DNB Working Paper
No. 613 / October 2018

Value-at-Risk prediction using option-implied risk measures

Kai Schindelhauer and Chen Zhou

DeNederlandscheBank
EUROSYSTEEM

Value-at-Risk prediction using option-implied risk measures

Kai Schindelhauer and Chen Zhou *

* Views expressed are those of the authors and do not necessarily reflect official positions of De Nederlandsche Bank.

Working Paper No. 613

October 2018

De Nederlandsche Bank NV
P.O. Box 98
1000 AB AMSTERDAM
The Netherlands

Value-at-Risk prediction using option-implied risk measures

Kai Schindelhauer [a,b] and Chen Zhou [a,b]

October 2018

Abstract

This paper investigates the prediction of Value-at-Risk (VaR) using option-implied information obtained by the maximum entropy method. The maximum entropy method provides an estimate of the risk-neutral distribution based on option prices. Besides commonly used implied volatility, we obtain implied skewness, kurtosis and quantile from the estimated risk-neutral distribution. We find that using the implied volatility and implied quantile as explanatory variables significantly outperforms considered benchmarks in predicting the VaR, including the commonly used GARCH(1,1)-model. This holds for all considered VaR prediction models and VaR probability levels. Overall, a simple quantile regression model performs best for all considered VaR probability levels and forecast horizons.

Keywords: Implied Quantile, GARCH, Quantile Regression, Comparative Backtest
JEL Classifications: C14, G17

[a] De Nederlandsche Bank
[b] Erasmus School of Economics, Erasmus University Rotterdam

1 Introduction

Value-at-Risk (VaR), defined as a high quantile of future portfolio losses, is a standard risk measure used by financial institutions to quantify market risk. Due to a growing risk awareness and the tightening of regulatory requirements, there is a strong need for accurate risk estimation and consequently, well-performing and reliable VaR prediction strategies. Most existing VaR prediction models perform weakly in practice (see, e.g., Kuester et al., 2006; Bao et al., 2006). This challenge becomes even more pronounced when considering forecasting VaR for portfolio losses over a longer horizon.

One potential reason for the difficulty in VaR prediction stems from the fact that the existing approaches have mainly focused on using historical return information to predict the VaR in the future. If the return distribution of the future differs significantly from the considered past, historical return information might not be sufficient to accurately forecast the VaR. This situation is similar to the prediction of future realized volatility using historical volatilities. To overcome this difficulty, in the context of forecasting future stock volatility, several studies show that using forward-looking information, such as the option-implied volatility, outperforms historical volatility (see, e.g., Latané and Rendleman, 1976; Chiras and Manaster, 1978; Christensen and Prabhala, 1998). In a similar spirit, we expect that option-implied information, such as implied quantile, can be valuable for forecasting VaR. Therefore, the aim of this paper is to investigate the improvement of VaR predictions by using option-implied information, in particular, the implied quantile.

A few studies have considered various ways to incorporate option-implied information for predicting VaR. Aït-Sahalia and Lo (2000) introduced an VaR prediction method that is based on option information instead of historical return information only. They show that their option-based VaR forecast captures certain market risk aspects that a conventional VaR prediction could not capture. Another method to predict the VaR based on option prices is to first estimate the quantile of the corresponding parametric risk-neutral probability distribution and then change the risk-neutral probability measure to the physical measure, as proposed by Barone Adesi (2016). Lastly, Giot (2003, 2005), Jeon and Taylor (2013), Louzis et al. (2013) and Kim and Ryu (2015) incorporate option-implied volatility into different established VaR prediction models. They show that the resulting predictions are not worse and in some settings better than commonly used VaR prediction approaches. Compared to these studies, we follow a similar approach by incorporating not only the implied volatility, but also higher implied moments as well as implied quantiles into established VaR prediction models and test whether they help to improve the performance of VaR prediction. For that purpose, we need to extract these implied measures from the corresponding risk-neutral probability distribution of returns. We achieve this goal by applying the maximum entropy method (MEM).

Based on market prices of European options, we use the MEM (see Jaynes, 1957) to estimate the risk-neutral probability distribution of asset returns. The main advantage of this approach is that it yields a non-parametric distribution with an unrestricted shape. Furthermore, it is completely determined by the available options traded in the market. Lastly, the method performs well without requiring many options with different strike prices (Xiao and Zhou, 2016). The estimated distribution allows us to extract implied moments and quantiles. Xiao and Zhou (2016) show that MEM-implied moments are good predictors for future realized moments. Similarly, we investigate whether the quantile of the estimated risk-neutral distribution can help to predict the realized quantile in the future.

To estimate the risk-neutral probability distribution of asset returns, we follow the approach by Buchen and Kelly (1996), which shows that applying the MEM on market prices of options can reproduce the risk-neutral probability function. The MEM method is based on the concept of entropy. A possible intuition behind entropy is the measurement of the amount of order and disorder (Zhou et al., 2013). We maximize the entropy of the distribution subject to constraints given by arbitrage-free derivative pricing. Solving this optimization problem yields the risk-neutral probability measure, such that the expected value of the discounted option payoffs is equal to the current option market price. This method yields a discrete distribution that is related to the asset returns over the option's remaining time to maturity. From this distribution we can derive the implied moments as well as quantiles, which we then incorporate as explanatory variables in different VaR prediction models, such as covariate, GARCH (see, e.g., Engle, 1982; Bollerslev, 1986) and quantile regression models (see, e.g., Koenker and Bassett, 1978; Engle and Manganelli, 2004).

We find that using option-implied information can significantly improve the VaR prediction performance. The prediction performance across different models are compared by the comparative backtest suggested by Nolde and Ziegel (2017). Prediction models based on implied volatility and quantile significantly outperform considered benchmarks, which are commonly used models solely based on historical return information. This is consistent for different model classes and VaR probability levels. In addition, we find that when incorporating implied measures, a quantile regression model significantly outperforms the GARCH-type models for all considered VaR probability levels.

Our study contributes to the literature in three ways. First, our study is related to the literature on extracting forward-looking information from options and forecasting based on the information extracted. We extract option-implied information based on the entropy concept. This method has mainly been used for portfolio selection (see, e.g., Philippatos and Wilson, 1972; Ou, 2005; Huang, 2008) and asset pricing, e.g. pricing of stock options (Gulko, 1999) and bond options (Gulko, 2002). In the context of forecasting, the existing studies mostly use implied volatility (see, e.g., Latané and Rendleman, 1976; Chiras and Manaster, 1978; Christensen and Prabhala, 1998; Guo, 1998; Bakshi and Kapadia, 2003). We consider other higher implied moments as well as implied quantiles, and show that they can be beneficial for forecasting risk measures.

Second, we contribute to the literature of forecasting VaR. Our study provides a comprehensive comparison across different VaR prediction models with and without incorporating implied measures. Compared to the studies of Giot (2003, 2005), Jeon and Taylor (2013), Louzis et al. (2013) and Kim and Ryu (2015), our result further improves on the VaR prediction performance by incorporating option-implied information.

Third, our results point to the direction that quantile regression models are well suited for predicting the VaR when using option-implied information. Quantile regressions were introduced for predicting the quantile of a dependent variable given the levels of the independent variables (Koenker and Bassett, 1978). However, with historical information only, Kuester et al. (2006) and Bao et al. (2006) show that GARCH-type models perform better than quantile regression models for VaR prediction. We show that when adding option-implied information this ranking reverses. This result re-establishes the usefulness of quantile regression in VaR prediction.

This paper is organized as follows. We explain the methodology to extract option-implied information in Sections 2. Section 3 provides the models for VaR predictions. In Section 4 we discuss the results of the empirical study conducted. We investigate the robustness of these results in Section 5 and Section 6 concludes the paper.

2 Option-Implied Measures from Maximum Entropy Method

This section explains the methodology of the MEM. First, Section 2.1 explains the procedure used to estimate the risk-neutral probability distribution. Second, the derivation of the considered implied measures is described in Section 2.2.

2.1 Estimation Procedure

Following the approach described by Buchen and Kelly (1996), we use the MEM to estimate the risk-neutral probability distribution of asset returns given market prices of corresponding European options. Based on the assumption of arbitrage-free derivative pricing, we use these prices to formulate constraints for the distribution.

We consider the option prices on day t with a corresponding option maturity of h. In order to determine the distribution of the gross asset returns over the time horizon t to $t+h$, denoted by $R_{t:t+h} \in \mathbb{R}^+$, we use the market prices of m options at time t, denoted by $c_{t,j}$, for $j = 1,\dots,m$.

Arbitrage-free pricing assumes that the current option price equals the expectation of its discounted payoff at maturity under the risk neutral measure, which is given by

$$\text{Call options:} \qquad c_{t,j} = \mathrm{E}\left[\max\left(S_t R_{t:t+h} - K_j, 0\right)/r_{f,t:t+h}\right], \quad \text{for } j = 1,\ldots,m_1 \tag{1}$$

$$\text{Put options:} \qquad c_{t,j} = \mathrm{E}\left[\max\left(K_j - S_t R_{t:t+h}, 0\right)/r_{f,t:t+h}\right], \quad \text{for } j = m_1+1,\ldots,m_1+m_2, \tag{2}$$

where m_1 and m_2 denote the number of call and put options, respectively, such that $m_1 + m_2 = m$. $K_j \in \mathbb{R}$ is the strike price of option j, S_t is the underlying asset price at time t, $r_{f,t:t+h} \in \mathbb{R}$ denotes the gross risk-free rate over the considered time horizon and E denotes the expectation under the risk-neutral measure.

For the implementation of this method, let $w_i \in \mathbb{R}^+$, for $i = 1,\ldots,S$, be the possible values of the gross returns over this horizon. We consider a range of such possible values as $w_1 = 0.5$, $w_{i+1} = w_i + 0.001$ for $i = 1,2,\ldots,1000$. Let $p_{t,i} \in \mathbb{R}^+$ denote the unknown risk neutral probability corresponding to the considered future states, i.e. $p_{t,i} = \Pr(R_{t:t+h} = w_i)$, such that $0 < p_{t,i} < 1$ and $\sum_{i=1}^{S} p_{t,i} = 1$. Then we can rewrite the pricing equations (1) and (2) as

$$\begin{aligned}
\text{Call options:} \qquad & c_{t,j} = \sum_{i=1}^{S} p_{t,i}\left(\max\left(S_t w_i - K_j, 0\right)/r_{f,t:t+h}\right), && \text{for } j = 1,\ldots,m_1 \\
\Leftrightarrow \quad & 0 = \sum_{i=1}^{S} p_{t,i}\left(\max\left(S_t w_i - K_j, 0\right)/r_{f,t:t+h} - c_{t,j}\right), && \text{for } j = 1,\ldots,m_1 \\
\text{Put options:} \qquad & c_{t,j} = \sum_{i=1}^{S} p_{t,i}\left(\max\left(K_j - S_t w_i, 0\right)/r_{f,t:t+h}\right), && \text{for } j = m_1+1,\ldots,m \\
\Leftrightarrow \quad & 0 = \sum_{i=1}^{S} p_{t,i}\left(\max\left(K_j - S_t w_i, 0\right)/r_{f,t:t+h} - c_{t,j}\right), && \text{for } j = m_1+1,\ldots,m.
\end{aligned}$$

We summarize these equations in the following manner:

$$\sum_{i=1}^{S} p_{t,i}\, g_j(w_i) = 0, \qquad \text{for } j = 1,\ldots,m.$$

Since usually $S > m$, there is no unique distribution that solves these constraints. We choose the distribution that implies the maximum uncertainty and is least prejudiced under the given constraints (Buchen and Kelly, 1996). To estimate this distribution, we maximize the entropy $l \in \mathbb{R}^+$, as defined by Shannon (1948):

$$l = -\sum_{i=1}^{S} p_{t,i} \ln\left(p_{t,i}\right).$$

This method provides us with a unique optimal risk-neutral asset return distribution.

To find the maximum entropy for our optimization problem, we use the Lagrange multiplier method. The optimum results from the stationary point of the corresponding Lagrange function, which is given by

$$\mathscr{L} = \sum_{i=1}^{S} p_{t,i} \ln(p_{t,i}) + \lambda_0 \left(\sum_{i=1}^{S} p_{t,i} - 1 \right) + \lambda' \left(\sum_{i=1}^{S} p_{t,i} \, g(w_i) \right),$$

where $g(w_i) = (g_1(w_i), \ldots, g_m(w_i))'$ and $\lambda_0 \in \mathbb{R}$ and $\lambda \in \mathbb{R}^m$ are the Lagrange multipliers. From the first order conditions of $\mathscr{L}$, we can solve the optimization problem as

$$\hat{p}_{t,i} = \frac{\exp\left(\hat{\lambda}' g(w_i)\right)}{\sum_{i=1}^{S} \exp\left(\hat{\lambda}' g(w_i)\right)}, \qquad \text{for } i = 1, \ldots, S,$$

$$\hat{\lambda}' = \left(\hat{\lambda}_1, \ldots, \hat{\lambda}_m\right) = \arg\min \sum_{i=1}^{S} \exp\left(\lambda' g(w_i)\right), \tag{3}$$

where ˆ indicates the optimized solution.

We solve the optimization problem in (3) numerically by applying a multidimensional Newton-Raphson method, which follows the approach of Agmon et al. (1979). As a sum of strictly convex functions, equation (3) is strictly convex and hence has a unique global minimum. Thus, the obtained distribution based on MEM is unique.

