An Approximate Long-Memory Range-Based Approach for Value at Risk Estimation

Xiaochun Meng and James W. Taylor

Saïd Business School, University of Oxford

International Journal of Forecasting, 2018, 34(3), 377-388.

Abstract

This paper proposes new approximate long-memory VaR models that incorporate intra-day price ranges. These models use lagged intra-day range with the feature of considering different range components calculated over different time horizons. We also investigate the impact of the market overnight return on the VaR forecasts, which has not yet been considered with the range in VaR estimation. Model estimation is performed using linear quantile regression. An empirical analysis is conducted on 18 market indices. In spite of the simplicity of the proposed methods, the empirical results show that they successfully capture the main features of the financial returns and are competitive with established benchmark methods. The empirical results also show that several of the proposed range-based VaR models, utilizing both the intra-day range and the overnight returns, are able to outperform GARCH-based methods and CAViaR models.

Key words: Value at Risk; CAViaR; Realized Volatility; Intra-day Range; Quantile Regression.

1. Introduction

Since the first appearance of value at risk (VaR) in the 1980s, it has become the most prevalent risk measure, and is currently a standard tool for risk management in financial and insurance institutions (Berkowitz et al. 2011, Nieto and Ruiz 2016). The VaR is the quantile of the conditional distribution of the return on a portfolio. Accurate forecasting of VaR is of great importance for internal risk control and financial regulation. Although the concept of VaR is not complex, its measurement has proved to be challenging. There has been a variety of approaches proposed for the forecasting of VaR, and yet no consensus has been reached as to the best method.

Classical approaches to VaR forecasting, such as the use of generalized autoregressive conditional heteroskedasticity (GARCH) models, historical simulation and conditional autoregressive value at risk (CAViaR) models (Engle and Manganelli 2004), use only the historical returns. Intra-day data is becoming increasingly available, and it has been found to provide useful information for the estimation of the distribution of the daily returns (Corsi et al. 2008). Therefore, efforts have been made to use intra-day data in the forecasting of VaR for daily returns. Realized volatility, which is a nonparametric measure of unobservable volatility, calculated using intra-day data, has been widely used as a basis for forecasting daily volatility (see, for example, Andersen and Bollerslev 1998, Andersen et al. 2001b, Barndorff-Nielsen 2002). However, intra-day data tends to be expensive, and often a long time series of observations is not available. Moreover, the effort and resources required to process the high-frequency data may prove excessive (Rogers and Zhou 2008).

In contrast, the daily opening, daily closing, intra-day low and intra-day high series for the last thirty years are readily available for most tradable assets. Instead of using intra-day data to produce the realized volatility for VaR estimation, we consider an alternative use of intra-day data, which is much easier to implement and yet highly efficient. We base VaR estimation on the intra-day range, which is the difference between the daily high and the daily low log prices. Despite the fact that the intra-day range has been widely studied in volatility forecasting (Parkinson 1980, Andersen and Bollerslev 1998, Brandt and Jones 2006), little attention has been devoted to utilizing the intra-day range in VaR estimation. The only such literature, that the authors are aware of, are the studies of Brownlees and Gallo (2010), Chen et al. (2012) and Fuertes and Olmo (2013). Brownlees and Gallo (2010) use the intra-day range in a parametric framework. Chen et al. (2012) consider the use of the intra-day range in CAViaR models. Fuertes and Olmo (2013) includes the intra-day range in a GARCH model. Another variable that we consider in this paper is the market overnight return. It has been pointed out that the overnight return is useful for volatility forecasting, because, while the market is closed, a great amount of highly relevant information arrives from markets abroad, and public announcements might be made after the previous day's closing time (Tsiakas 2008). The work of Brownlees and Gallo (2010) is the only study that has compared the performance of the intra-day range and realized volatility for VaR estimation, with only parametric methods being considered. The authors are not aware of any study that evaluates the performance of the overnight return for VaR estimation.

This paper has two contributions. First, we propose a number of new quantile regression models based on realized volatility, the intra-day range and the overnight return. Second, we carry out an empirical comparison between the proposed methods and a large number of benchmark methods. Moreover, this paper is the first study that compares VaR estimation performance of a large set of methods based on the intra-day range and methods based on realized volatility for VaR estimation.

The remainder of the paper is structured as follows. Section 2 reviews intra-day volatility measures. Section 3 gives a brief review of the established VaR methods, with particular focus on those that are closely related to our new proposals. Section 4 introduces our new VaR models. Section 5 uses 18 series of stock returns to evaluate the performance of the new models, and to compare their VaR estimation accuracy to established methods. Section 6 provides a summary and some concluding remarks.

2. Intra-day Volatility Measures

In this section, we introduce the intra-day volatility measures that are used in this study. A very popular intra-day volatility measure is realized volatility. The realized volatility of a certain stock is defined as follows:

$$\begin{aligned} RV_t &= \sqrt{\sum_{i=1}^{M}(P_{t,i\cdot\Delta} - P_{t,(i-1)\cdot\Delta})^2} \qquad (1) \\ \Delta &= \frac{S}{M} \end{aligned}$$

where RV_t denotes the realized volatility within day t, S denotes the interval span of market opening hours, Δ divides S equally into M intervals, and $P_{t,i\cdot\Delta}$ denotes the log price at time $i\cdot\Delta$ of day t. It has been theoretically shown that, if the prices $P_{t,i\cdot\Delta}$ are observed without noise, then expression (1) is a consistent estimate of the daily volatility as M tends to infinity (Andersen et al. 2001b, Barndorff-Nielsen 2002). It has been found that models incorporating the realized volatility can significantly improve daily volatility forecasts in comparison to the conventional GARCH models, which are applied to daily returns data (see, for example, Corsi et al. 2008, Hansen et al. 2012, Martens et al. 2009, Shephard and Sheppard 2010).

The calculation of realized volatility clearly requires access to high-frequency data. Although such data is gradually becoming available, it is still expensive and is not available for a long time series of observations. The calculation of expression (1) requires significant computational power. Moreover, microstructure noise might contaminate the data, making expression (1) an inconsistent estimate for daily volatility (Andersen et al. 2011). This prompts consideration of, as an alternative intra-day volatility measure, the intra-day range, which is readily available for most tradable assets and requires little computing resource. The intra-day range is defined as follows:

$$Range_t = H_t - L_t \qquad (2)$$

where H_t and L_t denote respectively the highest log price and the lowest log price of the day.

The intra-day range has been widely studied in volatility estimation. Parkinson (1980) shows that the properly scaled intra-day range is an unbiased estimator of daily volatility, and is more efficient than the squared daily return. Brandt and Jones (2006) show that the efficiency of the intra-day range is even comparable to that of realized variance calculated using 3-hour to 6-hour returns. Alizadeh et al. (2002) show that the intra-day range is more robust to market microstructure noise, in comparison with realized volatility.

