{Q}_\tau]$  can be interpreted as the ordinal dominance curve of the measures  $\mathbb{P}_t$  and  $\tilde{\mathbb{P}}_t$  ([Hsieh et al., 1996](#)). If  $\mathbb{P}_t = \tilde{\mathbb{P}}_t$ , agents are risk-neutral and the dominance curve evaluates to  $\phi(\tau) = \tau$ . In that case the quantile bound vanishes. The ordinal dominance curve has been used before in the finance literature, e.g. [Beare and Schmidt \(2016\)](#) to test for pricing kernel monotonicity. The quantile bound in (2.2) is essentially a studentized version of the ordinal dominance curve.

### 2.3 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. Let  $U \sim \mathbf{UNIF}[0, 1]$  (Uniform distribution on  $[0, 1]$ ) and consider the following specification

$$M_{t+1} = AU^\alpha, \quad R_{t+1} = BU^\beta \quad \text{with } \alpha > 0 > \beta \text{ and } A, B > 0.$$

Assume that  $[M_{t+1}, R_{t+1}]^\top$  is i.i.d. 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

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

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

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

Routine calculations show that the mean and variance of  $R_{t+1}$  are given by (provided  $\beta > -1/2$ )

$$\mathbb{E}[R_{t+1}] = \frac{B}{1+\beta} \quad \sigma^2(R_{t+1}) = \frac{B^2}{1+2\beta} - \left(\frac{B}{1+\beta}\right)^2.$$

In this case the Sharpe ratio is given by

$$\frac{\mathbb{E}[R_{t+1}] - R_{f,t+1}}{\sigma(R_{t+1})} = \frac{\frac{B}{1+\beta} - \frac{\alpha+1}{A}}{\sqrt{\frac{B^2}{1+2\beta} - \left(\frac{B}{1+\beta}\right)^2}}. \quad (2.4)$$

Likewise, the distribution of the SDF  $M_{t+1}$  follows from

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

In this case,  $M_{t+1}$  is said to have a Pareto lower tail. The expectation and variance are respectively given by

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

The constraint  $\mathbb{E}[M_{t+1}R_{t+1}] = 1$  forces

$$\frac{AB}{\alpha+\beta+1} = 1. \quad (2.5)$$

In addition from  $\mathbb{E}[M_{t+1}] = \frac{1}{R_{f,t+1}}$  it follows

$$\frac{A}{\alpha+1} = \frac{1}{R_{f,t+1}}. \quad (2.6)$$

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

- (i) Under  $\tilde{\mathbb{P}}$ ,  $R_{t+1} \sim \mathbf{PAR}\left(B, -\frac{\alpha+1}{\beta}\right)$ .
- (ii) The quantile bound depends only on the (left) tail index  $\alpha$  of  $M_{t+1}$ . In particular

$$\frac{\left|\tau - \mathbb{P}(R_{t+1} \leq \tilde{Q}_\tau)\right|}{R_{f,t+1}\sigma\left(\mathbb{1}\left(R_{t+1} \leq \tilde{Q}_\tau\right)\right)} = \frac{A}{1+\alpha} \frac{\left|\tau - 1 + (1-\tau)^{\frac{1}{\alpha+1}}\right|}{\sqrt{(1 - (1-\tau)^{\frac{1}{\alpha+1}})(1-\tau)^{\frac{1}{\alpha+1}}}.$$

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

*Proof.* (i) Since  $R_{f,t+1}M_{t+1}$  is the Radon-Nikodym derivative that induces a change of measure from  $\mathbb{P}$  to  $\tilde{\mathbb{P}}$ , it follows that

$$\begin{aligned} \tilde{\mathbb{P}}(R_{t+1} \leq x) &= R_{f,t+1}\mathbb{E}[M_{t+1}\mathbb{1}(R_{t+1} \leq x)] = R_{f,t+1} \int_0^1 AU^\alpha \mathbb{1}(Bu^\beta \leq x) du \\ &= R_{f,t+1}A \int_0^1 u^\alpha \mathbb{1}\left(u \geq \left(\frac{x}{B}\right)^{\frac{1}{\beta}}\right) du \quad (\text{since } \beta < 0) \\ &= \frac{R_{f,t+1}A}{\alpha+1} \left(1 - \left(\frac{x}{B}\right)^{\frac{\alpha+1}{\beta}}\right) \\ &= 1 - \left(\frac{x}{B}\right)^{\frac{\alpha+1}{\beta}}. \end{aligned}$$

The last line follows from (2.6).

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

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

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

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

As a result

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

Hence, the quantile bound evaluates to

$$\frac{\left|\tau - \mathbb{P}(R_{t+1} \leq \tilde{Q}_\tau)\right|}{R_{f,t+1}\sigma\left(\mathbb{1}\left(R_{t+1} \leq \tilde{Q}_\tau\right)\right)} = \frac{A}{1+\alpha} \frac{\left|\tau - 1 + (1-\tau)^{\frac{1}{\alpha+1}}\right|}{\sqrt{(1 - (1-\tau)^{\frac{1}{\alpha+1}})(1-\tau)^{\frac{1}{\alpha+1}}}. \quad (2.7)$$

- (iii). The HJ bound, as given by the Sharpe ratio in (2.4), goes to 0 as  $\beta \downarrow -1/2$  since  $\sigma(R_{t+1}) \uparrow \infty$ . ■

Properties (ii) and (iii) provide some intuition when the quantile bound is stronger than the HJ bound. Namely, heavier tails of the distribution of  $R_{t+1}$  (as measured by  $\beta$ ) lead to a lower Sharpe ratio. However, the quantile bound is unaffected by  $\beta$  since it only depends on the tail index  $\alpha$ . Therefore, when  $\beta$  gets close to  $-1/2$ , the HJ bound is rather uninformative, whereas the quantile bound may still render a good bound. Moreover, I do not need to impose any restrictions on the parameter space to calculate the quantile bound, whereas the HJ bound requires  $\beta > -1/2$ . The latter restriction is not unreasonable for asset returns though, since typical tail index estimates suggest  $\beta \in \{-1/3, -1/4\}$  (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. Typically, the Pareto exponent for the market return is between 3–4 (see e.g. Danielsson and De Vries (2000)), hence I set  $\beta = -1/3$ . For simplicity, assume  $R_f = 1$  and to match the observed equity premium of 8% (Cochrane, 2005) 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. The parameters  $A, \alpha$  from the SDF distribution are solved for using Equations (2.5) and (2.6). 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, the parameters  $A, \alpha$  are solved using (2.5) and (2.6). Table 1 summarizes the resulting parameter values for reference, together with the corresponding Sharpe ratio and SDF volatility.

Table 1: Model calibration

	$A$	$\alpha$	$B$	$\beta$	$\sigma(R_{t+1})$	Sharpe ratio	$\sigma(M_{t+1})$
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  $R_{f,t+1} = 1$ .  $\sigma(R_{t+1})$  denotes the return volatility and  $\sigma(M_{t+1})$  the 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 bound is better whenever the distribution of asset returns is more heavy tailed. The green line is our proposed estimator of the quantile bound (using 5,000 returns) and the shaded area depict the uniform confidence intervals for this estimated function.<sup>3</sup> In the left display of Figure 1, the sup quantile bound is slightly stronger than the HJ bound, but only marginally so. In particular, the confidence bounds for the quantile bound estimator contain the HJ bound at the supremum, meaning that a researcher cannot reject the null hypothesis that the sup quantile bound is stronger than the HJ bound. In contrast, the display on the right of Figure 1 highlights that the sup quantile bound is much stronger than the HJ bound, owing to the fatter tails of the return distribution.

<sup>3</sup>The estimator and construction of the confidence intervals are further discussed in Section 3.



Moreover, a researcher would be able to reject the null hypothesis that the quantile bound is equal to the HJ bound, as there is a range of values where the confidence intervals do not contain the HJ bound. Figure 2 compares the quantile bound to the HJ bound when  $\beta$  varies. The parameters  $A, \alpha$  are solved for implicitly, as in the calibration of Table 1. The graph shows that the quantile bound is better when  $\beta \notin [-0.28, -0.19]$ . 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 deliver descent approximations if returns are very heavy tailed ( $\beta < -0.28$ ) or more light tailed ( $\beta > -0.19$ ).

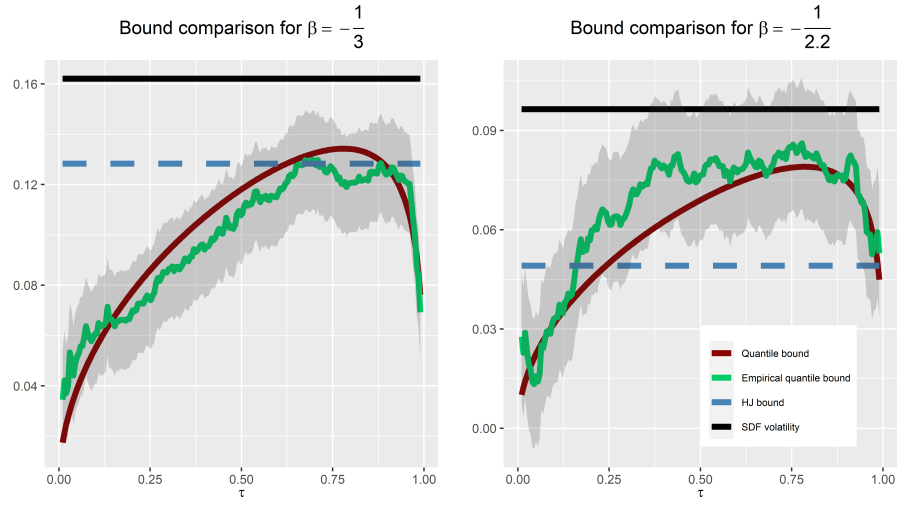


Figure 1: Plots of the quantile bound (in red), HJ bound (blue), true SDF volatility (black) and estimated quantile bound (green) for different values of  $\beta$ . The shaded area denote the uniform confidence intervals.

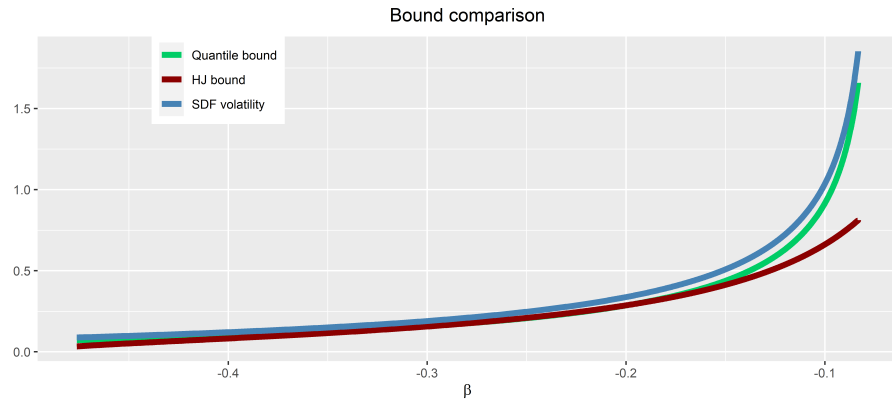


Figure 2: Plot of the Quantile bound (green) and HJ bound (red) when  $\beta$  varies. The true SDF volatility is denoted by the blue line.

## 2.4 Comparison of Quantile bound 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 F contains similar results in this direction using other well known asset pricing bounds.

**Example 2.1 (CAPM).** The Capital Asset Pricing Model (CAPM) posits the following SDF specification

$$M_{t+1} = \alpha - \beta R_{m,t+1}.$$

Here,  $R_{m,t+1}$  denotes the return on the market portfolio. Since the HJ bound is derived by applying the Cauchy-Schwarz inequality to  $\text{COV}_t(R_{m,t+1}, M_{t+1})$ , the inequality binds if  $M_{t+1}$  is a linear combination of  $R_{m,t+1}$ . Hence, under CAPM, the HJ bound is always stronger than the quantile bound regardless of the distribution of  $R_{m,t+1}$ .

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

**Lemma 2.4 (Stein's Lemma).** *If  $f, R$  are bivariate normal,  $g(f)$  is differentiable and  $\mathbb{E}|g'(f)| < \infty$ , then*

$$\text{COV}(g(f), R) = \mathbb{E}[g'(f)] \text{COV}(f, R).$$

**Example 2.2 (Joint normality).** Suppose that  $M_{t+1}$  and  $R_{t+1}$  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

$$\frac{\left| \tau - \mathbb{P}(R_{t+1} \leq \tilde{Q}_\tau) \right|}{R_{f,t+1}} = \left| \text{COV} \left( \mathbb{1} \left( R_{t+1} \leq \tilde{Q}_\tau \right), M_{t+1} \right) \right|$$

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

$$\left| \text{COV} \left( \mathbb{1} \left( R_{t+1} \leq \tilde{Q}_\tau \right), M_{t+1} \right) \right| = f(\tilde{Q}_\tau) |\text{COV}(R_{t+1}, M_{t+1})|. \quad (2.8)$$

Here,  $f(\cdot)$  is the marginal density of  $R_{t+1}$ . Standard SDF properties also yield the well known identity

$$\frac{|\mathbb{E}(R_{t+1}) - R_{f,t+1}|}{R_{f,t+1}} = |\text{COV}(R_{t+1}, M_{t+1})|.$$

To get a feeling which bound is stronger, consider the relative efficiency

$$\frac{\text{HJ bound}}{\text{Quantile bound}} = \frac{\frac{|\mathbb{E}[R_{t+1}] - R_{f,t+1}|}{\sigma(R_{t+1})R_{f,t+1}}}{\frac{|\tau - \mathbb{P}(R_{t+1} \leq \tilde{Q}_\tau)|}{\sqrt{\mathbb{P}(R_{t+1} \leq \tilde{Q}_\tau)(1 - \mathbb{P}(R_{t+1} \leq \tilde{Q}_\tau))}R_{f,t+1}}} \stackrel{(2.8)}{=} \frac{\sqrt{\mathbb{P}(R_{t+1} \leq \tilde{Q}_\tau)(1 - \mathbb{P}(R_{t+1} \leq \tilde{Q}_\tau))}}{\sigma(R_{t+1})f(\tilde{Q}_\tau)}. \quad (2.9)$$

Minimizing the RHS of (2.9) leads to a first order condition which implies that the relative efficiency is minimized by choosing  $\tilde{Q}_\tau = \mu_R$ . For this choice,  $\mathbb{P}(R_{t+1} \leq \tilde{Q}_\tau) = 1/2$  and  $f(\tilde{Q}_\tau) = 1/\sqrt{2\pi\sigma(R_{t+1})^2}$ . Therefore,

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

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

$$R_{t+1} = e^{(\mu_R - \frac{\sigma_R^2}{2})\lambda + \sigma\sqrt{\lambda}Z_R}$$

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

Both  $Z_R$  and  $Z_M$  are standard normal random variables with correlation  $\rho$ . Moreover,  $\lambda$  governs the time scale. To satisfy  $\mathbb{E}[R_{t+1}M_{t+1}] = 1$ , it follows  $\mu_R - r_f = -\rho\sigma_R\sigma_M$ . It is hard to find an analytical solution for the relative efficiency between the HJ and quantile bound in this case, but a linearization approach leads to a closed form expression which is quite accurate in simulations. The details are described in Appendix B.2, where I prove that

$$\frac{\text{HJ bound}}{\text{Quantile bound}} = \frac{1}{2} \sqrt{\frac{2\pi\sigma_R^2\lambda}{\exp(\sigma_R^2\lambda) - 1}}.$$

This expression is independent of  $\mu_R$ . An application of l'Hôpital's rule reveals that the relative efficiency converges to  $\sqrt{2\pi}/2$  if  $\lambda \rightarrow 0^+$ . This is the same relative efficiency in Example 2.2, which is unsurprising as the linearization becomes exact in the limit as  $\lambda \rightarrow 0^+$ . 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** (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 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

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

Here,  $\Delta c_{t+1}$  and  $\Delta d_{t+1}$  denote log consumption and dividend growth, while  $\sigma_t$  depicts conditional volatility. The parameter  $\rho$  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. \(2012\)](#). This renders an annual SDF volatility of 0.72, HJ bound of 0.53 and quantile bound of 0.33. The results do not meaningfully change on a monthly basis. Section 4 of this paper shows that the quantile bound applied to monthly return data is significantly stronger than the HJ bound. The LRR models fails to reconcile this feature of the data.

