Finance Research Letters 58 (2023) 103916

Contents lists available at ScienceDirect

Finance Research Letters

journal homepage: www.elsevier.com/locate/frl

Sample frequency robustness and accuracy in forecasting Value-at-Risk for Brent Crude Oil futures

Christian Ewald [a,b], Jelena Hadina [a], Erik Haugom [a,*], Gudbrand Lien [a], Ståle Størdal [a], Muhammad Yahya [c]

[a] *Department of Business Administration, Inland School of Business and Social Sciences, Inland Norway University of Applied Sciences, Lillehammer, 2604, Norway*
[b] *University of Glasgow, Adam Smith Business School, Glasgow, G12 8QQ, UK*
[c] *School of Economics and Business, Norwegian University of Life Sciences, Ås, 1433, Norway*

A R T I C L E I N F O

Keywords:
Realized volatility
Sample frequency
Value-at-Risk forecasting
HAR-RV
HAR-QREG

A B S T R A C T

In this paper we examine how sensitive Value-at-Risk (VaR) forecasts based on simple linear quantile regressions are to the sampling frequency used to calculate realized volatility. We use sampling frequencies from one to 108 min for ICE Brent Crude Oil futures and test the out-of-sample performance of a set of quantile regression models using formal coverage tests. The results show that a one-factor model performs exceptionally well for most sampling frequencies used to calculate realized volatility. In comparison with the well-known Heterogeneous Autoregressive Model of Realized Volatility (HAR-RV) and a quantile regression version of the HAR model (HAR-QREG), we also find that the one-factor model is much less sensitive to the sampling frequency used to calculate realized volatility.

1. Introduction

Value-at-Risk (VaR) is one of the most commonly used measures to quantify and manage financial risk and it has entrenched itself as a common benchmark for empirical checks, and to control and measure the underlying uncertainty of a portfolio. VaR forecasts have typically been obtained using more or less advanced time series models of the volatility of the assets or portfolio under consideration using returns at daily frequencies (see e.g., Mabrouk, 2009; Shao et al., 2009). Over the past decades, however, research has shown that more precise volatility estimates can be obtained by using high-frequency (intra-daily) data and the concept of realized volatility (Andersen and Bollerslev, 1998; Andersen et al., 2001a,b).

Using quantile regression to forecast VaR is not new. Haugom et al. (2016) for example use a quantile version of the HAR-model of Corsi (2009) combined with volatility estimates obtained from daily data with good results when compared to a set of alternative models. High-frequency data and realized volatility have also been used to forecast VaR in several studies (see Louzis et al., 2014 for a review). Combining the use of high-frequency data and realized volatility with quantile regression to forecast VaR has received less attention though. This is somewhat puzzling given the well-documented properties of realized volatility as a good estimator of the true (unobserved) volatility of financial assets, and because VaR is simply the conditional quantile of a portfolio return distribution. A rare example of previous research combining realized volatility with quantile regression is the seminal paper by Žikeš and Baruník (2015). They calculate realized volatility for S&P500- and WTI Crude Oil futures and use simple linear quantile regression models to forecast conditional daily returns. Their results show that the quantile regression models perform well compared to more advanced benchmark models and the findings hold across assets and the examined quantiles.

* Corresponding author.
E-mail address: erik.haugom@inn.no (E. Haugom).

https://doi.org/10.1016/j.frl.2023.103916
Received 26 January 2023; Received in revised form 11 April 2023; Accepted 17 April 2023
Available online 19 May 2023
1544-6123/

Our study builds on the work by Žikeš and Baruník (2015) but differs in several ways. First, Žikeš and Baruník (2015) use a 5-minute sampling frequency for the returns to calculate realized volatility. Hence, they do not examine how the sampling frequency affects the performance of the various VaR forecast. Though much extant research on realized volatility uses a 5-minute sampling frequency to balance the impact of microstructure noise with efficiency (see e.g., Andersen et al. (2007), Corsi et al. (2010)), the so-called signature plot cannot necessarily be used to assess how sensitive the VaR forecasts are to the sampling frequency. Though access to high-frequency data may not be an issue for most agents in the industry, the principle of parsimony still holds. That is, if equally good VaR forecasts can be obtained using a lower sampling frequency to calculate realized volatility, the simpler approach (lower frequency) should be used. Our main objective in this study is therefore to assess the sensitivity in the VaR forecasts based on quantile regression across sampling frequency used to calculate realized volatility. To assess this sensitivity we use sampling frequencies of between one and 108 min in our empirical analysis.

Second, Žikeš and Baruník (2015) focus solely on the main trading hours (9:30–16:00 EST) when calculating the realized volatility measures used in the quantile regressions to forecast VaR. We also construct volatility estimates and perform VaR predictions for the whole day (24 h) as suggested by Hansen and Lunde (2005). By doing so, we can also assess how sensitive the 24 h VaR forecasts are to the sampling frequency used to calculate the realized volatility measure for the whole day.

Third, while Žikeš and Baruník (2015) test their models when forecasting 5%- and 10% VaR in both tails, we also include both the 1%- and 2.5% VaR levels. Methodology-wise we focus on the quantile version of the HAR model proposed by Corsi (2009) and a simple one-factor model where only the current day estimate of realized volatility is used to generate the next day's VaR forecast. We then compare these models with the corresponding OLS alternatives combined with Gaussian critical quantile values to form the VaR forecasts as done in Haugom et al. (2014). We also include the results of the Dynamic Quantile Regression proposed by Laporta et al. (2018) and reported to perform well compared with a set of other models, including the CAViAR models of Engle and Manganelli (2004).

Our article thus contributes to answering a few conceptual questions:

- How does sampling frequency affect the quality of VaR forecasts for Brent Crude oil based on simple linear quantile regression models?
- How important is information about historical volatility beyond the one-day horizon when forming the VaR forecasts based on linear quantile regression models?
- Are there differences in how sensitive the VaR forecasts are across sampling frequencies for techniques including historical weekly- and monthly volatility estimates compared with techniques using just the last available one-day estimate of volatility?

Our results show that the relatively simple one-factor model can compete with more complex models in terms of accuracy and provides an advantage in terms of robustness against changes in the sampling frequency. The VaR forecasts based on quantile regression models are generally less sensitive to the sampling frequency used to calculate realized volatility compared with forecasts based on standard regression models. Among the quantile regression models, the one-factor model is the least sensitive to the sampling frequency used to calculate realized volatility. We also find that our simple one-factor model performs better both within a standard regression- and a quantile regression framework when compared with a corresponding three-factor (HAR) model. Finally, the DQR model used by Laporta et al. (2018) performs poorly in our analysis and is only able to pass the conditional coverage test for the 10% VaR level for the 24-hour returns.

While our emphasis is on one-day-ahead VaR forecasting, our paper naturally connects to the literature on forecasting (realized) volatility in general. Examples are (Chen et al., 2020) who use auto-regressive models to forecast realized volatility of crude oil futures prices, and Thomakos and Wang (2003) who use high-frequency data and realized volatility for various currency and index futures. We stress at this point that our focus is entirely on one-day ahead forecasts. Ghysels et al. (2019) assess direct against iterated multi-period forecasts on returns' variance within the GARCH and RV families with a similar agenda, but focusing on the longer term and not including quantile regression-based methods. However the latter indicates that the approach can be extended to include multi-period quantile forecasts using the methodology developed in Ghysels et al. (2016). In fact Žikeš and Baruník (2015) include a 5- and 10-day horizon in their forecasts of conditional returns and volatilities. A natural extension could therefore be to examine how sensitive longer-horizon VaR forecasts are to the sampling frequency used to calculate realized volatility. We leave this as potential future work. Such work could focus on testing the usefulness of using high-frequency data for longer-horizon VaR forecasts in general, in line with the work done by Ghysels and Valkanov (2012). Nevertheless, we note that these studies are thematically aligned with ours by emphasizing that a relatively simple methodology can produce excellent results when using high-frequency data.

Our work also contributes more generally to the literature that looks at how econometric methodology is affected by sample frequency and the transition from discrete to continuous time. Key contributions in this literature include McCrorie (2009), Chambers (2011), and Chambers et al. (2018). Robustness against changes in sampling frequency is an attractive attribute of an estimator and in our case selects a simple quantile regression among the alternatives.