It should be noted that both the realized volatility and the intra-day range ignore the market overnight return, which is defined as follows:

$$y_{N,t} = open_t - close_{t-1} \tag{3}$$

where $open_t$ denotes the log opening price, and $close_{t-1}$ denotes the log closing price on the previous day. The overnight return has raised the interest of volatility forecasters. Hansen and Lunde (2006) find that an intra-day volatility measure that ignores the overnight return might not be a good proxy for the true daily volatility. Ahoniemi and Lanne (2013) find that incorporating the overnight return can lead to a more accurate realized volatility measure and can influence the relative performance of different volatility forecasting models. Wang et al. (2015) find that including the overnight return along with other explanatory variables improves volatility forecasts from the heterogeneous autoregressive model of realized volatility (HAR-RV) of Corsi (2009). In this paper, we consider the approach of Blair et al. (2001) and Hua and Manzan (2013), which incorporates the overnight return in the realized volatility as follows:

$$RV_{N,t} = \sqrt{(RV_t)^2 + (y_{N,t})^2} \tag{4}$$

We can also incorporate the overnight return in the intra-day range, as in expression (5). To our knowledge, this has not previously been considered for VaR and volatility estimation:

$$Range_{N,t} = \sqrt{Range_t^2 + y_{N,t}^2} \tag{5}$$

Another way of incorporating the overnight information in the intra-day range is to replace the intra-day low and high log prices in expression (2) by the lowest and highest log prices in the daily close-to-close period (Gerlach and Wang 2016):

$$Range_{NC,t} = max(H_t, close_{t-1}) - min(L_t, close_{t-1}) \tag{6}$$

In this paper, we focus on the use of the intra-day range of expression (2), and the measure based on the intra-day range and overnight return in expressions (5)-(6).

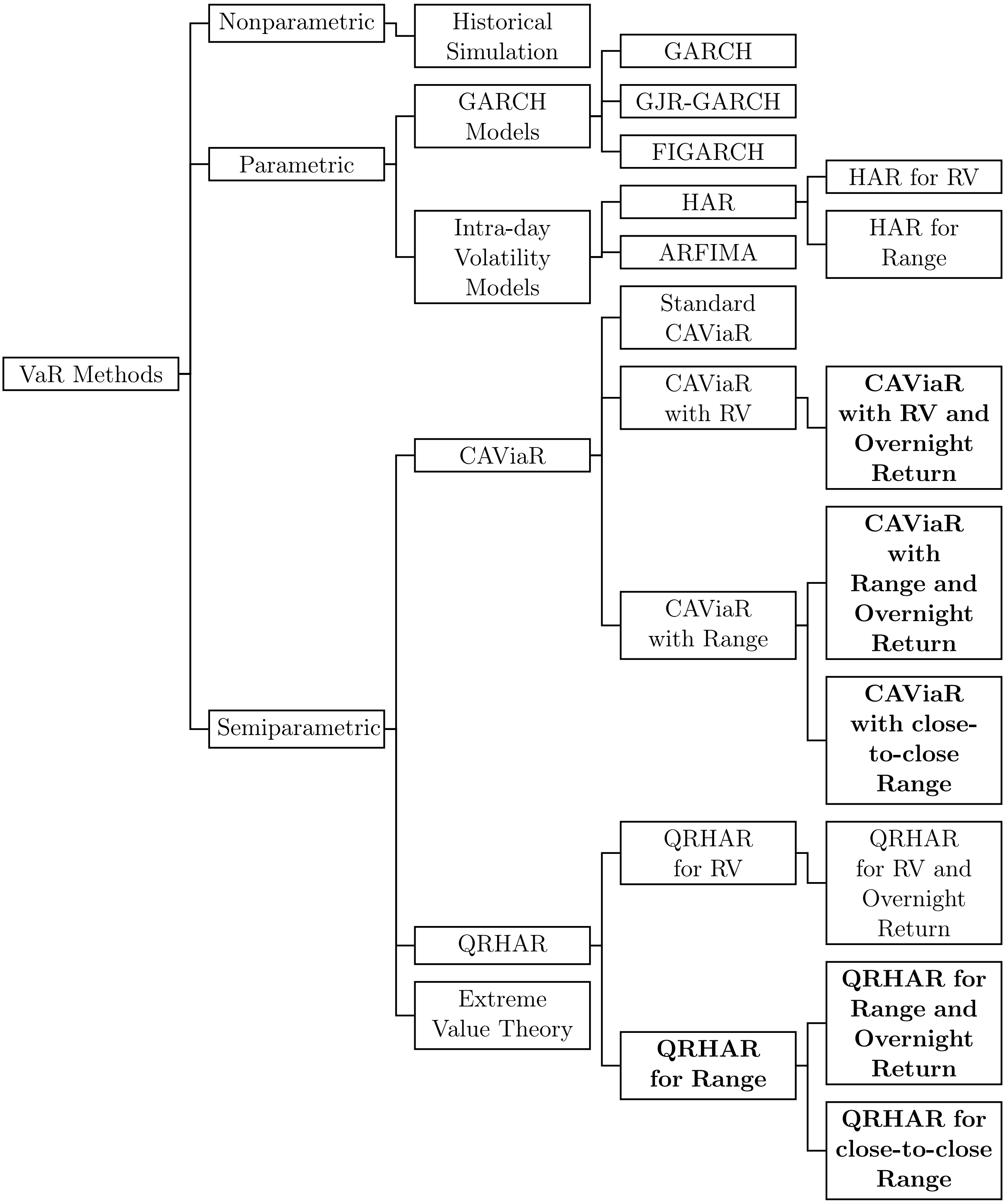
3. Review of Established VaR Methods

In this section, we review the established VaR methods. Following the classification of Manganelli and Engle (2004), we divide the VaR literature into three categories: parametric, nonparametric and semi-parametric. Figure 1 is a chart of the classification of the major established VaR methods. The methods in bold are our proposed new models, which will be discussed in Section 4. A recent comprehensive review of the VaR literature can be found in Nieto and Ruiz (2016). We now discuss the established methods in more detail.

3.1. Nonparametric Methods

Nonparametric methods make no assumption about the distribution of the returns. The most common example in this class is historical simulation, which estimates the VaR as the sample quantile of the returns over a certain window length (Pritsker 2006). Nonparametric methods have the advantages of being model free and easy to implement (Manganelli and Engle 2004). However, historical simulation requires an implicit assumption that the distribution of the returns remains at least roughly the same within the window of observations, which may not be true (Manganelli and Engle 2004). Moreover, it is not easy to determine the window length: a short window might incur large sampling error while a long window may cause the method to be less responsive to the changes in the true distribution. In empirical comparisons, nonparametric methods have not performed well (Manganelli and Engle 2004).

Figure 1: Classification of Established VaR Methods with Proposed New Methods. The methods in bold are our proposed new models.



3.2. Parametric Methods

In a parametric method, a quantile estimate is constructed by assuming a certain dynamic for the conditional volatility, $\sigma_t = \sqrt{var(y_t|I_{t-1})}$, of the log return y_t, at time t conditional upon I_{t-1}, the information set of all past information up to time $t-1$, and a specific conditional distribution for the log return. This can be viewed as the volatility of an error term $\epsilon_t = y_t - E(y_t|I_{t-1})$, where $E(y_t|I_{t-1})$ is the conditional mean, which is often assumed to be zero or a small constant. The error term ϵ_t is usually referred to as the price 'shock'. The distribution of y_t is not observable, and so various distributional assumptions have been proposed. A Gaussian distribution is computationally convenient, but the fact that the distribution of the return is often fat-tailed has prompted the use of other distributions, such as the Student's t distribution.