**Example 2.5** (Disaster risk). The disaster risk model of [Barro \(2006\)](#) posits that risk-premia are driven by extreme events that affect consumption growth. Assuming that the representative agent has power utility, the log pricing kernel is given by

$$\log M_{t+1} = \log(\beta) - \gamma \Delta c_{t+1}.$$

Innovations in consumption growth are driven by two independent shocks

$$\Delta c_{t+1} = \varepsilon_{t+1} + \eta_{t+1}. \quad (2.10)$$

Here,  $\varepsilon_{t+1} \sim N(\mu, \sigma^2)$  and

$$\eta_{t+1} | (J = j) \sim N(j\theta, j\nu^2), \quad J \sim \mathbf{Poisson}(\kappa).$$

The interpretation of  $\eta$  is that of a jump component (disaster) which induces negative shocks to consumption growth.  $\kappa$  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  $C_t^\lambda$ .

This implies an SDF volatility of 0.82, the HJ bound is about 0.39, whereas the quantile bound peaks around 0.6. Hence, the disaster risk model is able to match the empirical fact that the quantile bound is sharper than the HJ bound. The left panel of Figure 3 compares the HJ and quantile bound as a function of  $\tau$ . The quantile bound is stronger for very small values of  $\tau$ , but is generally lower for  $\tau > 0.065$ . The reason for this result can be understood from the right panel of Figure 3, which shows the physical and risk-neutral

distribution of return on equity.<sup>4</sup> The risk-neutral distribution displays a heavy left tail, owing to the implied disaster risk embedded in the SDF. As a result, it is extremely profitable to sell digital put options which pay out in case of a disaster. These put options must have high Sharpe ratios as their prices are high (insurance against disaster risk), but the actual probability of disaster is so low that the risk associated to selling such insurance is limited. It is hard to verify whether this is in line with the data that we observe, since the put option market is highly illiquid for extreme events. However, the quantile bound I obtain in Section 4 shows a steady increase when  $\tau \in [0.15, 0.75]$  and is above the HJ bound for all  $\tau \in [0.12, 0.88]$ . This in contrast to the quantile bound in the left panel of Figure 3, which shows a monotonic decline for all  $\tau$  greater than 0.065. In conclusion, adding disaster risk to a vanilla representative agent model increases the risk premium, but as a side effect it also creates a fat left tail in the risk neutral distribution. This discrepancy leads to a monotonically declining quantile bound for  $\tau \geq 0.065$ , which is not in line with actual return data.

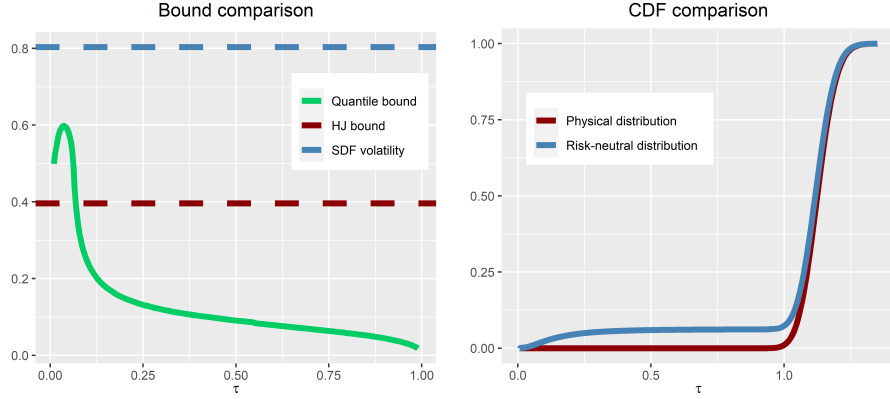


Figure 3: Left panel compares the HJ bound and quantile bound in disaster risk model. The right panel denotes the physical and risk-neutral distribution.

### 3 Quantile bound estimator and sampling properties

In this section I discuss estimators for the quantile bound (2.2). Additionally, to get an econometrically disciplined way of assessing the sampling error of the quantile estimator, I derive uniform confidence intervals in case the risk-neutral quantile function is known. A bootstrap procedure is used in case the risk-neutral quantile function is unknown. This is discussed in further detail in Section 4. Throughout this Section I assume that  $\tau \in [\varepsilon, 1 - \varepsilon] \subseteq [0, 1]$  for some small  $\varepsilon > 0$ .

<sup>4</sup>Appendix E.1 shows how to calculate the physical and risk-neutral CDF using cumulant generating functions.

### 3.1 Observed risk-neutral quantiles

Here, I make the simplifying assumption that the unconditional risk-neutral quantile curve  $\tilde{Q}(\tau)$  is known for all  $\tau \in [\varepsilon, 1 - \varepsilon]$ . This serves as a benchmark case if we believe the world is *i.i.d.* and digital put prices are directly observed. Moreover, the results derived under this assumption are useful when considering the additional estimation uncertainty that results if the risk-neutral quantile function is estimated, as discussed in Section 3.2.

The convergence results all take place in  $D[\varepsilon, 1 - \varepsilon]$ ; the space of càdlàg functions equipped with the Skorokhod topology and the Borel  $\sigma$ -field  $\mathcal{D}[\varepsilon, 1 - \varepsilon]$ . For further discussion about the properties of this space and basic empirical process terminology I refer to Billingsley (2013) and Van Der Vaart and Wellner (1996). The standard empirical process is denoted by

$$\mathbb{P}_T = \frac{1}{T} \sum_{t=1}^T \delta_{R_t},$$

where  $\delta_{R_t}$  are the Dirac measures at the observations.<sup>5</sup> If  $\tilde{Q}(\tau)$  is known, a natural estimator for the quantile bound (2.2) is

$$\hat{\theta} := \sup_{\varepsilon \leq \tau \leq 1 - \varepsilon} \frac{\tau - \mathbb{P}_T(\tilde{Q}(\tau))}{\sqrt{\mathbb{P}_T(\tilde{Q}(\tau))(1 - \mathbb{P}_T(\tilde{Q}(\tau)))}}. \quad (3.1)$$

Before presenting the main result I make the following assumptions.

**Assumption 3.1.** *The return process  $\{R_t\}_{t=1}^T$  is i.i.d.*

**Assumption 3.2.** *The unconditional risk-neutral quantile curve  $\tilde{Q}(\tau)$  is known and continuous on  $\tau \in [\varepsilon, 1 - \varepsilon]$ .*

*Discussion of Assumptions:* Assumption 3.1 allows me to use many classical results in empirical process theory to prove Theorem 3.3. Assumption 3.2 is satisfied in many popular option pricing models, e.g. Black-Scholes or the Heston (1993) model. The next theorem establishes a weak convergence result that permits the construction of uniform confidence intervals for (3.1).<sup>6</sup>

**Theorem 3.3.**

$$\begin{aligned} \sup_{\varepsilon \leq \tau \leq 1 - \varepsilon} \sqrt{T} \left| \frac{\tau - \mathbb{P}_T(\tilde{Q}(\tau))}{\sqrt{\mathbb{P}_T(\tilde{Q}(\tau))(1 - \mathbb{P}_T(\tilde{Q}(\tau)))}} - \frac{\tau - \mathbb{P}(\tilde{Q}(\tau))}{\sqrt{\mathbb{P}(\tilde{Q}(\tau))(1 - \mathbb{P}(\tilde{Q}(\tau)))}} \right| \\ \rightsquigarrow \sup_{\left\{ \frac{\mathbb{P}(\tilde{Q}(\varepsilon))}{1 - \mathbb{P}(\tilde{Q}(\varepsilon))} \leq t \leq \frac{\mathbb{P}(\tilde{Q}(1 - \varepsilon))}{1 - \mathbb{P}(\tilde{Q}(1 - \varepsilon))} \right\}} \frac{1}{\sqrt{t}} |\mathbb{W}(t)| =: \mathcal{K}. \end{aligned}$$

In this expression,  $\mathbb{W}(t)$  is a standard Brownian motion. The process  $|\mathbb{W}(t)| / \sqrt{t}$  is a rescaled version of a Bessel process of index 1 (Protter, 2005). From here one can construct uniform confidence intervals of the form

$$\left[ \frac{\tau - \mathbb{P}_T(\tilde{Q}(\tau))}{\sqrt{\mathbb{P}_T(\tilde{Q}(\tau))(1 - \mathbb{P}_T(\tilde{Q}(\tau)))}} \pm \frac{k(\alpha)}{\sqrt{T}} \right].$$

<sup>5</sup>This should not be confused with our notation for conditional probability  $\mathbb{P}_t$ .

<sup>6</sup>Following Van Der Vaart and Wellner (1996), weak convergence is denoted by  $\rightsquigarrow$ .

Here  $k(\alpha)$  is the  $\alpha$ -quantile of  $\mathcal{K}$ .

*Proof.* See Appendix B.3. ■

### 3.2 Unobserved quantiles

The result in the previous Section can only be applied if the unconditional risk-neutral quantile curve is directly observed. In reality, one can only estimate the conditional risk-neutral curve, using the result of Breeden and Litzenberger (1978). This introduces measurement error, since the formula of Breeden and Litzenberger (1978) requires differentiation of the put option price curve, and we only observe prices at a discrete range of strike values. Moreover, to obtain the unconditional curve requires additional averaging. In view of the empirical application in Section 4, it is desirable to get a hand on the sampling distribution of the quantile bound when  $\tilde{Q}(\tau)$  is estimated in this way. This Section presents some results in that direction. First, I describe the steps to estimate the quantile bound in this more general setup.

- (i) At the beginning of each time period  $t$ , I calculate the implied risk-neutral CDF (denoted by  $\tilde{\mathbb{P}}_t(\cdot)$ ) using the formula of Breeden and Litzenberger (1978). In particular, I find the closest two maturities in the market to the maturity of interest. I use smooth cubic splines (Wahba, 1990) to get the CDF for all  $x \in [a, b]$ . Thereafter, the two CDF curves are linearly interpolated to get the curve for the maturity of interest (e.g. 30 days).
- (ii) Calculate the unconditional risk-neutral CDF from

$$\tilde{\mathbb{P}}_T := \frac{1}{T} \sum_{t=1}^T \tilde{\mathbb{P}}_t.$$

Get the risk neutral quantile curve via

$$\tilde{Q}_T(\tau) := \inf \left\{ x \in \mathbb{R} : \tau \leq \tilde{\mathbb{P}}_T(x) \right\}. \quad (3.2)$$

- (iii) Calculate the empirical CDF based on the return data

$$\mathbb{P}_T = \frac{1}{T} \sum_{t=1}^T \delta_{R_t}.$$

- (iv) Calculate the quantile bound

$$\hat{\theta}_{\text{emp}}(\tau) := \frac{\tau - \mathbb{P}_T(\tilde{Q}_T(\tau))}{\sqrt{\mathbb{P}_T(\tilde{Q}_T(\tau))(1 - \mathbb{P}_T(\tilde{Q}_T(\tau)))}} \quad \varepsilon \leq \tau \leq 1 - \varepsilon. \quad (3.3)$$

I refer to the above estimator as the empirical quantile bound. In simulation, as well as in the empirical application in Section 4, the resulting estimator is too jagged due to the non-smooth nature of the empirical distribution function. The jaggedness implies that the estimator in 3.3 is too

volatile in small samples. I therefore also consider a smoothed version of [3.3](#) (denoted by  $\hat{\theta}_{\text{smooth}}$ ), and defined by

$$\hat{\theta}_{\text{smooth}}(\tau) := \frac{\tau - \mathbb{P}_{T,\text{smooth}}(\tilde{Q}_T(\tau))}{\sqrt{\mathbb{P}_{T,\text{smooth}}(\tilde{Q}_T(\tau))(1 - \mathbb{P}_{T,\text{smooth}}(\tilde{Q}_T(\tau)))}} \quad \varepsilon \leq \tau \leq 1 - \varepsilon. \quad (3.4)$$

Here,  $\mathbb{P}_{T,\text{smooth}}$  is a kernel estimator for the CDF. I use the Epanechnikov kernel, as this is shown to have descent performance in simulations (see [Section D.1](#)).

The estimator in [3.3](#) now carries two sources of randomness, namely in the estimation of  $\mathbb{P}_T$ , but also in the estimation of  $\tilde{Q}_T$ . I can use the results from the previous Section to argue that the extra randomness can be separated as follows. For any  $\tau \in [\varepsilon, 1 - \varepsilon]$

$$\begin{aligned} \frac{\sqrt{T} [\mathbb{P}_T(\tilde{Q}_T(\tau)) - \mathbb{P}(\tilde{Q}(\tau))]}{\sqrt{\mathbb{P}_T(\tilde{Q}_T(\tau))(1 - \mathbb{P}_T(\tilde{Q}_T(\tau)))}} &= \frac{\sqrt{T} [\mathbb{P}_T(\tilde{Q}_T(\tau)) - \mathbb{P}(\tilde{Q}(\tau))]}{\sqrt{\mathbb{P}_T(\tilde{Q}(\tau))(1 - \mathbb{P}_T(\tilde{Q}(\tau)))}} + o_p(1) \\ &= \overbrace{\frac{\sqrt{T} [\mathbb{P}_T(\tilde{Q}_T(\tau)) - \mathbb{P}_T(\tilde{Q}(\tau))]}{\sqrt{\mathbb{P}_T(\tilde{Q}(\tau))(1 - \mathbb{P}_T(\tilde{Q}(\tau)))}}}^{I_1 = \text{Estimation error}} + \overbrace{\frac{\sqrt{T} [\mathbb{P}_T(\tilde{Q}(\tau)) - \mathbb{P}(\tilde{Q}(\tau))]}{\sqrt{\mathbb{P}_T(\tilde{Q}(\tau))(1 - \mathbb{P}_T(\tilde{Q}(\tau)))}}}^{\text{(B.2)} \frac{1}{\sqrt{T}} W(\tau)} + o_p(1). \quad (3.5) \end{aligned}$$

The first equality uses the continuous mapping theorem, provided that  $\tilde{Q}_T(\tau)$  is uniformly consistent. The second term in Equation (3.5) comes from the proof of Theorem 3.3. Given the various stages to estimate  $\tilde{Q}_T$  it is hard to characterize the limiting distribution of  $I_1$ . This is even more complicated for the smooth estimator in (3.4), as it will also depend on the bandwidth. As a robust alternative to the uniform confidence bands based on asymptotic theory, one can use bootstrap instead. This is discussed in more detail in [Section 4](#).

## 4 Empirical application

This Section presents estimates of the quantile bound using a combination of forward looking option data and historical market returns. The main result shows that the quantile bound for monthly returns is stronger than the HJ bound, which implies a much more volatile SDF. Moreover, the shape of the quantile bound contains features that common asset pricing models cannot reconcile. As byproduct of these calculations, I note that the conditional risk-neutral quantile curves behave different depending on market conditions. This is further explored in [Sections 5 and 6](#).

### 4.1 Data description and estimation

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.<sup>7</sup> Finally, stock price data on the closing price of the S&P 500 are obtained via WRDS.

I use the following procedure to calculate the unconditional risk-neutral quantile curve at maturities 30, 60, 90, 180 and 360 days. Prior to constructing the bound, 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. I calculate the put or call option price as the average of the highest closing bid and lowest closing ask. Subsequently, the conditional and unconditional quantile curves are estimated using Steps (i) and (ii) described in Section 3.2. In more detail, for Step (i), I use both put and call option data to calculate the price of digital put option  $P_{m,t+1}(K)$  via

$$P_{m,t+1}(K) = \begin{cases} \left. \frac{\partial}{\partial K} \text{put}_{m,t+1}(K) \right|_{K=K} & \text{if } 0 \leq K \leq F_{m,t+1} \\ \left. \frac{1}{R_{f,t+1}} + \frac{\partial}{\partial K} \text{call}_{m,t+1}(K) \right|_{K=K} & \text{if } K > F_{m,t+1}, \end{cases} \quad (4.1)$$

where  $\text{put}_{m,t+1}(K)$  (resp.  $\text{call}_{m,t+1}(K)$ ) is the price of a European put (resp. call) option with strike  $K$  expiring at time  $t+1$  and  $F_{m,t+1}$  is the price of a forward contract on the market return.<sup>8</sup> All these prices are known at time  $t$ . By the results of [Breedon and Litzenberger \(1978\)](#)

$$R_{f,t+1} P_{m,t+1}(K) = \tilde{\mathbb{P}}_t(R_{m,t+1} \leq K). \quad (4.2)$$

Since strike prices come in discrete units, I approximate 4.1 by a forward finite difference scheme. Using linear interpolation (see Step (i)), the conditional quantile curves for maturities 30, 60, 90, 180 and 360 days are obtained. The unconditional quantile curves are obtained using Step (ii).