2.2 Implied Measures for Value-at-Risk Forecasting

Having obtained the probability distribution using the MEM, we extract information that we use for VaR forecasting. First, we derive the following moments of the estimated discrete probabilities $\hat{p}_{t,i}$:

Implied mean: $$\mu_{\text{MEM},\,t} = \sum_{i=1}^{S} \hat{p}_{t,i} \ln(w_i),$$

Implied volatility: $$\sigma_{\text{MEM},\,t} = \sqrt{\sum_{i=1}^{S} \hat{p}_{t,i} \left(\ln(w_i) - \mu_{\text{MEM},\,t}\right)^2},$$

Implied skewness: $$s_{\text{MEM},\,t} = \frac{1}{\sigma^3_{\text{MEM},\,t}} \sum_{i=1}^{S} \hat{p}_{t,i} \left(\ln(w_i) - \mu_{\text{MEM},\,t}\right)^3,$$

Implied kurtosis: $$k_{\text{MEM},\,t} = \frac{1}{\sigma^4_{\text{MEM},\,t}} \sum_{i=1}^{S} \hat{p}_{t,i} \left(\ln(w_i) - \mu_{\text{MEM},\,t}\right)^4.$$

We incorporate these implied moments into VaR prediction models, which we describe in Section 3. Since MEM estimates the risk-neutral probability distribution of asset returns, theoretically, its mean approximately equals the risk-free rate over the considered time period. As the constraints already contain the risk-free rate, the implied mean should not add additional information and hence, we do not use it as an explanatory variable in the VaR prediction models.

In order to make the implied moments of different option maturities comparable, we annualize them by following the algorithm described by Meucci (2010) and assume 365 calendar days a year. Meucci shows how to project moments of general distributions to arbitrary horizons.

In addition to the implied moments, we use the quantile of the obtained distribution for predicting the VaR. The implied $(1-\alpha)$-quantile of the distribution is given by

$$q_{\text{MEM},\,t}^{(1-\alpha)} = \inf\left\{ \ln(w_I) \;\middle|\; \sum_{i=1}^{I} \hat{p}_{t,i} \geq 1-\alpha,\, I \in \mathbb{N} \right\},$$

which will also be used as an explanatory variable in VaR prediction models. This implied measure uses the tail information contained in the MEM-estimated probability distribution and therefore might be superior to the implied moments of the distribution for VaR prediction.

3 VaR Prediction Models Using Option-Implied Measures

We aim at predicting the h trading days ahead VaR at a probability level of $\alpha \in (0,1)$. The VaR is defined as follows. Denote that, conditional on all available information until time period t, the distribution function of an asset's cumulative return over the next h trading days as $F_{R_{t:t+h}}(r) = \Pr(R_{t:t+h} \leq r \,|\, \mathscr{F}_t)$. Then the corresponding VaR at time t at probability level α is given by

$$\text{VaR}_{t:t+h}^{\alpha} = -\inf\left\{\ln(r) \;\middle|\; F_{R_{t:t+h}}(r) \geq 1-\alpha\right\},$$

where $R_{t:t+h} \in \mathbb{R}^+$ denotes the asset return over a horizon of h, starting at time t, and $\mathscr{F}_t$ is the set of available information until time t.

Furthermore, let $r_t \in \mathbb{R}$ denote the daily return at time t, which is defined as $r_t = \ln(S_t/S_{t-1})$. Similarly, cumulative returns over h trading days are defined as $r_{t:t+h} = \ln(S_{t+h}/S_t) = \sum_{j=1}^{h} r_{t+j}$. The forecasting models are based on the time series of r_t. To estimate the model parameters, we use a moving window with a length of $T = 1000$ trading days preceding to the time period t, i.e. approximately four years of daily data. To predict the VaR, we use implied measures obtained from option prices. The general notation is given as follows. Let $x_t \in \mathbb{R}^N$ be a set of the implied measures at time t, where N denotes the number of measures used.

In the rest of this section, we introduce the different model families applied in VaR forecasting. We also show how to incorporate the obtained implied measures as explanatory variables. Section 3.1, Section 3.2 and Section 3.3 explain the model classes of covariate, GARCH and quantile regression models, respectively. The methodology applied to evaluate and compare the performance of the different VaR prediction models is described in Section 3.4.

3.1 Covariate Models

3.1.1 Model Specification

The covariate model class uses a simple model specification, which describes daily returns by an intercept and residuals with a possibly time-varying variance. This conditional variance, h_t, is modeled by a regression model on lagged implied measures. Giot (2003, 2005) and Kim and Ryu (2015) use this type of covariate model to incorporate implied volatility. We additionally incorporate other implied measures to such a model. The model is given by

$$
\begin{aligned}
r_t &= \mu + \sqrt{h_t}\varepsilon_t, \qquad && \text{for } t = 2, \ldots, T, \\
h_t &= \omega + \theta x_{t-1},
\end{aligned}
$$

where $\mu \in \mathbb{R}$, $\omega \in \mathbb{R}$ and $\theta \in \mathbb{R}^{1\times N}$ are the model parameters to be estimated, $\varepsilon_t \in \mathbb{R}$ denotes the residual at time t with mean of 0 and variance of 1. $h_t \in \mathbb{R}^+$ is the (latent) conditional variance of returns at time t, i.e. $h_t = \mathrm{Var}(r_t \mid \mathscr{F}_{t-1})$. Since the model does not include the lagged conditional variance nor lagged innovations in the conditional variance, it is not a GARCH model. However, as the conditional variance h_t has to be positive for all t, we impose similar restrictions as in GARCH models (Bollerslev, 1986). We require $\omega > 0$ and $\theta_i \geq 0$ for $i = 1, \ldots, N$. Furthermore, the explanatory variables have to be non-negative to ensure positivity of the conditional variance. Hence, for this model we use implied variance, absolute implied skewness [1], implied kurtosis and negative implied quantile as implied measures.

If no implied measures x_t are added, the model simplifies to a simple independent and identically distributed univariate model for the distribution of r_t. We use this simple model as a benchmark to investigate the improvement by adding implied measures as regressors.

3.1.2 Model Parameter Estimation

To obtain estimates of the model parameters, we apply the maximum likelihood method based on the observations r_t and x_{t-1} for $t = 2, \ldots, T$. Therefore, we assume that the residuals ε_t follow a specific parametric distribution and are conditionally independent, i.e. given the history $\mathscr{F}_{t-1}$, the residuals ε_t are independent. We follow the studies of Giot (2003, 2005) and Louzis et al. (2013) and opt for the skewed t-distribution (Fernández and Steel, 1998). The skewed t-distribution is able to reproduce the stylized facts of asset returns, namely negative skewness and heavy tails. Giot and Laurent (2003, 2004) and Giot (2003, 2005) show that using this distribution improves the performance of VaR predictions by GARCH-type models compared to forecasts based on the normal distribution. We assume that $\varepsilon_t \mid \mathscr{F}_{t-1} \sim skt(0, 1, \xi, \nu)$, which has mean 0 and variance 1.

[1] Higher skewness indicates a skewed perception of future returns, which potentially leads to a higher uncertainty, irrespectively of its sign.

The parameters ξ and ν are related to the distribution's skewness and kurtosis, respectively. For details of the density of the skewed t-distribution, see Appendix A.1.

The maximum likelihood method involves an optimization over a large number of parameters. To decrease the number of parameters and to increase the convergence speed of the optimization algorithm, we apply the variance targeting technique (see, e.g., Francq et al., 2011). For GARCH models, variance targeting can be superior to the maximum likelihood method alone when predicting VaR, especially when the model is misspecified, as shown by Francq et al. (2011). We adopt the two-step variance targeting approach, introduced by Engle and Mezrich (1996). Firstly, we estimate the intercept of the conditional variance, ω. Secondly, estimates the remaining parameters by maximum likelihood. In the first step, we set the model's unconditional variance $\mathrm{E}(h_t)$ equal to the sample variance of the residuals and obtain a closed form solution of ω, given by

$$\mathrm{E}(h_t) = \omega + \theta\, \mathrm{E}(x_t),$$

where we use the sample mean of regressors as the estimate for $\mathrm{E}(x_t)$. As it depends on the value of θ, we update this ω estimate in each iteration of the maximum likelihood procedure.

As the model specification requires $\omega > 0$, we obtain an additional restriction on θ from the variance targeting approach: $\theta \frac{\mathrm{E}(x_t)}{\mathrm{E}(h_t)} < 1$. To ensure that h_t is positive for all $t = 2, \ldots, T$, we impose this additional restriction when estimating the model parameters θ by maximum likelihood.

3.1.3 Value-at-Risk Prediction

To forecast $\mathrm{VaR}^{\alpha}_{T:T+h}$ for $h = 10$, we predict a sample of future returns and then use their empirical quantile as VaR forecast. We use the estimated model parameters to obtain multi-period forecasts of returns by iteratively performing Monte-Carlo simulations. We obtain $S = 10{,}000$ predictions for each $r_{T+1}, \ldots, r_{T+h}$ by drawing samples of $\varepsilon_{T+1}, \ldots, \varepsilon_{T+h}$ from the underlying probability distribution and plugging them into the estimated model. We calculate the cumulative returns corresponding to these simulated paths, $\hat{r}_{T:T+h,i} = \sum_{t=1}^{h} \hat{r}_{T+t,i}$, for $i = 1, \ldots, S$. Finally, we use the negative of the empirical $(1-\alpha)$-quantile of the obtained cumulative return sample as an estimate for $\mathrm{VaR}^{\alpha}_{T:T+h}$.

For this purpose, we need to predict h_t for the next ten days by using the implied measures $x_T, x_{T+1}, \ldots, x_{t+h-1}$, as we predict the VaR for the next ten trading days. Since only x_T is observed, we need to predict the implied measures for the other $h-1$ days. For simplicity, we keep the implied measures constant over the forecasting period, i.e. $\hat{x}_{T+1}, \ldots, \hat{x}_{T+h-1} := x_T$.

3.2 GARCH Models

The second model class considered consists of GARCH models. We investigate the VaR prediction performance improvement by adding option-implied measures to a GARCH model, which is commonly used and performs well for VaR forecasting in practice (see, e.g., Kuester et al., 2006; Bao et al., 2006). We follow the GARCH-type model studied by Giot (2003, 2005) and Kim and Ryu (2015), who incorporate implied volatility into it. We adjust their models to include more implied measures as explanatory variables.

We study an extended GARCH(1,1) model (Bollerslev, 1986), given by

$$\begin{aligned} r_t &= \mu + \sqrt{h_t}\varepsilon_t, & \text{for } t = 2, \ldots, T, \\ h_t &= \omega + \gamma h_{t-1}\varepsilon_{t-1}^2 + \delta h_{t-1} + \theta x_{t-1}. \end{aligned}$$

Compared to the covariate models, it has two additional autoregressive terms to describe the conditional variance h_t, which are essential for a GARCH-type model. Additionally, we estimate the corresponding model parameters, γ, $\delta \in \mathbb{R}$, and impose restrictions on the parameters to ensure positivity of h_t. Following Bollerslev (1986), we require $\omega > 0$ and γ, $\delta \geq 0$ and for stationarity, $\gamma + \delta < 1$. Furthermore, we impose that $\theta_i \geq 0$ for $i = 1, \ldots, N$.

We use the same non-negative explanatory variables as listed for the covariate models in Section 3.1.1. We compare these models to the benchmark model in which no implied measures x_t are involved. The benchmark model is therefore the standard GARCH(1,1) model.

To estimate the model parameters, we recursively compute the conditional variance h_t and follow the same procedure as that for covariate models, explained in Section 3.1.2. There are three technical differences. Firstly, we set the initial value of the conditional variance, h_1, to the sample variance of r_t. Secondly, we adjust the variance targeting procedure to account for the fact that the unconditional variance is different for the covariate models. Lastly, similar to the covariate models, we obtain an additional parameter restriction, given by $\theta \frac{\mathrm{E}(x_t)}{\mathrm{Var}(r_t)} + \gamma + \delta < 1$.

Having obtained the parameter estimates for the models, we estimate the VaR by following the same procedure as that for the covariate models, outlined in Section 3.1.3.

3.3 Quantile Regression Models

While the first two model classes are similar, the quantile regression models follow a different approach. Firstly, this model class does not assume a specific probability distribution of the residuals. Secondly, in contrast to a typical (linear) regression model, which models the conditional mean of the dependent variable, quantile regression estimates the $(1-\alpha)$-quantile of the dependent variable conditional on the explanatory variables. The main advantage of this approach is that some of the considered implied measures might be better suited for explaining the quantile of a distribution than its mean or variance. A downside of quantile regression models is the potentially large standard errors, especially for extreme quantiles close to zero or one (see, e.g., Krause, 2003) that are determined by only a few observations.

3.3.1 Model Specifications

We consider a simple quantile regression model proposed by Koenker and Bassett (1978), which allows a simple incorporation of all considered option-implied measures. We model the conditional $(1-\alpha)$-quantile of the cumulative asset returns over the horizon of h trading days as follows

$$Q_{r_{t:t+h}\,|\,x_t}(1-\alpha) = \mu + \beta x_t,$$

where $\mu \in \mathbb{R}$ and $\beta \in \mathbb{R}^{1\times N}$ are the model parameters to be estimated to fit the linear relationship between conditional quantile and implied measures. To our best knowledge, no existing study considers incorporating option-implied information in a simple quantile regression model yet. Nevertheless, Chen and Chen (2005) show that incorporating historical volatility into this quantile regression yields better VaR predictions than parametric models, especially at a probability level of 99% and for holding periods of more than five days.

We incorporate the implied measures as described in Section 2.2. We set a benchmark model using quantile regression by replacing x_t by lagged return information available at time t, i.e. for the benchmark model $x_t = r_{t-h:t}$.

3.3.2 Model Parameter Estimation

To estimate the model parameters, we follow the method explained by Koenker and Bassett (1978) based on asymmetric penalties to the residuals. The parameter estimates are obtained from the following optimization problem:

$$\min_{\mu,\beta_0,\beta} \quad \frac{1}{T-h}\left[\sum_{t=1}^{T-h}\left((1-\alpha) - \mathscr{I}_{\left\{r_{t:t+h}<Q_{r_{t:t+h}|x_t}(1-\alpha)\right\}}\right)\left(r_{t:t+h} - Q_{r_{t:t+h}\,|\,x_t}(1-\alpha)\right)\right],$$

where $\mathscr{I}_{\left\{r_{t:t+h}<Q_{r_{t:t+h}|x_t}(1-\alpha)\right\}} = 1$ if $r_{t:t+h} < Q_{r_{t:t+h}\,|\,x_t}(1-\alpha)$ and 0 otherwise.