For the conditional volatility σ_t, there are two widely-used types of models. Firstly, there are models based on the daily return, in which the daily volatility appears as a latent variable. Typical examples are the GARCH models. For example, the GARCH and GJR-GARCH models (Glosten et al. 1993) have been widely studied in VaR estimation (Gerlach et al. 2012). A slightly different type of GARCH model, the long-memory fractionally integrated generalized autoregressive (FIGARCH) model, is proposed by Baillie et al. (1996) to capture the high persistence found in daily volatility.

The second type of volatility model induces an estimate of a certain intra-day volatility measure. Realized volatility RV_t, defined in expression (1), is undoubtedly the most popular intra-day volatility measure. Among the established realized volatility models, the HAR-RV model proposed by Corsi (2009), is a simple and pragmatic approach, whereby the forecast is constructed from the realized volatility over different time horizons. The HAR-RV model has been shown to perform well for volatility estimation, and its success and simplicity have made it popular among researchers (Corsi 2009, Bollerslev et al. 2016, Maheu and McCurdy 2011). These models have also been considered for VaR forecasting. Brownlees and Gallo (2010) show that the HAR-RV model improves VaR forecast accuracy in comparison with various GARCH models. Brownlees and Gallo (2010) also consider the HAR model, but with realized volatility replaced by the intra-day range, which is defined in expression (2). They find that the HAR model for the intra-day range is not outperformed by the HAR model for the realized volatility, and that both models improve VaR estimation accuracy compared to various GARCH models. We refer to the HAR model for the intra-day range as HAR-Range. Another class of models based on realized volatility is the realized GARCH model of Hansen et al. (2012). These model the realized volatility and the volatility simultaneously, where the realized volatility is assumed to be dependent on the contemporaneous volatility, and the volatility is assumed to be modeled using its own lagged values and the lagged value of the estimated realized volatility. Hansen et al. (2012) find that the realized GARCH model outperforms a standard GARCH model. Gerlach and Wang (2016) consider the use of the intra-day range instead of the realized volatility in the realized volatility model. The model is shown to be competitive compared to the realized GARCH model. We refer to the realized GARCH model with realized volatility as the Realized-GARCH-RV model, and we refer to the realized GARCH model with the intra-day range as the Realized-GARCH-Range model.

3.3. Semiparametric Methods

The semiparametric methods include applications of extreme value theory and applications based on the use of quantile regression (Manganelli and Engle 2004). We focus on the semiparametric methods based on quantile regression, which estimates the desired quantile directly. A generic quantile regression model can be expressed in the following expression:

$$q_t(\theta) = q(\boldsymbol{\Psi}_{t-1}, q_{t-1}(\theta), \ldots, \boldsymbol{\Psi}_1, q_1(\theta); \boldsymbol{\beta}, \theta) \tag{7}$$

where θ is the probability level of interest, $q_t(\cdot)$ is the estimated θth quantile of some daily return y_t, $\boldsymbol{\beta}$ is a vector of parameters, and $\boldsymbol{\Psi}_t$ is a vector of explanatory variables. The parameter vector $\boldsymbol{\beta}$ is chosen to minimize the following quantile regression objective function:

$$\sum_{t=1}^{T} [y_t - q_t(\theta)][\theta - I(y_t < q_t(\theta))] \tag{8}$$

where T is the number of observations in the in-sample period. Different quantile regression models involve different selections of $\mathbf{\Psi}_t$ and formulations of the $q_t(\cdot)$ in expression (7). Next we review several benchmark quantile regression models.

3.3.1. CAViaR Models and Quantile Regression

Engle and Manganelli (2004) propose four CAViaR models that use an autoregressive framework to model the evolution of the desired quantile directly. CAViaR models have several advantages. They are free from distributional misspecification since no distributional assumption is needed; they allow the conditional distribution to be time-varying; and the models can have different parameters for the quantiles associated with different probability levels, so that different quantiles are able to have distinct dynamics. To save space, we present only three CAViaR models here:

CAViaR Symmetric Absolute Value (CAViaR-SAV):

$$q_t(\theta) = \beta_1 + \beta_2 q_{t-1}(\theta) + \beta_3 |y_{t-1}| \tag{9}$$

CAViaR Asymmetric Slope (CAViaR-AS):

$$q_t(\theta) = \beta_1 + \beta_2 q_{t-1}(\theta) + \beta_3 (y_{t-1})^+ + \beta_4 (y_{t-1})^- \tag{10}$$

CAViaR Indirect GARCH (CAViaR-IndG):

$$q_t(\theta) = sgn(\theta - 0.5)(\beta_1 + \beta_2 q_{t-1}(\theta)^2 + \beta_3 y_{t-1}^2)^{\frac{1}{2}} \tag{11}$$

where $(x)^+ = \max(x, 0)$, and $(x)^- = -\min(x, 0)$ in expression (10).

The lagged values of the returns in the CAViaR models serve as the innovation terms, which can be replaced or augmented by other explanatory variables, such as the intra-day volatility measures. Žikeš and Baruník (2015) use the realized volatility in the CAViaR models, and show that the proposed models improve VaR forecasting accuracy. Similar in spirit to this, Chen et al. (2012) consider several CAViaR models with the intra-day range, and show that the proposed models perform well when compared to a variety of VaR models. For illustrative purposes, we present here the CAViaR-SAV model with realized volatility and with the intra-day range:

CAViaR with Realized Volatility (CAViaR-RV):

$$q_t(\theta) = \beta_1 + \beta_2 q_{t-1}(\theta) + \beta_3 RV_t \tag{12}$$

CAViaR Range Value (CAViaR-Range):

$$q_t(\theta) = \beta_1 + \beta_2 q_{t-1}(\theta) + \beta_3 Range_{t-1} \tag{13}$$

3.3.2. Quantile Regression HAR Models Based on Realized Volatility

Quantile models with HAR-type specifications have also been considered. To distinguish these semiparametric HAR models from the parametric HAR models in Section 3.2, we refer to these approaches as quantile regression HAR (QRHAR) models. Hua and Manzan (2013) consider four QRHAR models based on realized volatility and the overnight return to estimate the quantiles of the returns. They find that the QRHAR models with intra-day volatility measures outperform benchmark GARCH models. However, they do not find that the inclusion of the overnight return brings any improvement. We present two QRHAR models here:

QRHAR with Realized Volatility (QRHAR-RV):

$$q_t(\theta) = \beta_0 + \beta_1 RV_{t-1} + \beta_2 RV^w_{t-1} + \beta_3 RV^m_{t-1} \tag{14}$$

$$RV^w_{t-1} = \frac{1}{5}\sum_{i=1}^{5} RV_{t-i} \tag{15}$$

$$RV^m_{t-1} = \frac{1}{22}\sum_{i=1}^{22} RV_{t-i} \tag{16}$$

QRHAR with Realized Volatility and overNight return (QRHAR-RV-N):

$$q_t(\theta) = \beta_0 + \beta_1 RV_{N,t-1} + \beta_2 RV^w_{N,t-1} + \beta_3 RV^m_{N,t-1} \tag{17}$$

$$RV^w_{N,t-1} = \frac{1}{5}\sum_{i=1}^{5} RV_{N,t-i} \tag{18}$$

$$RV^m_{N,t-1} = \frac{1}{22}\sum_{i=1}^{22} RV_{N,t-i} \tag{19}$$

where RV_t is the realized volatility; RV_t^w and RV_t^m in expressions (14)-(16) denote the average realized volatility over a week and a month respectively; and $RV_{N,t}$ in expressions (17)-(19) is defined as in expression (4).