Figure 4 shows the unconditional quantile curves. 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.

## 4.2 Quantile bound for 30-day returns

I now turn to the computation of the unconditional quantile bound for 30-day returns. Recall from Equation (2.2) that the estimand is given by

$$\theta(\tau) := \frac{\tau - \mathbb{P}(\tilde{Q}(\tau))}{\sqrt{\mathbb{P}(\tilde{Q}(\tau))(1 - \mathbb{P}(\tilde{Q}(\tau)))}} \quad \varepsilon \leq \tau \leq 1 - \varepsilon. \quad (4.3)$$

I use the methodology described in Section 3.2 to estimate this function. In particular, I estimate  $\theta$  in two different ways, using  $\hat{\theta}_{\text{emp}}$  and  $\hat{\theta}_{\text{smooth}}$  described

<sup>7</sup>See [http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data\\_library.html#Research](http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html#Research)

<sup>8</sup>I do so since out of the money call options are more liquid than in the money put options.

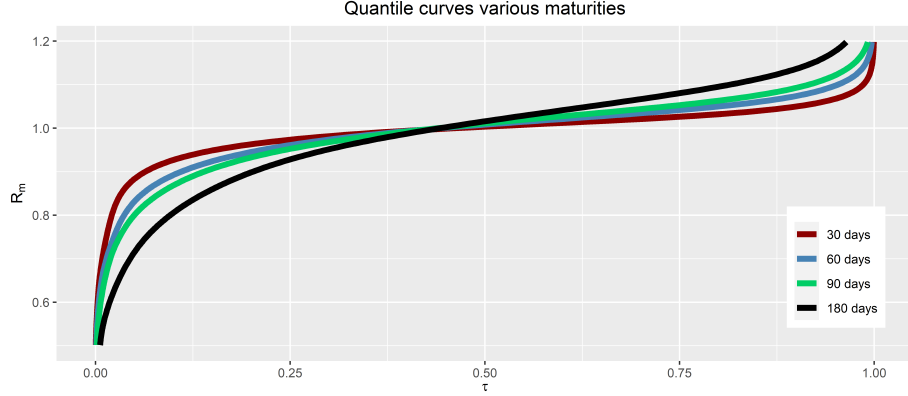


Figure 4: Plots of the unconditional (risk-neutral) quantile function for various maturities.

in Step (iv). The unconditional CDF estimate in both cases 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 obtained from Equation 3.2, 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 empirical quantile bound (in blue), as well as the smoothed version (in red). The smoothed curve is my preferred estimator, since it is less prone to discretization errors and thus less volatile than the empirical quantile bound. It is apparent that the smoothed curve is above the HJ bound at almost any probability level  $\tau \in [0.12, 0.88]$ . Moreover, the supremum peaks around 0.186, which is higher than the in-sample HJ bound.

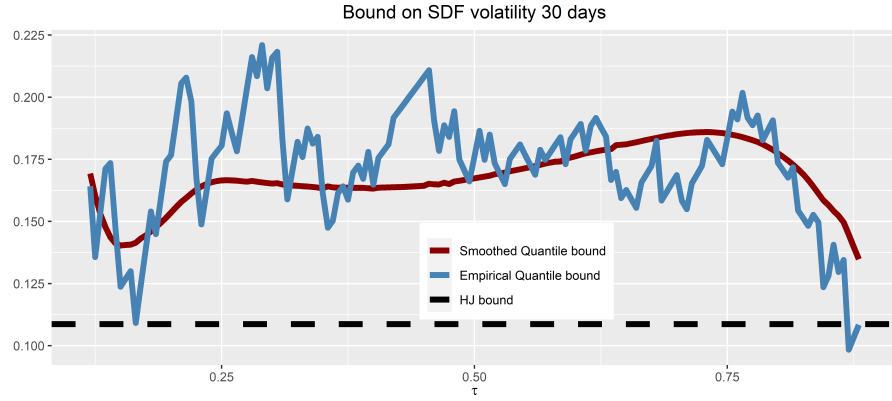


Figure 5: Plot of the quantile bound as function of  $\tau$ . Solid blue line denotes the empirical quantile bound as a function of  $\tau$ . The solid red line is the smoothed version of the quantile bound. The dashed black line depicts the HJ bound.

I now discuss if this is a statistically significant improvement over the HJ

bound. By definition, the sup quantile bound estimator is given by

$$\hat{\theta}_{\text{smooth}} := \sup_{\varepsilon \leq \tau \leq 1-\varepsilon} \frac{\tau - \mathbb{P}_{T,\text{smooth}}(\tilde{Q}_T(\tau))}{\sqrt{\mathbb{P}_{T,\text{smooth}}(\tilde{Q}_T(\tau))(1 - \mathbb{P}_{T,\text{smooth}}(\tilde{Q}_T(\tau)))}}. \quad (4.4)$$

The proposed SUP statistic essentially iterates over many test statistics and is therefore prone to the data snooping critique of [White \(2000\)](#). In order to address this issue I propose a bootstrap scheme that produces confidence intervals, which takes the “data mining” into account.

First, I sample months  $t$  (with replacement) to generate 100,000 bootstrap samples, using the stationary bootstrap of [Politis and Romano \(1994\)](#). Each sample consists of  $T = 288$ , 30-day returns. For each bootstrap sample, I calculate the implied (smoothed) empirical process  $\mathbb{P}_{T,\text{smooth}}^{(n)}$  (here  $n$  refers to the  $n$ th bootstrap sample). For the same time data, I calculate  $\tilde{\mathbb{P}}_T^{(n)}$  and obtain  $\tilde{Q}_T^{(n)}(\tau)$  via [\(3.2\)](#). The analogue of [\(4.4\)](#) is then calculated for each bootstrap sample

$$\hat{\theta}_{\text{smooth}}^{(n)} = \sup_{\varepsilon \leq \tau \leq 1-\varepsilon} \frac{\tau - \mathbb{P}_{T,\text{smooth}}^{(n)}(\tilde{Q}_T^{(n)}(\tau))}{\sqrt{\mathbb{P}_{T,\text{smooth}}^{(n)}(\tilde{Q}_T^{(n)}(\tau))(1 - \mathbb{P}_{T,\text{smooth}}^{(n)}(\tilde{Q}_T^{(n)}(\tau)))}}$$

I set  $\varepsilon = 0.12$  to avoid volatile estimates in the tail. The resulting confidence intervals are given by

$$[\hat{\theta}_{\text{smooth}} + \hat{\xi}_{\alpha/2}, \hat{\theta}_{\text{smooth}} + \hat{\xi}_{1-\alpha/2}].$$

Here,  $\hat{\xi}_{\alpha}$  is the  $\alpha$ -quantile of  $\hat{\theta}_{\text{smooth}}^{(n)} - \hat{\theta}_{\text{smooth}}$ .<sup>9</sup> Table 2 shows the results. Apparently, the bootstrap confidence intervals imply that the quantile bound is significantly tighter than the HJ bound at the 5% level.

Table 2: Bootstrap results

Sup quantile Bound	$CI_{1-\alpha}$	HJ bound	$H_0$	Annualized vol.
0.1860	[0.1171, 0.3700]	0.1086	0.0081	0.6487

Note: *Bootstrap results of the maximum statistic of the quantile bound. Results are based on 100,000 bootstrap samples with sample size of 288 each. Sup quantile bound denotes the supremum of the smoothed quantile bound estimator in Figure 5.  $CI_{1-\alpha}$  denotes the 95% (bootstrapped) confidence interval. The third column is the HJ bound.  $H_0$  is the  $p$ -value corresponding to the null hypothesis that the Sup quantile bound equals the HJ bound. Annualized vol. scales the Sup quantile bound to annual units.*

What are the economic implications of this finding? First, in Section 2.3 I argued that under CAPM, the HJ bound is always stronger than the quantile bound. Hence, Table 2 is evidence against CAPM. The LRR model of [Bansal et al. \(2012\)](#) can neither reconcile this feature of the data under common parameter calibration. In contrast, the disaster risk model of [Barro \(2006\)](#)

<sup>9</sup>Simulation evidence in Appendix D.1 shows that the bootstrap approach is reasonable in a Black-Scholes setting, despite undercoverage of the resulting confidence intervals. Moreover, the point estimates are upward biased.

does reconcile this phenomenon, but in that model the quantile bound is only stronger for very small values of  $\tau$ . The implied quantile bound is decreasing in the disaster risk model for  $\tau \geq 0.065$ , which is not in line with Figure 5. Hence, the quantile bound provides a rich set of implications which are hard to satisfy for popular asset pricing models.

Secondly, the finding that the SDF is more volatile than what could be concluded based on the HJ bound exacerbates the so-called equity premium puzzle in the macro finance literature. Consider a consumption based SDF of the form  $\beta g_c^{-\gamma}$ , where  $\beta$  is a time discount factor,  $g_c$  is consumption growth and  $\gamma$  is the representative agent's degree of risk aversion. Resembling the back of the envelope calculations of [Cochrane \(2005\)](#), it follows that

$$\gamma \sigma(g_c) \geq \text{HJ bound.}$$

Here  $\sigma(g_c)$  is the volatility of consumption growth. Using the annualized HJ-bound in Table 2 and  $\sigma(g_c) = 0.01$  ([Cochrane, 2005](#), Chapter 21), it follows that  $\gamma \geq 38$ . Using the annualized of the quantile bound instead, leads to  $\gamma \geq 65$ , a significant worsening of the (already) disproportionate risk-aversion coefficient based on the HJ bound.

### 4.3 Conditional quantile curves

A key ingredient of the quantile bound in the previous Section is the unconditional risk-neutral quantile curve. However, the conditional risk-neutral curves also contain important information, as I now show. In particular, I estimate the conditional quantile curves during a "normal" and "crisis" week. This provides insight that risk-neutral quantiles are different under changing market conditions and hence embed information that could potentially be useful to estimate the physical quantile curve. Figure 6 displays the forward looking quantile curves for 30,60,90 and 180 days during the 3<sup>rd</sup> week of September in 2005 and 2008. The 3<sup>rd</sup> week in 2005 was a relatively tranquil week without major events happening, whereas the 3<sup>rd</sup> week of September in 2008 was a week of major financial turmoil, with the bankruptcy announcement of Lehman Brothers and the liquidity crisis of AIG among others. The curves obtained represent a weekly average. As expected, the quantile curve obtained during the 3<sup>rd</sup> week of 2008 displays a heavier negative tail compared to the 3<sup>rd</sup> week in 2005. More interestingly, the upper tail is also heavier during the crisis week, signaling that investors expect swings in either direction. Since these quantile curves can be computed daily and exhibit noticeable different behavior during different market conditions, they might carry important information that can be incorporated to make forecasts about the (conditional) physical quantile curve. I analyze this further in Section 5.

## 5 Relation between physical and risk-neutral quantile

Section 4 documents that the conditional quantile function behaves different under varying market conditions. The risk-neutral quantile function is forward looking, as it is obtained from option data. In this Section, I further explore

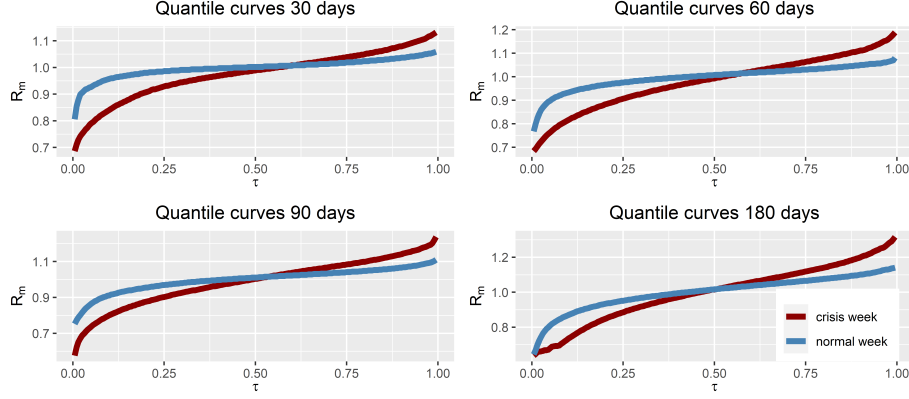


Figure 6: Plots of the conditional (risk-neutral) quantile function for various maturities. “normal week” and “crisis week” denote the 3rd week of September in 2005 and 2008 respectively.

how the risk-neutral quantile function relates to the physical quantile function. The main purpose is to obtain a proxy of the latent physical quantile function using market observable data.

### 5.1 Lower bound on physical quantile

I start by deriving a market observable lower bound on the physical quantile function. The starting point is the quantile relation derived in (B.1)

$$\mathbb{P}_t(R_{t+1} \leq \tilde{Q}_\tau) = \tau - R_{f,t+1} \text{COV}_t \left( M_{t+1}, \mathbb{1} \left( R_{t+1} \leq \tilde{Q}_\tau \right) \right). \quad (5.1)$$

Write  $F_t(x) := \mathbb{P}_t(R_{t+1} \leq x)$  for the conditional distribution function of  $R_{t+1}$  and  $f_t(\cdot)$  for the corresponding pdf. Then rearranging (5.1) tells us that

$$\tau = F_t(Q_\tau) = F_t(\tilde{Q}_\tau) + R_{f,t+1} \text{COV}_t \left( M_{t+1}, \mathbb{1} \left( R_{t+1} \leq \tilde{Q}_\tau \right) \right).$$

The idea is to invert this equation to derive a link between the risk-neutral quantile and physical quantile. This can be achieved by means of a first order Taylor expansion of the physical quantile function around  $F_t(\tilde{Q}_\tau)$ :

$$\begin{aligned} Q_\tau &= F_t^{-1} \left( F_t(\tilde{Q}_\tau) + R_{f,t+1} \text{COV}_t \left( M_{t+1}, \mathbb{1} \left( R_{t+1} \leq \tilde{Q}_\tau \right) \right) \right) \\ &\approx \tilde{Q}_\tau + \underbrace{\frac{1}{f_t(\tilde{Q}_\tau)} R_{f,t+1} \text{COV}_t \left( M_{t+1}, \mathbb{1} \left( R_{t+1} \leq \tilde{Q}_\tau \right) \right)}_{\text{risk adjustment}} \\ &= \tilde{Q}_\tau + \frac{1}{f_t(\tilde{Q}_\tau)} \left[ \tau - \mathbb{E}_t \left[ \mathbb{1} \left( R_{t+1} \leq \tilde{Q}_\tau \right) \right] \right]. \end{aligned} \quad (5.2)$$

The last line follows from writing out the covariance term. Equation (5.2) cannot directly be used for forecasting purposes since the conditional density  $f_t(\tilde{Q}_\tau)$  is unknown, as well as the conditional expectation on the right. Under additional assumptions the conditional expectation can be inferred (or at least

approximated) with market data. The idea that the physical expectation can be inferred from the data is related to recent work of [Ross \(2015\)](#) and [Borovička et al. \(2016\)](#) on the recovery of subjective beliefs. To show this, I link the expectation to a risk-neutral covariance term, similar to [Kremens and Martin \(2019\)](#) and [Chabi-Yo and Loudis \(2020\)](#). Use the reciprocal of the SDF to pass from physical to risk-neutral measure

$$\begin{aligned}\mathbb{E}_t \left[ \mathbb{1} \left( R_{t+1} \leq \tilde{Q}_\tau \right) \right] &= \tilde{\mathbb{E}}_t \left[ R_{t+1} \leq \tilde{Q}_\tau \right] \frac{\mathbb{E}_t [M_{t+1}]}{M_{t+1}} \\ &= \widetilde{\text{COV}}_t \left[ \mathbb{1} \left( R_{t+1} \leq \tilde{Q}_\tau \right), \frac{\mathbb{E}_t [M_{t+1}]}{M_{t+1}} \right] + \tau.\end{aligned}\quad (5.3)$$

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

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

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

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

The  $\theta_k$ -coefficients depend on the specific utility representation employed, but are conditionally non-random. I substitute the above in Equation (5.3)

$$\begin{aligned}&\widetilde{\text{COV}}_t \left[ \mathbb{1} \left( R_{m,t+1} \leq \tilde{Q}_\tau \right), \frac{\tilde{\mathbb{E}}_t (M_{t+1})}{M_{t+1}} \right] \\ &= \frac{\sum_{k=1}^{\infty} \theta_k \left( \tilde{\mathbb{E}}_t \left[ \mathbb{1} \left( R_{m,t+1} \leq \tilde{Q}_\tau \right) (R_{m,t+1} - R_{f,t+1})^k \right] - \tau \tilde{\mathbb{E}}_t \left[ (R_{m,t+1} - R_{f,t+1})^k \right] \right)}{1 + \sum_{k=1}^{\infty} \theta_k \tilde{\mathbb{E}}_t \left[ (R_{m,t+1} - R_{f,t+1})^k \right]},\end{aligned}\quad (5.4)$$

since  $\tilde{\mathbb{E}}_t(\mathbb{1}(R_{m,t+1} \leq \tilde{Q}_\tau)) = \tau$ . Results in Appendix A 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