3.3.3 Value-at-Risk Prediction

We use the out-of-sample forecast of the conditional quantile as a direct estimate, given as

$$\widehat{\text{VaR}}^{\alpha}_{T:T+h} = -\widehat{Q}_{r_{T:T+h}\,|\,x_T}(1-\alpha) = -\left(\hat{\mu} + \hat{\beta} x_T\right).$$

Note that with this approach, we obtain a direct forecast on the $h-$day VaR by running the quantile regression based on $h-$day returns. Using cumulative returns over a horizon of multiple days (e.g. ten trading days) leads to autocorrelated residuals, which may potentially yield large standard errors in the estimates of the model parameters and hence imprecise forecasts. Alternatively, we run the quantile regression based on daily returns to obtain a $1-$day VaR forecast and then scale the forecast to the target horizon. The $h-$day VaR forecast is obtained by applying the square-root-rule and given by

$$\widehat{\text{VaR}}^{\alpha}_{T:T+h} = \widehat{\text{VaR}}^{\alpha}_{T:T+1}\sqrt{h}.$$

Even though this method assumes identically distributed returns over the horizon, which is possibly incorrect in practice (see, e.g., Diebold et al., 1997; Odening and Hinrichs, 2003), this method might still improve the VaR predictions compared to using cumulative returns.

3.4 Forecast Evaluation

By using a moving time window for model estimations of $T = 1000$ trading days, we obtain out-of-sample VaR predictions for the whole time period of the testing sample. To evaluate the prediction performance of the VaR forecasting models, we apply two methods: the unconditional coverage ratio and the comparative backtest approach. They are explained in Section 3.4.1 and Section 3.4.2, respectively.

3.4.1 Unconditional Coverage Ratio

We follow the approach described by Christoffersen (1998) and count the number of occurrences that the realized ten days cumulative loss is higher than the corresponding predicted VaR. Let $I_t(\alpha) \in \{0,1\}$ be the indicator whether the VaR forecast at time t is violated. For a given model, this indicator is defined as

$$I_t(\alpha) = \mathscr{I}_{\left\{r_{t:t+h} < -\widehat{\text{VaR}}^{\alpha}_{t:t+h}\right\}}, \qquad \text{for } t = 1, \dots, n,$$

where $\mathscr{I}_{\left\{r_{t:t+h} < -\widehat{\text{VaR}}^{\alpha}_{t:t+h}\right\}} = 1$ if $r_{t:t+h} < -\widehat{\text{VaR}}^{\alpha}_{t:t+h}$ and 0 otherwise; and $n \in \mathbb{N}$ is the number of out-of-sample predictions.

Let $V \in \mathbb{N}$ be the total number of VaR violations of a given prediction model, i.e. $V = \sum_{t=1}^{n} I_t(\alpha)$. Under the assumption that the VaR forecasts have correct unconditional coverage ratio and the violations are identically and independently distributed, the number of violations, V, follows a binomial distribution with parameters n and $1-\alpha$ (Christoffersen, 1998). Following Christoffersen (1998), we test whether a model's VaR predictions differ significantly from the accurate unconditional coverage ratio using the following likelihood ratio test:

$$\text{LR} = 2\ln\left(\frac{\left(1-\frac{V}{n}\right)^{n-V}\left(\frac{V}{n}\right)^{V}}{\alpha^{n-V}\left(1-\alpha\right)^{V}}\right), \tag{4}$$

which is asymptotically $\chi^2(1)$ distributed.

Due to the multi-period VaR predictions studied in this paper, the VaR violation indicators $I_t(\alpha)$ are autocorrelated. Hence, the assumption of independent forecasts does not hold and we cannot apply the test procedure directly on the entire series of $I_t(\alpha)$. Therefore, we follow Diebold et al. (1998) and consider the h sub-series $\{I_1(\alpha), I_{1+h}(\alpha), I_{1+2h}(\alpha), \ldots\}$, $\{I_2(\alpha), I_{2+h}(\alpha), I_{2+2h}(\alpha), \ldots\}$, $\ldots$, $\{I_h(\alpha), I_{2h}(\alpha), I_{3h}(\alpha), \ldots\}$. Using Bonferroni bounds, we obtain a formal unconditional coverage ratio test by combining the h tests based on each sub-series. For a hypothesis test with size $p \in (0,1)$, we apply the described likelihood ratio test (equation (4)) on each sub-series individually. If any of the corresponding h p-values is smaller than p/h, we reject the hypothesis of an accurate unconditional coverage ratio for the considered model (Diebold et al., 1998; Louzis et al., 2013).

The unconditional coverage ratio gives a good indication of a model's VaR prediction performance, but cannot be applied to compare the performance of two different models. As it only tests the frequency of violations but not the magnitude of VaR predictions, it could be that two models have identical unconditional coverage ratios, but one of them yields more conservative risk forecasts than the other model in most time periods. This would lead to unequal capital requirements in practice, but the coverage ratio test cannot distinguish the two models. Therefore, we consider comparative backtest for comparing the VaR prediction performance of two models, which is explained in the following subsection.

3.4.2 Comparative Backtest

To account for both frequency of violations and the difference of VaR forecasts to realized losses when evaluating models, we follow the comparative backtest approach of Nolde and Ziegel (2017). Using a scoring function, this method assigns a score to each VaR prediction. Based on these scores, we can then compare two different models.

Let $S\left(\widehat{\text{VaR}}^{\alpha}_{t:t+h},\, l_{t:t+h}\right)$ be a scoring function. By assigning a score to the forecast, it assesses the VaR prediction of the examined model regarding the realized cumulative loss over the horizon h, $l_{t:t+h} := -r_{t:t+h}$. As proposed by Nolde and Ziegel (2017), we apply the classical 1-homogeneous scoring function, given by

$$S\left(\widehat{\text{VaR}}^{\alpha}_{t:t+h},\, l_{t:t+h}\right) = (1-\alpha)\ \widehat{\text{VaR}}^{\alpha}_{t:t+h} + \mathscr{I}_{\left\{l_{t:t+h} > \widehat{\text{VaR}}^{\alpha}_{t:t+h}\right\}} \left(l_{t:t+h} - \widehat{\text{VaR}}^{\alpha}_{t:t+h}\right).$$

The first term of this scoring function penalizes for large VaR predictions and the second term for the difference between VaR prediction and realized loss in the event of a violation. Hence, an optimal model would yield VaR predictions with an unconditional coverage ratio equal to the target probability level of $1-\alpha$, while being as low as possible.

We apply the following test statistic to evaluate whether one model significantly outperforms the other:

$$\Psi = \frac{\Delta \bar{S}}{\hat{\sigma}_{\Delta S}/\sqrt{n}},$$

where $\Delta\bar{S} = \frac{1}{n}\sum_{t=1}^{n}\left(S\left(\widehat{\text{VaR}}_{t:t+h}^{\alpha},\, l_{t:t+h}\right) - S\left(\widehat{\text{VaR}}_{t:t+h}^{\alpha,*},\, l_{t:t+h}\right)\right)$ is the average score difference, where $\widehat{\text{VaR}}_{t:t+h}^{\alpha,*}$ denotes the VaR forecast of the model that we use as a reference model, $n \in \mathbb{N}$ denotes the number of out-of-sample predictions and $\sigma_{\Delta S}^2 = \text{Var}\left(\sqrt{n}\,\Delta\bar{S}\right)$ is the asymptotic variance of the score differences. To correct for potential autocorrelation and heteroscedasticity, we use a heteroscedasticity and autocorrelation consistent (HAC) estimator for this variance (see, e.g., Andrews, 1991). Following Nolde and Ziegel (2017), we use the Parzen kernel density, which we truncate at the rounded up square-root of n, to obtain the HAC estimate $\hat{\sigma}_{\Delta S}^2$.

Nolde and Ziegel (2017) show that Ψ asymptotically follows the standard normal distribution and explain how to interpret this test statistic. Let $\Phi(\cdot)$ be the distribution function of the standard normal distribution and $p \in (0,1)$ be the targeted significance level. Then we reject the hypothesis that the examined model forecasts the VaR at least as well as the reference model if $1 - \Phi(\Psi) \leq p$. In this case we would conclude that the reference model significantly outperforms the examined model. If $\Phi(\Psi) \leq p$, we reject the hypothesis that the examined model predicts at most as well as the reference model. This yields the conclusion that the examined model significantly outperforms the reference model. There is no clear evidence for out-performance, if we cannot reject either of the two hypotheses.

We use a color code to visualize the outcome of the model comparison. Green means that the examined model outperforms the reference model, whereas red indicates the opposite. Yellow shows that both models have a similar performance and do not differ significantly.

4 Empirical Study

To investigate whether incorporating implied measures improves the VaR prediction performance of the considered models, we assess and compare different cases in an empirical study. For each model class, we consider a benchmark model that does not use any implied measures and is solely based on historical return information. We compare the VaR prediction performance of this benchmark to the case of individually and simultaneously incorporating MEM-implied measures into the model. Finally, we compare the prediction performance across the best model stemming from the different model classes.

4.1 Constructing Implied Measures

For the empirical analysis, we use exchange-traded European equity options on the S&P 500, which are amongst the most liquid options. We work with the option information provided by OptionMetrics IvyDB US database. The data contains maturity date, open interest and strike, bid and ask prices of options from 04 January 1996 to 29 April 2016.

To choose the set of options we apply the MEM on, we follow the practical approach used by Xiao and Zhou (2016). We consider out-of-the-money and at-the-money call and put options within a moneyness [2] range from 0.85 to 1.15. Within this range we only consider the options that are steps of 0.025 apart, i.e. having a moneyness of 0.850, 0.875, 0.900, 0.925 etc. This is done to obtain a stable numerical solution for the optimization problem of the maximum entropy method. Buchen and Kelly (1996) illustrate that closely spaced options can lead to an ill-conditioned [3] or even singular Jacobian matrix, which is not suitable in the optimization procedure. As there are usually no strike prices available that yield exactly those moneyness steps, we allow for an absolute tolerance around those steps of $\frac{0.025}{4}$. The choice of this tolerance is a trade-off between numerical stability and an equal distribution of used strike prices over the considered moneyness range. The latter has a positive influence on the accuracy of the estimated distribution due to the incorporation of more information. Furthermore, we only consider options that have a positive bid price and a positive open interest, since this indicates active trading of those options [4]. For all calculations we use the option's mid-price, which is defined as the average of bid and ask price at a given date.

We obtain daily close price information of the underlying asset over the same time period from OptionMetrics IvyDB US database. Table 1 shows the summary statistics of the S&P 500 returns. Due to the moving window with a length of $T = 1000$ trading days, we start the out-of-sample period for VaR predictions in January 2000. The first prediction is estimated on 03 January 2000, so that the first ten trading days ahead VaR forecasts corresponds to 18 January 2000. As the last forecast corresponds to 29 April 2016, we obtain 4097 out-of-sample VaR predictions for all considered models. The simulation study by Nolde and Ziegel (2017) suggests that this is a sufficient sample size to obtain stable rankings of different model performances.

As a proxy for the risk-free rate we use the zero-coupon bond rate available in the OptionMetrics IvyDB US database. If zero-coupon bond rates are not available at a specific trading day, we use the rates of the last trading day they were available before. To obtain the risk-free rate, we select the bond that has a maturity closest to the expiration date of the corresponding set of options. We then scale the chosen bond rate to the same time horizon as the option maturity and use it as risk-free rate, $r_{f,t:t+h}$, in the described methods to obtain option-implied measures.

The obtained implied measures correspond to the market expectations until the maturity date of the underlying options used in the MEM. Therefore, the time-to-maturity of the considered options and the corresponding implied measures changes over time. We solve this issue by considering

[2] Moneyness is defined as the option's strike price over the current price of the underlying.

[3] According to Süli and Mayers (2003), the condition number of a non-singular matrix A and a norm $\|\cdot\|$ is defined as $\kappa(A) = \|A\|\|A^{-1}\|$. The matrix A is said to be ill-conditioned if $\kappa(A) \gg 1$. The condition number can be used to measure how sensitive the output of a function is to changes or errors in the input values.

[4] On three trading days in the considered time period all options in the database have an open interest of 0. We assume that this is a data error and neglect the constraint of positive open interest in these cases.

Table 1: Summary statistics of S&P 500 returns

	Whole Sample		Training Sample		Testing Sample	
	Daily	10 Days	Daily	10 Days	Daily	10 Days
Mean	0.0003	0.0030	0.0009	0.0092	0.0002	0.0015
Volatility	0.0123	0.0335	0.0109	0.0310	0.0126	0.0339
Skewness	−0.0536	−0.6641	−0.3595	−0.3377	0.0040	−0.7151
Kurtosis	10.7476	7.4742	7.1770	3.7539	11.1264	8.0734
Minimum	−0.0903	−0.2588	−0.0687	−0.1166	−0.0903	−0.2588
Maximum	0.1158	0.2164	0.0512	0.1241	0.1158	0.2164
Start	03 Jan 1996	16 Jan 1996	03 Jan 1996	16 Jan 1996	04 Jan 2000	18 Jan 2000
End	29 Apr 2016	29 Apr 2016	03 Jan 2000	03 Jan 2000	29 Apr 2016	29 Apr 2016
Obs.	5117	5108	1011	1002	4106	4097

Note: This table shows the summary statistics for daily returns and cumulative returns over ten days. Start and end denote the first and last day of the time series and obs. is the number of observations.