Overall, despite the significant importance of VaR, research focusing on VaR forecasting based on high-frequency data for Brent Crude oil is still scarce (Brownlees and Gallo, 2010; Grigoletto and Lisi, 2009; Nieto and Ruiz, 2016). We focus our analysis on Brent Crude oil futures as it is, by far, the most important commodity traded globally (Sadorsky, 2006). In fact, understanding the variability of crude oil prices is of fundamental importance and has been found to predict a significant number of industry portfolios, as shown by Wang et al. (2016). As noted by Žikeš and Baruník (2015), oil prices also exhibit substantially higher volatility than stock indices or FX and thus serve as a good test of methodology on less well-behaved financial time series.

The rest of the paper is organized as follows: In Section 2 we briefly discuss the theory of quadratic variation and describe our approach to forecasting VaR. Section 3 presents the data and preliminary analyses while our main results are presented in Section 4. Finally, we conclude with a brief summary and outlook in Section 5, illustrating various extensions of our approach which can be examined further in future work.

2. Quadratic variation and realized volatility

2.1. Realized volatility

Let $r_{i,t}$ be the continuously compounded return for sub-period i on day t and let M be the sampling frequency in one time unit (i.e., one day in this case). Realized variance is then defined as,

$$RV_t = \sum_{i=1}^{M} r_{i,t}^2. \quad (1)$$

Hence, the term realized variance is the sum of the intra-daily squared returns sampled at short intervals. The mathematical theory of quadratic variation suggests that the following holds if the discretely sampled returns are serially uncorrelated and the sample paths for σ_t are continuous (see e.g., Karatzas and Shreve, 1988; Poon and Granger, 2003; Andersen et al., 2003):

$$p - \lim_{M\to\infty} \left(\int_0^1 \sigma_{t+\tau}^2 d\tau - \sum_{i=1}^{M} r_{i,t}^2 \right) \to 0, \quad (2)$$

where the limit above denotes the limit in probability. This result shows that the latent (unobservable) variance can be measured (almost) perfectly using discrete data if the sampling frequency is high enough.[1] But what exactly is high enough? Previous research has shown that micro-structure noise induces bias in the variance estimate obtained by realized variance (see e.g., Bandi and Russell, 2008) when using very high sampling frequencies. In practice, there are also computational costs and costs associated with storing excessive amounts of high-frequency data. These costs must be compared with the potential benefits of more accurate VaR forecasts. To make a valid assessment of the trade-off between sampling frequency and forecast accuracy, the risk modeler needs decision support based on the application the data are meant for.

Forecast accuracy should of course remain as the primary objective, but if a forecast is highly sensitive to the sampling frequency in the data, then it disqualifies itself, if the true underlying model is in fact in continuous time. Sample frequency robustness should therefore be used as an additional criterion when choosing a specific forecasting method, and in our case this secondary criterion selects our simple one factor quantile regression model against some of the more advanced models we investigated.

2.2. Overnight returns and whole-day realized volatility

For some markets, trading takes place around the clock and a variance measure for the whole day can be readily obtained from high-frequency data. For ICE Brent Crude oil futures the trading hours are from 12:00 AM to 10.00 PM London time (GMT). One could therefore use data from this 22-hour trading window and probably obtain good estimates of the whole-day-variance (24-h). The information 'lost' in the two-hour window is likely low for most trading days. From both a practical and an academic perspective, however, there are reasons to choose another approach. First, liquidity is relatively low during night-time GMT, and particularly so for the early sample period. Low liquidity will induce many zero-returns at the highest sampling frequencies and market microstructure noise can influence the volatility estimates (see e.g., Andersen et al. (2003); (Hansen and Lunde, 2006)). Second, many financial institutions require VaR forecasts to be handed in by the end of the 'regular' trading day. Waiting until 10.00 PM (GMT) to obtain all the intra-daily information to form a next-day VaR forecast is then not an option.[2] We therefore construct a whole-day variance measure by properly scaling the realized-variance measure for regular business hours (from 8.00 AM to 5.00 PM London time) and the last available overnight return variance. The same intra-day sampling period for Brent Crude oil futures was used by Haugom et al. (2014) and Haugom and Ray (2017). In a similar way, Hansen and Lunde (2006) also discarded observations outside "regular trading hours" (from 9.30 AM to 4.00 PM) when studying transaction prices from NYSE and NASDAQ.

Several ways of constructing such a 24-h 'whole-day-variance' measure have been suggested in the literature, but we rely on the procedure described by Hansen and Lunde (2005), which is simple to implement and has been shown to perform well in empirical applications.

---

[1] Over recent years, volatility estimation using such high-frequency, intra-daily, data has been subject to growing interest among researchers and practitioners and the concept is by now well established in the literature. A full review will not be provided here. See the following references for detailed descriptions and/or empirical examination of realized volatility using high-frequency data: Dacorogna et al. (2001), Andersen and Bollerslev (1998), McAleer and Medeiros (2008) and Kambouroudis et al. (2016).

[2] Another issue is that in periods with extreme volatility huge additional margin requirements could be imposed by the clearing house within a trading day and with a short time limit. Institutions are therefore in need of such VaR forecasts to estimate the financial needs for margin requirements.

The aim is to find optimal weights, ω_1 and ω_2, for the overnight and business-hours variance by minimizing the variance of the whole day. If we define $r^2_{co,t}$ as the overnight variance[3] and RV_t business-hours variance - both for day t - we can, according to Hansen and Lunde (2005), calculate a realized-variance measure for the whole day as follows:

$$RV_t^* = \hat{\omega}_1 r^2_{co,t} + \hat{\omega}_2 RV_t \quad (3)$$

The approach to calculating the optimal weights involves estimating the expected values, the variance, and covariance of the whole-day variance ($r^2_{co,t} + RV_t$), the overnight variance ($r^2_{co,t}$) and the business-hours variance (RV_t). The calculation steps are thoroughly described by Hansen and Lunde (2005) and Haugom et al. (2014) and will not be repeated here.

3. Methodology

3.1. Value-at-Risk and the proposed model

Value-at-Risk (VaR) is one of the most widely used risk measures in financial risk management. It can be defined as the loss, in present value terms, that we are α% confident will not be exceeded if the portfolio is held static over a certain period of h days (Alexander, 2009). Formally, if we let $\{r_t\}_{t=1}^T$ denote a time series of portfolio returns, the challenge is to forecast VaR_{t+1} such that:

$$P[r_{t+1} < \text{VaR}_{t+1} | \Omega_t] = \alpha \quad (4)$$

where Ω_t is the information set available at the time the VaR forecast is made. The information set will vary among various players in the marketplace and trying to include all possible variables that investors may be using when making trading decisions is an impossible task. All of the information that the various market players possess will, however, ultimately be reflected in the price movements of the asset under consideration, as suggested by the efficient market hypothesis (Fama, 1970). The current volatility of the portfolio returns is therefore an obvious choice if the objective is to operationalize the information set in an efficient and coherent way. The problem is then to forecast VaR_{t+h} such that:

$$P[r_{t+h} < \text{VaR}_{t+h} | \sigma_t] = \alpha, \quad (5)$$

where σ_t is the day t volatility which can be accurately estimated by realized volatility. The current day volatility reflects not only the information becoming available on day t, but also what investors believe is relevant historical information, and expectations about the future.