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

This means (5.4) can be rewritten to

$$\widetilde{\text{COV}}_t \left[ \mathbb{1} \left( R_{m,t+1} \leq \tilde{Q}_\tau \right), \frac{1}{M_{t+1} R_{f,t+1}} \right] = \frac{\sum_{k=1}^{\infty} \theta_k \left( \tilde{\mathbb{M}}_{t+1}^{(k)}[\tilde{Q}_\tau] - \tau \tilde{\mathbb{M}}_{t+1}^{(k)} \right)}{1 + \sum_{k=1}^{\infty} \theta_k \tilde{\mathbb{M}}_{t+1}^{(k)}}.\quad (5.5)$$

Combining Equation (5.5) and (5.3) in (5.2) leads to the first order approximation

$$Q_\tau \approx \tilde{Q}_\tau + \frac{1}{f_t(\tilde{Q}_\tau)} \left( \frac{\sum_{k=1}^{\infty} \theta_k \left( \tau \tilde{\mathbb{M}}_{t+1}^{(k)} - \tilde{\mathbb{M}}_{t+1}^{(k)}[\tilde{Q}_\tau] \right)}{1 + \sum_{k=1}^{\infty} \theta_k \tilde{\mathbb{M}}_{t+1}^{(k)}} \right). \quad (5.6)$$

The right hand side still depends on the unknown density  $f_t(\tilde{Q}_\tau)$  and unknown parameters  $\theta_k$ . However, Chabi-Yo and Loudis (2020) show that we can make assumption about  $\theta_k$  that lead to a lower bound. I adopt the following assumptions of Chabi-Yo and Loudis (2020):

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

$$\begin{aligned} \tilde{\mathbb{M}}_{t+1}^{(1)}[k_0] &\leq 0, & \tilde{\mathbb{M}}_{t+1}^{(2)}[k_0] &\geq 0 \\ \tilde{\mathbb{M}}_{t+1}^{(3)}[k_0] &\leq 0, & \tilde{\mathbb{M}}_{t+1}^{(4)}[k_0] &\geq 0. \end{aligned}$$

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

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

For Corollary 5.5 below, I strengthen Assumption 5.2 as follows.

**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.

I will continue to assume that the approximation in (5.6) holds with equality. The following Lemma and Corollary summarize the lower bound on the discrepancy between physical and risk-neutral quantile.

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

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

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

$$Q_\tau - \tilde{Q}_\tau \geq \frac{1}{f_t(\tilde{Q}_\tau)} \left( \frac{\sum_{k=1}^3 \theta_k \left( \tau \tilde{\mathbb{M}}_{t+1}^{(k)} - \tilde{\mathbb{M}}_{t+1}^{(k)}[\tilde{Q}_\tau] \right)}{1 + \sum_{k=1}^3 \theta_k \tilde{\mathbb{M}}_{t+1}^{(k)}} \right). \quad (5.7)$$

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

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

*Proof.* See Appendix B.4. ■

## 5.2 Predictive theory

Recall that  $Q_\tau$  is the conditional physical quantile of  $R_{m,t+1}$ , which is latent and thus unknown at time  $t$ . For risk management purposes, it is important to forecast  $Q_\tau$ , since it corresponds to the Value-at-Risk (VaR) if a company is fully invested in the market portfolio. Corollary 5.5 suggest that if the lower bound is tight, one could predict  $Q_\tau$  by means of

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

This is promising since  $\tilde{Q}_\tau$  and  $LRB_{t+1}(\tau)$  are both observed at time  $t$ . I still need to get a hand on  $f_t(\tilde{Q}_\tau)$  to make this operational. To do so I argue as follows.

From the proof of Corollary 5.5,  $LRB_{t+1}(\tau)$  approximates  $\tau - F_t(\tilde{Q}_\tau)$ . This can be rewritten to

$$F_t(\tilde{Q}_\tau) = \tau - LRB_{t+1}(\tau). \quad (5.8)$$

Since  $f_t(\cdot)$  is the derivative of  $F_t(\cdot)$ , differentiate (5.8) to obtain

$$f_t(\tilde{Q}_\tau) = \tilde{f}_t(\tilde{Q}_\tau) - LRB_{t+1}(\tau)', \quad (5.9)$$

$$\text{where} \quad LRB_{t+1}(\tau)' := \frac{\partial LRB_{t+1}(\tau)}{\partial \tilde{Q}_\tau}.$$

$\tilde{f}_t(\tilde{Q}_\tau)$  denotes the risk-neutral pdf evaluated at the risk-neutral quantile.<sup>10</sup> Finally, define

$$\Xi_{t+1}(\tau) := \frac{\tau - LRB_{t+1}(\tau)}{\tilde{f}_t(\tilde{Q}_\tau) - LRB_{t+1}(\tau)'} \approx \frac{\tau - F_t(\tilde{Q}_\tau)}{f_t(\tilde{Q}_\tau)}. \quad (5.10)$$

Appendix C.2 presents evidence that the approximation in (5.10) is accurate in a Black-Scholes environment. Given the accuracy of this approximation, I consider a quantile regression specification of the form

$$Q_\tau(R_{m,t+1}) = \beta_0(\tau) + \left[ \tilde{Q}_\tau + \Xi_{t+1}(\tau) \right] \beta_1(\tau). \quad (5.11)$$

If the approximations above are accurate one expects

$$\begin{aligned} H_0 : \quad & \beta_0(\tau) = 0 \\ & \beta_1(\tau) = 1. \end{aligned} \quad (5.12)$$

---

<sup>10</sup>This formula follows from the relation  $\tau = \tilde{F}_t(\tilde{Q}_\tau)$ , which implies  $\frac{\partial \tau}{\partial \tilde{Q}_\tau} = \tilde{f}_t(\tilde{Q}_\tau)$ .



To test this theory driven conjecture, I use the same data as in Section 4. Firstly, the (un)truncated moments appearing in  $LRB_{t+1}(\tau)$  are computed using the estimated quantile curves from Section 4 and using the results in Appendix A.<sup>11</sup> I calculate the lower bound for 30,60, and 90 days, as these tend to have enough liquidity so that the risk-neutral quantile curve is well approximated in the tail and the estimation error is minimal. Thereafter,  $\Xi_{t+1}(\tau)$  in (5.10) is calculated using a finite difference scheme. Subsequently, the realized  $n$ -day returns are regressed on a constant and the forward looking variable  $\tilde{Q}_\tau + \Xi_{t+1}(\tau)$  (calculated at time  $t$ ) using the quantile regression approach of Koenker and Bassett (1978). I narrow down the sample period to 2006-2019, since it is harder to estimate the risk-neutral quantile function in the period before due to insufficient option data. In addition, I impose the mild economic constraint that  $\Xi_{t+1}(\tau) \in [0.0, 0.4]$ . I throw out the small part of the sample for which this condition is violated.<sup>12</sup> This occurs due to inaccurate finite difference approximations of  $\tilde{f}_t(\tilde{Q}_\tau)$ .

The resulting estimates are shown in Table 3. The estimates line up reasonably well with my theoretical prediction, as 4 out of the 9 models do not reject the joint restriction in (5.12).<sup>13</sup> Moreover, the fit seems to improve over longer time horizons. The standard errors are obtained via bootstrap using the `boot.rq` function in R, since overlapping returns are used which likely generate autocorrelation in the error term.<sup>14</sup> To show that including  $\Xi_{t+1}(\tau)$  in the quantile regression leads to genuine information gain, I also report the quantile regression estimates, when including  $\tilde{Q}_\tau$  only. The coefficients are shown in the bottom of Table 3. The coefficients  $\beta_0(\tau)$  tend to be higher and  $\beta_1(\tau)$  lower compared to the estimates that result when including  $\Xi_{t+1}(\tau)$ . Moreover, for every quantile and day, the null hypothesis in (5.12) is rejected.

## 6 Long run VaR forecasts and evaluation

The quantile specification in Equation (5.11) lines up well with theoretical predictions. Since the specification in (5.11) is forward looking and can be calculated for longer periods of time (e.g. 30,60 or 90 days), I analyze how well the model performs over longer time horizons. This differs from the typical VaR literature which is focused on daily forecasts (e.g. Kuester et al. (2006)). As stressed by Engle (2009), firms' liabilities tend to extend over longer periods of time and hence the need for long term VaR forecasts. First, two tests are discussed which can be used to diagnose whether a VaR model produces accurate results. Thereafter, two potential alternative models are considered that can be used to predict long term VaR. I document that (5.11) shows promise in predicting long term VaR compared to extant procedures.

<sup>11</sup>Since the quantile curves are hard to estimate in the tail (due to insufficient put options available), the even moments are slightly underestimated.

<sup>12</sup>Typically this is about 1% of the sample size. The results are robust to different intervals.

<sup>13</sup>Notice it is not possible to construct simultaneous tests for multiple quantiles along the lines of Angrist et al. (2006), since their theory does not cover the case when covariates depend on the quantiles as well.

<sup>14</sup>Hansen and Hodrick (1980) and Hodrick (1992) propose solutions for this problem in standard OLS regressions.

Table 3: Quantile regression estimates

	30 days		60 days		90 days	
	$\widehat{\beta}_0(\tau)$	$\widehat{\beta}_1(\tau)$	$\widehat{\beta}_0(\tau)$	$\widehat{\beta}_1(\tau)$	$\widehat{\beta}_0(\tau)$	$\widehat{\beta}_1(\tau)$
$\tau = 0.01$	0.31 (0.1138)	0.72 (0.1332)	- 0.04 (0.2084)	1.14 (0.2778)	- 0.09 (0.2403)	1.18 (0.3437)
$H_0 : [\beta_0(\tau), \beta_1(\tau)] = [0, 1]$	0		0.67		0.59	
$\tau = 0.05$	0.51 (0.0673)	0.47 (0.0727)	0.29 (0.0925)	0.70 (0.1031)	0.17 (0.1020)	0.85 (0.1157)
$H_0 : [\beta_0(\tau), \beta_1(\tau)] = [0, 1]$	0		0		0.01	
$\tau = 0.1$	0.34 (0.0599)	0.65 (0.0630)	0.12 (0.1104)	0.89 (0.1181)	0.10 (0.0998)	0.92 (0.1095)
$H_0 : [\beta_0(\tau), \beta_1(\tau)] = [0, 1]$	0		0.13		0.18	
Quantile regression with regressor $\widetilde{Q}_\tau$						
$\tau = 0.01$	0.42 (0.1153)	0.59 (0.1386)	0.45 (0.1294)	0.47 (0.1948)	0.68 (0.1375)	0.08 (0.2211)
$\tau = 0.05$	0.60 (0.0556)	0.37 (0.0604)	0.57 (0.0632)	0.40 (0.0734)	0.51 (0.0733)	0.47 (0.0871)
$\tau = 0.1$	0.55 (0.0349)	0.43 (0.0373)	0.56 (0.0609)	0.43 (0.0665)	0.51 (0.0600)	0.48 (0.0664)

Note: Quantile regression estimates of (5.11) using 30, 60 and 90-day forward looking estimates. The bottom part of the table contains regression coefficients using only  $\tilde{Q}_\tau$  as covariate. Standard errors are obtained via bootstrap and displayed under the coefficients.

## 6.1 Tests for model performance

### 6.1.1 Unconditional coverage

Firstly, I assess whether the model has correct unconditional coverage, using the unconditional coverage test of Christoffersen (1998). A VaR forecast model is deemed efficient if

$$\mathbb{E}_t[H_{t+1}] = \tau, \quad H_{t+1} = \mathbb{1}\left(R_{m,t+1} < \hat{Q}_\tau(R_{m,t+1})\right). \quad (6.1)$$

$\hat{Q}_\tau(R_{m,t+1})$  is the quantile forecast produced at time  $t$ . Using iterated expectations one can test the null hypothesis for correct unconditional coverage

$$H_0 : \mathbb{E}(H_{t+1}) = \tau \quad H_a : \mathbb{E}(H_{t+1}) \neq \tau.$$

Christoffersen (1998) shows that the likelihood-ratio test takes the form

$$LR_{uc} = 2[\mathcal{L}(\hat{\tau}; H_1, \dots, H_T) - \mathcal{L}(\tau; H_1, \dots, H_T)] \stackrel{\text{asy}}{\sim} \chi_1^2,$$

where  $\mathcal{L}$  denotes the log-likelihood of the binomial distribution and  $\hat{\tau} = n_1/(n_0 + n_1)$  is the MLE estimator with  $n_1$  being the number of violations and  $n_0 + n_1 = T$ . This result is derived for *i.i.d.* data. Since my application deals with overlapping return data, the finite sample distribution will differ significantly from the asymptotic  $\chi_1^2$  distribution. Hence, I use the stationary bootstrap (Politis and Romano, 1994) to approximate the sampling distribution of  $\sqrt{T}(\hat{\tau} - \tau)$  by

$\sqrt{T}(\hat{\tau}^* - \hat{\tau})$ . Here  $\hat{\tau}^*$  denotes the bootstrapped statistic. Returns are sampled with replacement using 10,000 bootstrap samples. Appendix D.2 contains a more detailed description.

### 6.1.2 VQR test

Models may well have correct unconditional coverage, but fail to have correct coverage conditional on time  $t$ . For example, this happens if  $H_{t+1} = 1$  occurs in clusters, which is an undesirable feature (Christoffersen, 1998). Gaglianone et al. (2011) propose a test to deal with correct conditional coverage. To explain this test, suppose we have a candidate VaR estimate  $V_t$ . Under certain assumptions, the conditional quantile function obtains the form

$$Q_\tau(R_{m,t+1}) = \alpha_0(\tau) + \alpha_1(\tau)V_t.$$

If  $-V_t$  is indeed the  $\tau$ -conditional quantile, then  $\mathbb{P}_t(R_{m,t+1} < -V_t) = \tau$ . Moreover, under the null hypothesis that  $-V_t$  is the  $\tau$ -conditional quantile, the coefficients satisfy

$$H_0 : \begin{cases} \alpha_0(\tau) = 0 \\ \alpha_1(\tau) = -1. \end{cases}$$

This can be rewritten to the basic form  $H_0 : \theta(\tau) = 0$ , where  $\theta(\tau) = [\alpha_0(\tau), (\alpha_1(\tau) + 1)]'$ . Gaglianone et al. (2011) show that the following test statistic has an asymptotic chi-squared distribution with 2 degrees of freedom.

$$\zeta_{VQR} := T \left[ \hat{\theta}(\tau)' (\tau(1-\tau)Z_\tau^{-1}JZ_\tau^{-1})^{-1} \hat{\theta}(\tau) \right] \xrightarrow{d} \chi_2^2.$$

Here  $\hat{\theta}$  is the quantile regression estimator of  $\theta(\tau)$  (Koenker and Bassett, 1978) and

$$\begin{aligned} J &= \text{plim}_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T x_t x_t' \\ Z_\tau &= \text{plim}_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T x_t x_t' [f_t(Q_\tau(R_{m,t+1}|x_t))] \\ x_t &= [1 \quad V_t]'. \end{aligned}$$

Consistent estimates of  $J$  and  $H_\tau$  can be obtained using bootstrap (Koenker, 1994).

## 6.2 Alternative models of long term VaR

### 6.2.1 GARCH approach

The VaR literature typically focuses on daily estimates, but some methods have been proposed in the literature that deal with long term forecasts. One of these methods is described on the website of VLAB.<sup>15</sup> The idea is to fit a GARCH

<sup>15</sup><https://vlab.stern.nyu.edu/docs/lrrisk>

model on daily log return data  $r_{t+1} := \log R_{m,t+1}$ . In particular, the model specifies

$$\begin{aligned} r_{t+1} &= \sqrt{\sigma_{t+1}} \varepsilon_{t+1} \\ \sigma_{t+1} &= \omega + \alpha r_t^2 + \beta \sigma_t. \end{aligned}$$

After the parameters are estimated, I obtain the historical shocks from  $\varepsilon_t = r_t / \sqrt{\sigma_t}$ . The historical shocks are used to bootstrap future  $k$ -day returns

$$r_{t,t+k} = \sum_{i=1}^k r_{t+i}.$$

Finally, log returns are converted back to arithmetic returns. The  $k$ -day return VaR is then obtained from the quantile of the bootstrapped distribution.