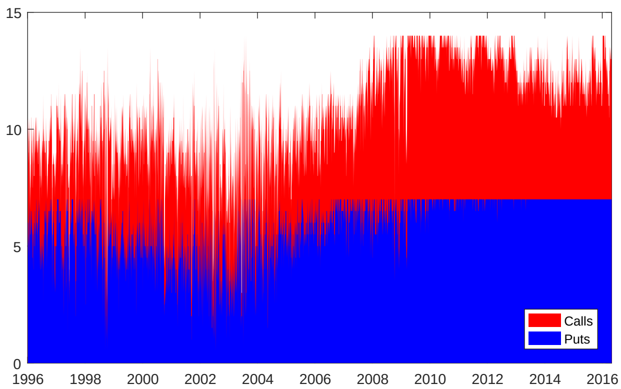

Figure 1: Time series of the number of options used for MEM. Plotted values are averages of numbers of call and put options, respectively, over both considered option sets at each trading day.

two sets of options and linearly interpolating or extrapolating their implied measures to the target horizon of $h = 21$ trading days. This approach is similar to the calculations of the well-known volatility indexes VXO and VIX (see, e.g., Carr and Wu, 2006), which are interpolated to a target of 30 calendar days. The horizon of $h = 21$ trading days matches the average number of trading days per month. We choose a length of one month, as we prefer interpolation over extrapolation and because options are mostly traded around this time to maturity. Due to less frequently traded options and possible mispricing, a shorter horizon may lead to inaccurate estimates.

We apply the MEM individually on two sets of options, which expire at the third Fridays of the following months [5]. Note that we do not use options with a time-to-maturity of less than ten

[5] Most exchange-traded options expire on the third Friday of a month. Thus, compared to options with different

Table 2: Correlation of implied measures

	σ_{VIX}	σ_{BSIV}	σ_{MEM}	s_{MEM}	k_{MEM}	$q_{MEM}^{10\%}$	$q_{MEM}^{5\%}$	$q_{MEM}^{1\%}$
σ_{VIX}	1	0.985	0.994	0.429	−0.437	−0.968	−0.976	−0.947
σ_{BSIV}		1	0.985	0.414	−0.408	−0.962	−0.966	−0.927
σ_{MEM}			1	0.400	−0.412	−0.970	−0.977	−0.959
s_{MEM}				1	−0.902	−0.427	−0.403	−0.311
k_{MEM}					1	0.423	0.419	0.393
$q_{MEM}^{10\%}$						1	0.970	0.912
$q_{MEM}^{5\%}$							1	0.933
$q_{MEM}^{1\%}$								1

Note: This table shows the correlation matrix of the considered option-implied measures. The BSIV and MEM measures are interpolated for 21 trading days ahead. The statistics are based on 5116 observations of the period from 03 January 1996 to 29 April 2016.

trading days, as those are usually not frequently traded anymore. Figure 1 displays the time series of the average number of call and put options over both considered sets used for the MEM. After obtaining the two implied probability distributions based on these two option sets, we compute the implied measures of both resulting distributions and interpolate or extrapolate them to the targeted horizon of $h = 21$ trading days [6].

Table 2 shows the correlation of the obtained MEM-implied measures as well as commonly used option-implied volatility measures, which are VIX and Black-Scholes formula based implied volatility (BSIV). We obtain the closing prices of the VIX over the considered time period from OptionMetrics IvyDB US database. The BSIV calculation (see, e.g., Hull, 2012) is based on the same set of options selected for the MEM. Table 2 shows that the three considered implied volatility measures and quantiles are highly correlated. The plots of their evolution over time in Figure 2 and 3 confirm their similarity. Figure 2 displays that the VIX and MEM-implied volatility are almost identical, whereas BSIV differs slightly in some periods. Even though the VIX calculation uses weekly options, which presumably improves the estimation and interpolation accuracy of the implied volatility, it yields almost the same estimates as that of the MEM, which only uses monthly options. The plot also shows that the MEM-implied volatility contains a few extreme values that do not match the other two measures.

Figure 3 shows the evolution of the implied quantiles of the estimated risk-neutral probability

maturity, options expiring on the third Friday of a month cover a higher number of strike prices and a longer trading period before maturity.

[6] We use the following formula for linear interpolation and extrapolation: $M(h) = M(h_1) + (h - h_1)\frac{M(h_2)-M(h_1)}{h_2-h_1}$, where h is the target horizon, h_1 and h_2 are the time to maturities of the two considered option sets and $M(\cdot)$ denotes the implied measure for a given horizon.

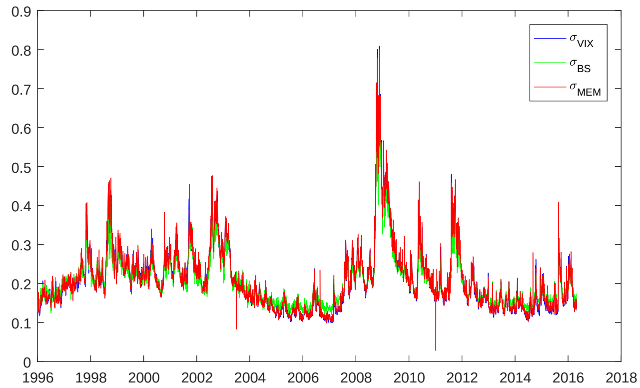

Figure 2: Time series of annualized option-implied volatility measures.

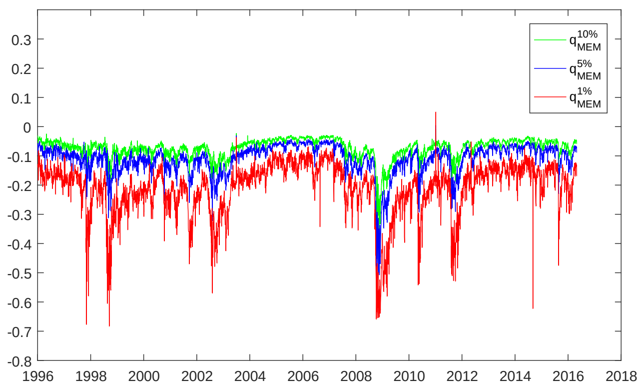

Figure 3: Time series of MEM-implied quantiles.

function, $q^{(1-\alpha)}_{\mathrm{MEM},\,t}$, for different levels of α. The lower $1-\alpha$, the more fluctuating the implied quantiles become. The changing differences between the quantiles over time also highlights the change of the left distribution tail in some periods. As options closer to at-the-money are more frequently traded, the accuracy of the MEM-estimated distribution is more accurate around the mean. Options further out-of-the-money are traded less frequently and hence yield less precise estimates for the distribution tails. To avoid potential inaccurate implied quantile estimates, we only use the 10%-implied quantile as explanatory variable for the VaR prediction models regardless of the VaR probability level.

4.2 Results

We investigate the VaR prediction performance of the considered implied measures and models in three steps. First, in Section 4.2.1 we test whether the predicted VaR corresponds to the designated unconditional coverage ratio for all models considered. A model is regarded as having a poor performance, if the VaR violation frequency is significantly different from the target probability level. Second, in Section 4.2.2, we compare the prediction performance within each model class between models with and without incorporating implied measures. Lastly, Section 4.2.3 compares the VaR prediction performance across the considered model classes using the best performing implied measures for each model class.

4.2.1 Unconditional Coverage Ratio Test

The first two panels of Table 3 show the VaR prediction performance for the covariate and GARCH models, respectively. For most models the hypothesis of accurate conditional coverage ratio is not rejected. However, for all considered covariate and GARCH models, the VaR predictions at a 95% probability level are too conservative (i.e. overestimating the risk), whereas those at a 99% level are too optimistic (i.e. underestimating the risk). There are at least two possible explanations for such a result. First, the skewed t-distribution may not be able to capture the tail behavior of the daily returns sufficiently, even though it allows for non-zero skewness and heavy tails. A possible solution is to model the distribution tail individually, e.g. by Extreme Value Theory (EVT) (see, e.g, McNeil and Frey, 2000; Longin, 2000). A second possible explanation follows from portfolio theory. The potential negative autocorrelation of daily returns and the resulting mean reversion may make stock investments less risky for a longer investment horizon (see, e.g., Campbell and Viceira, 2002). Since the considered models are based on daily returns without capturing the autocorrelation, they may result in an overestimation of tail risk. However, this does not explain the underestimated tail risk for the most extreme quantile corresponding to a probability level of 99%.

The third panel of Table 3 displays the number of VaR violations for the considered simple quantile regression model when using cumulative returns to forecast the VaR directly. The model underestimates the risk: the number of VaR violations is too high for the considered target probability levels. The high number of violations for the benchmark model does not substantially decrease towards the target level when adding implied measures. For the 99% probability level, the deviation from the target probability level is significant at a 5% level for all considered implied measures. This poor performance is most likely caused by the autocorrelation among the cumulative returns. In contrast, when using the alternative approach that scales the daily VaR obtained from the simple quantile regression model to $10-$day VaR, we obtain a better VaR prediction performance. This is indicated by the fewer VaR violations, shown in the last panel of Table 3. Except for the benchmark model at a 99% probability level, the number of violations does not differ significantly from the targeted VaR probability level. It should be noted that, in contrast to the benchmark model, we do not incorporate lagged returns when adding implied measures as explanatory variables [7].

[7] For each considered quantile regression model, the performance without lagged returns is superior. The corresponding results when including lagged returns are worse and are available upon request.

Table 3: Value-at-Risk violations of estimated models

	α in %		Benchmark	σ_{VIX}	σ_{BS}	σ_{MEM}	s_{MEM}	k_{MEM}	$q_{MEM}^{10\%}$	σ_{MEM}, s_{MEM}	$\sigma_{MEM}, s_{MEM}, k_{MEM}$	$\sigma_{MEM}, s_{MEM}, k_{MEM}, q_{MEM}^{10\%}$	$\sigma_{MEM}, q_{MEM}^{10\%}$
				Incorporated Implied Measures									
Covariate	95.0	#	189	158	158	158	190	188	157	153	154	151*	156*
		%	4.61	3.86	3.86	3.86	4.64	4.59	3.83	3.73	3.76	3.69	3.81
	97.5	#	114	86	93	89	113	114	90	88	90	89	90
		%	2.78	2.10	2.27	2.17	2.76	2.78	2.20	2.15	2.20	2.17	2.20
	99.0	#	72*	50	48	49	73*	71	51	48	49	49	49
		%	1.76	1.22	1.17	1.20	1.78	1.73	1.24	1.17	1.20	1.20	1.20
GARCH	95.0	#	180	158	158	158	184	184	156	153	154	150*	156*
		%	4.39	3.86	3.86	3.86	4.49	4.49	3.81	3.73	3.76	3.66	3.81
	97.5	#	103	86	93	89	103	108	91	88	90	87	89
		%	2.51	2.10	2.27	2.17	2.51	2.64	2.22	2.15	2.20	2.12	2.17
	99.0	#	61	50	47	49	62	61	51	48	49	49	48
		%	1.49	1.22	1.15	1.20	1.51	1.49	1.24	1.17	1.20	1.20	1.17
Quantile Regression (cumulative returns)	95.0	#	255	260	248	259	274	283	255	281	276	273	265
		%	6.22	6.35	6.05	6.32	6.69	6.91	6.22	6.86	6.74	6.66	6.47
	97.5	#	164**	148	147	153	169**	169**	146	163	173*	161*	152
		%	4.00	3.61	3.59	3.73	4.12	4.12	3.56	3.98	4.22	3.93	3.71
	99.0	#	91**	101*	91*	96*	92**	89**	86*	103**	112**	100**	92*
		%	2.22	2.47	2.22	2.34	2.25	2.17	2.10	2.51	2.73	2.44	2.25
Quantile Regression (daily returns)	95.0	#	193	182	178	178	194	199	169	173	180	184	181
		%	4.71	4.44	4.34	4.34	4.74	4.86	4.12	4.22	4.39	4.49	4.42
	97.5	#	114	98	87	89	113	112	89	90	109	105	91
		%	2.78	2.39	2.12	2.17	2.76	2.73	2.17	2.20	2.66	2.56	2.22
	99.0	#	62*	55	49	54	59	59	48	54	56	53	51
		%	1.51	1.34	1.20	1.32	1.44	1.44	1.17	1.32	1.37	1.29	1.24

Note: This table shows the absolute (#) and relative number (%) of violations among 4097 out-of-sample ten trading days ahead VaR predictions for different probability levels α and model classes.

*, ** Hypothesis of accurate unconditional coverage ratio is rejected at 5% and 1% significance level, respectively.

Due to its importance for financial institutions (see, e.g., Basle Committee on Banking Supervision, 1996), we only consider the prediction of VaR at the probability level of 99% for the remainder of Section 4. For the analysis of the other probability levels of 95% and 97.5%, we refer to the robustness check in Section 5.1.

4.2.2 Comparison Within Model Classes

First, we evaluate the VaR prediction performance when adding implied volatility measures as explanatory variables into the covariate and GARCH models. In line with the findings of Giot (2003, 2005) and Kim and Ryu (2015), we find that adding implied volatility measures as explanatory variables yields better VaR prediction than the benchmark model, as indicated by fewer VaR violations in Table 3. All three considered implied volatility measures yield similar VaR prediction performances and the comparative backtest results in Figures 4 and 5 show that all of these measures improve the VaR prediction performance compared to the corresponding benchmark models. Likely due to their high correlation, not only the number of violations but also their comparative backtest results do not differ significantly.

The good performance of the implied volatility measures shows that they are able to describe the conditional variance of returns well. However, this is not the case for the implied skewness and kurtosis: incorporating them yields the highest number of violations for all considered covariate models. Moreover, their VaR forecasts are not better than the benchmark as shown in Figures 4 and 5. As these implied measures appear to be good predictors of the corresponding realized moments in the future (Xiao and Zhou, 2016), we expect that they will improve VaR predictions. However, the linear relationship used to explain the conditional variance does not seem to be suitable for incorporating these implied information. We conduct a robustness check in Section 5.4 to further examine this issue.