4. Proposed New Models for VaR Forecasting

In this section, we propose four new semiparametric models. Each is a quantile model estimated using quantile regression. Quantile regression modelling is appealing because it avoids the need for a distributional assumption, it allows the quantiles associated with different probability levels to have different dynamics, and such models have performed well in empirical studies (Manganelli and Engle 2004, Taylor 2008, Chen et al. 2012). The new quantile regression models that we consider are CAViaR and QRHAR formulations. Our new models incorporate the overnight return and intra-day range in these models.

4.1. Two New CAViaR Models Incorporating the Overnight Return

We first consider the use of the overnight return. As we explained in Section 2, information regarding public announcements and relevant information arriving from markets abroad, after the previous day's closing time, can be useful for forecasting the distribution of daily returns. Intra-day volatility measures, such as the intra-day range and the realized volatility, contain only the information during the market opening hours of the day. In general, the use of the overnight return has been neglected in the VaR literature. To the authors' knowledge, the only study that has used the overnight return for VaR forecasting methods is the paper of Hua and Manzan (2013), who incorporate the overnight return in QRHAR models based on realized volatility. In that study, the overnight return did not lead to improved VaR forecast accuracy. In this paper, we revisit the use of the overnight return, but instead consider its use within CAViaR models. As the overnight return is, in essence, a measure based on intra-day data, we opt to use the versions of CAViaR models that are based on intra-day data. The first is based on the CAViaR-RV model of expression (12) and the other two are based on the CAViaR-Range model of expression (13):

CAViaR with Realized Volatility and overNight return (CAViaR-RV-N):

$$q_t(\theta) = \beta_1 + \beta_2 q_{t-1}(\theta) + \beta_3 RV_{t-1} + \beta_4 |y_{N,t-1}| \tag{20}$$

CAViaR with intra-day Range and overNight return (CAViaR-Range-N):

$$q_t(\theta) = \beta_1 + \beta_2 q_{t-1}(\theta) + \beta_3 Range_{t-1} + \beta_4 |y_{N,t-1}| \tag{21}$$

CAViaR with intra-day close-to-close Range (CAViaR-Range-C):

$$q_t(\theta) = \beta_1 + \beta_2 q_{t-1}(\theta) + \beta_3 Range_{NC,t-1} \tag{22}$$

where $y_{N,t}$ is the overnight return, as defined in expression (3). Note that, we could also add the overnight return to the other CAViaR models of Section 3.3.1, but for simplicity we only consider the above two models in this paper.

4.2. Two New QRHAR Models Incorporating the Intra-day Range

Next, we consider the use of the intra-day range within a QRHAR model. The HAR model has been shown to perform well for volatility estimation, and its success and simplicity have made it popular among researchers (Corsi 2009, Brownlees and Gallo 2010, Maheu and McCurdy 2011, Bollerslev et al. 2016). The HAR model considers the use of intra-day volatility aggregated over different horizons to capture the asymmetry and long-memory property in volatility dynamics. The motivation for this is that it has been observed that the volatility over a long period has a strong impact on the volatility over a short period, and there is significant evidence of long-memory in daily volatility (Baillie et al. 1996, Andersen et al. 2001a). The HAR model is able to capture these features in a simple formulation, which contrasts with standard GARCH models (Corsi et al. 2008). Given the success of the HAR models in volatility estimation, it seems somewhat surprising that the QRHAR models, which use the HAR structure within a quantile regression framework, have not attracted more attention in VaR forecasting. The only such study, of which the authors are aware, is the work of Hua and Manzan (2013), which, as we described in the previous section, is based on realized volatility. In this paper, we introduce the idea of incorporating the intra-day range within a QRHAR model. Our motivation for this is threefold. Firstly, the intra-day range is an efficient estimator of the daily volatility, and so it is a suitable candidate explanatory variable in a VaR model. Secondly, various studies have shown the intra-day range has competitive predictive power compared to the realized volatility in terms of volatility and VaR forecasting (Brownlees and Gallo 2010, Chen et al. 2012, Gerlach and Wang 2016). Thirdly, a practical advantage of the intra-day range is that it is free and readily available for most tradable assets over long historical periods, which contrasts with realized voaltility, which has the added inconvenience of needing to be processed prior to use (Rogers and Zhou 2008). We propose the following two new QRHAR models based on the intra-day range:

QRHAR with intra-day Range (QRHAR-Range):

$$q_t(\theta) = \beta_1 + \beta_2 Range_{t-1} + \beta_3 Range^w_{t-1} + \beta_4 Range^m_{t-1} \tag{23}$$

$$Range^w_{t-1} = \frac{1}{5}\sum_{i=1}^{5} Range_{t-i} \tag{24}$$

$$Range^m_{t-1} = \frac{1}{22}\sum_{i=1}^{22} Range_{t-i} \tag{25}$$

QRHAR with intra-day Range and overNight return (QRHAR-Range-N):

$$q_t(\theta) = \beta_1 + \beta_2 Range_{N,t-1} + \beta_3 Range^w_{N,t-1} + \beta_4 Range^m_{N,t-1} \tag{26}$$

$$Range^w_{N,t-1} = \frac{1}{5}\sum_{i=1}^{5} Range_{N,t-i} \tag{27}$$

$$Range^m_{N,t-1} = \frac{1}{22}\sum_{i=1}^{22} Range_{N,t-i} \tag{28}$$

QRHAR with close-to-close Range (QRHAR-Range-C):

$$q_t(\theta) = \beta_1 + \beta_2 Range_{NC,t-1} + \beta_3 Range^w_{NC,t-1} + \beta_4 Range^m_{NC,t-1} \tag{29}$$

$$Range^w_{NC,t-1} = \frac{1}{5}\sum_{i=1}^{5} Range_{NC,t-i} \tag{30}$$

$$Range^m_{NC,t-1} = \frac{1}{22}\sum_{i=1}^{22} Range_{NC,t-i} \tag{31}$$

where $Range_{N,t}$ in expressions (26)–(28) is defined as in expression (5), $Range_{Nc,t}$ in expressions (29)–(31) is defined as in expression (6), $Range^w_t$, $Range^w_{N,t}$ and $Range^w_{NC,t}$ denote the average of $Range_t$, $Range_{N,t}$ and $Range_{NC,t}$ over a week, and $Range^m_t$, $Range^m_{N,t}$ and $Range^m_{NC,t}$ denote the average of $Range_t$, $Range_{N,t}$ and $Range_{NC,t}$ over a month. As we stated in Section 2, we are not aware of the previous use of $Range_{N,t}$ in VaR forecasting.