As VaR is the conditional quantile of the portfolio return distribution, we can use the quantile regression method by Koenker and Bassett (1978) to estimate the h-day α% VaR. If we let $q = \alpha$ and $Y_q = \text{VaR}_{q,t+1}$, we can use the definition given in Eq. (5) to express the conditional quantile function as:

$$\hat{Y}_q | \sigma_t = \hat{\alpha}_q + \hat{\beta}_q RV_t. \quad (6)$$

The model is easily estimated for all relevant conditional q quantiles ($0 < q < 1$) by solving the minimization problem presented in Koenker and Bassett (1978):

$$\min_{\alpha_q, \beta_q} \sum_{t=1}^{T} (q - 1_{Y_q \le \alpha_q + \beta_q RV_t})(Y_q - (\alpha_q + \beta_q RV_t)), \quad (7)$$

where

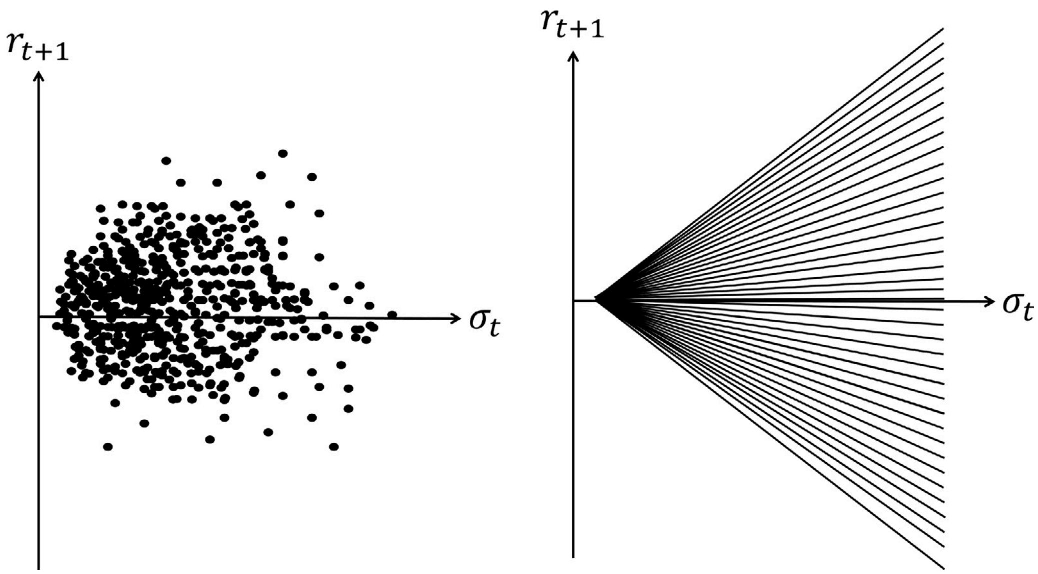
$$1_{Y_q \le \alpha_q + \beta_q RV_t} = \begin{cases} 1 & \text{if } Y_q \le \alpha_q + \beta_q RV_t \\ 0 & \text{otherwise.} \end{cases} \quad (8)$$

An illustration of the basic properties of the model is presented in Fig. 1. Panel 1(a) shows how a scatterplot of next day's return (r_{t+1}) and an estimate of current day volatility (σ_t) could look like. Panel 1(b) illustrates conditional quantile estimates based on Eq. (7). The line with the steepest positive slope would reflect a conditional quantile in the far right tail of the return distribution, and vice versa. The illustration highlights a few key points. First, the estimated slope for a conditional quantile will increase in absolute terms the further away from the conditional median. Second, the estimate of the conditional quantile will increase with the current day volatility, except for the conditional median and conditional quantiles very close to that. The latter point reflects the aspect of volatility clustering as high current day volatility will induce more extreme next day tail predictions compared with low current day volatility levels. Hence, small price changes will tend to be followed by small price changes of either sign, and vice versa.

To assess the performance of the proposed one-factor model, we also include VaR forecasts based on the Heterogeneous Autoregressive Model of Realized Variance (HAR-RV) as proposed by Corsi (2009):

$$\widehat{RV}_{t+1} = \hat{\alpha} + \hat{\beta}_d RV_t + \hat{\beta}_w RV_{t,t-4} + \hat{\beta}_m RV_{t,t-19} + \varepsilon_t \quad (9)$$

[3] The subscript co stands for close-open.



(a) Next day return and current volatility (b) Conditional quantile estimtates

Fig. 1. Illustration of the potential relationship between next day return and current daily volatility (left panel) and conditional quantile estimates (right panel). The conditional quantile lines with the steepest positive slopes would reflect the right tail of the return distribution and vice versa.

where RV_{t+1} denotes next day realized volatility, RV_t is the observed realized volatility today, $RV_{t,t-4}$ is the average observed realized volatility over the last five business days, and $RV_{t,t-19}$ is the average realized volatility over the last four business weeks (i.e., the last 20 business days).[4]

The volatility forecasts from this model are then combined with Gaussian critical quantile values to form the VaR forecasts. One may argue that return distributions are rarely Gaussian and that the above approach therefore is doomed to fail. Andersen et al. (2001b), however, have shown that daily exchange-rate returns standardized by realized volatility are close to Gaussian. The approach was also used by Haugom et al. (2014) when using high-frequency data of Crude oil futures with good results. Nevertheless, this approach is less direct as it first needs a forecast of the volatility and then make a distributional assumption to obtain the VaR forecast, while quantile regression directly targets VaR.

One could argue though that the more or less static approach based on Eq. (6) is missing dynamic properties of the realized volatility process and is therefore less suitable to make longer term or multi-stage forecasts as compared to (9). We therefore include a quantile version of the three factor HAR-model (called HAR-QREG) and an OLS version of the one factor model combined with Gaussian quantile values in our analysis.

$$\widehat{Y}_q|\sigma_t = \hat{\alpha}_q + \hat{\beta}_{d,q} RV_t + \hat{\beta}_{w,q} RV_{t,t-4} + \hat{\beta}_{m,q} RV_{t,t-19} + \varepsilon_{q,t}, \tag{10}$$

$$\widehat{RV}_{t+1} = \hat{\alpha} + \hat{\beta}_d RV_t + \varepsilon_t, \tag{11}$$

Additionally, we test the performance of the DQR model which has recently been used by Laporta et al. (2018) and reported to perform well:

$$\widehat{Y}_q|r_t = \hat{\alpha}_q + \hat{\beta}_q r_t + \varepsilon_{q,t}. \tag{12}$$

That is, the conditional quantile the next business day is predicted using only the latest available return (r_t).

3.2. Evaluating the value-at-Risk forecasts

To assess the performance of the VaR forecasts based on the various models using sampling frequencies of between one and 108 min we apply both an unconditional- and a conditional coverage test. The accuracy of a given VaR forecast can simply be assessed by comparing the fraction of the ex-post returns that exceed the forecasted VaR level:

$$I_{t+1} = \begin{cases} 1 & \text{if } r_{t+1} < \text{VaR}_{t+1,\alpha} \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$

[4] Note that the HAR-RV model is in a sense a so-called MIDAS regression with step-functions. See Ghysels and Valkanov (2012) and Ghysels et al. (2007) for details.

Kupiec (1995) then formulated a test to assess whether the number of violations is sufficiently close to the VaR level of interest:

$$LR_{uc} = -2log\left[\frac{\alpha^{n1}(1-\alpha)^{n0}}{\hat{\pi}^{n1}(1-\hat{\pi})^{n0}}\right] \sim \chi^2(1) \quad (14)$$

where n_0 and n_1 represent the number of non-VaR violations and VaR violations at the given α quantile level and $\hat{\pi} = n_1/(n_0 + n_1)$ is thus the observed proportion of violations. The test assumes independent violations and the obtained test statistic is compared to the critical value of the chi-square distribution with one degree of freedom at a chosen significance level.

Christoffersen (1998) correctly points out, however, that the assumption of independent violations is not always fulfilled in practice. He, therefore, proposed a joint likelihood ratio test statistic of unconditional coverage and independence as follows:

$$LR_{cc} = -2log\frac{(1-\alpha^{n0})\alpha^{n1}}{(1-\hat{\pi}_{n01})^{n00}\hat{\pi}_{01}^{n01}(1-\hat{\pi}_{11})^{n10}\hat{\pi}_{11}^{n11}} \sim \chi^2(2) \quad (15)$$

where π_{ij} are the transition probabilities and n_{ij} is the number of observations with value i followed by j. As the test of Christoffersen (1998) included both an assessment of unconditional coverage and independence we focus on this when reporting the results for expository reasons. The unconditional coverage test results are available upon request.