Different from incorporating the implied moments, incorporating the implied 10%-quantile as an explanatory variable in the covariate and GARCH model leads to an out-performance of the benchmark at a 6% significance level. Based on the comparative backtests and the number of violations, it does not perform better than the individual implied volatility measures.

Next, we investigate the performance of combining different MEM-implied measures for predicting the VaR at a probability level of 99%. We cannot draw a conclusion solely based on the number of violations shown in Table 3, since they are similar for all considered combinations and also similar to the implied volatility only. The comparative backtests in Figures 4 and 5 provide more insights. All considered combinations outperform the benchmark at a similar significance level. The combinations containing both implied volatility and the implied quantile perform best.

The comparative backtests show that the performance of VaR predictions do not differ sig-

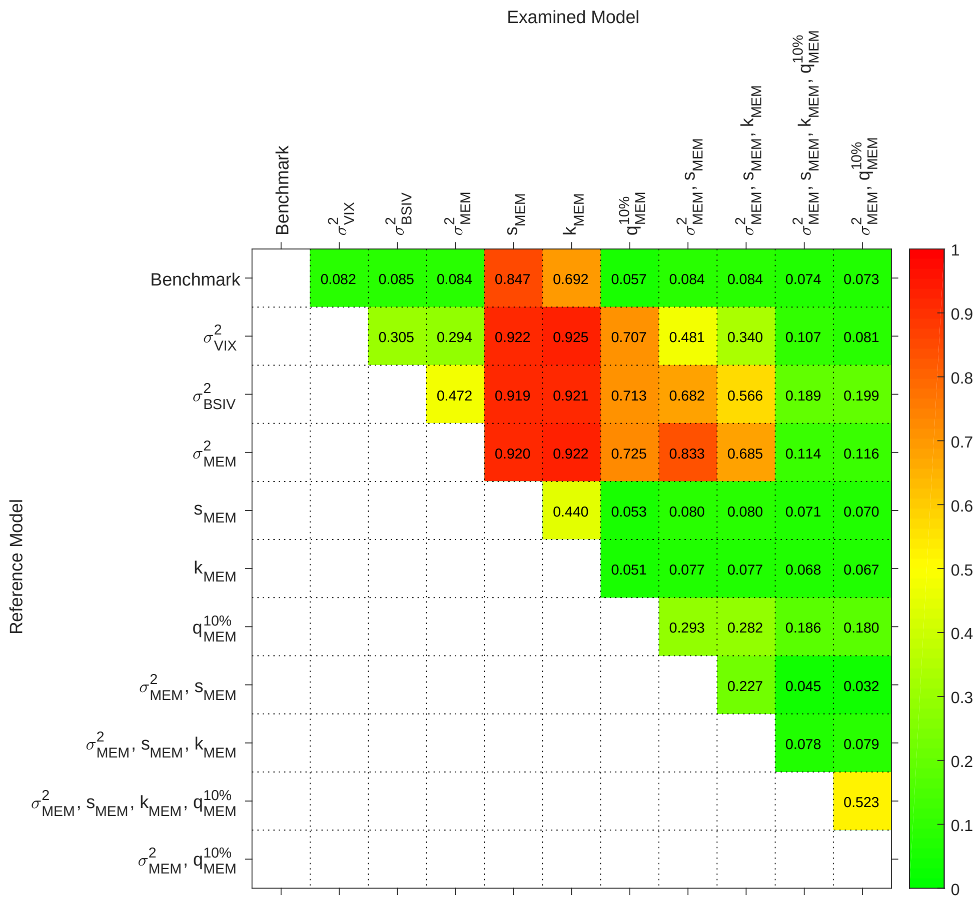

Figure 4: Comparative backtest results of covariate models for the VaR probability level of $\alpha = 99\%$. The plotted values correspond to $\Phi(\Psi)$, which can be used to determine p-values regarding the test hypothesis. If $\Phi(\Psi) \leq p$, then the examined model significantly outperforms the reference model. If $1 - \Phi(\Psi) \leq p$, then the reference model significantly outperforms the examined model.

nificantly between the combination of all implied measures and the combination of only implied volatility and quantile. We investigate this finding further by plotting the estimated coefficients of the covariate model. Figure 6 shows that the estimated model parameters θ_i, corresponding to skewness and kurtosis, do not differ significantly from zero for almost all moving windows over the testing sample period. The estimated coefficients of the GARCH models show a similar pattern. Therefore, these models incorporating all implied measures yield essentially the same forecasts as the models using implied volatility and quantile only, because implied skewness and kurtosis do not provide additional information in this setting.

We conclude that a model with incorporated implied volatility and quantile yields the best VaR prediction performance for both the considered covariate and GARCH model class for the VaR at a 99% probability level. The implied skewness and kurtosis do not contain additional information for VaR prediction on top of the implied volatility and quantile.

Lastly, we investigate the VaR prediction performance within the quantile regression model

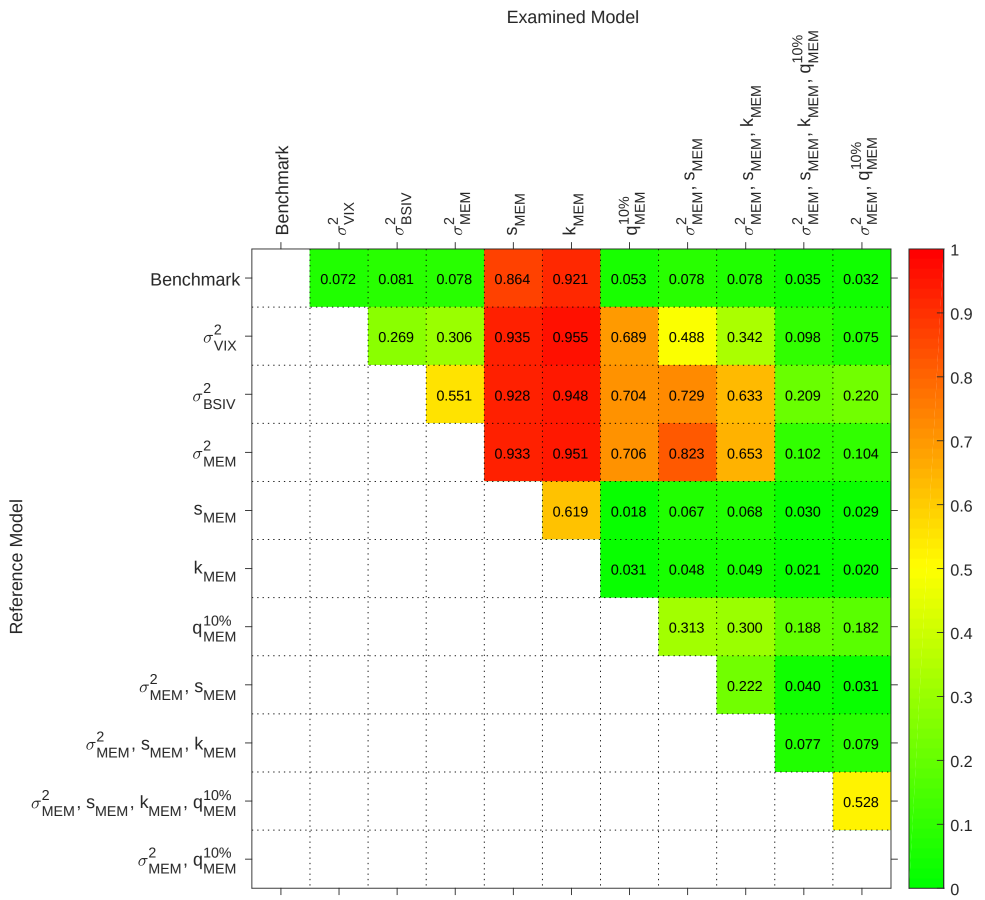

Figure 5: Comparative backtest results of extended GARCH(1,1) model at the probability level of $\alpha = 99\%$. The plotted values correspond to $\Phi(\Psi)$, which can be used to determine p-values regarding the test hypothesis. If $\Phi(\Psi) \leq p$, then the examined model significantly outperforms the reference model. If $1 - \Phi(\Psi) \leq p$, then the reference model significantly outperforms the examined model.

class. The comparative backtest results in Figure 7 confirm the significantly better performance of using daily returns compared to cumulative returns for each implied measure incorporated; see also the unconditional coverage ratio test results in Table 3. Although applying the square-root rule to scale the VaR forecast to the target horizon may not be theoretically correct, it yields better VaR forecasts in our empirical study. Therefore, we focus on this alternative approach that scales the daily VaR to 10-day VaR in the remainder of this paper.

Comparing the prediction performance of incorporating different implied measures by the comparative backtest results in Figure 8, we find a similar pattern as for covariate and GARCH models. Incorporating each of the different implied volatility measures and implied quantile individually outperforms the benchmark model. For VaR at the 99% probability level, all considered implied volatility measures perform similarly well. In contrast, incorporating implied skewness and kurtosis leads to a worse performance compared to the other implied measures. When incorporating the implied measures simultaneously, the combinations containing implied skewness and kurtosis

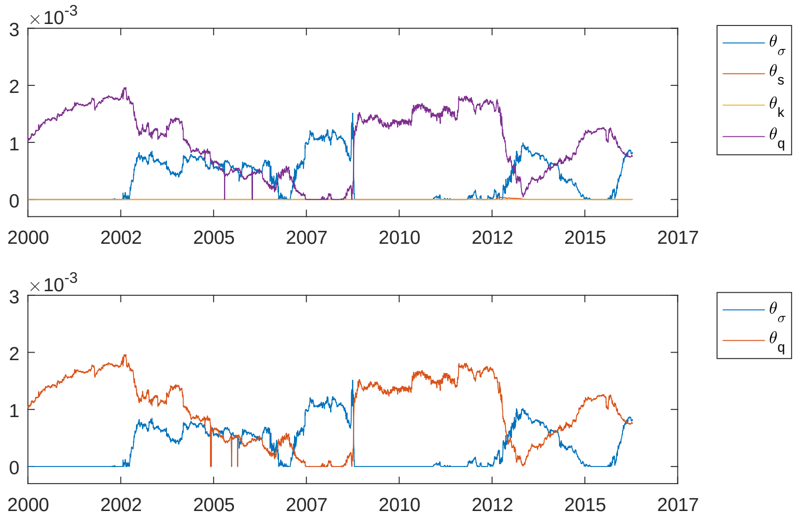

Figure 6: Estimated model parameters corresponding to the MEM-implied measures in covariate models. The upper plot corresponds to the model combining all implied measures and the lower one to a combination of implied volatility and quantile.

perform worst. Incorporating the combination of implied volatility and quantile appears to be the best, but does not perform significantly better than individually incorporating implied volatility or quantile measures.

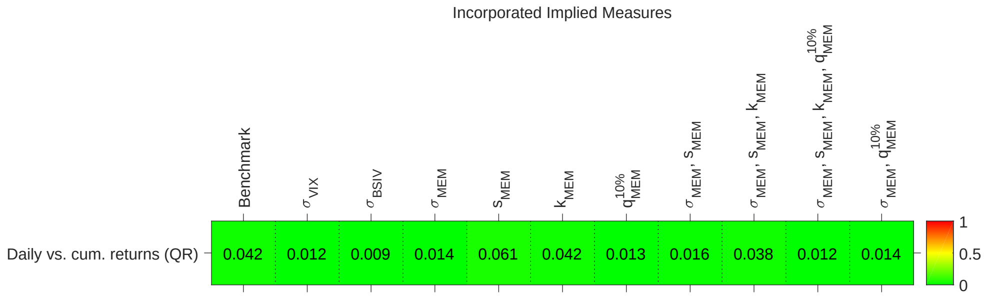

Figure 7: Comparative backtest results for comparing the prediction performance between using daily and cumulative returns in the simple quantile regression (QR) model at the probability level of $\alpha = 99\%$. The plotted values correspond to $\Phi(\Psi)$, which can be used to determine p-values regarding the test hypothesis. If $\Phi(\Psi) \leq p$, then the examined model significantly outperforms the reference model. If $1 - \Phi(\Psi) \leq p$, then the reference model significantly outperforms the examined model. The model using daily returns represents the examined model and the one using cumulative returns the reference model.

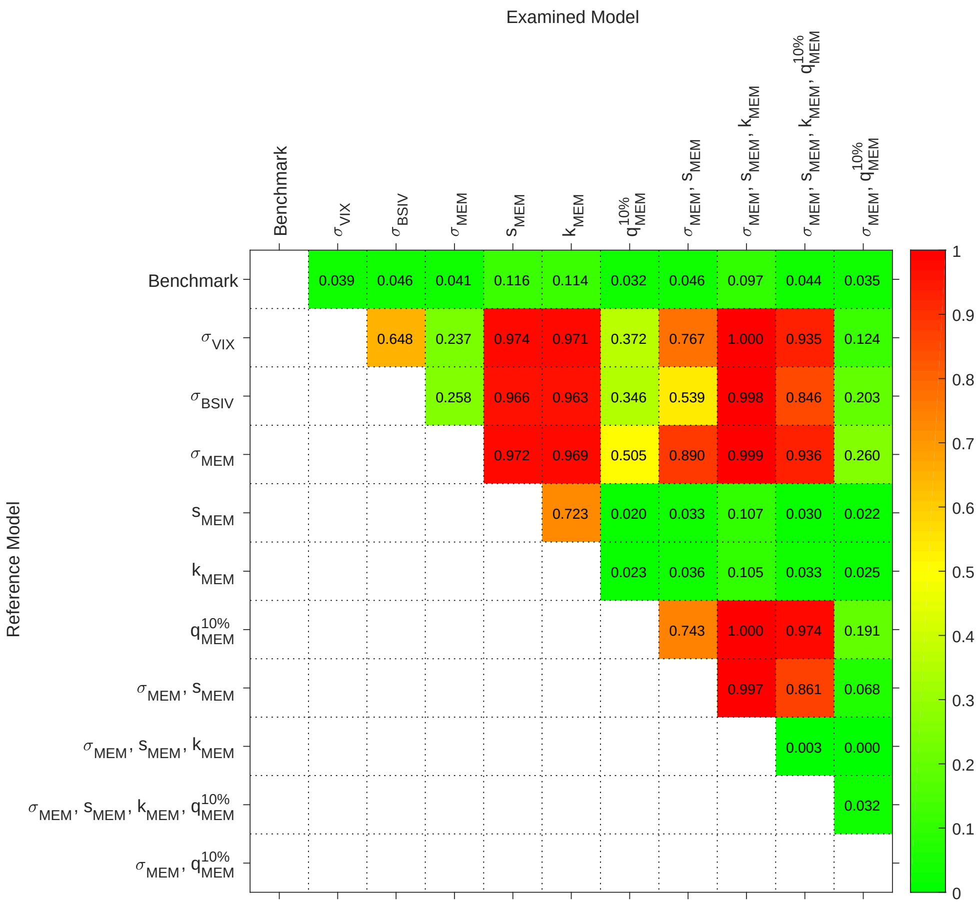

Figure 8: Comparative backtest results of the simple quantile regression model at the probability level of $\alpha = 99\%$. The estimated conditional quantiles are based on daily returns, which are then scaled to the VaR target horizon. The plotted values correspond to $\Phi(\Psi)$, which can be used to determine p-values regarding the test hypothesis.