### 6.2.2 CAViaR approach

The CAViaR model of [Engle and Manganelli \(2004\)](#) posits an autoregressive specification for the quantiles, potentially enriched with covariates such as the absolute value of lagged returns. These models are typically estimated on daily returns (e.g. [Kuester et al. \(2006\)](#)). I explore whether the CAViaR model performs accurately over longer term horizons, using the adaptive specification

$$\text{VaR}_{t+1} = \text{VaR}_t + \beta \left( [1 + \exp(K[R_{m,t} - \text{VaR}_t])]^{-1} - \tau \right). \quad (6.2)$$

Here,  $K$  is a tuning parameter that smooths the indicator function. [Kuester et al. \(2006\)](#) find that this model has decent unconditional coverage based on daily return data.<sup>16</sup>

## 6.3 Empirical results

I now turn to the actual evaluation of the proposed model using the performance criteria discussed above. In doing so, I consider the following four model specifications

1. The full quantile regression specification (5.11), using  $\tilde{Q}_\tau + \Xi_{t+1}(\tau)$  as covariate. This is referred to as **Model I**.
2. Quantile regression which only includes  $\tilde{Q}_\tau$  as covariate. This (benchmark) model is referred to as **Model II**.
3. The adaptive CAViaR specification in Equation 6.2. I set  $K = 10$  and use the same optimization procedure in [Engle and Manganelli \(2004\)](#). This specification is referred to as **Model III**.
4. The GARCH approach described in Section 6.2.1 using 10,000 bootstrap samples. This is referred to as **Model IV**.

<sup>16</sup>[Kuester et al. \(2006\)](#) consider the specification  $\text{VaR}_{t+1} = \text{VaR}_t + \beta[1(R_{m,t} \leq -\text{VaR}_t) - \tau]$ . This corresponds to letting  $K \rightarrow \infty$  in (6.2).

As in [Kuester et al. \(2006\)](#), I consider a rolling window size 1,000 to allow for changing parameters over time. The VaR forecasts are analyzed for 30,60 and 90-day returns. Given the forward looking information of option data one might conjecture that **Model I** and **II** could still perform well over longer time horizons. This is still of great interest, for example to risk managers. Daily VaR forecast are useful for answering the question: how much capital do I need tomorrow to withstand a shock of size  $\alpha$  or less? A 90 day VaR forecast answers the question: how much capital do I need in 90 days to withstand a shock of size  $\alpha$  or less? The answer to the second question carries mutually distinct information from the answer to the first question. Note, however, that the answer to the second question cannot be interpreted as having enough capital to withstand a shock of size  $\alpha$  during the entire 90 day period, since our analysis applies only to a fixed time horizon and not to the period in between.

Table 4 shows that **Models I,II** and **III** have good unconditional coverage properties for any quantile. However, neither of the models passes the *VQR* test. This evidences that none of the first three models is able approximate the conditional quantile function well. More revealing information can be distilled from Figures 7, 8 and 9, which show the quantile forecasts against the realized returns over time for **Model I,II** and **III** respectively. Evidently, neither of the 3 models tracks the quantile of the conditional return distribution well. However, **Model I** and **II** show that when violations of  $R_{m,t+1} < \hat{Q}_\tau(R_{m,t+1})$  occur, the quantile forecast is usually not far off. In contrast, Figure 9 shows that violations of  $R_{m,t+1} < \hat{Q}_\tau(R_{m,t+1})$  are accompanied by aberrant quantile forecasts, in the sense that they grossly overestimate the conditional quantile function. From this perspective **Models I** and **II** are clearly preferred. **Model IV** performs poorly under our evaluation criteria, as can be seen from the bottom of Table 4.

Altogether, forward looking proxies based on option data in combination with quantile regression seem to render descent VaR forecasts over longer time horizons, although the *VQR* statistic shows that the forecasts are still not ideal. In the next Section I show how this can be interpreted as model free evidence against conditional lognormal models.

## 6.4 Prediction conditional lognormal environment

I now show that the risk-neutral quantile is, in a certain sense, the optimal predictor of VaR, if returns are conditionally lognormal. Consider the following discretized version of the Black-Scholes model, with a riskless asset that offers a certain return  $R_{f,t+1} = e^{r_f}$  and a risky asset with return  $R_{t+1} = \exp(\mu - \frac{1}{2}\sigma_t^2 + \sigma_t Z_{t+1})$ , where  $Z_{t+1}$  is standard normal and  $\sigma_t$  is the conditional ( $\mathcal{F}_t$ -measurable) volatility of returns. In this setup,  $M_{t+1} := \exp(-r_f - \xi_t^2/2 - \xi_t Z_{t+1})$ , 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

$$R_{t+1} = \exp(r_f - \frac{1}{2}\sigma_t^2 + \sigma_t Z_{t+1}).$$

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

Table 4: VaR prediction performance

	30 days			60 days			90 days		
Model I									
100 $\tau$	1	2.5	5	1	2.5	5	1	2.5	5
% Viol.	1.64	3.24	5.33	1.26	3.81	5.69	2.33	3.36	5.65
$LR_{uc}$	0.39	0.40	0.55	0.66	0.56	0.42	0.32	0.46	0.60
$VQR$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Model II									
100 $\tau$	1	2.5	5	1	2.5	5	1	2.5	5
% Viol.	1.64	3.15	5.25	1.39	3.64	6.11	1.82	3.63	5.21
$LR_{uc}$	0.61	0.85	0.88	0.47	0.76	0.65	0.52	0.78	0.92
$VQR$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Model III									
100 $\tau$	1	2.5	5	1	2.5	5	1	2.5	5
% Viol.	1.30	4.17	4.57	1.28	1.98	5.12	1.73	2.24	3.16
$LR_{uc}$	0.58	0.07	0.77	0.68	0.61	0.94	0.42	0.83	0.22
$VQR$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Model IV									
100 $\tau$	1	2.5	5	1	2.5	5	1	2.5	5
% Viol.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$LR_{uc}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$VQR$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Note: VaR evaluation using a rolling window of length 1,000. % Viol. denotes the percentage of violations of the predicted VaR at level  $\tau$ .  $LR_{uc}$  denotes the p-value of the unconditional coverage test, obtained from 10,000 bootstrap samples. Row  $VQR$  contains p-values of the VQR test.

**Theorem 6.1.** Consider the setup described above with return observations  $\{R_{t+1}\}_{t=1}^T$  stacked in the  $T \times 1$  vector  $R$ . Let  $\tilde{X}_t(\tau) := [1 \ \tilde{Q}_\tau(R_{t+1})]^\top$  and denote the  $T \times 2$  matrix of stacked  $\tilde{X}_t(\tau)$  by  $\tilde{X}(\tau)$ . Define the regression quantile  $\hat{\beta}(\tau; R, \tilde{X}(\tau))$  as the solution to the quantile regression with the risk-neutral quantile as a covariate

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

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

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

$$\hat{\beta}(\tau; R, X(\tau)) \in \arg \min_{\beta \in \mathbb{R}^2} \sum_{t=1}^T \rho_\tau \left( R_{t+1} - X_t(\tau)^\top \beta \right). \quad (6.3)$$

Then

$$\tilde{X}_{T+1}(\tau)^\top \hat{\beta}(\tau; R, \tilde{X}(\tau)) = X_{T+1}(\tau)^\top \hat{\beta}(\tau; R, X(\tau)). \quad (6.4)$$

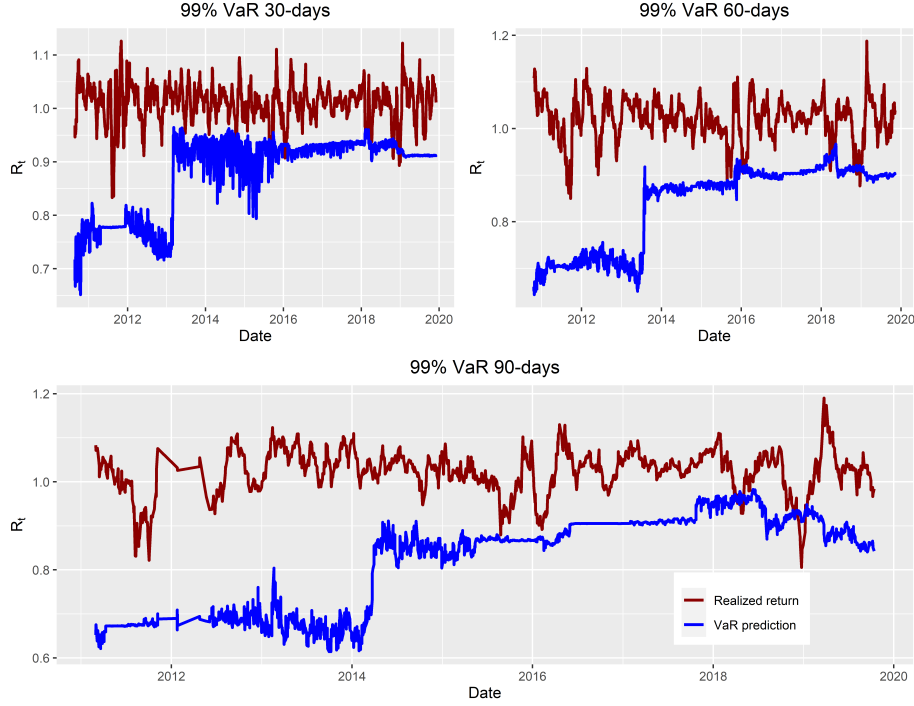


Figure 7: VaR prediction based on **Model I**. Red line denotes the realized  $n$ -day return, whereas the blue line depicts the corresponding VaR prediction.

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

*Proof.* See Appendix B.5. ■

The importance of Theorem 6.1 comes from the fact that  $\tilde{Q}_\tau(R_{t+1})$  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

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

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

$$\hat{\beta}_0(\tau) + \hat{\beta}_1(\tau)Q_\tau(R_{T+1}) = Q_\tau(R_{T+1}) + o_p(1). \quad (6.5)$$

Theorem 6.1 details that the predicted value on the left hand side of (6.5) is the same as the prediction obtained using quantile regression with observable  $\tilde{Q}_\tau(R_{t+1})$  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

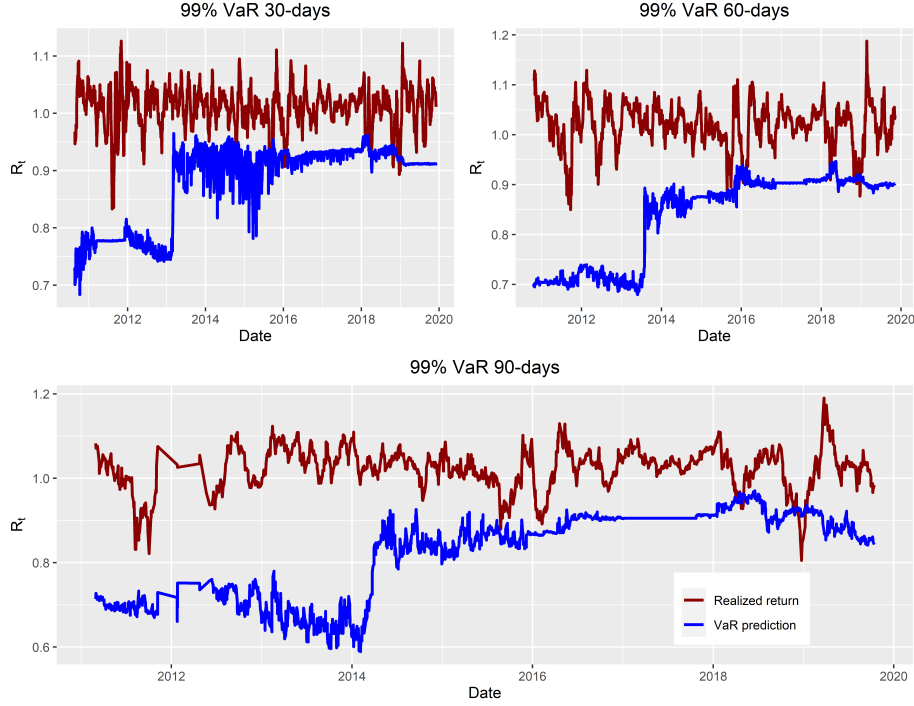


Figure 8: VaR prediction based on **Model II**. Red line denotes the realized  $n$ -day return, whereas the blue line depicts the corresponding VaR prediction.

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 6.1 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.

In fact, Theorem 6.1 can be used as model free evidence against the conditional lognormal assumption. Table 4 shows that the VQR test for the model including  $\tilde{Q}_\tau$  as the only regressor (**Model II**) decisively rejects  $H_0$ . Hence, the conditional lognormal assumption is not plausible since  $H_0$  in the VQR test posits that we are estimating with the true conditional quantile function. This is consistent with Martin (2017, Result 4), even though his conclusion is reached using different tools and is more general, since he also allows for time variation in the mean and risk-free rate.

## 7 Crash risk premium

The measure  $LRB_{t+1}(\tau)$  defined in Equation (5.8) approximates  $\tau - \mathbb{P}_t(R_{m,t+1} \leq \tilde{Q}_\tau)$ . This can also be interpreted as a crash risk premium, which I formally



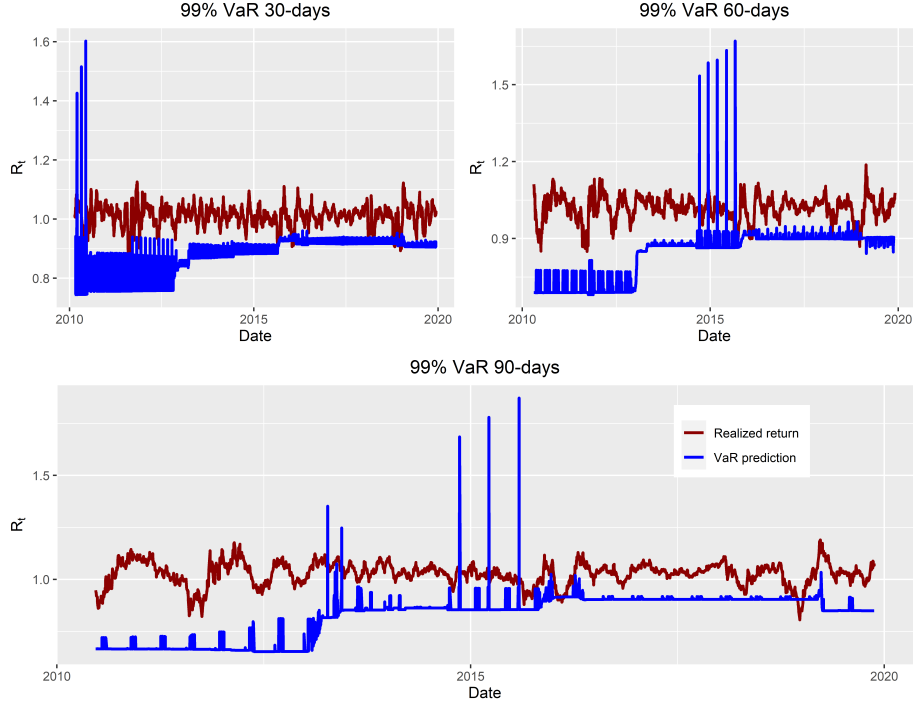


Figure 9: CAViaR prediction based on **Model III**. Red line denotes the realized  $n$ -day return, whereas the blue line depicts the corresponding VaR prediction.

define as

$$CRP_{t+1} := \frac{1}{T} \left( \mathbb{E}_t [\mathbb{1} (R_{m,t+1} \leq \alpha)] - \tilde{\mathbb{E}}_t [\mathbb{1} (R_{m,t+1} \leq \alpha)] \right). \quad (7.1)$$

Here,  $T$ , is a scaling factor which converts the premium to annual units. For example, if the 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 (7.1) 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  $\tilde{Q}_\tau = \alpha$ , so that  $\tau = \tilde{\mathbb{P}}_t(\alpha)$ , I get

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

By construction, this proxy is completely forward looking and is not prone to 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. I prove this in [Appendix E.2](#). [Figure 10](#) shows the crash risk premium for  $\alpha \in \{0.95, 0.9, 0.85\}$ . Consistent with the theory in [Lemma 5.4](#), the premium is everywhere negative. 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.<sup>17</sup> Hence my forward looking crash premium measure aligns well with periods that are associated with more uncertainty.<sup>18</sup> Also of interest is that the premium for 30 days is more negative than the 60 day premium, signaling that investors are willing to pay relatively more for insurance against a crash over shorter time horizons.