4. Data, analysis, and results

The data used in our empirical analysis consist of almost 16 years of high-frequency transaction level data for the front-month Brent Crude oil futures contracts traded at the Intercontinental Exchange (ICE). The sample period extends from January 3, 2006 to October 29, 2021, for a total of 4048 trading days. The contract price is given in US dollars and cents per barrel and the contract size is 1,000 barrels (42,000 US gallons).

For our sample period, trading in each contract ceases at the end of the designated period of settlement on the Business Day (a trading day that is not a public holiday in England and Wales) immediately preceding: (i) Either the fifteenth day before the first day of the contract month, if this fifteenth day is a Business Day, or (ii) if this fifteenth day is not a Business Day, the next preceding Business Day (ICE, 2018).[5]

The front-month contract is by far the most liquid contract, and it is also the price movements of this contract that is reported in the news worldwide. In the empirical analysis, we therefore focus on the front-month contract. Haugom et al. (2014) note, however, that the trading volume on the last trading day of the front-month contract is only one tenth of the average trading volume for the rest of the trading days for ICE Brent Crude futures. We therefore follow their procedure and roll over to the second-position contract for the last trading day of the first-position contract. We employ this rollover strategy over the entire sample period.

Fig. 2 shows the daily realized volatility estimates based on sampling frequencies from one minute to one hour (60 min).[6] We note that the sample includes the very volatile period associated with the Covid-19 outbreak when daily realized volatility reached a maximum of approximately 20%. From the plots it is also evident that the variance of the realized volatility estimates decreases with sampling frequency. The signature plot presented in Fig. 3 confirms this: The standard deviation of realized volatility is negatively related to the sampling frequency (right y-axis) and is more than 13% higher when using the lowest sampling frequency (~ 0.0085) compared with using the highest sampling frequency (~ 0.0075) to calculate it.[7] The same figure also shows that the average realized volatility estimate is highly dependent on the sampling frequency. Using a sampling frequency of one minute induce an average realized volatility estimate that is more than 10% higher compared with the realized volatility estimates obtained with sampling frequencies of one hour or lower.

Andersen et al. (1999) suggest that the signature plot of realized volatility itself can be used as a simple approach to assess the sampling frequency. The aim is to find the highest sampling frequency for which realized volatility is approximately constant. The corresponding measure of realized volatility sampled at this frequency is fairly accurate and free of micro-structure bias (Bollerslev et al., 2008). In our case, the signature plot suggest that at a sampling frequency of approximately 20 min should be used to balance between the theoretical result of improved precision of realized volatility at higher sampling frequencies and problems related to market micro-structure noise occurring in practice.

In the current study, however, we are first and foremost interested in examining how sensitive the VaR forecasts are to the sampling frequency used to calculate realized volatility. We therefore do not focus on the optimal sampling frequency of realized volatility per se.

[5] A new expiration date procedure was implemented by ICE for contract months beginning March 2016. Specifically, for these contracts, trading ceases at the end of the designated settlement period on the last Business Day of the second month preceding the relevant contract month (e.g. the March contract will expire on the last Business Day of January). If the day on which trading is due to cease would be either: (i) the Business Day preceding Christmas Day, or (ii) the Business Day preceding New Year's Day, then trading shall cease on the next preceding Business Day. See ICE (2018) for further details.

[6] To simplify language and notation we refer to sampling frequency in minutes rather than per minute or per hour. A sampling frequency of 60 min means that one sample is taken every sixty minutes. A higher sampling frequency corresponds to a lower number of minutes between the sampled prices. Note that all the sampling frequencies we employ ensure that the last sampled price is exactly at 5:00 PM. E.g., when using a sampling frequency of 108 min, we sample prices at 8.00 AM; 9.48 AM; 11.36 AM; 1.24 PM; 3.12 PM; 5.00 PM.

[7] Note: The estimator of each days realized volatility has a standard deviation and based on the law of large numbers, this standard deviation will clearly decrease when the sampling frequency goes up, i.e. the more data are used. This is reflected in Eq. (2). However, this is not the standard deviation that Fig. 2 reflects upon. Here we are looking at the estimates of realized volatility on each day and the standard deviation of these as in the time series of daily realized volatilities.

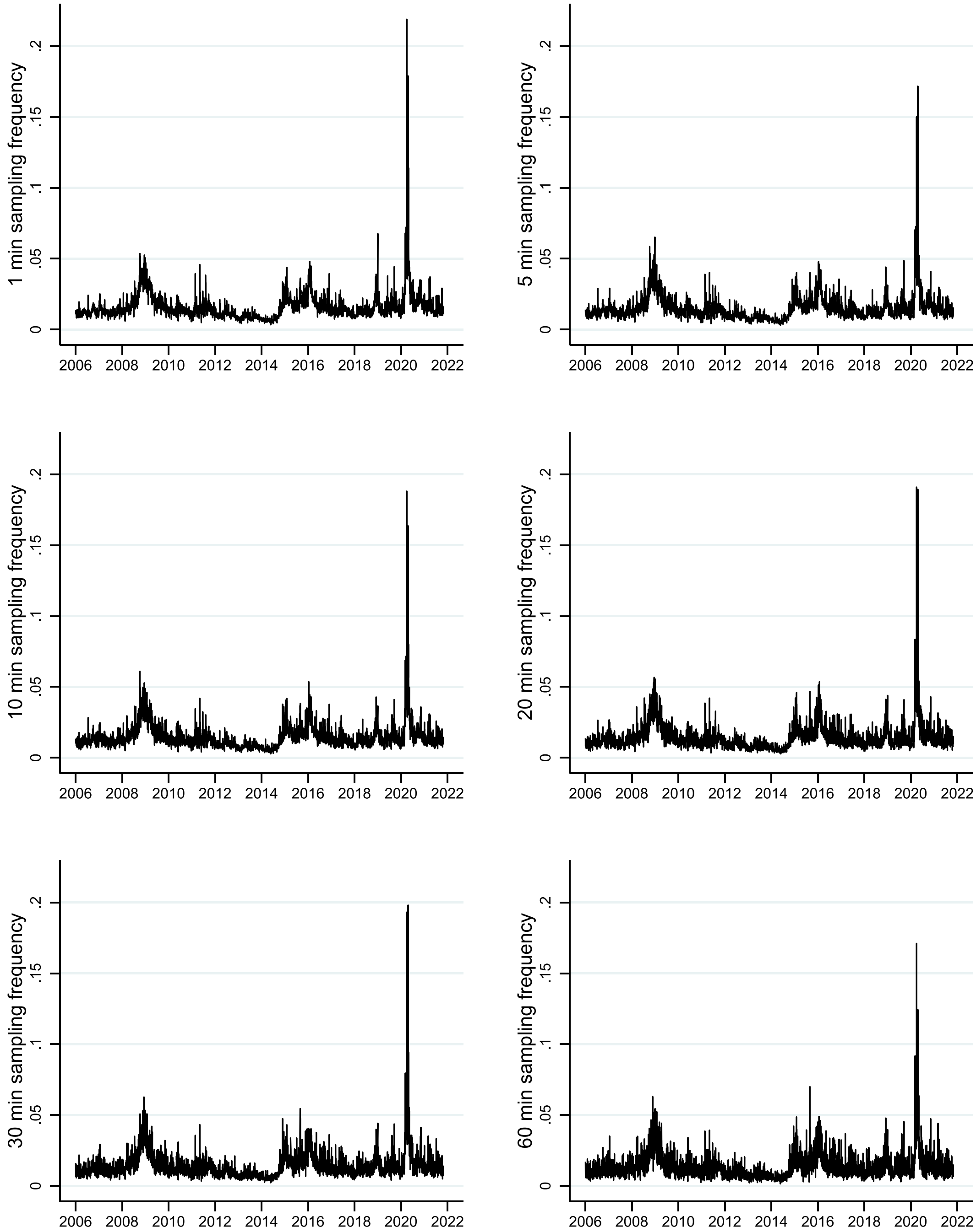


Fig. 2. Realized volatility calculated from various sampling frequencies.

The empirical results for the models using realized volatility are presented in Tables 1, 2, 3, and 4. The two former focus on business hours returns, while the two latter show the results for 24 h returns. Table 5 presents the results for the DQR model as proposed by Laporta et al. (2018). In all cases we estimate the models using a rolling window of 500 business days (corresponding to approximately two years) and make one-day ahead VaR forecasts out of the estimation sample. The estimates are thus solely obtained with the information available at the time the forecast is made.