4.2.3 Comparison Across Model Classes

To investigate which model class benefits the most from incorporating implied measures when forecasting the VaR at a probability level of 99%, we compare the VaR prediction performances of the best performing model of each class. That means we analyze the models that use the combination of MEM-implied volatility and quantile as explanatory variables. Even though all three models have similar low numbers of VaR violations, their performance differs when applying the comparative backtest procedure. The corresponding results are shown in Figure 9.

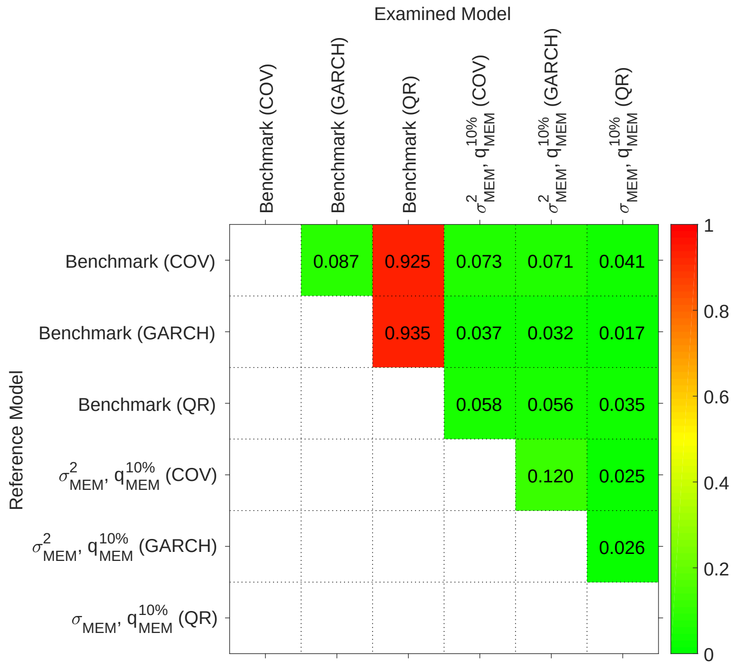

Figure 9: Comparative backtest results comparing the best performing models of each model class and their corresponding benchmarks at the VaR probability level of $\alpha = 99\%$. These covariate (COV), GARCH and simple quantile regression (QR) models combine MEM-implied volatility and quantile as explanatory variables and are based on daily returns. The plotted values correspond to $\Phi(\Psi)$, which can be used to determine p-values regarding the test hypothesis. If $\Phi(\Psi) \leq p$, then the examined model significantly outperforms the reference model. If $1 - \Phi(\Psi) \leq p$, then the reference model significantly outperforms the examined model.

We find that all models incorporating the implied volatility and quantile significantly outperform all benchmark models at a significance level of 10%. Hence, incorporating option-implied information can improve the performance of established models solely based on historical return information.

It should be noted that the performance of the different benchmark models is in line with the existing results for VaR prediction, see Kuester et al. (2006) and Bao et al. (2006). Quantile regression models based on historical returns perform worse than GARCH models. Additionally, the GARCH(1,1) model outperforms the covariate benchmark model, which is probably due to the capability to account for volatility clustering.

Comparing across different model classes that incorporate MEM-implied information, we find that the extended GARCH model performs better than the extended covariate model. This relation is similar to the benchmark comparison, but only at a significance level of 12%. It is notable that the extended quantile regression model performs the best, as it outperforms the extended covariate and GARCH models at a 3% significance level.

A possible reason for the performance differences is that the GARCH model predictions do not adjust as quickly to large losses and events as the quantile regression model does. This leads to lower VaR predictions by the quantile regression over long time periods, which leads to a better outcome in the comparative backtest. Notice that the two models incorporate the implied measures differently. The quantile regression explains the VaR directly using the implied information. Hence, it can use all information contained in the implied measures that is relevant for tail estimation. By contrast, the GARCH model only takes into account the information useful for conditional variance estimation. Additionally, the GARCH model is a parametric model, which is not as flexible as the non-parametric quantile regression model. Hence, we conclude that for VaR forecasting the simple quantile regression model is better suited to incorporate the information contained in the option-implied measures and yields a superior VaR prediction performance.

5 Robustness Analysis

This section investigates the robustness in the results of Section 4 by considering alternative VaR probability levels (Section 5.1), prediction horizons (Section 5.2) and testing periods (Section 5.3). Additionally, we investigate another approach of estimating the quantile of the risk-neutral probability distribution in Section 5.4, which we then use for VaR forecasting. Lastly, Section 5.5 analyzes alternative VaR prediction models. In most of the robustness checks we provide a description of the results without detailed tables and figures. The detailed tables and figures are available upon request.

5.1 Alternative Probability Levels

Section 4.2.2 shows that incorporating implied volatility and quantile yields the best performing VaR predictions. The performance comparison is focused on applying the comparative backtest approach and considering a VaR probability level of 99%. If we consider probability levels of 95% and 97.5%, we obtain a similar pattern as shown in Figures 4, 5 and 8. For lower probability levels, though, the VIX-implied volatility performs better than the other two considered implied volatility measures. Our main conclusion that the combination containing both implied volatility and the implied quantile performs best remains valid for all considered probability levels. Moreover, the result of the model comparison in Section 4.2.3 does not change qualitatively for different VaR probability levels.

5.2 Alternative Prediction Horizons

In the empirical study (Section 4) we focus on a VaR prediction horizon of ten trading days. To investigate whether the results are robust for different horizons, we assess the VaR prediction performance for alternative horizons of 1, 5, 15, 20 and 25 trading days. For all considered prediction horizons, we use the same training and testing sample. Note that the number of out-of-sample predictions differs per horizon, but includes the same trading days.

Firstly, we compare the prediction performance of models incorporating implied volatility and quantile to their corresponding benchmark models, which are solely based on historical return information. Figure 10 shows that adding implied information is useful for most considered prediction horizons. The level of outperformance relative to the benchmark model decreases as the prediction horizon increases. The implied volatility and quantile do not add significant value to VaR predictions by the GARCH-type models for a horizon longer than one month. Likely due to the poor performance of its benchmark, the quantile regression model is improved the most by incorporating the implied measures. For the other considered probability levels, the results displayed in Figure 10 do not change qualitatively.

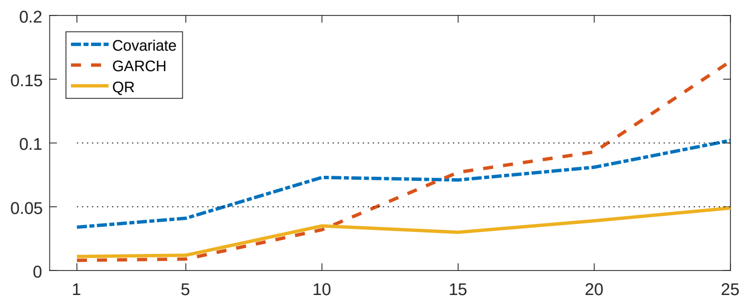

Figure 10: P-values of the comparative backtests when comparing the performance of models using MEM-implied volatility and quantile to their corresponding benchmark models. The horizontal axis indicates different VaR forecast horizons for a probability level of 99%. The test hypothesis is that the model incorporating implied measures is not better than the corresponding benchmark model.

Secondly, when comparing the VaR prediction performance of different model classes for different prediction horizons, Figure 11 shows that the performance of different models becomes more similar for a longer horizon. Similar to the result for a ten trading days horizon, the quantile regression model significantly outperforms the other two models for all considered VaR prediction horizons. This is consistent for all considered probability levels.

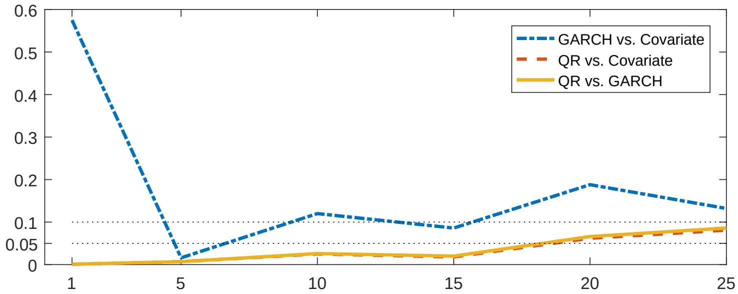

Figure 11: P-values of the comparative backtests when comparing the performance of considered model classes (using MEM-implied volatility and quantile). The horizontal axis indicates different VaR forecast horizons for a probability level of 99%. The first model mentioned indicates the examined model and the second the reference model. The test hypothesis is that the model mentioned first is not better than the second model.

5.3 Alternative Testing Periods

As an alternative testing sample, we consider the periods of US recession, obtained from the National Bureau of Economic Research (NBER). We compare the VaR prediction performance of the different models during these periods. The results are shown in Figure 12. Our qualitative conclusion remains valid: the quantile regression model incorporating implied volatility and quantile outperforms the other models when only evaluating the VaR forecasts during the periods of recession. However, the p-value increases slightly when considering lower VaR probability levels.

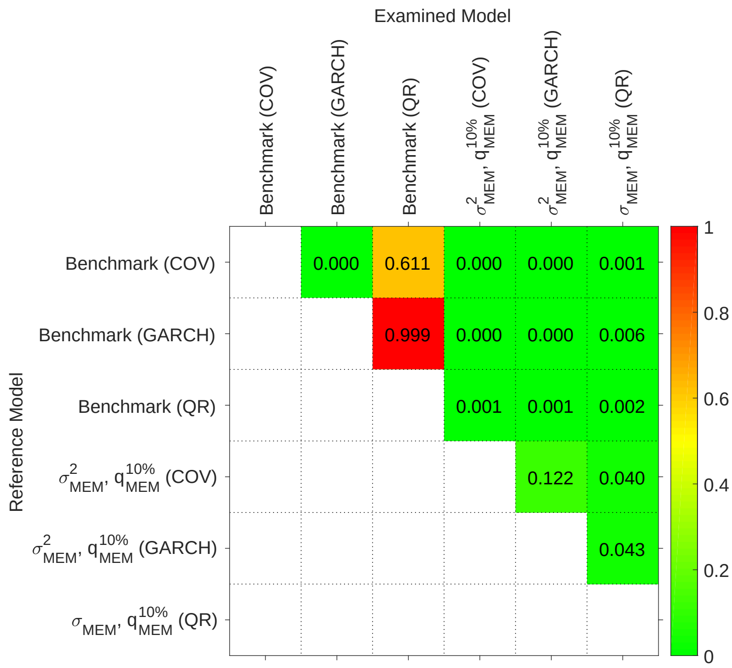

Figure 12: Comparative backtest results comparing the best performing models of each model class and their corresponding benchmarks at the VaR probability level of $\alpha = 99\%$ during periods of recessions according to NBER. These covariate (COV), GARCH and simple quantile regression (QR) models combine MEM-implied volatility and quantile as explanatory variables and are based on daily returns. The plotted values correspond to $\Phi(\Psi)$, which can be used to determine p-values regarding the test hypothesis. If $\Phi(\Psi) \leq p$, then the examined model significantly outperforms the reference model. If $1 - \Phi(\Psi) \leq p$, then the reference model significantly outperforms the examined model.

5.4 Alternative Implied Quantile Estimation

Section 4.2.2 shows that using implied skewness and kurtosis leads to the worst performance in VaR predictions. These implied measures are linearly incorporated into the different VaR prediction models. A possible reason for the poor performance is that the linear structure is not an efficient way to employ the implied skewness and kurtosis. As a robustness check, we investigate an alternative approach, which utilizes implied skewness and kurtosis in a structural way. We first estimate the quantile of a parametric risk-neutral probability distribution based on implied skewness and kurtosis. Then we incorporate the obtained parametric implied quantile into the models. With such a structural approach, the implied skewness and kurtosis are incorporated into the VaR prediction models in a non-linear way. This could potentially improve the VaR prediction.

We opt for the skewed t-distribution as the parametric model for the risk-neutral distribution, since we require a unimodal parametric distribution that allows for skewness and heavy tails. Applying the method of moments, we fit the skewed t-distribution to the moments of the risk-neutral probability distribution estimated from the MEM. This method uses the MEM-implied moments to obtain estimates for parameters ν and ξ of the skewed t-distribution. Having fitted this distribution, we derive its quantile, which we use as an explanatory variable in VaR prediction models. These steps are explained in more detail in Appendix A.2.