$Range_{NC,t}$ has been considered in parametric models in Gerlach and Wang (2016), but has not been considered in a semiparametric quantile regression model. Expressions (23)-(25), expressions (26)-(28) and expressions (29)-(31) are essentially the QRHAR-RV model of expressions (14)-(16), and the QRHAR-RV-N model of expressions (17)-(19), with the realized volatility terms replaced by the intra-day range. We will show in the empirical study that the use of the intra-day range leads to better forecasting accuracy, in spite of the similarity between model structures.

In order to illustrate our contribution, consider again Figure 1. It can be seen from Figure 1 that we focus on the use of the intra-day range and the overnight return for semiparametric quantile regression models. We are proposing two new CAViaR models that include the overnight return in two existing CAViaR models, and two new QRHAR models that differ from the existing QRHAR models in that they are based on the intra-day range, rather than the realized volatility.

5. Empirical Results

5.1. Data

To explore the forecasting performance of our four proposed new VaR models, and a set of benchmark methods, we used the daily opening, closing, high and low prices of the following 18 global stocks: S&P500, FTSE100, Nikkei225, DAX30, Russel2000, All Ordinaries, DJIA, Nasdaq100, CAC40, KOSPI Composite Index, AEX Index, Swiss Market Index, IBEX35, IPC Mexico, Bovespa Index, Euro STOXX 50, FT Straits Times Index, and FTSE Mib Index. We used the realized volatility of all the global stocks sampled at a 5-min sampling frequency. We obtained the data from the Oxford-Man Institute's realized library Version 0.2 (Heber et al. 2009). We excluded only two indices from the Oxford-Man Institute's realized library, the Hang Seng and S&P/TSX Composite Index, due to their inadequate numbers of observations. The sample period used in our study was from 3 January 2000 to 20 May 2014. We chose a post-sample period consisting of 1500 days to evaluate day-ahead quantile estimates. We employed a rolling window approach, where each method was re-estimated for each day in the post-sample period using data from the previous 1800 days. Note that different markets tend to have different public holidays, and so they can have different numbers of observations within a certain calendar period. We chose our post-sample period in such a way that the observations from all the markets shared the same end date, which was 20 May 2014. Following common practice, we did not estimate the conditional mean of the returns, and assumed it to be zero (see, for example, Žikeš and Baruník 2015).

5.2. VaR Forecasting Methods

For each method considered, one-day-ahead forecasts of VaR were generated for each day in the post sample period. We considered the 1% and the 5% quantiles, as is common in the VaR literature. We implemented the historical simulation nonparametric method with a variety of different window lengths. However, as the results were poor, for simplicity, we do not discuss the results further.

As parametric benchmarks, we implemented the GARCH(1,1) model, the GJR-GARCH(1,1) model, the FIGARCH(1,d,1) model, where d is a parameter between 0 and 1, the HAR-RV model, the HAR-Range model, the Realized-GARCH-RV model, and the Realized-GARCH-Range model. For each model, we used a Student's t distribution for parameter estimation and for VaR forecasting.

In terms of semiparametric methods, we included a variety of quantile regression benchmark methods in our empirical study. These are models from the literature, rather than new models proposed by us. We first implemented the standard CAViaR models based on the daily return: the CAViaR-SAV model of expression (9), the CAViaR-AS model of expression (10), and the CAViaR-IndG model of expression (11). For the CAViaR models based on realized volatility, we implemented all the models proposed by Žikeš and Baruník (2015). We found that these models performed similarly to each other in our study. Therefore, we only report the results of the CAViaR-RV model of expression (12), which uses just the realized volatility alongside the lagged quantile. For the CAViaR models based on the intra-day range, we also implemented all the models proposed by Chen et al. (2012). We once again found that these models performed similarly to each other in our empirical study. In view of this, to save space, we only present the results for the CAViaR-Range model of expression (13). Moving on to the QRHAR models, we implemented the QRHAR-RV model of expressions (14)-(16), and the QRHAR-RV-N model of expressions (17)-(19).

With regard to our proposed new methods, we implemented the six models in Section 4: the CAViaR-RV-N model based on realized volatility and the overnight return of expression (20); the CAViaR-Range-N model based on the intra-day range and the overnight return of expression (21); the CAViaR-Range-C model based on the close-to-close range of expression (22); the QRHAR-Range model based on the intra-day range of expressions (23)-(25); the QRHAR-Range-N model based on the intra-day range and the overnight return of expressions (26)-(28); and the QRHAR-Range-N model based on the intra-day range and the overnight return of expressions (29)-(31).

5.3. In-sample Parameters

Table 1 presents parameters for our new CAViaR-Range-N model of expression (21) and our new QRHAR-Range-N model of expressions (26)-(28) estimated using the FTSE100 data. Note that we estimated 1500 different parameter vectors, since the parameters are re-estimated as the rolling window passes through the post-sample period. In Table 1, we present the estimated parameters derived using the first moving window, together with their p-values. For the CAViaR-Range-N model, the parameters of $Range_{t-1}$ and $|y_{N,t-1}|$ are negative, indicating that larger values of these variables will result in a lower estimated quantile, which is intuitive. The parameters of q_{t-1} are positive, which is consistent with the well-known volatility clustering effect. For the QRHAR-Range-N model, all the parameters except β_4 for the 5% quantiles are negative, indicating that larger values of $Range_{t-1}$, $Range^{w}_{t-1}$, and $Range^{m}_{t-1}$ will result in a lower value of the estimated quantile. The exception of β_4 might stem from multicollinearity between $Range_{t-1}$, $Range^{w}_{t-1}$, and $Range^{m}_{t-1}$. We will discuss this issue and the p-values in more detail later.

In Table 2, we report the number of times, out of 1500, that each parameter was statistically significantly different from zero at the 5% significance level for the CAViaR-Range-N model and QRHAR-Range-N models. For the CAViaR-Range-N model, it can be seen from Table 2 that the coefficient β_2 of the lagged quantile is almost always statistically significant; the coefficient β_3 of the intra-day range was statistically significant for more than half the periods; and the coefficient β_4 for the overnight return was significant reasonably often.

The results for the QRHAR-Range-N model in Table 2 are somewhat more complex. The coefficient β_3 of $Range^{w}_{N,t-1}$ was always significant for both quantiles; and the coefficient β_4 of $Range^{m}_{N,t-1}$ was significant for a reasonable proportion of the periods, especially for the 5% quantile, which suggests the presence of long-memory in VaR. The coefficient β_2 of $Range_{N,t-1}$ was always significant for the 1% quantiles, but it was never significant for the 5% quantiles. Our explanation for this is that $Range_{N,t-1}$ and $Range^{w}_{N,t-1}$ are highly correlated. In fact, the correlation between these two variables for our FTSE100 data is 0.85. However, multicollinearity is not necessarily a problem for forecasting (see, for example, Brooks 2014). In fact, we estimated the QRHAR-Range-N model without the $Range_{N,t}$ term, but this worsened the post-sample VaR forecasting performance. Thus, we conclude that $Range_{N,t-1}$, $Range^{w}_{N,t-1}$ and $Range^{m}_{N,t-1}$ should all be included in the model.