The benefit of this approach is that the crash risk premium estimate is forward looking, whereas many papers rely on estimates coming from historical data. Some work has been done in this direction, for example [Backus et al. \(2011\)](#) use equity index options to analyze consumption growth disasters and [Martin \(2017\)](#) derives the forward looking crash probability perceived by a representative agent with log preferences (see also Section E.2). However, neither of these papers analyze the premium that is associated to crash risk. More closely related is [Bates \(1991\)](#), who calibrates a jump-diffusion model from option prices and notes that the expected number of negative jumps spikes prior to Black Monday, thus lending support for the hypothesis that a crash was expected. [Bollerslev and Todorov \(2011\)](#) is also similar in spirit, as they document evidence for the variance risk premium by translating ex-post measures into forward-looking  $\mathbb{P}_t$  counterparts. [Bollerslev and Todorov \(2011\)](#) find that the variance risk premium has a notable spike downwards around the same period a spike in Figure 10 is observed. Their findings add to the variance risk premium puzzle by documenting significant fluctuation in the premium, which is consistently negative.

My estimates are akin to this finding. First, the crash risk premium is consistently negative. This is to be expected, since the physical probability measure stochastically dominates the risk-neutral measure.<sup>19</sup> Second, and more puzzling, is the significant fluctuation observed in the premium, especially during crises. This fluctuation cannot be reconciled with representative agent models that posit i.i.d. shocks, such as the disaster risk model, as the conditional premium is constant. The LRR model of [Bansal et al. \(2012\)](#) does produce time variation in  $CRP_{t+1}$ , but the premium is *negatively* correlated with consumption growth and the market return.<sup>20</sup> This seemingly contradicts the data in Figure 10, even though definitive evidence would require formal statistical testing. I leave this for future work.

## 8 Conclusion

This paper proposes a new 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  $\tau$ -quantile level. I show that this bound compares favorably

<sup>17</sup>See <https://www.thebalance.com/stock-market-crash-of-2008-3305535#:~:text=The%20stock%20market%20crash%20of%202008%20occurred%20on%20Sept.,largest%20point%20drop%20in%20history>. Recall also that the COVID pandemic is not in our sample.

<sup>18</sup>These estimates are conservative, since from the proof of Corollary 5.5  $\tau - \mathbb{P}_t(\tilde{Q}_\tau) \geq LRB_{t+1}(\tau)$ .

<sup>19</sup>This is may not be true if the risk-neutral measure is projected on a hedging asset, like gold.

<sup>20</sup>Unreported simulations (based on monthly data) show that the correlation between consumption growth and  $CRP_{t+1}(-0.1)$  is about  $-0.06$ , whereas the correlation between  $CRP_{t+1}(-0.1)$  and  $\log R_{m,t+1}$  is  $-0.04$ .

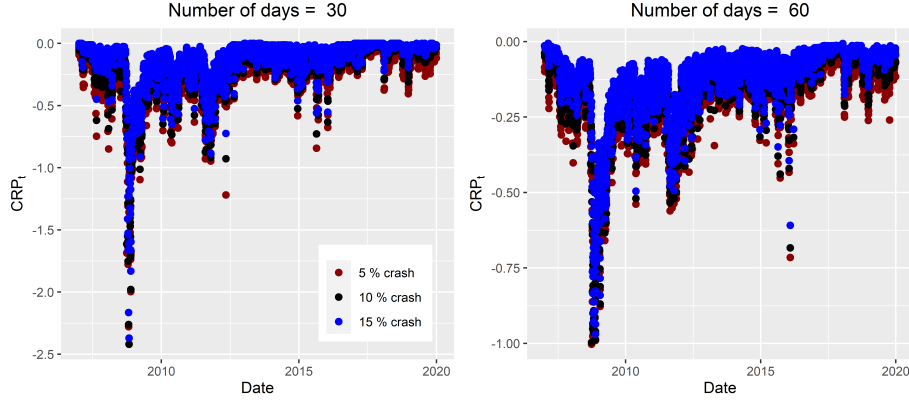


Figure 10: Plots of the crash risk premium  $CRP_{t+1}$ , where  $\alpha \in \{0.95, 0.9, 85\}$ .

to the HJ bound in scenarios where return data are heavy tailed. Using a combination of historical returns and option price data, Section 4 shows that the quantile bound is statistically stronger than the HJ bound. Two implications for asset pricing models are immediate: (i) This is counter evidence to CAPM that the SDF is a linear combination of the market return, since the HJ bound binds in that case (ii) The classical equity premium puzzle exacerbates: the HJ bound implies a (annualized) risk-aversion coefficient of 38, whereas estimates based on the quantile bound suggest a (annualized) risk-aversion coefficient of 65. Moreover, the LRR model fails to reconcile this feature of the data under typical calibration. The disaster risk model is able to replicate this empirical observation, but falls short of replicating other features of the quantile bound, due to the implied fat tail in the risk-neutral distribution.

A critical ingredient in my analysis is the risk-neutral quantile function, which is shown to be useful in other contexts as well. A Taylor expansion is used to shed light on the difference between the risk-neutral and physical quantile function. Theoretical results show that the first order Taylor expansion provides a good approximation to the physical quantile function and quantile regression estimates confirm that theoretical predictions line up well with empirical estimates. Two applications serve to illustrate the usefulness of this approach. First, I estimate a forward looking premium on crash risk and document that it dropped significantly during the height of the financial crisis. Secondly, I use the risk-neutral quantile function as explanatory variable in a quantile regression on the physical quantile function. This produces estimates of long term VaR that are more accurate than a vanilla CAViaR or GARCH type model. Finally, I identify an environment under which the risk-neutral quantile function is the optimal predictor, namely if returns are conditionally lognormal. In my empirical application I find that the risk-neutral quantile is not a good predictor of the physical quantile, as the null hypothesis for the  $VQR$  test is rejected. This serves as model-free evidence against the conditional lognormal assumption of the return distribution.

## A Formulas for market moments

I present 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 \(2020, Appendix B\)](#).

**Proposition A.1.** *The higher order risk-neutral moments can be computed directly from the estimated quantile functions*

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

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

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

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

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

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

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

This follows straightforward from the substitution  $x = Q(\tau)$ . Now use that for any constant  $a \in \mathbb{R}$ ,  $Q_{X-a}(\tau) = Q_X(\tau) - a$  to derive (A.1). The truncated formula follows similarly. ■

## B Proofs

### B.1 Proof of Theorem 2.1

*Proof.* I suppress the dependence of the  $\tau$ -quantile on  $R_{t+1}$  and write  $\tilde{Q}_\tau := \tilde{Q}_\tau(R_{t+1})$ . Start from the definition of a risk-neutral quantile

$$\begin{aligned} \tau &= \tilde{\mathbb{P}}_t [R_{t+1} \leq \tilde{Q}_\tau] = \tilde{\mathbb{E}}_t [\mathbb{1}(R_{t+1} \leq \tilde{Q}_\tau)] = R_{f,t+1} \mathbb{E}_t [M_{t+1} \mathbb{1}(R_{t+1} \leq \tilde{Q}_\tau)] \\ &= R_{f,t+1} [\text{COV}_t(M_{t+1}, \mathbb{1}(R_{t+1} \leq \tilde{Q}_\tau)) + \mathbb{E}_t[M_{t+1}] \mathbb{E}_t[\mathbb{1}(R_{t+1} \leq \tilde{Q}_\tau)]] \\ &= R_{f,t+1} \underbrace{\text{COV}_t(M_{t+1}, \mathbb{1}(R_{t+1} \leq \tilde{Q}_\tau)) + \mathbb{E}_t[\mathbb{1}(R_{t+1} \leq \tilde{Q}_\tau)]}_{=\mathbb{P}_t(R_{t+1} \leq \tilde{Q}_\tau)}. \end{aligned} \quad (\text{B.1})$$

Rearranging then yields

$$\frac{\tau - \mathbb{P}_t(R_{t+1} \leq \tilde{Q}_\tau)}{R_{f,t+1}} = \text{COV}_t(M_{t+1}, \mathbb{1}(R_{t+1} \leq \tilde{Q}_\tau)).$$

Using Cauchy-Schwarz renders the inequality

$$\begin{aligned} \frac{\left| \tau - \mathbb{P}_t(R_{t+1} \leq \tilde{Q}_\tau) \right|}{R_{f,t+1}} &\leq \sigma_t(M_{t+1}) \sigma_t \left( \mathbb{1} \left( R_{t+1} \leq \tilde{Q}_\tau \right) \right) \\ \frac{\left| \tau - \mathbb{P}_t(R_{t+1} \leq \tilde{Q}_\tau) \right|}{\sigma_t \left( \mathbb{1} \left( R_{t+1} \leq \tilde{Q}_\tau \right) \right) R_{f,t+1}} &\leq \sigma_t(M_{t+1}). \end{aligned} \quad (\text{B.2})$$

■

## B.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

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

Both  $Z_R$  and  $Z_M$  are standard normal random variables with correlation  $\rho$ . 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

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

Again by Stein's Lemma

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

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

$$\left| \frac{\mathbb{E}[R_{t+1}] - e^{\lambda r_f}}{\tau - \mathbb{P}(R_{t+1} \leq \tilde{Q}_\tau)} \right| \approx \frac{\sigma_R \sqrt{\lambda} e^{\mu_R \lambda}}{f(x)}. \quad (\text{B.3})$$

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

$$\begin{aligned} \frac{\text{HJ bound}}{\text{Quantile bound}} &= \frac{\frac{|\mathbb{E}[R_{t+1}] - R_{f,t+1}|}{\sigma(R_{t+1})R_{f,t+1}}}{\frac{|\tau - \mathbb{P}(R_{t+1} \leq \tilde{Q}_\tau)|}{\sqrt{\mathbb{P}(R_{t+1} \leq \tilde{Q}_\tau)(1 - \mathbb{P}(R_{t+1} \leq \tilde{Q}_\tau))}R_{f,t+1}}} \\ &\stackrel{\text{(B.3)}}{\approx} \frac{\sqrt{\mathbb{P}(r_{t+1} \leq x)(1 - \mathbb{P}(r_{t+1} \leq x))}}{\sigma(R_{t+1})} \times \frac{\sigma_R \sqrt{\lambda} e^{\mu_R \lambda}}{f(x)}. \quad (\text{B.4}) \end{aligned}$$

Here,  $r_{t+1} = \log R_{t+1}$  and  $x = \log \tilde{Q}_\tau$ . Using the same reasoning as in Example 2.2, the expression on the RHS of (B.4) is minimized by choosing  $x = \log \tilde{Q}_\tau^*$  s.t.  $\mathbb{P}(R_{t+1} \leq \tilde{Q}_\tau^*) = 1/2$ . In that case the relative efficiency equals

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

### B.3 Proof of Theorem 3.3

Theorem 3.3 follows immediately from the continuous mapping theorem and Lemma B.2 below. To prove Lemma B.2, I use the following result.

**Lemma B.1.** *Let assumptions 3.1–3.2 hold. Then, under Skorokhod construction, there is a standard Brownian bridge  $\mathbb{G}(\tau)$ , such that, as  $T \rightarrow \infty$*

$$\sup_{\varepsilon \leq \tau \leq 1-\varepsilon} \left| \sqrt{T} \left[ \mathbb{P}_T(\tilde{Q}(\tau)) - \mathbb{P}(\tilde{Q}(\tau)) \right] - \mathbb{G} \left( \mathbb{P} \left[ \tilde{Q}(\tau) \right] \right) \right| \rightarrow 0 \quad \text{almost surely.}$$

*Proof.* Assumption 3.1 means the classical Donsker theorem (Van der Vaart, 2000, Theorem 19.3) can be applied, which states

$$\sqrt{T} [\mathbb{P}_T(\tau) - \mathbb{P}(\tau)] \rightsquigarrow \mathbb{G} \circ \mathbb{P}(\tau).$$

By Skorokhod's representation theorem (Wichura, 1970, Theorem 1) there exists a suitable probability space so that almost sure convergence holds

$$\sup_{\varepsilon \leq \tau \leq 1-\varepsilon} \left| \sqrt{T} (\mathbb{P}_T(\tau) - \mathbb{P}(\tau)) - \mathbb{G}(\mathbb{P}(\tau)) \right| \rightarrow 0 \quad \text{almost surely.}$$

The result now follows from the time change  $\tau \rightarrow \tilde{Q}(\tau)$ . ■

**Lemma B.2.** *The following weak convergence holds for  $\tau \in [\varepsilon, 1 - \varepsilon]$*

$$\begin{aligned} \sqrt{T} \left[ \frac{\tau - \mathbb{P}_T(\tilde{Q}(\tau))}{\sqrt{\mathbb{P}_T(\tilde{Q}(\tau))(1 - \mathbb{P}_T(\tilde{Q}(\tau)))}} - \frac{\tau - \mathbb{P}(\tilde{Q}(\tau))}{\sqrt{\mathbb{P}(\tilde{Q}(\tau))(1 - \mathbb{P}(\tilde{Q}(\tau)))}} \right] \\ \rightsquigarrow \frac{\mathbb{G}(\mathbb{P} \circ \tilde{Q}(\tau))}{\sqrt{\mathbb{P}(\tilde{Q}(\tau))(1 - \mathbb{P}(\tilde{Q}(\tau)))}} \stackrel{d}{=} \frac{1}{\sqrt{t}} \mathbb{W}(t), \\ \text{where } t \in \left[ \frac{\mathbb{P}(\tilde{Q}(\varepsilon))}{1 - \mathbb{P}(\tilde{Q}(\varepsilon))}, \frac{\mathbb{P}(\tilde{Q}(1 - \varepsilon))}{1 - \mathbb{P}(\tilde{Q}(1 - \varepsilon))} \right]. \quad (\text{B.5}) \end{aligned}$$

Here  $\mathbb{W}(t)$  is a standard Wiener process.

*Proof.* Notice that the denominator terms are bounded away from zero on  $\tau \in [\varepsilon, 1 - \varepsilon]$  since  $\mathbb{P}$  and  $\tilde{\mathbb{P}}$  are equivalent measures. From Lemma B.1 and the Slutsky theorem in  $D[\varepsilon, 1 - \varepsilon]$

$$\begin{aligned} \sqrt{T} \left[ \frac{\tau - \mathbb{P}_T(\tilde{Q}(\tau))}{\sqrt{\mathbb{P}_T(\tilde{Q}(\tau))(1 - \mathbb{P}_T(\tilde{Q}(\tau)))}} - \frac{\tau - \mathbb{P}(\tilde{Q}(\tau))}{\sqrt{\mathbb{P}_T(\tilde{Q}(\tau))(1 - \mathbb{P}_T(\tilde{Q}(\tau)))}} \right] \\ \rightsquigarrow \frac{\mathbb{G}(\mathbb{P} \circ \tilde{Q}(\tau))}{\sqrt{\mathbb{P}(\tilde{Q}(\tau))(1 - \mathbb{P}(\tilde{Q}(\tau)))}}. \end{aligned}$$

Under Skorokhod construction

$$\begin{aligned} \sup_{\varepsilon \leq \tau \leq 1 - \varepsilon} \left| \sqrt{T} \left[ \frac{\tau - \mathbb{P}_T(\tilde{Q}(\tau))}{\sqrt{\mathbb{P}_T(\tilde{Q}(\tau))(1 - \mathbb{P}_T(\tilde{Q}(\tau)))}} - \frac{\tau - \mathbb{P}(\tilde{Q}(\tau))}{\sqrt{\mathbb{P}_T(\tilde{Q}(\tau))(1 - \mathbb{P}_T(\tilde{Q}(\tau)))}} \right] \right. \\ \left. - \frac{\mathbb{G}(\mathbb{P} \circ \tilde{Q}(\tau))}{\sqrt{\mathbb{P}(\tilde{Q}(\tau))(1 - \mathbb{P}(\tilde{Q}(\tau)))}} \right| \rightarrow 0 \quad \text{almost surely.} \end{aligned}$$

The first part of the Lemma is proved provided

$$\sup_{\varepsilon \leq \tau \leq 1 - \varepsilon} \sqrt{T} \left| \frac{\tau - \mathbb{P}(\tilde{Q}(\tau))}{\sqrt{\mathbb{P}_T(\tilde{Q}(\tau))(1 - \mathbb{P}_T(\tilde{Q}(\tau)))}} - \frac{\tau - \mathbb{P}(\tilde{Q}(\tau))}{\sqrt{\mathbb{P}(\tilde{Q}(\tau))(1 - \mathbb{P}(\tilde{Q}(\tau)))}} \right| \rightarrow 0 \quad \text{almost surely}$$