The results of the comparative backtest, displayed in Figure 13, show that firstly, using this parametric implied quantile as explanatory variable does not significantly improve the VaR prediction performance, compared to solely using historical return information for any model class. Secondly, compared to the implied quantile directly obtained from the MEM, it performs worse at a 6% significance level for all considered models.

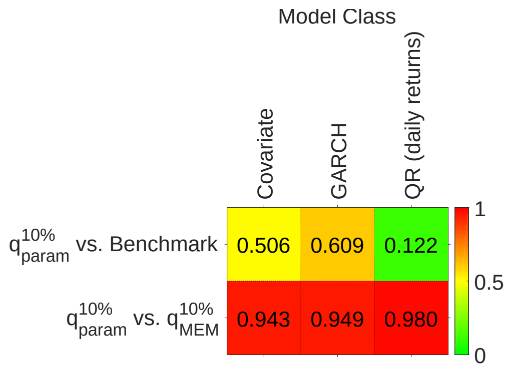

Figure 13: Comparative backtest results of using the parametric implied quantile in the considered model classes at the probability level of $\alpha = 99\%$. The corresponding VaR prediction performance is compared to the benchmarks and the MEM-implied quantiles. The plotted values correspond to $\Phi(\Psi)$, which can be used to determine p-values regarding the test hypothesis. If $\Phi(\Psi) \leq p$, then the examined model significantly outperforms the reference model. If $1 - \Phi(\Psi) \leq p$, then the reference model significantly outperforms the examined model. The first model mentioned ($q^{10\%}_{\text{param}}$) indicates the examined model and the second one represents the reference model.

Hence, we conclude that the worse performance of implied skewness and kurtosis, shown in Section 4.2.2, is not caused by the linear incorporation of these measures into the VaR prediction models. The results of the superior performance of the MEM-implied quantile stems from its more relevant information content for VaR prediction. By contrast, the information content of implied skewness and kurtosis is genuinely not superior to historical return information in terms of VaR prediction performance.

5.5 Alternative Value-at-Risk Prediction Models

Additional to the standard GARCH and quantile regression model discussed in this paper, one may also consider extensions of these models that are often used for VaR forecasting (see, e.g., Kuester et al., 2006; Bao et al., 2006). We opt for an extension of the GJR-GARCH (Glosten et al., 1993) and the CAViaR model (Engle and Manganelli, 2004) and compare their VaR prediction performance when incorporating implied measures to the results of the corresponding base models discussed in Section 4.

5.5.1 GJR-GARCH

The GJR-GARCH model distinguishes between the impact of negative and positive innovations to returns, ε_{t-1}, on the conditional variance h_t. An intuition Glosten et al. (1993) provide is that a negative innovation to today's stock return will change the capital structure of the corresponding firm by increasing its leverage. This results in a higher expected variance in the future and can be accounted for by this model. We consider the following model specification:

$$\begin{aligned} r_t &= \mu + \sqrt{h_t}\varepsilon_t, && \text{for } t = 2,\dots,T, \\ h_t &= \omega + \left(\gamma + \rho \mathscr{I}_{\{\varepsilon_{t-1}<0\}}\right) h_{t-1}\varepsilon_{t-1}^2 + \delta h_{t-1} + \theta x_{t-1}, \end{aligned}$$

where $\mathscr{I}_{\{\varepsilon_{t-1}<0\}} = 1$ if $\varepsilon_{t-1} < 0$ and 0 otherwise. To ensure positive conditional variance and stationarity, we impose the following parameter restrictions: $\omega > 0$ and γ, ρ, $\delta, \theta_i \geq 0$, for all $i = 1,\dots,N$, and $\gamma + \rho \Pr(\varepsilon_{t-1} < 0) + \delta < 1$. In other words, we scale ρ to take into account how many residuals are expected to be negative. With commonly used symmetric distributions, such as normal or standard t-distribution, this scaling factor is $1/2$. However, since we allow for non-zero skewness, we have to estimate the proportion of negative residuals and therefore, use the sample residuals for that purpose.

We find that this extended GJR-GARCH(1,1) model yields a similar VaR prediction performance pattern but does not perform better compared to the discussed GARCH model (see Figure 14 and the upper panel of Table 4). Our conclusion that incorporating implied volatility and quantile yields the best VaR prediction performance remains qualitatively unchanged. The reason for these similar results is that the parameter estimate corresponding to the lagged residual is negligibly small so that it only has a minor effect on the VaR prediction performance of the considered model. Therefore, the VaR predictions become almost identical to the ones of the considered GARCH(1,1) models.

5.5.2 CAViaR

The second model we apply is the general CAViaR specification, which is often used for VaR predictions (see, e.g., Kuester et al., 2006). The standard CAViaR model uses the estimated quantile from the previous period as explanatory variable. We add implied measures as extra explanatory variables. Hence, the conditional quantile at time t is given as

$$Q_{r_{t:t+h}|x_t}(1-\alpha) = \mu + \beta_0\, Q_{r_{t-1:t-1+h}|x_{t-1}}(1-\alpha) + \beta x_t,$$

where $\beta_0 \in \mathbb{R}$ is an additional model parameter.

Similar to the quantile regression model, we apply the alternative approach of using daily returns when comparing the prediction performance across the different considered quantile regression models. The comparison of VaR violations of simple quantile regression and CAViaR model does not clearly indicate which model performs better (see the lower panel of Table 4). However, the comparative backtest results in Figure 14 show that the CAViaR does not perform better than the simple quantile regression model when incorporating any implied measure. This conclusion is consistent for all considered probability levels. Hence, the lagged conditional quantile in the CAViaR specification does not lead to a better VaR prediction. A possible explanation for this is that we do not perform iterative multi-period out-of-sample forecasts to predict the VaR. The CAViaR procedure might benefit from the additional lagged quantile. This would require a time series mode

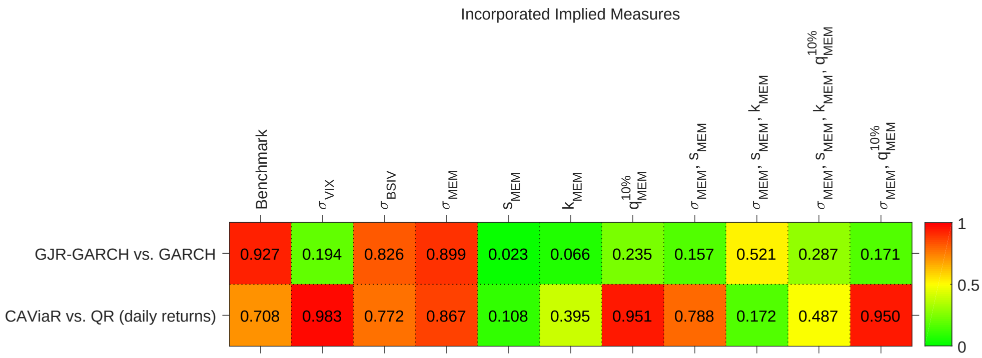

Figure 14: Comparative backtest results of comparing prediction performance of GJR-GARCH and GARCH as well as CAViaR and quantile regression (QR) model using daily returns for a probability level of $\alpha = 99\%$. The plotted values correspond to $\Phi(\Psi)$, which can be used to determine p-values regarding the test hypothesis. If $\Phi(\Psi) \leq p$, then the examined model significantly outperforms the reference model. If $1 - \Phi(\Psi) \leq p$, then the reference model significantly outperforms the examined model. The first model mentioned (i.e. GJR-GARCH and CAViaR, respectively) indicates the examined model and the second one represents the reference model.

Table 4: Value-at-Risk violations of estimated models

				Incorporated Implied Measures									
	α in %		Benchmark	σ_{VIX}	σ_{BS}	σ_{MEM}	s_{MEM}	k_{MEM}	$q^{10\%}_{MEM}$	σ_{MEM}, s_{MEM}	$\sigma_{MEM}, s_{MEM}, k_{MEM}$	$\sigma_{MEM}, s_{MEM}, k_{MEM}, q^{10\%}_{MEM}$	$\sigma_{MEM}, q^{10\%}_{MEM}$
GJR-GARCH	95.0	#	180	158	158	158	181	180	156	153	154	150*	156*
		%	4.39	3.86	3.86	3.86	4.42	4.39	3.81	3.73	3.76	3.66	3.81
	97.5	#	103	86	93	89	101	106	91	88	90	87	89
		%	2.51	2.10	2.27	2.17	2.47	2.59	2.22	2.15	2.20	2.12	2.17
	99.0	#	61	50	47	49	61	59	51	48	49	49	48
		%	1.49	1.22	1.15	1.20	1.49	1.44	1.24	1.17	1.20	1.20	1.17
CAViaR (daily returns)	95.0	#	173	172	175	169	203	220	173	170	180	178	166
		%	4.22	4.20	4.27	4.12	4.95	5.37	4.22	4.15	4.39	4.34	4.05
	97.5	#	110	102	92	94	110	131	90	95	101	98	90
		%	2.68	2.49	2.25	2.29	2.68	3.20	2.20	2.32	2.47	2.39	2.20
	99.0	#	61	65	56	58	59	74	53	57	66	62	51
		%	1.49	1.59	1.37	1.42	1.44	1.81	1.29	1.39	1.61	1.51	1.24

Note: This table shows the absolute (#) and relative number (%) of violations of 4097 out-of-sample ten trading days ahead VaR predictions for different probability levels α and model classes.

*, ** Hypothesis of accurate unconditional coverage ratio is rejected at 5% and 1% significance level, respectively.

6 Conclusion

This paper investigates the impact of incorporating option-implied information into different VaR prediction models on their prediction performance. To obtain option-implied measures, we apply the MEM, which estimates the risk-neutral probability distribution of future returns based on the market prices of options. We compute the implied moments and quantiles of the risk-neutral distribution and incorporate them into three different types of VaR prediction models: covariate, GARCH and quantile regression models. We consider using the implied measures individually or in combination as explanatory variables and predict the VaR for different target horizons and probability levels. To evaluate the forecasting performance, we use the comparative backtest procedure of Nolde and Ziegel (2017).

For all considered models, we find that adding implied volatility yields a significantly better performance than the benchmark model. Compared to the two conventional implied volatility measures VIX and BSIV, the MEM-implied volatility performs equally well. Using the implied 10%-quantile of the risk-neutral distribution as explanatory variable yields a similar performance. Combining both MEM-implied volatility and quantile yields the best VaR prediction performance for all model classes, which is consistent for all considered probability levels.

Comparing the prediction performance across different model classes, we find that the simple quantile regression model yields the best VaR forecasts and is best suited to incorporate option-implied information. This is in contrast to the findings in existing studies based on historical returns that the GARCH model class performs best (see, e.g., Kuester et al., 2006; Bao et al., 2006) [8]. This is possibly due to the fact that the quantile regression model does not model the whole return distribution, but only focuses on the quantile conditional on the implied measures.

Due to the information content contained in the MEM-implied measures and the estimated risk-neutral distribution, further research may investigate other applications using such information. Furthermore, the models used in this paper do not allow time-varying skewness and kurtosis. By further modeling the the time variation of skewness and kurtosis, the corresponding implied measures may help to improve the forecast of these characteristics. In turn, it may further improve the VaR prediction.

A major limitation of our approach and the MEM-algorithm is the dependence on available option data. We use options on the S&P 500, which are among the most liquid options. However, even with these frequently traded options, the range of strike prices that can be used for the MEM is limited. Hence, for assets with less liquid options the number of suitable options and sufficient strike prices is lower. Although MEM does not require many options (Xiao and Zhou, 2016), the approach of incorporating the corresponding implied measures based on fewer options could lead to a worse prediction performance than shown in the current study.

Due to the strong performance of incorporating MEM-implied measures into VaR forecasting models, the proposed approach is a beneficial supplement to the commonly used risk estimation methods. However, as our approach heavily depends on available option data, it should not substitute methods that are based on historical returns, but may provide additional insights for risk assessment.

[8] One limitation of this comparison should be noted. Kuester et al. (2006) conclude that a GARCH model combined with EVT performs best. We also analyzed the performance of applying EVT to our models. Even though the number of VaR violations is closer to the target probability level, the comparative backtest indicated a lower performance of the EVT method. The corresponding results are available upon request.

References

Agmon, N., Alhassid, Y., and Levine, R. (1979). An algorithm for finding the distribution of maximal entropy. *Journal of Computational Physics*, 30(2):250 – 258.

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

Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, 59(3):817–858.

Bakshi, G. and Kapadia, N. (2003). Delta-hedged gains and the negative market volatility risk premium. *The Review of Financial Studies*, 16(2):527.

Bao, Y., Lee, T.-H., and Saltoglu, B. (2006). Evaluating predictive performance of Value-at-Risk models in emerging markets: a reality check. *Journal of Forecasting*, 25(2):101–128.

Barone Adesi, G. (2016). VaR and CVaR implied in option prices. *Journal of Risk and Financial Management*, 9(1).

Basle Committee on Banking Supervision (1996). *Overview of the amendment to the capital accord to incorporate market risks*. Bank for International Settlements.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307 – 327.

Buchen, P. W. and Kelly, M. (1996). The maximum entropy distribution of an asset inferred from option prices. *The Journal of Financial and Quantitative Analysis*, 31(1):143–159.

Campbell, J. Y. and Viceira, L. M. (2002). *Strategic asset allocation: portfolio choice for long-term investors*. Clarendon Lectures in Economic.

Carr, P. and Wu, L. (2006). A tale of two indices. *The Journal of Derivatives*, 13(3):13–29.

Chen, M.-Y. and Chen, J.-E. (2005). Application of quantile regression to estimation of Value-at-Risk. *Review of Financial Risk Management*, pages 1–15.

Chiras, D. P. and Manaster, S. (1978). The information content of option prices and a test of market efficiency. *Journal of Financial Economics*, 6(2):213 – 234.