5.4. Post-Sample Evaluation

5.4.1. Evaluation Measures

We evaluated unconditional and conditional coverage in the post-sample VaR forecasts. The unconditional coverage test assesses the hit percentage defined as the proportion of the returns below the estimated quantile corresponding to probability level θ. Significant difference from the ideal value of θ is tested using a binomial distribution. To evaluate conditional coverage,

Table 1: Estimated parameters of the CAViaR-Range-N model and the QRHAR-Range-N model for the FTSE100, estimated using the first moving window of 1800 days.

	CAViaR-Range-N		QRHAR-Range-N	
		1% Quantiles		
	Parameters	p-values	Parameters	p-values
β_1	-0.116	0.113	-0.182	0.078
β_2	0.588	0.003	-0.229	0.014
β_3	-0.622	0.145	-2.045	0.000
β_4	-0.460	0.067	0.414	0.009
		5% Quantiles		
β_1	-0.035	0.276	-0.227	0.014
β_2	0.792	0.000	-0.089	0.305
β_3	-0.217	0.151	-0.967	0.000
β_4	-0.133	0.241	-0.053	0.662

Table 2: Number of times, out of 1500, that each estimated parameter, was significantly different from zero at the 5% significance level, for the CAViaR-Range-N model and the QRHAR-Range-N model applied to the FTSE100.

	CAViaR-Range-N		QRHAR-Range-N	
	1%	5%	1%	5%
β_1	0	0	102	778
β_2	1498	1500	1500	0
β_3	1043	823	1500	1500
β_4	202	107	226	608

we used the dynamic quantile test proposed by Engle and Manganelli (2004), which is a joint test of correct unconditional coverage and independence of violations. It evaluates whether the sequence of the hit variable, defined as $Hit_t = I(y_t < q_t(\theta)) - \theta$, is distributed i.i.d. Bernoulli with probability θ, and is independent of lags of the Hit_t variable and the conditional quantile forecast. Following Engle and Manganelli (2004), we used four lags of the Hit_t variable in the test's regression.

We also evaluated the VaR forecasts using the quantile score, which is the quantile regression loss function used in the summation of expression (8). The quantile score is a proper scoring rule for evaluating quantile forecasts, implying that the expectation of the score is minimized by the perfect set of quantile forecasts (Gneiting and Raftery 2007). For each method, and each of the 18 stock indices, we computed the average quantile score. As the resulting values have no intuitive interpretation, for each method and each series, we calculated the ratio of the score to that of a benchmark method, then subtracted this ratio from one, and multiplied the result by 100. We term this the quantile skill score. For the calculation of the skill score, we used the GARCH-t model as the benchmark. To summarize performance across the 18 indices, we calculated the geometric mean of the ratios of the score for each method to the score for the GARCH-t benchmark method, then subtracted this from one, and multiplied the result by 100. Higher values are preferable for the resulting geometric mean values, and indeed the skill scores, reflecting superiority over the benchmark method.

Finally, we use the model confidence set evaluation method to select the best-performing models (Hansen et al. 2011). In practice, limitations of data can decrease the power of statistical tests. Two models might not be distinguishable when there is not a sufficiently large number of post-sample observations. The model confidence set approach does not assume that the confidence sets contain the best method, instead, the approach selects a set of best-performing models that will contain the best model with a given level of confidence. We use the quantile score as the loss function, and select methods for the confidence set based on the Diebold-Mariano test for statistical significance between the mean of the quantile scores of pairs of methods (Diebold and Mariano 2002). Following Hansen et al. (2011), we used 75% and 90% as the confidence levels for the model confidence sets.[1]

5.4.2. Evaluation Results

We now consider the post-sample evaluation results. For the FTSE100, Table 3 presents the hit percentage, the p-value for the dynamic quantile test, and the quantile skill score for each method. Larger p-values are preferred for the dynamic quantile test. The results for our six new proposed models are presented in the bottom six rows of the table. We present this table merely for illustrative purposes, as we must consider the results for all 18 stock indices in order

[1] To implement the model confidence set approach to evaluation, we used the MATLAB code provided by Kevin Sheppard in his MFE Toolbox, which is freely available on https://www.kevinsheppard.com/MFE_Toolbox.

to compare the methods. This is done in Table 4, which presents the number of rejections for the unconditional coverage test and the dynamic quantile test at the 5% significance level for each model applied to the 18 indices, as well as the skill score averaged across all 18 indices. Our discussion focuses on this summary table.

Let us first consider the results for unconditional coverage in Table 4. We find that all the parametric methods and the four semiparametric methods based on realized volatility have large numbers of rejections. All the other models perform reasonably well. The poor performance of the parametric methods suggests misspecification of the volatility model or inappropriate choice of distribution. The poor performance of the semiparametric methods utilizing the realized volatility is somewhat surprising, since the other semiparametric methods all perform well. One explanation is that the realized volatility is contaminated by microstructure noise. We next consider the dynamic quantile test results. The proposed new models have fewer rejections than the benchmark methods, with the CAViaR-Range-N model, the CAViaR-Range-C model, the QRHAR-Range-N model and the QRHAR-Range-C model performing the best. Now we move on to the skill scores using the GARCH-t model as the benchmark. A positive skill score indicates the model has outperformed the benchmark method. Naturally, the skill scores for the benchmark method are zero. The six proposed new methods have positive skill scores for both the 1% and 5% quantiles. One general observation is that the models including the realized volatility or the intra-day range perform better than the methods based only on the daily returns. Despite their relatively poor coverage test results, the realized GARCH models have the best skill scores for the 1% quantile. The proposed QRHAR-Range-C and QRHAR-Range-N models perform the best for the 5% quantiles. Overall, the proposed QRHAR-Range-N and the CAViaR-Range-N model perform very competitively.

To further illustrate the benefit of including the intra-day range and the overnight return in quantile regression models, we compare the numbers of rejections between the models based on realized volatility and the corresponding models based on the intra-day range, and the numbers

Table 3: Evaluation of 1% and 5% VaR forecast accuracy for the FTSE100. The hit percentage, the p-values for the dynamic quantile test, and the skill scores using the GARCH-t model as the benchmark.

	1% Quantiles			5% Quantiles		
	Hit%	Dynamic Quantile	Skill Score%	Hit%	Dynamic Quantile	Skill Score%
Parametric Benchmarks						
GARCH-t	1.5	0.003**	0.0	6.7**	0.029*	0.0
GJR-GARCH-t	1.5	0.200	1.7	6.1	0.266	1.3
FIGARCH-t	1.1	0.169	0.4	6.4*	0.035*	0.4
HAR-RV-t	1.7*	0.101	-1.5	6.7**	0.043*	1.3
HAR-Range-t	1.5	0.003**	-1.9	6.6**	0.031*	-0.9
Realized-GARCH-RV-t	1.6*	0.078	0.6	6.2*	0.132	1.8
Realized-GARCH-Range-t	1.5	0.002**	-0.7	6.6**	0.04**	0.1
Semiparametric Benchmarks						
CAViaR-SAV	0.9	0.061	-5.9	5.6	0.131	-0.4
CAViaR-AS	1.2	0.536	2.6	5.3	0.597	1.0
CAViaR-IndG	0.7	0.154	-1.7	5.3	0.170	0.0
CAViaR-RV	1.0	0.991	0.2	5.1	0.636	1.9
CAViaR-Range	1.0	0.471	-1.3	5.2	0.937	1.1
QRHAR-RV	1.1	0.986	-4.1	5.3	0.505	2.1
QRHAR-RV-N	1.0	0.992	-2.6	5.2	0.884	0.1
New Semiparametric Models						
CAViaR-RV-N	1.1	0.975	2.4	5.1	0.594	2.1
CAViaR-Range-N	1.1	0.532	1.1	4.9	0.982	1.0
CAViaR-Range-C	1.1	0.470	2.0	5.3	0.953	1.1
QRHAR-Range	0.9	0.426	0.4	5.6	0.727	1.5
QRHAR-Range-N	1.1	0.986	1.3	5.3	0.924	1.8
QRHAR-Range-C	1.4	0.557	-0.5	5.4	0.512	1.2