The latter statement follows from the Glivenko-Cantelli theorem (Van der Vaart, 2000, Theorem 19.1), which states that one can find  $\Omega_0 \subseteq \Omega$  with  $\mathbb{P}(\Omega_0) = 1$ , s.t. for all  $\omega \in \Omega_0$

$$\sup_{\varepsilon \leq \tau \leq 1 - \varepsilon} \left| \mathbb{P}_T(\tilde{Q}(\tau)) - \mathbb{P}(\tilde{Q}(\tau)) \right| \rightarrow 0.$$

Since  $\tau$  is restricted to a compact set,  $\mathbb{P}(\tilde{Q}(\tau)) \in [a, b] \subseteq [0, 1]$  with  $0 < a < b < 1$ . The function

$$g(\tau) := \frac{1}{\sqrt{\tau(1 - \tau)}}, \quad \tau \in [a, b]$$

is continuous. Composing a uniformly convergent sequence with a continuous function on a compact set preserves uniform convergence. Hence, for all  $\omega \in \Omega_0$

$$\sup_{\varepsilon \leq \tau \leq 1 - \varepsilon} \sqrt{T} \left| \frac{1}{\sqrt{\mathbb{P}_T(\tilde{Q}(\tau))(1 - \mathbb{P}_T(\tilde{Q}(\tau)))}} - \frac{1}{\sqrt{\mathbb{P}(\tilde{Q}(\tau))(1 - \mathbb{P}(\tilde{Q}(\tau)))}} \right| \rightarrow 0.$$

The distributional equality in the end of (B.5) follows from the following well known properties of Brownian motion. First, recall that if  $\mathbb{G}(t)$  is a Brownian bridge, then, from the defining properties of Brownian motion

$$\mathbb{W}(t) := (t + 1)\mathbb{G}\left(\frac{t}{t + 1}\right).$$

Here,  $\mathbb{W}(\cdot)$  is a standard Brownian motion. Substituting  $x = t/(1+t)$  gives

$$\frac{\mathbb{W}\left(\frac{x}{1-x}\right)}{\sqrt{\frac{x}{1-x}}} = \frac{\mathbb{G}(x)}{\sqrt{x(1-x)}}.$$

The final result comes from the substitution  $x \rightarrow \mathbb{P}(\tilde{Q}(\tau))$ . ■

#### B.4 Proof of Lemma 5.4 and Corollary 5.5

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

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

$$\Lambda := \underbrace{\left( \mathbb{1} \left( X_1 \leq \tilde{Q}_\tau \right) - \mathbb{1} \left( X_2 \leq \tilde{Q}_\tau \right) \right)}_{=I} \underbrace{\left( (X_1 - R_{f,t+1})^k - (X_2 - R_{f,t+1})^k \right)}_{=II}. \quad (\text{B.6})$$

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 (B.6), using the non-positivity of  $\Lambda$  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

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

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

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

$$\tilde{\mathbb{E}}_t \left[ \left( \mathbb{1} \left( R_{m,t+1} \leq \tilde{Q}_\tau \right) - \tau \right) (R_{m,t+1} - R_{f,t+1})^k \right]. \quad (\text{B.7})$$



Consider the associated function

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

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

$$\begin{aligned}\frac{\partial}{\partial \tau} \Gamma(\tau) &= (\tilde{Q}_\tau - R_{f,t+1})^k \tilde{f}_{R_{m,t+1}}(\tilde{Q}_\tau) \frac{\partial \tilde{Q}_\tau}{\partial \tau} - \mathbb{E}_t [(R_{m,t+1} - R_{f,t+1})^k] \\ &= (\tilde{Q}_\tau - R_{f,t+1})^k - \mathbb{E}_t [(R_{m,t+1} - R_{f,t+1})^k], \quad (\text{B.8})\end{aligned}$$

since, by the rules for derivatives of inverses

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

Because I assume that  $\sup_k \|R_{m,t+1}\|_k < \infty$ , it follows that (B.8) is positive for all  $\tau \in [0, \tau^*]$ , where  $\tau^*$  solves

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

In conclusion, I have shown that the covariance (B.7) vanishes when  $\tau \rightarrow 0^+$  and the slope of (B.7) is positive for all  $\tau \leq \tau^*$ . This means that (B.7) is positive for all  $\tau \leq \tau^*$ . Thus for all such  $\tau \in (0, \tau^*]$

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

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

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

*Part 3: Combining both cases*

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

$$\begin{aligned}Q_\tau - \tilde{Q}_\tau &\approx \frac{1}{f_t(\tilde{Q}_\tau)} \left( \frac{\sum_{k=1}^{\infty} \theta_k \left( \tau \tilde{\mathbb{M}}_{t+1}^{(k)} - \tilde{\mathbb{M}}_{t+1}^{(k)}[\tilde{Q}_\tau] \right)}{1 + \sum_{k=1}^{\infty} \theta_k \tilde{\mathbb{M}}_{t+1}^{(k)}} \right) \\ &\geq \frac{1}{f_t(\tilde{Q}_\tau)} \left( \frac{\sum_{k=1}^3 \theta_k \left( \tau \tilde{\mathbb{M}}_{t+1}^{(k)} - \tilde{\mathbb{M}}_{t+1}^{(k)}[\tilde{Q}_\tau] \right)}{1 + \sum_{k=1}^3 \theta_k \tilde{\mathbb{M}}_{t+1}^{(k)}} \right). \quad (\text{B.9})\end{aligned}$$

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 (B.9) gives

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

■

## B.5 Proof of Theorem 6.1

*Proof.* By definition

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

Similarly

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

As a result

$$e^{(\mu-r_f)} \tilde{Q}_\tau = Q_\tau. \quad (\text{B.10})$$

Recall that the quantile regression estimator is equivariant to reparametrization of design: for any  $2 \times 2$  nonsingular matrix  $A$

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

By Equation (B.10)

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

Therefore

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

Hence, the predicted quantile using the physical quantile regression (6.3) equals

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

This is exactly (6.4). ■

## C Simulation results for quantile approximation

In this section I assess the accuracy of the physical quantile approximation developed in Section 5, as well as the use of  $LRB_{t+1}(\tau)$  as a predictor variable for VaR.

### C.1 Quantile function approximation

First, I show that the first order approximation (5.2) to the physical quantile function is tight in a typical Black-Scholes model. In this case, the stock price  $S_t$  follows a stochastic differential equation

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t^{\mathbb{P}}.$$

$W_t^{\mathbb{P}}$  is a standard Brownian motion under physical measure,  $\mu$  is the mean return and  $\sigma$  is the volatility. Under risk neutral measure the dynamics of the stock price are given by

$$\frac{dS_t}{S_t} = r dt + \sigma dW_t^{\tilde{\mathbb{P}}},$$

where  $W_t^{\mathbb{P}}$  denotes a standard Brownian motion under risk neutral measure and  $r$  is the risk-free rate. In this case, I can compute the physical CDF and PDF explicitly, as they are given as the respective CDF and PDF of a lognormal distribution with mean  $(\mu - \frac{1}{2}\sigma^2)\lambda$  and variance  $\sigma^2\lambda$ , where  $\lambda$  is the time to maturity. The same is true for the risk-neutral distribution, however in that case the mean is  $(r - \frac{1}{2}\sigma^2)\lambda$ . Figure 11 shows how a first order correction to the risk-neutral quantile function leads to a fairly accurate approximation of the true quantile function. Unreported simulations affirm this for a large variety of different parameter settings, as well as for the alternative option pricing model of Heston (1993).<sup>21</sup>

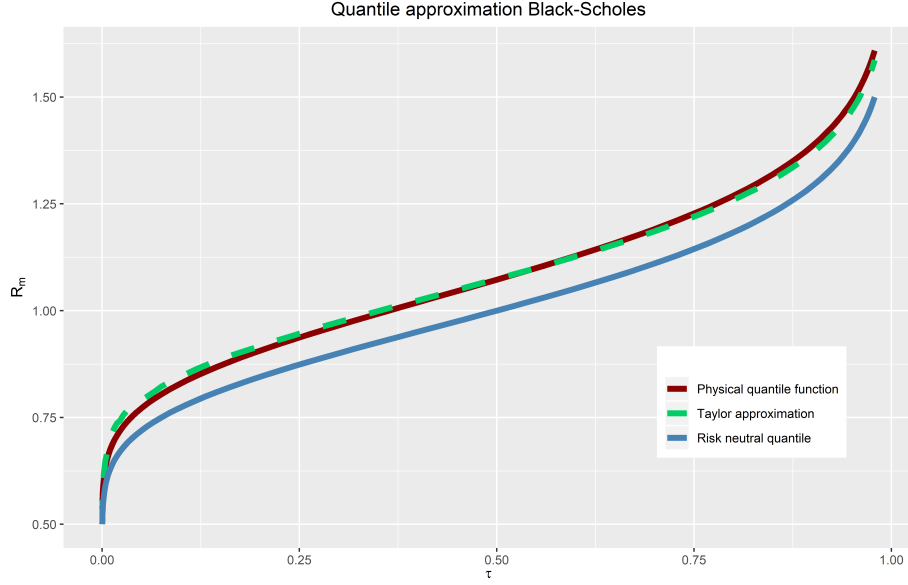


Figure 11: Black-Scholes model with  $\mu = 0.09$ ,  $r = 0.02$ ,  $\sigma = 0.2$  and  $\lambda = 1$ . The red line denotes the physical quantile function, whereas the blue line denotes the risk neutral quantile function. The green line is the quantile function obtained by using the first order correction in the Taylor expansion of the physical quantile.

## C.2 Covariance approximation and quantile regression

This section shows that quantile regression is an appropriate tool for VaR prediction in a Black-Scholes environment. In addition, I establish that term II in the Taylor approximation below can be estimated accurately given market data.

$$Q_\tau \approx \underbrace{\tilde{Q}_\tau}_I + \underbrace{\frac{\tau - \mathbb{P}_t(R_{t+1} \leq \tilde{Q}_\tau)}{f_t(\tilde{Q}_\tau)}}_{II} \quad (\text{C.1})$$

<sup>21</sup>The code for the simulation results is available from the author upon request.

Recall from the previous section that the physical returns in a Black-Scholes environment are distributed according to

$$R_{t+1} \sim \mathbf{Lognormal} \left( \left( \mu - \frac{1}{2} \sigma^2 \right) \lambda, \sigma \sqrt{\lambda} \right), \quad \lambda = T - t.$$

In order to create time variation in the distribution of returns I randomly generate volatility each period according to

$$\sigma_t = \sigma + U, \quad U \sim \mathbf{UNIF}[\underline{\sigma}, \bar{\sigma}], \quad \underline{\sigma} \geq -\sigma.$$

This implies  $R_{t+1} \stackrel{i.i.d.}{\sim} \mathbf{Lognormal}([\mu - \frac{1}{2} \sigma_t^2] \lambda, \sigma_t \sqrt{\lambda})$ . I approximate the numerator of term II in (C.1) by our measure from Corollary 5.5

$$\tau - \mathbb{P}_t(R_{t+1} \leq \tilde{Q}_\tau) \approx \underbrace{\frac{\sum_{k=1}^3 \frac{(-1)^{k+1}}{R_{f,t+1}^k} \left( \tau \tilde{\mathbb{M}}_{t+1}^{(k)} - \tilde{\mathbb{M}}_{t+1}^{(k)}[\tilde{Q}_\tau] \right)}{1 + \sum_{k=1}^3 \frac{(-1)^{k+1}}{R_{f,t+1}^k} \tilde{\mathbb{M}}_{t+1}^{(k)}}}_{LRB_{t+1}(\tau)}. \quad (\text{C.2})$$

To gauge the accuracy of (C.2) I rearrange the above and solve for  $\mathbb{P}_t(R_{t+1} \leq \tilde{Q}_\tau)$ , since  $\mathbb{P}_t(R_{t+1} \leq \tilde{Q}_\tau)$  and  $\tilde{Q}_\tau$  are explicitly known in the Black-Scholes model. The upper left panel of Figure 12 reveals that the approximation is quite accurate. This motivates the approximation of  $f_t(\tilde{Q}_\tau)$  by the numerical derivative of  $LRB_{t+1}(\tau)$  (seen as a function of the risk-neutral quantile). The result is shown in the upper right panel of Figure 12. Clearly, the approximation is rather accurate as well. Finally, I use the two approximations to approximate the ratio  $\frac{\tau - \mathbb{P}_t(\tilde{Q}_\tau)}{f_t(\tilde{Q}_\tau)}$ . The result is shown in the bottom panel of Figure 12. Unsurprisingly, the approximation is rather close and only the right-end tail of the distribution seems to be (slightly) overestimated. Unreported simulations affirm this for a large variety of different parameter settings. Additionally, I find that the approximation is superior to approximating the physical distribution by the risk-neutral distribution.<sup>22</sup>

The foregoing discussion suggests that we can obtain a market observable proxy of  $\frac{\tau - \mathbb{P}_t(\tilde{Q}_\tau)}{f_t(\tilde{Q}_\tau)}$ , given by

$$\Xi_{t+1}(\tau) := \frac{\tau - LRB_{t+1}(\tau)}{LRB_{t+1}(\tau)'},$$

where  $LRB_{t+1}(\tau)'$  denotes the (numerical) derivative of  $LRB_{t+1}(\tau)$  seen as a function of  $\tilde{Q}_\tau$ .<sup>23</sup> Since the first order Taylor approximation in (C.1) was shown to be tight in Figure 11, a natural candidate to estimate the physical quantile would be a quantile regression of the form

$$Q_\tau(R_{t+1}) = \beta_0(\tau) + \beta_1(\tau) \left( \Xi_{t+1}(\tau) + \tilde{Q}_\tau \right). \quad (\text{C.3})$$

Below is a succinct description of the simulation steps, with parameters chosen to mimic the empirical application.

<sup>22</sup>The risk-neutral distribution would be a good proxy for the physical distribution if investors are approximately risk-neutral. This case serves as a useful benchmark.

<sup>23</sup>Note that  $\Xi_{t+1}(\tau)$  is  $\mathcal{F}_t$ -measurable (known at time  $t$ ).

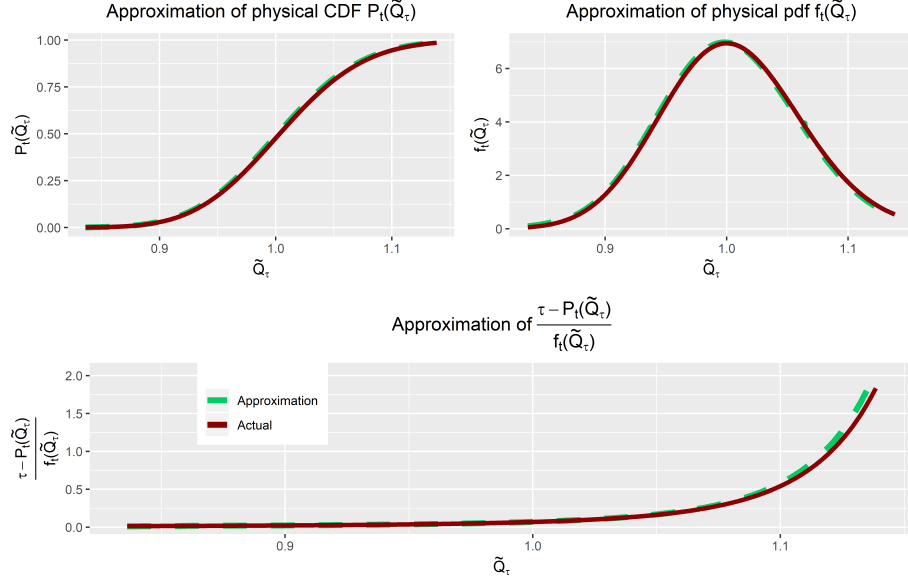


Figure 12: Approximation of physical pdf, cdf and first order correction term II in (C.1) in Black-Scholes environment with  $\mu = 0.06, r = 0, \sigma = 0.2$  and  $\lambda = \frac{30}{365}$ . The green dashed line denotes the approximation, whereas the solid red line depicts the actual quantity. The left upper panel denotes the approximation of  $\mathbb{P}_t(\tilde{Q}_\tau)$ , the upper right panel denotes the approximation of  $f_t(\tilde{Q}_\tau)$ . The bottom panel considers the approximation to  $\frac{\tau - \mathbb{P}_t(\tilde{Q}_\tau)}{f_t(\tilde{Q}_\tau)}$ .