Focus first on the results presented in Table 1 and note that the quantile regression version of the one factor model (right panel) in general performs better than the OLS version of the same model, based on the conditional coverage test. For the highest sampling frequencies between one and ten minutes, the quantile regression version passes almost all the examined VaR levels. Though the

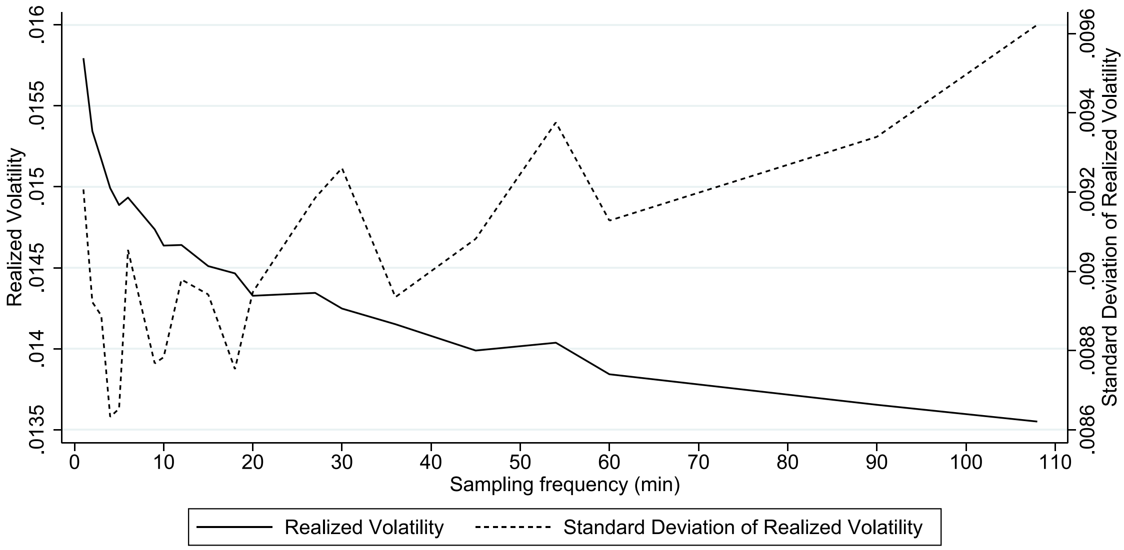


Fig. 3. Average realized volatility and its standard deviation across sampling frequencies. The sampling frequency used to calculate daily realized volatility is given on the x-axis and ranges between one and 108 min.

Table 1
VaR forecast based on one-factor models using daily realized volatility.

	Business hours returns (open-close)															
	HAR-RV: One factor model								HAR-QREG: One factor model							
	Left tail				Right tail				Left tail				Right tail			
Sampling	**1%**	**2.50%**	**5%**	**10%**	**10%**	**5%**	**2.50%**	**1%**	**1%**	**2.50%**	**5%**	**10%**	**10%**	**5%**	**2.50%**	**1%**
1	1.72%	3.17%	5.37%	9.78%	7.86%	4.01%	2.09%	0.90%	1.39%	2.71%	4.81%	8.99%	10.29%	5.09%	2.86%	1.22%
2	1.92%	3.73%	5.65%	10.15%	8.17%	4.13%	2.32%	0.99%	1.27%	2.71%	4.89%	9.05%	10.12%	5.12%	2.77%	1.16%
3	2.01%	3.79%	5.94%	10.57%	8.48%	4.41%	2.35%	1.10%	1.30%	2.71%	4.81%	9.08%	10.38%	5.32%	2.66%	1.33%
4	2.09%	3.99%	6.19%	10.55%	8.51%	4.38%	2.52%	1.27%	1.27%	2.69%	4.86%	9.13%	10.09%	5.26%	2.77%	1.19%
5	2.12%	3.82%	6.45%	10.72%	8.68%	4.50%	2.60%	1.24%	1.13%	2.71%	4.89%	9.10%	10.15%	5.12%	2.74%	1.24%
6	2.21%	3.90%	6.50%	10.97%	8.93%	4.44%	2.63%	1.33%	1.22%	2.71%	4.92%	9.30%	10.23%	5.06%	2.66%	1.16%
9	2.26%	4.04%	6.42%	11.20%	9.05%	4.83%	2.83%	1.33%	1.24%	2.80%	4.78%	9.30%	10.29%	5.06%	2.69%	1.13%
10	2.23%	3.99%	6.53%	11.22%	9.08%	4.86%	2.88%	1.30%	1.24%	2.71%	4.98%	9.44%	10.35%	5.15%	2.71%	1.24%
12	2.29%	3.99%	6.47%	11.48%	8.88%	4.64%	2.86%	1.47%	1.36%	2.77%	5.00%	9.41%	10.29%	5.20%	2.63%	0.90%
15	2.37%	4.27%	6.50%	11.45%	9.19%	4.86%	2.80%	1.47%	1.47%	2.77%	4.95%	9.41%	10.15%	5.17%	2.83%	1.13%
18	2.40%	4.16%	6.64%	11.51%	8.99%	4.95%	2.88%	1.67%	1.24%	2.80%	5.12%	9.50%	9.95%	5.15%	2.54%	1.33%
20	2.54%	4.35%	6.70%	11.90%	9.61%	5.17%	3.00%	1.58%	1.36%	2.86%	4.89%	9.44%	10.07%	5.15%	2.69%	1.19%
27	2.40%	4.44%	6.96%	11.99%	9.36%	5.15%	3.14%	1.78%	1.39%	2.86%	4.92%	9.67%	10.32%	5.06%	2.83%	1.36%
30	2.52%	4.33%	6.81%	11.73%	9.47%	5.06%	3.19%	1.67%	1.36%	2.86%	4.78%	9.56%	10.01%	5.17%	2.77%	1.19%
36	2.77%	4.47%	7.24%	12.10%	9.75%	5.09%	3.03%	1.78%	1.53%	3.05%	5.00%	9.47%	10.29%	5.34%	2.80%	1.27%
45	2.63%	4.86%	7.04%	12.41%	9.67%	5.46%	3.42%	2.04%	1.39%	2.77%	5.03%	9.50%	10.40%	5.15%	2.57%	1.44%
54	2.54%	5.00%	7.35%	12.04%	9.84%	5.26%	3.56%	1.98%	1.39%	3.08%	5.09%	9.98%	10.40%	5.32%	2.66%	1.44%
60	3.00%	4.98%	7.75%	12.72%	9.78%	5.57%	3.65%	2.21%	1.22%	2.97%	5.00%	9.75%	10.18%	5.15%	2.66%	1.39%
90	3.19%	5.29%	7.80%	12.92%	10.09%	5.57%	3.82%	2.21%	1.47%	3.00%	5.09%	9.73%	10.09%	5.34%	2.60%	1.58%
108	3.39%	5.57%	8.14%	13.12%	10.74%	5.88%	3.85%	2.52%	1.36%	3.14%	5.17%	10.07%	10.15%	5.32%	2.60%	1.27%
Min	1.72%	3.17%	5.37%	9.78%	7.86%	4.01%	2.09%	0.90%	1.13%	2.69%	4.78%	8.99%	9.95%	5.06%	2.54%	0.90%
Max	3.39%	5.57%	8.14%	13.12%	10.74%	5.88%	3.85%	2.52%	1.53%	3.14%	5.17%	10.07%	10.40%	5.34%	2.86%	1.58%
Range	1.67	2.40	2.77	3.34	2.88	1.87	1.75	1.61	0.40	0.45	0.40	1.07	0.45	0.28	0.31	0.68
Pass	0%	0%	10%	35%	60%	90%	70%	40%	70%	95%	100%	100%	100%	100%	95%	85%

Notes. Green color indicates that the VaR forecasts at the given level pass the conditional coverage test.

OLS version of the one factor model (left panel) also performs well, it consistently fails to pass the conditional coverage test at the 10% VaR level in the right tail — with only a few exceptions.