Christensen, B. and Prabhala, N. (1998). The relation between implied and realized volatility. *Journal of Financial Economics*, 50(2):125 – 150.

Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review*, 39(4):841–862.

Diebold, F., Hickman, A., Inoue, A., and Schuermann, T. (1997). Converting 1-day volatility to h-day volatitlity: Scaling by root-h is worse than you think. Center for financial institutions working papers, Wharton School Center for Financial Institutions, University of Pennsylvania.

Diebold, F. X., Gunther, T. A., and Tay, A. S. (1998). Evaluating density forecasts. *International Economic Review*, 39:863–883.

Engle, R. and Mezrich, J. (1996). Garch for groups: A round-up of recent developments in garch techniques for estimating correlation. *Risk*, 9:36–40.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50(4):987–1007.

Engle, R. F. and Manganelli, S. (2004). CAViaR. *Journal of Business & Economic Statistics*, 22(4):367–381.

Fernández, C. and Steel, M. F. J. (1998). On bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association*, 93(441):359–371.

Francq, C., Horváth, L., and Zakoïan, J.-M. (2011). Merits and drawbacks of variance targeting in GARCH models. *Journal of Financial Econometrics*, 9(4):619–656.

Giot, P. (2003). The information content of implied volatility in agricultural commodity markets. *Journal of Futures Markets*, 23(5):441–454.

Giot, P. (2005). Implied volatility indexes and daily Value-at-Risk models. *The Journal of Derivatives*, 12(4):56–64.

Giot, P. and Laurent, S. (2003). Value-at-Risk for long and short trading positions. *Journal of Applied Econometrics*, 18(6):641–663.

Giot, P. and Laurent, S. (2004). Modelling daily Value-at-Risk using realized volatility and ARCH type models. *Journal of Empirical Finance*, 11(3):379 – 398.

Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, 48(5):1779–1801.

Gulko, L. (1999). The entropy theory of stock option pricing. *International Journal of Theoretical and Applied Finance*, 02(03):331–355.

Gulko, L. (2002). The entropy theory of bond option pricing. *International Journal of Theoretical and Applied Finance*, 05(04):355–383.

Guo, D. (1998). The risk premium of volatility implicit in currency options. *Journal of Business & Economic Statistics*, 16(4):498–507.

Huang, X. (2008). Mean-entropy models for fuzzy portfolio selection. *IEEE Transactions on Fuzzy Systems*, 16(4):1096–1101.

Hull, J. (2012). *Options, Futures, and Other Derivatives (Eighth Edition)*. Prentice Hall.

Jaynes, E. T. (1957). Information theory and statistical mechanics. *Physical Review*, 106:620–630.

Jeon, J. and Taylor, J. W. (2013). Using CAViaR models with implied volatility for Value-at-Risk estimation. *Journal of Forecasting*, 32(1):62–74.

Kim, J. S. and Ryu, D. (2015). Are the KOSPI 200 implied volatilities useful in Value-at-Risk models? *Emerging Markets Review*, 22(C):43–64.

Koenker, R. and Bassett, G. (1978). Regression quantiles. *Econometrica*, 46(1):33–50.

Krause, A. (2003). Exploring the limitations of Value-at-Risk: How good is it in practice? *The Journal of Risk Finance*, 4(2):19–28.

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.

Lambert, P. and Laurent, S. (2001). Modelling financial time series using garch-type models with a skewed student distribution for the innovations. *Discussion Paper 01-25, Institut de Statistique, Université Catholique de Louvain, Louvain-la-Neuve, Belgium*.

Latané, H. A. and Rendleman, R. J. (1976). Standard deviations of stock price ratios implied in option prices. *The Journal of Finance*, 31(2):369–381.

Longin, F. M. (2000). From Value-at-Risk to stress testing: The extreme value approach. *Journal of Banking & Finance*, 24(7):1097–1130.

Louzis, D. P., Xanthopoulos-Sisinis, S., and Refenes, A. P. (2013). The role of high-frequency intra-daily data, daily range and implied volatility in multi-period Value-at-Risk forecasting. *Journal of Forecasting*, 32(6):561–576.

McNeil, A. J. and Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of empirical finance*, 7(3):271–300.

Meucci, A. (2010). Quant nugget 4: annualization and general projection of skewness, kurtosis and all summary statistics. *Working Paper (available at SSRN: https://ssrn.com/abstract=1635484)*.

Nolde, N. and Ziegel, J. (2017). Elicitability and backtesting: Perspectives for banking regulation. *Annals of Applied Statistics*.

Odening, M. and Hinrichs, J. (2003). Using extreme value theory to estimate ValueatRisk. *Agricultural Finance Review*, 63(1):55–73.

Ou, J. (2005). Theory of portfolio and risk based on incremental entropy. *The Journal of Risk Finance*, 6(1):31–39.

Philippatos, G. C. and Wilson, C. J. (1972). Entropy, market risk, and the selection of efficient portfolios. *Applied Economics*, 4(3):209–220.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423.

Süli, E. and Mayers, D. (2003). *An Introduction to Numerical Analysis*. Cambridge University Press.

Xiao, X. and Zhou, C. (2016). Entropy-based implied volatility and its information content. *Working Paper (available at SSRN: http://dx.doi.org/10.2139/ssrn.2667489)*.

Zhou, R., Cai, R., and Tong, G. (2013). Applications of entropy in finance: A review. *Entropy*, 15(11):4909 – 4931.

Appendix

A Skewed t-Distribution

A.1 Probability Density Function

The probability density function of a skewed t distribution with mean zero and unit variance, $\varepsilon_t \sim skt(0,1,\xi,\nu)$, is given by

$$f^{skt}(\varepsilon_t \mid \nu,\xi) = \begin{cases} \frac{2s}{\xi+\frac{1}{\xi}} f^t\left[\xi(s\varepsilon_t+m) \mid \nu\right], & \text{if} \quad \varepsilon_t < -\frac{m}{s} \\ \frac{2s}{\xi+\frac{1}{\xi}} f^t\left[\frac{1}{\xi}(s\varepsilon_t+m) \,\middle|\, \nu\right], & \text{if} \quad \varepsilon_t \geq -\frac{m}{s} \end{cases},$$

where $m = \frac{\Gamma\left(\frac{\nu-1}{2}\right)\sqrt{\nu-2}}{\sqrt{\pi}\Gamma\left(\frac{\nu}{2}\right)}\left(\xi - \frac{1}{\xi}\right)$, $s = \sqrt{\left(\xi^2 + \frac{1}{\xi^2} - 1\right) - m^2}$ and $f^t(x \mid \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi(\nu-2)}\Gamma\left(\frac{\nu}{2}\right)}\left(1 + \frac{x^2}{\nu-2}\right)^{-\frac{\nu+1}{2}}$ is the density of the standardized t-distribution with $\nu > 2$ degrees of freedom (see, e.g., Giot and Laurent, 2003). Here $\xi \in \mathbb{R}^+$ is the asymmetry coefficient of the distribution, which describes its skewness and is defined as $\xi^2 = \frac{\Pr(\varepsilon_t \geq 0)}{\Pr(\varepsilon_t < 0)}$. Hence, if $\xi > 1$ ($\xi < 1$), the probability mass above (below) zero is greater, which corresponds to positive (negative) skewness. For the symmetric case, $\xi = 1$, the skewed t-distribution simplifies to the standard t-distribution.

A.2 Method of Moments to Calculate Implied Quantile

To calculate the moments of a skewed t-distribution, we follow Fernández and Steel (1998), who show that the r-th raw moment of the skewed t-distribution is given by

$$\mathrm{E}(\varepsilon_t^r \mid \nu,\xi) = M_{r\mid\nu}\,\frac{\xi^{r+1} + \frac{(-1)^r}{\xi^{r+1}}}{\xi + \frac{1}{\xi}},$$

where $M_{r\mid\nu}$ is the r-th non-central moment of the standardized t-distribution truncated to positive values, i.e.

$$M_{r\mid\nu} = \int_0^\infty 2\,\varepsilon_t^r\, f^t(\varepsilon_t \mid \nu)\,\mathrm{d}\varepsilon_t = \frac{\Gamma\left(\frac{\nu-r}{2}\right)\Gamma\left(\frac{1+r}{2}\right)(\nu-2)^{\frac{1+r}{2}}}{\sqrt{\pi(\nu-2)}\,\Gamma\left(\frac{\nu}{2}\right)},$$

see, e.g., Lambert and Laurent (2001). Given the raw moments, we can derive skewness and kurtosis of the skewed t-distribution as functions of the distribution parameters ν and ξ.

We estimate ν and ξ by setting the skewed t-distribution's skewness and kurtosis equal to the corresponding MEM-implied measures and solving the resulting equation system with the assumption that the fourth moment exists, i.e. $\nu > 4$. Having obtained parameter estimates for ν and ξ, we scale and shift the skewed t-distribution using the MEM-implied volatility and mean, respectively.

Note that, we obtain the MEM-implied measures used in this approach by following the procedure in Section 2.2. However, we do not annualize the implied measures, since in the method of moments context we aim at estimating the risk-neutral distribution of returns until option maturity.

After obtaining the estimated skewed t-distribution by fitting to the MEM moments, we follow the same approach as in Section 4.1 to obtain the 10%-quantile of the risk-neutral distribution over the following 21 trading days. This means, we apply the described method of moments on two sets of options and obtain the corresponding risk-neutral probability distributions and their 10%-quantiles. We interpolate these quantiles to the target of 21 trading days ahead and use the resulting implied quantile as an explanatory variable in the VaR prediction models of Section 3.

Previous DNB Working Papers in 2018

No. 583 **Dorinth van Dijk, David Geltner and Alex van de Minne**, Revisiting supply and demand indexes in real estate
No. 584 **Jasper de Jong**, The effect of fiscal announcements on interest spreads: Evidence from the Netherlands
No. 585 **Nicole Jonker**, What drives bitcoin adoption by retailers?
No. 586 **Martijn Boermans and Robert Vermeulen,** Quantitative easing and preferred habitat investors in the euro area bond market
No. 587 **Dennis Bonam, Jakob de Haan and Duncan van Limbergen,** Time-varying wage Phillips curves in the euro area with a new measure for labor market slack
No. 588 **Sebastiaan Pool**, Mortgage debt and shadow banks
No. 589 **David-Jan Jansen**, The international spillovers of the 2010 U.S. flash crash
No. 590 **Martijn Boermans and Viacheslav Keshkov**, The impact of the ECB asset purchases on the European bond market structure: Granular evidence on ownership concentration
No. 591 **Katalin Bodnár, Ludmila Fadejeva, Marco Hoeberichts, Mario Izquierdo Peinado, Christophe Jadeau and Eliana Viviano**, Credit shocks and the European labour market
No. 592 **Anouk Levels, René de Sousa van Stralen, Sînziana Kroon Petrescu and Iman van Lelyveld**, CDS market structure and risk flows: the Dutch case
No. 593 **Laurence Deborgies Sanches and Marno Verbeek**, Basel methodological heterogeneity and banking system stability: The case of the Netherlands
No. 594 **Andrea Colciago, Anna Samarina and Jakob de Haan**, Central bank policies and income and wealth inequality: A survey
No. 595 **Ilja Boelaars and Roel Mehlkopf**, Optimal risk-sharing in pension funds when stock and labor markets are co-integrated
No. 596 **Julia Körding and Beatrice Scheubel**, Liquidity regulation, the central bank and the money market
No. 597 **Guido Ascari, Paolo Bonomolo and Hedibert Lopes**, Walk on the wild side: Multiplicative sunspots and temporarily unstable paths
No. 598 **Jon Frost and René van Stralen,** Macroprudential policy and income inequality
No. 599 **Sinziana Kroon and Iman van Lelyveld**, Counterparty credit risk and the effectiveness of banking regulation
No. 600 **Leo de Haan and Jan Kakes**, European banks after the global financial crisis: Peak accumulated losses, twin crises and business models
No. 601 **Bahar Öztürk, Dorinth van Dijk, Frank van Hoenselaar and Sander Burgers**, The relation between supply constraints and house price dynamics in the Netherlands
No. 602 **Ian Koetsier and Jacob Bikker**, Herding behavior of Dutch pension funds in asset class investments
No. 603 **Dirk Broeders and Leo de Haan,** Benchmark selection and performance
No. 604 **Melanie de Waal, Floor Rink, Janka Stoker and Dennis Veltrop**, How internal and external supervision impact the dynamics between boards and Top Management Teams and TMT reflexivity
No. 605 **Clemens Bonner, Eward Brouwer and Iman van Lelyveld,** Drivers of market liquidity - Regulation, monetary policy or new players?
No. 606 **Tanja Artiga Gonzalez, Iman van Lelyveld and Katarina Lucivjanska,** Pension fund equity performance: Patience, activity or both?
No. 607 **Jasper de Jong en Niels Gilbert,** Fiscal discipline in EMU? Testing the effectiveness of the Excessive Deficit Procedure
No. 608 **Carin van der Cruijsen, Maurice Doll and Frank van Hoenselaar,** Trust in other people and the usage of peer platform markets
No. 609 **Jon Frost, Patty Duijm, Clemens Bonner, Leo de Haan and Jakob de Haan,** Spillovers of monetary policy across borders: International lending of Dutch banks, insurers and pension funds
No. 610 **Randall Hanegraaf, Nicole Jonker, Steven Mandley and Jelle Miedema**, Life cycle assessment of cash payments

No. 611 **Carin van der Cruijsen and Joris Knober,** Ctrl+C Ctrl+pay: Do people mirror payment behaviour of their peers?

No. 612 **Rob Bauer, Matteo Bonetti and Dirk Broeders,** Pension funds interconnections and herd behavior

DeNederlandscheBank

EUROSYSTEEM

De Nederlandsche Bank N.V.
Postbus 98, 1000 AB Amsterdam
020 524 91 11
dnb.nl