**Step 1:** Initialize the parameters:  $\mu = 0.06, r = 0, \sigma = 0.2, \underline{\sigma} = -0.1, \bar{\sigma} = 0.1, \lambda = 30/365$ . The sample size  $n$  equals  $n = 3,000$ . Finally, a rolling window of size  $w = 1,000$  is used to allow for time variation in the regression estimates.<sup>24</sup>

**Step 2:** Simulate 3,000 returns  $R_{t+1} \stackrel{i.i.d.}{\sim} \text{Lognormal}([\mu - \frac{1}{2}\sigma_t^2]\lambda, \sigma_t\sqrt{\lambda})$  with

$$\sigma_t = \sigma + U_t, \quad U_t \stackrel{i.i.d.}{\sim} \text{Unif}(\underline{\sigma}, \bar{\sigma}).$$

**Step 3:** Compute the (forward looking) quantities  $\tilde{Q}_\tau$  and  $\Xi_{t+1}(\tau)$  using the observable Black-Scholes call and put option prices.

**Step 4:** Estimate the coefficients  $\beta_0(\tau), \beta_1(\tau)$  using the quantile regression specification (C.3) using a rolling window of size  $w = 1,000$ .

As an additional check I repeat the same steps above but only use the risk-neutral quantile  $\tilde{Q}_\tau$  in the quantile regression (C.3). This is referred to as model I. Model II concerns the full model described above with  $\Xi_{t+1}(\tau)$  as an additional regressor. The simulation results are shown in Table 5 for several quantiles.<sup>25</sup> It is apparent that both models perform well for the various probability levels. Model II performs slightly worse than the benchmark model I in

<sup>24</sup>In this setup a rolling window does not make much sense since returns are *i.i.d.*, but I do this to mimic the empirical application.

<sup>25</sup>I do not need a bootstrap approach to obtain critical values, since in this simulation I assume that data are *i.i.d.*. This means that the results of [Gaglianone et al. \(2011\)](#) and [Christoffersen \(1998\)](#) apply asymptotically.

terms of VaR violations for  $\tau = 0.01$ . However, the *VQR* test of [Gaglianone et al. \(2011\)](#) illustrates that the conditional quantile estimates obtained using Model II are more accurate than those obtained using model I. In particular,  $H_0$  is rejected using a 5% significance level, whereas I cannot reject  $H_0$  under model II. The models both pass the (un)conditional coverage tests of [Christoffersen \(1998\)](#). Overall, the results are promising and underscore the potential gains of using forward looking regressors in predicting VaR. Also notice that in this (albeit contrived) *i.i.d.* setup, backward looking models like CAViaR would be completely useless.

Table 5: Predictive performance VaR

	Model	% Viol.	$LR_{uc}$	$LR_{cc}$	$VQR$
$\tau = 0.01$	I	1.05	0.05	0.50	<b>7.53</b>
	II	1.20	0.76	1.34	1.30
$\tau = 0.025$	I	2.65	0.18	0.32	1.89
	II	2.65	0.18	0.32	0.00
$\tau = 0.05$	I	5.25	0.26	0.31	1.59
	II	5.05	0.01	0.01	0.48

Note: Predictive ability of 2 different VaR prediction models. Model I uses the risk-neutral quantile  $\hat{Q}_\tau$  as instrument in the quantile regression, so  $\beta_1(\tau) = 0$  in (C.3). Model II uses the full specification of (C.3). Results are obtained using a rolling window of length 1,000. % Viol. denotes the number of VaR violations for a given probability level  $\tau$ .  $LR_{uc}$  is the likelihood ratio test statistic of unconditional coverage.  $LR_{cc}$  is the likelihood ratio test statistic of conditional coverage and independence (see [Christoffersen \(1998\)](#)).  $VQR$  denotes the test statistic of the *VQR* test. Test statistics in boldface reject the null hypothesis at 5% significance.

## D 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. Section D.2 explains the bootstrap approach to calculate critical values of the unconditional coverage test of [Christoffersen \(1998\)](#).

### D.1 Simulation results

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

$$R_{t+1} = \exp\left(\left(\mu - \frac{1}{2}\sigma_t^2\right)\lambda + \sigma_t\sqrt{\lambda}Z\right), \quad Z \sim N(0, 1).$$

I generate  $\sigma_t$  independent of  $Z$  according to

$$\sigma_t \sim \text{UNIF}[0.06, 0.26].$$

It is assumed that  $\sigma_t$  is known at the start of period  $t$ , so that the distribution of  $R_{t+1}$  is given by

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

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

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

The conditional SDF is given by

$$M_{t+1} = \exp\left(-\left(r + \frac{1}{2}\xi_t^2\right)t - \xi_t\sqrt{t}Z\right), \quad Z \sim N(0, 1).$$

Here,  $\xi_t$  is the conditional Sharpe ratio

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

To mimic the empirical application, I use  $t = 30/365$  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

$$\begin{aligned} \text{call}(S_0, t) &= \Phi(d_1)S_0 - \Phi(d_2)Ke^{-rt} \\ d_1 &= \frac{1}{\sigma_t\sqrt{t}}\left[\ln\left(\frac{S_0}{K}\right) + \left(r - \frac{\sigma_t^2}{2}\right)t\right] \\ d_2 &= d_1 - \sigma_t\sqrt{t}. \end{aligned} \tag{D.1}$$

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 CDF's corresponding to maturities 20 and 50 days respectively. These CDF's are obtained from the [Breedon and Litzenberger \(1978\)](#) formula, assuming 1,000 strike values per maturity. This is consistent with the later 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.<sup>26</sup> For convenience the main parameter settings are summarized in Table 6.

Figure 13 shows that my interpolation approach gives a very accurate approximation to the risk-neutral CDF with maturity 30 days. The maximum approximation error in this sample is about 0.0055.

To assess the accuracy of the quantile bound estimator and the resulting confidence intervals I repeat the estimation procedure 1,000 times. The results are summarized in Table 7. The quantile bound estimator is upward biased. Moreover, the coverage properties of the 95% confidence interval are also quite far below 0.95, as they contain the true quantile bound only 72% 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 bound of 0.1378, but the coverage doesn't improve much. In particular, the coverage is almost identical as in the case with only 300 return observations.

<sup>26</sup>I use the build in *R* function `tsbootstrap` for this calculation.

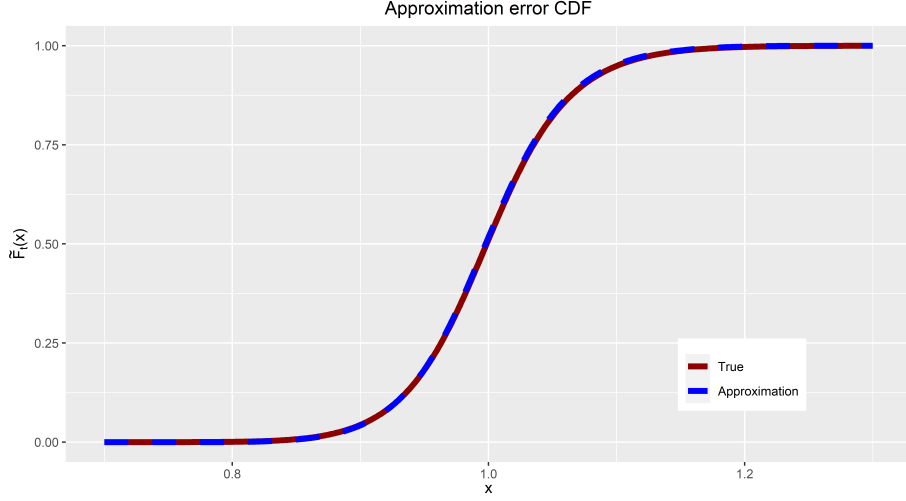


Figure 13: Red line denotes unconditional risk-neutral CDF. Blue line results from the interpolation of risk-neutral CDF's using maturities 20 and 50 days. The risk-neutral CDF's are estimated using the [Breedon and Litzenberger \(1978\)](#) formula.

Table 6: Parameter settings

$r$	$\mu$	$\lambda$	$\bar{\sigma}$	$T$	$S_0$	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,  $\mu$  is the growth rate,  $\lambda$  is time to maturity,  $\bar{\sigma}$  is the average volatility with  $\sigma_t \sim \text{UNIF}[\bar{\sigma} - 0.1, \bar{\sigma} + 0.1]$ ,  $T$  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 20 and 50 days.*

## D.2 Bootstrap Christoffersen test

This Section details the bootstrap approach needed to adjust the unconditional coverage test of [Christoffersen \(1998\)](#) for overlapping returns. Under  $H_0$  we have  $\mathbb{E}(H_{t+1}) = \tau$ , with

$$H_{t+1} = \mathbb{1} \left( R_{m,t+1} \leq \hat{Q}_\tau(R_{m,t+1}) \right).$$

Here,  $\hat{Q}_\tau(R_{m,t+1})$  is the quantile predictor of  $Q_\tau(R_{t+1})$  at time  $t$ . Since returns are overlapping and strongly correlated, the following bootstrap procedure is implemented. Bootstrap samples are created using the stationary bootstrap of [Politis and Romano \(1994\)](#). Let  $\hat{\tau}$  and  $\hat{\tau}^*$  be the sample and bootstrap estimates respectively and  $T$  is the number of return observations  $\{R_{m,t+1}\}_{t=1}^T$ . For every bootstrap sample, I compute  $\sqrt{T}(\hat{\tau}^* - \tau)$ , which ought to approximate  $\sqrt{T}(\hat{\tau} - \tau)$  under  $H_0$ . The  $p$ -value is obtained by inverting the null hypothesis.



Table 7: Simulation results

Time length	Coverage	Mean	Median	True quantile bound	SDF vol
300	0.7200	0.1496	0.1483	0.1177	0.1626
1,000	0.7270	0.1385	0.1378	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.*

The algorithm below describes the main steps.

---

**Algorithm 1:** Bootstrap Christoffersen test

---

**Result:**  $p$ -value

Compute  $\hat{\tau}$ ;

Generate 10,000 bootstrap samples;

**for**  $i$  in bootstrap sample **do**

    Compute  $\tau^*$ ;

    Compute  $\Gamma_i = \sqrt{T}(\hat{\tau}^* - \hat{\tau})$ ;

**end**

Compute  $\mathbb{P}_I(x) = \frac{1}{10,000} \sum_{i=1}^{10,000} \mathbb{1}(|\Gamma_i| \leq x)$ ;

Compute  $p$ -value:  $1 - \mathbb{P}_I(|\sqrt{T}(\hat{\tau} - \tau)|)$ .

---

## E 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 E.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).

### E.1 Disaster risk probabilities

Consumption growth is i.i.d. by assumption (see Equation (2.10)). 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  $\Delta c$  be log consumption growth and define

$$k(s; \Delta c) := \log \mathbb{E} [e^{s\Delta c}].$$

$k$  is the CGF of the random variable  $\Delta c$ . Due to the Poisson mixture assumption, the CGF obtains the explicit form

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

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

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

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

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

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

$$\varphi(s) := \exp(k(is)), \quad \tilde{\varphi}(s) := \exp(\tilde{k}(is)).$$

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

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

$\tilde{\mathbb{P}}(\lambda \Delta c \leq x)$  can be obtained in the same fashion, replacing  $\varphi(\cdot)$  with  $\tilde{\varphi}(\cdot)$  in (E.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.

## E.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

$$\mathbb{P}_t(R_{m,t+1} < \alpha) = \alpha \left[ P_{t+1}(\alpha S_t) - \frac{\text{put}_{t+1}(\alpha S_t)}{\alpha S_t} \right]. \quad (\text{E.3})$$

Under log preferences, it follows using (5.3) that

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

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

## F 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 2.1. For all the results to follow it is well known under which conditions the bounds are tight. For example, the log bound in Section F.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 Section 2, which is used repeatedly in this Section to analyze other SDF bounds

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

### F.1 Bound of Bansal and Lehmann (1997)

Here I consider a bound on the logarithm of the SDF. Recall that as an application of Jensen's inequality

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

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 log transformation to (F.1) we obtain for any  $\tau \in (0, 1)$

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

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

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

Taking expectations on both sides also renders an unconditional version.

### F.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

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

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

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

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

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

### F.3 Bound of Liu (2020)

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

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

The proof, as in Liu (2020), follows from an application of the *reverse Hölder inequality*.<sup>27</sup> 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 (F.2), since the reverse Hölder inequality requires almost sure positivity of  $\mathbb{1}(R_{t+1} \leq \tilde{Q}_\tau)$  to prove the lower bound. For  $p \in (1, \infty)$ , apply the reverse Hölder inequality to the relation (F.1)

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

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

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

<sup>27</sup>The reverse Hölder inequality states that for any  $p \in (1, \infty)$  and measure space  $(S, \Sigma, \mu)$  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  $\mu$ -almost all  $s \in S$ ,  $\|fg\|_1 \geq \|f\|_{\frac{1}{p}} \|g\|_{\frac{-1}{p-1}}$ .

## References

- 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.
- Angrist, J., Chernozhukov, V., and Fernández-Val, I. (2006). Quantile regression under misspecification, with an application to the US wage structure. *Econometrica*, 74(2):539–563.
- 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.
- Beare, B. K. and Schmidt, L. D. (2016). An empirical test of pricing kernel monotonicity. *Journal of Applied Econometrics*, 31(2):338–356.
- Billingsley, P. (2013). *Convergence of probability measures*. John Wiley & Sons.
- 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 Cochrane, J. H. (1999). By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of political Economy*, 107(2):205–251.
- Chabi-Yo, F. and Loudis, J. (2020). The conditional expected market return. *Journal of Financial Economics*.
- Christoffersen, P. F. (1998). Evaluating interval forecasts. *International economic review*, pages 841–862.

- Cochrane, J. H. (2005). *Asset pricing: Revised edition*. Princeton university press.
- Danielsson, J. and De Vries, C. G. (2000). Value-at-risk and extreme returns. *Annales d'Economie et de Statistique*, pages 239–270.
- Engle, R. F. (2009). The risk that risk will change. *Journal Of Investment Management (JOIM)*, Fourth Quarter.
- Engle, R. F. and Manganelli, S. (2004). Caviar: Conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics*, 22(4):367–381.
- 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.
- Gaglianone, W. P., Lima, L. R., Linton, O., and Smith, D. R. (2011). Evaluating value-at-risk models via quantile regression. *Journal of Business & Economic Statistics*, 29(1):150–160.
- Gil-Pelaez, J. (1951). Note on the inversion theorem. *Biometrika*, 38(3-4):481–482.
- 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.
- 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.
- Harrison, J. M. and Kreps, D. M. (1979). Martingales and arbitrage in multi-period securities markets. *Journal of Economic theory*, 20(3):381–408.
- Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The review of financial studies*, 6(2):327–343.
- 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.
- Koenker, R. (1994). Confidence intervals for regression quantiles. In *Asymptotic statistics*, pages 349–359. Springer.
- Koenker, R. and Bassett, G. (1978). Regression quantiles. *Econometrica: journal of the Econometric Society*, pages 33–50.

- Kremens, L. and Martin, I. (2019). The quanto theory of exchange rates. *American Economic Review*, 109(3):810–43.
- Kuester, K., Mittnik, S., and Paolella, M. S. (2006). Value-at-risk prediction: A comparison of alternative strategies. *Journal of Financial Econometrics*, 4(1):53–89.
- 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. W. and Wagner, C. (2019). What is the expected return on a stock? *The Journal of Finance*, 74(4):1887–1929.
- Politis, D. N. and Romano, J. P. (1994). The stationary bootstrap. *Journal of the American Statistical association*, 89(428):1303–1313.
- Protter, P. E. (2005). *Stochastic integration and differential equations*. Springer.
- Ross, S. (2015). The recovery theorem. *The Journal of Finance*, 70(2):615–648.
- Ross, S. A. (1976). Options and efficiency. *The Quarterly Journal of Economics*, 90(1):75–89.
- 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.
- Van Der Vaart, A. W. and Wellner, J. A. (1996). *Weak convergence and empirical processes*. Springer.
- Wahba, G. (1990). *Spline models for observational data*. SIAM.
- White, H. (2000). A reality check for data snooping. *Econometrica*, 68(5):1097–1126.
- Wichura, M. J. (1970). On the construction of almost uniformly convergent random variables with given weakly convergent image laws. *The Annals of Mathematical Statistics*, 41(1):284–291.