The second thing to note is that the one factor quantile regression model continuous to perform well also for the lower sampling frequencies used to calculate realized volatility. Even when using a one hour sampling frequency the model passes seven out of the eight VaR levels tested in the empirical exercise. The corresponding results for the OLS version is worse. The performance drops quickly when using a sampling frequency beyond 10 min. Using a sampling frequency of 20 min the model only passes three out of the eight defined VaR levels and from 54 min and lower frequencies the model only passes two out of eight. The finding is related to the accuracy of the VaR forecast themselves. For all the examined VaR levels the range of the predicted exceedances are generally much higher for the OLS version compared with the quantile regression model. For example, if using a sampling frequency of 108 min the forecasted exceedances at the 1% left tail VaR is at 2.91%. This is almost three times what it should be. The corresponding result for the quantile one factor model is 1.75%.

Let us now focus on the results in Table 2, where the results of the HAR-RV and HAR-QREG models are presented. We notice that the OLS version now actually performs best when using the highest sampling frequencies to calculate realized volatility. For a

Table 2
VaR forecast based on three-factor models using daily, weekly, and monthly realized volatility. Business hours returns.

Business hours returns (open-close)																
	HAR-RV: Three factor model								HAR-QREG: Three factor model							
	Left tail				Right tail				Left tail				Right tail			
Sampling	**1%**	**2.50%**	**5%**	**10%**	**10%**	**5%**	**2.50%**	**1%**	**1%**	**2.50%**	**5%**	**10%**	**10%**	**5%**	**2.50%**	**1%**
1	1.81%	3.14%	5.23%	9.75%	8.03%	4.07%	2.26%	0.76%	1.72%	3.11%	4.83%	9.02%	10.60%	5.60%	3.00%	1.53%
2	1.98%	3.42%	5.77%	10.09%	8.51%	4.24%	2.40%	0.90%	1.55%	3.00%	4.75%	9.19%	10.55%	5.37%	2.88%	1.58%
3	2.01%	3.56%	6.02%	10.40%	8.68%	4.55%	2.49%	1.02%	1.55%	3.03%	4.86%	9.13%	10.80%	5.40%	3.05%	1.61%
4	2.12%	3.53%	6.08%	10.55%	8.85%	4.55%	2.54%	1.10%	1.55%	2.97%	4.78%	9.27%	10.97%	5.40%	2.91%	1.67%
5	2.23%	3.73%	6.19%	10.80%	9.05%	4.64%	2.63%	1.22%	1.47%	3.03%	4.72%	9.10%	10.63%	5.48%	3.00%	1.50%
6	2.06%	3.68%	5.97%	10.72%	9.10%	4.78%	2.63%	1.10%	1.50%	3.08%	4.81%	9.16%	10.69%	5.57%	3.08%	1.70%
9	2.35%	3.93%	6.45%	11.11%	9.41%	4.89%	2.80%	1.27%	1.39%	3.11%	4.69%	9.10%	10.91%	5.51%	3.17%	1.33%
10	2.40%	3.99%	6.53%	11.17%	9.39%	4.89%	2.71%	1.24%	1.67%	3.05%	4.83%	9.22%	10.72%	5.63%	2.97%	1.64%
12	2.29%	3.90%	6.28%	11.34%	9.56%	4.86%	2.83%	1.30%	1.58%	3.11%	4.64%	9.16%	10.66%	5.51%	3.11%	1.53%
15	2.49%	3.99%	6.76%	11.39%	9.53%	5.03%	3.00%	1.44%	1.47%	3.00%	4.89%	9.05%	10.57%	5.37%	2.97%	1.55%
18	2.46%	4.07%	6.56%	11.45%	9.87%	5.00%	3.05%	1.47%	1.64%	2.83%	4.55%	9.08%	10.35%	5.40%	2.94%	1.50%
20	2.49%	3.96%	6.79%	11.73%	9.87%	5.17%	2.88%	1.55%	1.53%	2.97%	4.81%	9.08%	10.43%	5.43%	2.94%	1.70%
27	2.60%	4.01%	6.76%	11.62%	9.78%	5.29%	3.28%	1.61%	1.44%	2.86%	4.72%	9.22%	10.60%	5.60%	2.88%	1.47%
30	2.57%	3.96%	6.96%	11.56%	9.98%	5.40%	3.22%	1.61%	1.47%	2.77%	4.78%	9.13%	10.69%	5.65%	3.03%	1.55%
36	2.54%	4.01%	7.12%	12.02%	10.15%	5.17%	3.08%	1.58%	1.53%	2.94%	4.86%	8.96%	10.80%	5.40%	3.11%	1.55%
45	2.83%	4.13%	7.55%	12.19%	10.38%	5.40%	3.34%	1.95%	1.64%	2.80%	4.66%	8.99%	10.94%	5.43%	3.14%	1.50%
54	2.69%	4.30%	7.24%	12.07%	10.38%	5.46%	3.25%	1.64%	1.58%	2.80%	4.58%	9.02%	10.86%	5.20%	2.97%	1.61%
60	2.83%	4.35%	7.58%	12.50%	10.32%	5.82%	3.42%	1.92%	1.36%	2.83%	4.81%	9.36%	10.72%	5.37%	2.83%	1.33%
90	3.00%	4.50%	7.89%	12.98%	10.86%	6.02%	3.65%	1.95%	1.33%	2.94%	4.86%	9.16%	11.11%	5.43%	2.77%	1.33%
108	3.03%	4.64%	8.00%	12.92%	10.80%	6.16%	3.73%	2.18%	1.61%	3.00%	4.92%	9.27%	10.88%	5.32%	2.86%	1.47%
Min	1.81%	3.14%	5.23%	9.75%	8.03%	4.07%	2.26%	0.76%	1.33%	2.77%	4.55%	8.96%	10.35%	5.20%	2.77%	1.33%
Max	3.03%	4.64%	8.00%	12.98%	10.86%	6.16%	3.73%	2.18%	1.72%	3.11%	4.92%	9.36%	11.11%	5.65%	3.17%	1.70%
Range	1.22	1.50	2.77	3.22	2.83	2.09	1.47	1.41	0.40	0.34	0.37	0.40	0.76	0.45	0.40	0.37
Pass	0%	5%	10%	35%	80%	85%	65%	45%	10%	100%	100%	80%	65%	100%	95%	15%

Notes. Green color indicates that the VaR forecasts at the given level pass the conditional coverage test.

Table 3
VaR forecast based on one-factor models using daily realized volatility. 24 hours returns.