Note: Larger p-values are better. Significance at 5% and 1% levels is indicated by * and **, respectively. Larger skill scores are better.

of rejections between the models without the overnight return and the corresponding models with the overnight return. We provide this comparison for the unconditional coverage and dynamic quantile tests in Table 5. For each pair of models, the numbers are obtained by subtracting the numbers of test rejections of the latter models from those of the former models. A positive number indicates that the former model is outperformed by the latter. In Table 5, the upper block of results compares the performance between the quantile regression models using realized volatility and the intra-day range. The lower block of results compares the performance between the quantile regression models with and without the overnight return. The results suggest that the models based on the intra-day range outperform the corresponding models based on

Table 4: Summary of the VaR forecast accuracy results. Number of test rejections at the 5% significance level for the unconditional coverage and dynamic quantile tests, and the skill scores using the GARCH-t model as the benchmark.

	Unconditional Coverage			Dynamic Quantile			Skill Score	
	1%	5%	Total	1%	5%	Total	1%	5%
Parametric Benchmarks								
GARCH-t	6	7	13	9	9	18	0.0	0.0
GJR-GARCH-t	8	6	14	8	3	11	2.5	1.7
FIGARCH-t	2	4	6	8	10	18	1.0	0.6
HAR-RV-t	10	8	18	9	8	17	3.3	2.2
HAR-Range-t	9	8	17	6	9	15	2.7	1.1
Realized-GARCH-RV-t	10	9	19	10	7	17	3.7	2.1
Realized-GARCH-Range-t	6	7	13	7	9	16	3.7	1.5
Semiparametric Benchmarks								
CAViaR-SAV	1	0	1	8	3	11	-3.1	-0.9
CAViaR-AS	3	0	3	7	3	10	0.2	1.1
CAViaR-IndG	1	0	1	9	4	13	-1.1	-0.5
CAViaR-RV	4	4	8	6	6	12	3.0	2.3
CAViaR-Range	0	0	0	6	3	9	2.5	2.1
QRHAR-RV	6	4	10	9	5	14	1.2	2.0
QRHAR-RV-N	5	2	7	6	3	9	0.9	1.8
New Semiparametric Models								
CAViaR-RV-N	6	1	7	5	3	8	2.6	2.0
CAViaR-Range-N	1	0	1	1	1	2	2.5	2.0
CAViaR-Range-C	2	0	2	3	2	5	2.6	2.4
QRHAR-Range	3	1	4	6	1	7	1.4	2.2
QRHAR-Range-N	2	0	2	4	0	4	1.7	2.4
QRHAR-Range-C	2	0	2	3	1	4	1.5	2.2

Note: Smaller numbers of rejections and larger skill scores are better.

realized volatility, and the models using the overnight return outperform the corresponding models without the overnight return.

Finally, we report the number of series for which each model lies inside the model confidence set for the 75% and 90% confidence levels in Table 6. A large number is preferred as it indicates the corresponding model is not frequently eliminated in the selection process. A value of 18 is the best possible performance, and this indicates that the method was not eliminated from the model confidence set for any of the 18 indices. Although the results are not particularly helpful in differentiating between the accuracy of the methods, we can still draw some conclusions from the test results. Firstly, the proposed models are the best-performing models, with very few

Table 5: Comparison of the number of rejections for the unconditional coverage and dynamic quantile tests at the 5% significance level.

	Unconditional Coverage		Dynamic Quantile	
	1%	5%	1%	5%
Realized volatility and intra-day range				
CAViaR-RV - CAViaR-Range	4	4	0	3
CAViaR-RV-N - CAViaR-Range-N	5	1	4	2
QRHAR-RV - QRHAR-Range	3	3	3	4
QRHAR-RV-N - QRHAR-Range-N	3	2	2	3
Without and with the overnight return				
CAViaR-RV - CAViaR-RV-N	-2	3	1	3
CAViaR-Range - CAViaR-Range-N	-1	0	5	2
CAViaR-Range - CAViaR-Range-C	-2	0	3	1
QRHAR-RV - QRHAR-RV-N	1	2	3	2
QRHAR-Range - QRHAR-Range-N	1	1	2	1
QRHAR-Range - QRHAR-Range-C	1	1	3	0

Note: The numbers are obtained by subtracting the number of test rejections in Table 4 of the latter model from that of the former model. A positive number indicates that the former model is outperformed by the latter.

times being excluded from the model confidence set. Secondly, in general, the models including the intra-day data perform well, while the models solely depending on the historical returns are more often excluded from the model confidence set, with the GJR-GARCH-t model being the only exception.

To summarize, the results of the proposed new methods are promising. In terms of the unconditional coverage and dynamic quantile tests, the new CAViaR-Range-N and QRHAR-Range-N models perform the best. They have the lowest numbers of rejections for the dynamic quantile test and have very low numbers of rejections for the unconditional coverage test. The results were also good for our other two newly proposed methods, as well as the QRHAR-RV-N model proposed by Hua and Manzan (2013), and the CAViaR-Range model proposed by Chen et al. (2012). In terms of the quantile scores and the model confidence set evaluation method, the methods based on the intra-day data are better than the methods solely depending on the historical returns, and the methods using the intra-day range were generally more accurate than the methods using the realized volatility.

Table 6: Summary of the model confidence set results for the 1% and 5% quantiles using the Diebold-Mariano test on based on quantile scores. Number of series each model is inside the model confidence set for 75% and 90% confidence level.

	1% Quantiles		5% Quantiles	
Confidence Level	75%	90%	75%	90%
Parametric Benchmarks				
GARCH-t	16	18	12	16
GJR-GARCH-t	17	18	16	17
FIGARCH-t	17	18	13	17
HAR-RV-t	18	18	17	18
HAR-Range-t	18	18	17	18
Realized-GARCH-RV-t	18	18	17	18
Realized-GARCH-Range-t	18	18	18	18
Semiparametric Benchmarks				
CAViaR-SAV	12	16	9	11
CAViaR-AS	16	18	14	16
CAViaR-IndG	15	15	10	14
CAViaR-RV	18	18	17	18
CAViaR-Range	18	18	18	18
QRHAR-RV	18	18	18	18
QRHAR-RV-N	16	18	17	18
New Semiparametric Models				
CAViaR-RV-N	18	18	17	18
CAViaR-Range-N	18	18	17	18
CAViaR-Range-C	18	17	18	18
QRHAR-Range	17	18	18	18
QRHAR-Range-N	17	18	18	18
QRHAR-Range-C	16	17	18	18

Note: Large numbers are better.