24 hours returns (close-close)																
	HAR-RV: One factor model								HAR-QREG: One factor model							
	Left tail				Right tail				Left tail				Right tail			
Sampling	**1%**	**2.50%**	**5%**	**10%**	**10%**	**5%**	**2.50%**	**1%**	**1%**	**2.50%**	**5%**	**10%**	**10%**	**5%**	**2.50%**	**1%**
1	1.39%	2.71%	5.03%	9.16%	7.32%	3.99%	1.98%	0.96%	1.44%	2.69%	4.58%	9.73%	10.23%	5.40%	2.94%	1.44%
2	1.58%	3.00%	5.40%	9.39%	7.66%	4.13%	2.18%	1.05%	1.39%	2.54%	4.66%	9.73%	10.26%	5.57%	2.91%	1.44%
3	1.58%	3.11%	5.51%	9.78%	7.80%	4.07%	2.15%	1.13%	1.44%	2.54%	4.81%	9.73%	10.21%	5.29%	2.91%	1.44%
4	1.61%	3.17%	5.51%	9.92%	7.92%	4.33%	2.29%	1.10%	1.39%	2.43%	4.64%	9.58%	10.23%	5.37%	2.94%	1.55%
5	1.70%	3.11%	5.68%	10.12%	8.14%	4.35%	2.32%	1.24%	1.47%	2.54%	4.81%	9.92%	10.21%	5.37%	2.94%	1.47%
6	1.70%	3.11%	5.65%	10.12%	7.97%	4.47%	2.49%	1.19%	1.41%	2.60%	4.78%	9.73%	10.32%	5.23%	2.86%	1.44%
9	1.67%	3.48%	5.82%	10.26%	8.06%	4.55%	2.60%	1.22%	1.44%	2.54%	4.78%	9.75%	10.09%	5.29%	2.94%	1.41%
10	1.67%	3.42%	5.97%	10.57%	8.37%	4.55%	2.52%	1.39%	1.30%	2.66%	5.09%	9.92%	10.38%	5.48%	3.00%	1.33%
12	1.75%	3.56%	5.91%	10.49%	8.26%	4.52%	2.74%	1.33%	1.47%	2.52%	4.83%	9.84%	10.04%	5.29%	2.91%	1.47%
15	1.81%	3.31%	5.85%	10.69%	8.57%	4.52%	2.63%	1.44%	1.41%	2.57%	5.03%	10.23%	9.92%	5.34%	2.80%	1.55%
18	1.75%	3.59%	6.11%	10.63%	8.43%	4.72%	2.71%	1.36%	1.41%	2.54%	5.00%	10.12%	10.07%	5.40%	2.97%	1.41%
20	1.72%	3.65%	6.11%	10.91%	8.51%	4.61%	2.77%	1.58%	1.44%	2.57%	4.98%	9.98%	10.18%	5.17%	2.91%	1.41%
27	2.01%	3.90%	6.08%	10.63%	8.54%	4.81%	2.71%	1.50%	1.61%	2.63%	4.64%	9.84%	10.26%	5.29%	2.91%	1.47%
30	1.95%	3.70%	6.16%	10.88%	8.62%	4.81%	2.69%	1.64%	1.44%	2.57%	5.03%	10.18%	10.04%	5.43%	2.97%	1.58%
36	2.09%	4.16%	6.22%	10.83%	8.54%	4.95%	2.83%	1.64%	1.50%	2.66%	4.95%	9.75%	9.98%	5.43%	3.03%	1.61%
45	2.21%	3.90%	6.56%	10.94%	8.85%	5.23%	3.03%	1.53%	1.70%	2.54%	5.06%	9.95%	9.90%	5.23%	2.80%	1.53%
54	2.18%	4.13%	6.70%	11.28%	8.76%	5.23%	2.94%	1.75%	1.67%	2.71%	5.26%	10.07%	10.15%	4.95%	3.05%	1.53%
60	2.49%	4.38%	6.93%	11.48%	8.93%	5.37%	2.94%	1.78%	1.47%	2.71%	5.32%	9.70%	10.07%	4.95%	2.94%	1.55%
90	2.43%	4.10%	6.84%	11.82%	8.82%	5.29%	3.19%	1.81%	1.55%	2.54%	4.78%	10.04%	9.84%	5.26%	2.91%	1.44%
108	2.77%	4.81%	7.07%	11.59%	9.16%	5.68%	3.28%	1.89%	1.61%	2.80%	5.23%	10.12%	9.95%	5.26%	2.80%	1.47%
Min	1.39%	2.71%	5.03%	9.16%	7.32%	3.99%	1.98%	0.96%	1.30%	2.43%	4.58%	9.58%	9.84%	4.95%	2.80%	1.33%
Max	2.77%	4.81%	7.07%	11.82%	9.16%	5.68%	3.28%	1.89%	1.70%	2.80%	5.32%	10.23%	10.38%	5.57%	3.05%	1.61%
Range	1.39	2.09	2.04	2.66	1.84	1.70	1.30	0.93	0.40	0.37	0.74	0.65	0.54	0.62	0.25	0.28
Pass	0%	25%	25%	80%	5%	85%	90%	40%	20%	95%	100%	95%	100%	100%	90%	5%

Notes. Green color indicates that the VaR forecasts at the given level pass the conditional coverage test.

sampling frequency of between one and six minutes the HAR-RV model passes almost all the defined VaR levels. The corresponding results for the HAR-QREG show that it consistently fails to pass the 10%, 2.5%, and 1% VaR levels in the right tail. When using lower sampling frequencies to calculate realized volatility, however, the results are consistent with the findings from the one factor model: the performance of the quantile regression model is more stable across sampling frequencies compared with the OLS version.

When comparing the general performance of the proposed one factor model with the three factor model, the conclusion is clear: the one factor quantile regression model passes more of the conditional coverage tests compared with the three factor version. The results for the corresponding OLS versions are not that clear, but the results suggest that there is not a big gain in performance from going from the one factor model to the three factor model in this case either.

Tables 3 and 4 present the results for the 24 h (close–close) returns. The general findings are consistent with those reported for the business hours returns. The one factor quantile regression model performs well across all sampling frequencies while the corresponding OLS model is not able to keep up when the sampling frequency is reduced. Again the three factor models generally perform worse than the one factor models when using 24 h returns, but the main finding is in line with that reported earlier: the sampling frequency affects the performance of the OLS version more than it does affect the quantile regression version.

Table 4
VaR forecast based on three-factor models using daily, weekly, and monthly realized volatility. 24 hours returns.

24 hours returns (close-close)																
	HAR-RV: Three factor model								HAR-QREG: Three factor model							
	Left tail				Right tail				Left tail				Right tail			
Sampling	**1%**	**2.50%**	**5%**	**10%**	**10%**	**5%**	**2.50%**	**1%**	**1%**	**2.50%**	**5%**	**10%**	**10%**	**5%**	**2.50%**	**1%**
1	1.19%	2.54%	5.03%	9.08%	7.58%	3.76%	1.92%	0.96%	1.55%	2.60%	4.78%	9.39%	10.60%	5.51%	3.11%	1.78%
2	1.30%	2.74%	5.32%	9.44%	7.89%	3.99%	2.09%	1.05%	1.58%	2.66%	4.66%	9.56%	10.57%	5.48%	3.31%	1.84%
3	1.27%	3.00%	5.51%	9.70%	8.11%	4.16%	2.12%	1.05%	1.50%	2.43%	4.92%	9.56%	10.77%	5.43%	3.22%	1.81%
4	1.41%	3.08%	5.80%	9.78%	8.11%	4.24%	2.18%	1.13%	1.44%	2.52%	4.72%	9.47%	10.63%	5.65%	3.14%	1.78%
5	1.39%	3.08%	5.77%	9.98%	8.37%	4.35%	2.21%	1.13%	1.47%	2.49%	4.66%	9.61%	10.55%	5.54%	3.31%	1.72%
6	1.41%	3.19%	5.63%	10.01%	8.45%	4.27%	2.37%	1.16%	1.44%	2.57%	4.72%	9.64%	10.72%	5.71%	3.36%	1.84%
9	1.58%	3.25%	5.80%	10.12%	8.43%	4.47%	2.32%	1.22%	1.47%	2.69%	4.78%	9.41%	10.52%	5.29%	3.34%	1.87%
10	1.58%	3.36%	5.99%	10.29%	8.48%	4.35%	2.40%	1.16%	1.58%	2.66%	4.75%	9.47%	10.66%	5.68%	3.42%	1.92%
12	1.64%	3.48%	5.82%	10.21%	8.57%	4.38%	2.40%	1.19%	1.64%	2.71%	4.78%	9.50%	10.57%	5.37%	3.36%	1.78%
15	1.64%	3.19%	5.99%	10.18%	8.62%	4.47%	2.57%	1.39%	1.47%	2.74%	4.72%	9.50%	10.49%	5.54%	3.31%	1.70%
18	1.75%	3.51%	6.22%	10.21%	8.54%	4.44%	2.60%	1.27%	1.61%	2.74%	4.69%	9.47%	10.18%	5.40%	3.25%	1.78%
20	1.81%	3.53%	6.42%	10.38%	8.91%	4.61%	2.52%	1.33%	1.50%	2.83%	4.81%	9.56%	10.52%	5.34%	3.14%	1.87%
27	1.67%	3.59%	6.22%	10.60%	8.99%	4.72%	2.71%	1.33%	1.55%	2.71%	4.75%	9.58%	10.57%	5.26%	3.03%	1.67%
30	1.78%	3.51%	6.25%	10.43%	9.16%	4.83%	2.74%	1.44%	1.55%	2.57%	4.75%	9.64%	10.60%	5.60%	3.14%	1.72%
36	2.01%	3.68%	6.22%	10.60%	9.22%	4.83%	2.66%	1.41%	1.50%	2.74%	4.81%	9.53%	10.43%	5.26%	3.14%	1.75%
45	1.81%	3.96%	6.36%	10.74%	9.22%	5.09%	2.74%	1.41%	1.64%	2.69%	4.81%	9.27%	10.72%	5.51%	3.00%	1.67%
54	1.81%	3.90%	6.47%	10.88%	9.47%	5.03%	2.77%	1.44%	1.50%	2.71%	5.03%	9.41%	10.66%	5.54%	3.03%	1.72%
60	2.06%	4.18%	6.53%	10.88%	9.73%	5.34%	2.94%	1.47%	1.50%	2.57%	4.92%	9.47%	10.63%	5.26%	2.86%	1.61%
90	2.09%	4.27%	6.64%	10.94%	9.84%	5.32%	2.91%	1.67%	1.33%	2.63%	4.75%	9.50%	10.12%	5.40%	2.97%	1.72%
108	2.18%	4.50%	6.84%	11.05%	9.70%	5.46%	3.11%	1.58%	1.44%	2.86%	4.98%	9.36%	10.32%	5.46%	2.86%	1.36%
Min	1.19%	2.54%	5.03%	9.08%	7.58%	3.76%	1.92%	0.96%	1.33%	2.43%	4.66%	9.27%	10.12%	5.26%	2.86%	1.36%
Max	2.18%	4.50%	6.84%	11.05%	9.84%	5.46%	3.11%	1.67%	1.64%	2.86%	5.03%	9.64%	10.77%	5.71%	3.42%	1.92%
Range	0.99	1.95	1.81	1.98	2.26	1.70	1.19	0.71	0.31	0.42	0.37	0.37	0.65	0.45	0.57	0.57
Pass	15%	10%	0%	100%	45%	90%	95%	0%	5%	80%	95%	90%	100%	100%	45%	0%