6. Concluding Remarks

In this paper, we have considered the use of realized volatility, the intra-day range, and the overnight return in quantile regression models to improve VaR forecasting accuracy. Our proposed new methods delivered promising results in our empirical study. The proposed QRHAR-Range-N model, QRHAR-Range-C model, and the CAViaR-Range-N model, and the and the CAViaR-Range-C model, which use the intra-day range and the overnight return, performed particularly well. The empirical study also included previously proposed methods, and enabled a comparison of methods based on the intra-day range with methods based on realized volatility, and a comparison between methods including and excluding the overnight return. We found that the intra-day range is more useful than realized volatility for VaR forecasting, and that it is beneficial to include the overnight return. A further contribution of the paper is that we provide empirical support for the previous work of Chen et al. (2012) and Hua and Manzan (2013).

Acknowledgements

We are very grateful to an editor, an associate editor and two referees for providing comments that helped greatly to improve the paper.

References

Ahoniemi, K. and Lanne, M. (2013). Overnight stock returns and realized volatility. *International Journal of Forecasting*, 29(4):592–604.

Alizadeh, S., Brandt, M. W., and Diebold, F. X. (2002). Range-based estimation of stochastic volatility models. *Journal of Finance*, 57(3):1047–1091.

Andersen, T. G. and Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 39(4):885–905.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Ebens, H. (2001a). The distribution of realized stock return volatility. *Journal of Financial Economics*, 61(1):43–76.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2001b). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association*, 96(453):42–55.

Andersen, T. G., Bollerslev, T., and Meddahi, N. (2011). Realized volatility forecasting and market microstructure noise. *Journal of Econometrics*, 160(1):220–234.

Baillie, R. T., Bollerslev, T., and Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 74(1):3–30.

Barndorff-Nielsen, O. E. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(2):253–280.

Berkowitz, J., Christoffersen, P., and Pelletier, D. (2011). Evaluating value-at-risk models with desk-level data. *Management Science*, 57(12):2213–2227.

Blair, B. J., Poon, S.-H., and Taylor, S. J. (2001). Forecasting S&P 100 volatility: the incremental information content of implied volatilities and high-frequency index returns. *Journal of Econometrics*, 105(1):5–26.

Bollerslev, T., Patton, A. J., and Quaedvlieg, R. (2016). Exploiting the errors: A simple approach for improved volatility forecasting. *Journal of Econometrics*, 192(1):1–18.

Brandt, M. W. and Jones, C. S. (2006). Volatility forecasting with range-based egarch models. *Journal of Business & Economic Statistics*, 24(4):470–486.

Brooks, C. (2014). *Introductory Econometrics for Finance*. Cambridge University Press.

Brownlees, C. T. and Gallo, G. M. (2010). Comparison of volatility measures: A risk management perspective. *Journal of Financial Econometrics*, 8(1):29–56.

Chen, C. W., Gerlach, R., Hwang, B. B., and McAleer, M. (2012). Forecasting value-at-risk using nonlinear regression quantiles and the intra-day range. *International Journal of Forecasting*, 28(3):557–574.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, 7(2):174–196.

Corsi, F., Mittnik, S., Pigorsch, C., and Pigorsch, U. (2008). The volatility of realized volatility. *Econometric Reviews*, 27(1-3):46–78.

Diebold, F. X. and Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 20(1):134–144.

Engle, R. F. and Manganelli, S. (2004). CAViaR: Conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics*, 22(4):367–381.

Fuertes, A.-M. and Olmo, J. (2013). Optimally harnessing inter-day and intra-day information for daily value-at-risk prediction. *International Journal of Forecasting*, 29(1):28–42.

Gerlach, R. and Wang, C. (2016). Forecasting risk via realized garch, incorporating the realized range. *Quantitative Finance*, 16(4):501–511.

Gerlach, R. H., Chen, C. W., and Chan, N. Y. (2012). Bayesian time-varying quantile forecasting for value-at-risk in financial markets. *Journal of Business & Economic Statistics*, 9(4):481–492.

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. *Journal of Finance*, 48(5):1779–1801.

Gneiting, T. and Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378.

Hansen, P. R., Huang, Z., and Shek, H. H. (2012). Realized GARCH: A joint model for returns and realized measures of volatility. *Journal of Applied Econometrics*, 27(6):877–906.

Hansen, P. R. and Lunde, A. (2006). Consistent ranking of volatility models. *Journal of Econometrics*, 131(1):97–121.

Hansen, P. R., Lunde, A., and Nason, J. M. (2011). The model confidence set. *Econometrica*, 79(2):453–497.

Heber, G., Lunde, A., Shephard, N., and Sheppard, K. (2009). *Oxford-Man Institute's Realized Library, Version 0.2*. Oxford-Man Institute, University of Oxford.

Hua, J. and Manzan, S. (2013). Forecasting the return distribution using high-frequency volatility measures. *Journal of Banking & Finance*, 37(11):4381–4403.

Maheu, J. M. and McCurdy, T. H. (2011). Do high-frequency measures of volatility improve forecasts of return distributions? *Journal of Econometrics*, 160(1):69–76.

Manganelli, S. and Engle, R. F. (2004). Value at risk models in finance. In Szegö, G., editor, *Risk Measures for the 21st Century*, pages 123–143. Wiley, Chichester, UK.

Martens, M., Van Dijk, D., and De Pooter, M. (2009). Forecasting S&P 500 volatility: Long memory, level shifts, leverage effects, day-of-the-week seasonality, and macroeconomic announcements. *International Journal of Forecasting*, 25(2):282–303.

Nieto, M. R. and Ruiz, E. (2016). Frontiers in var forecasting and backtesting. *International Journal of Forecasting*, 32(2):475–501.

Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of Business*, 53(1):61–65.

Pritsker, M. (2006). The hidden dangers of historical simulation. *Journal of Banking & Finance*, 30(2):561–582.

Rogers, L. C. and Zhou, F. (2008). Estimating correlation from high, low, opening and closing prices. *Annals of Applied Probability*, 18(2):813–823.

Shephard, N. and Sheppard, K. (2010). Realising the future: Forecasting with high-frequency-based volatility (heavy) models. *Journal of Applied Econometrics*, 25(2):197–231.

Taylor, J. W. (2008). Using exponentially weighted quantile regression to estimate value at risk and expected shortfall. *Journal of Financial Econometrics*, 6(3):382–406.

Tsiakas, I. (2008). Overnight information and stochastic volatility: A study of european and us stock exchanges. *Journal of Banking & Finance*, 32(2):251–268.

Wang, X., Wu, C., and Xu, W. (2015). Volatility forecasting: The role of lunch-break returns, overnight returns, trading volume and leverage effects. *International Journal of Forecasting*, 31(3):609–619.

Žikeš, F. and Baruník, J. (2015). Semi-parametric conditional quantile models for financial returns and realized volatility. *Journal of Financial Econometrics*, 14(1):185–226.