Notes. Green color indicates that the VaR forecasts at the given level pass the conditional coverage test.

Table 5
Dynamic Quantile Regression (DQR): Business hours and 24 h returns.

	Left tail				Right tail				Pass
Period	**1%**	**2.50%**	**5%**	**10%**	**10%**	**5%**	**2.50%**	**1%**	
Business hours	1.78%	3.31%	6.16%	10.63%	10.07%	5.71%	3.31%	1.58%	0/8
24 hours	1.89%	3.48%	5.91%	10.60%	9.98%	5.29%	2.97%	1.67%	0/8

Notes. VaR forecasts based on the Dynamic Quantile Regression proposed by Laporta et al. (2018). Green color indicates that the VaR forecasts at the given level pass the conditional coverage test.

For comparison we include the performance of the DQR model proposed by Laporta et al. (2018) in Table 5. Our results show that the DQR model is only able to pass one out of the eight examined VaR levels (the 10% right tail VaR), and only so when considering 24 hours returns. Laporta et al. (2018) tested the model on Brent crude oil spot data using a sample period from 2002 to 2017, but only examined the 95% and 99% VaR levels. Their results show that the DQR model passed the conditional coverage test at both these VaR levels. Our analyses are based on futures prices and the results are therefore not directly comparable with those reported by Laporta et al. (2018). However, the differences in performance between the DQR model and our proposed one factor model based on realized volatility are so large that most likely we would find similar results for Brent crude spot prices.

5. Summary, implications, and directions for future research

We propose a one factor quantile regression model based on realized volatility to forecast Value-at-Risk and empirically examine the model using Brent Crude oil futures, evaluating its performance across various sampling frequencies used to calculate the daily realized volatility. The results are compared with VaR forecasts obtained from the well-known (three factor) HAR-RV model and the quantile regression version of this. We also examine a one factor regression model of realized volatility to further assess the difference in performance between using OLS and quantile regression to forecast VaR.

The results are summarized in Tables 6 and 7 and show that the one factor quantile regression model performs as good as, or better than the corresponding three factor model when aggregating the results across all sampling frequencies used to calculate realized volatility. The quantile regression models generally perform better than the corresponding one- and three factor OLS models. Among the OLS models, there is no clear answer to whether the one- or three factor model is best. The general findings hold for both business hours and 24 hours (close-close) returns.

Our results should be good news to practitioners: a very simple one factor quantile regression model can be used to forecast VaR with good results. In addition, the model is not very sensitive to the sampling frequency used to estimate the volatility. Even with realized volatility estimates obtained from sampling only once every 108 min, the model performs relatively well.

A number of avenues for future research emerge. First, we need to stress that our empirical observations may be specific to crude oil. The model performance and the sensitivity to sampling frequency used to calculate realized volatility should therefore be examined across more securities. Clements et al. (2008) looked at quantile forecasts of daily exchange rate returns based on forecasted realized volatility and it would be interesting to adapt our approach to this setting as well as investigate its applicability

Table 6
VaR forecasts passing conditional coverage test: Business hours returns.

	VaR	OLS		Quantile regression	
		One factor	Three factor	One factor	Three factor
Left tail	1.00%	0%	0%	70%	10%
	2.50%	0%	5%	95%	100%
	5.00%	10%	10%	100%	100%
	10.00%	35%	35%	100%	80%
Right tail	10.00%	60%	80%	75%	65%
	5.00%	90%	85%	100%	100%
	2.50%	70%	65%	95%	95%
	1.00%	40%	45%	85%	15%

Notes. Fraction of VaR forecasts passing the conditional coverage test across all sampling frequencies for one and three factor OLS- and quantile regression models using business hours returns.

Table 7
VaR forecasts passing the conditional coverage test: Business hours returns.

	VaR	OLS		Quantile regression	
		One factor	Three factor	One factor	Three factor
Left tail	1.00%	0%	15%	20%	5%
	2.50%	25%	10%	95%	80%
	5.00%	25%	0%	100%	95%
	10.00%	80%	100%	95%	90%
Right tail	10.00%	5%	45%	100%	100%
	5.00%	85%	90%	100%	100%
	2.50%	90%	95%	90%	45%
	1.00%	40%	0%	5%	0%

Notes. Fraction of VaR forecasts passing the conditional coverage test across all sampling frequencies for one and three factor OLS- and quantile regression models using 24 h (close-close) returns.

to other commodities such as natural gas and electricity. Nevertheless, our findings have important implications for financial risk managers with access to high-frequency data.

Second, disentangling the continuous variation from jumps has been subject to substantial interest in research on realized volatility over recent years (see e.g., Patton and Sheppard, 2015). The performance of the proposed one factor model should therefore be compared with models that separate the realized volatility into a continuous and jump component. If the performance of the one factor model is similar to that obtained with a more complicated model separating the total variance into a continuous- and jump component, one may ask if the extra effort required to properly extract the jumps is needed when it comes to VaR forecasting.

CRediT authorship contribution statement

Christian Ewald: Writing – original draft, Writing – review & editing. **Jelena Hadina:** Writing – review & editing. **Erik Haugom:** Conceptualization, Methodology, Data curation, Writing – original draft, Writing – review & editing, Formal analysis, Project administration, Software, Visualization. **Gudbrand Lien:** Writing – original draft, Writing – review & editing. **Ståle Størdal:** Writing – original draft, Writing – review & editing. **Muhammad Yahya:** Writing – original draft, Writing – review & editing, Visualization.

Data availability

Data will be made available on request.

References

Alexander, C., 2009. Market Risk Analysis, Value at Risk Models, vol. 4. John Wiley & Sons.

Andersen, T.G., Bollerslev, T., 1998. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. Internat. Econom. Rev. 39 (4), 885–905, URL http://www.jstor.org/stable/2527343.

Andersen, T.G., Bollerslev, T., Diebold, F.X., 2007. Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. Rev. Econ. Stat. 89 (4), 701–720.

Andersen, T.G., Bollerslev, T., Diebold, F.X., Ebens, H., 2001a. The distribution of realized stock return volatility. J. Financ. Econ. 61 (1), 43–76. http://dx.doi.org/10.1016/S0304-405X(01)00055-1, URL http://www.sciencedirect.com/science/article/pii/S0304405X01000551.

Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 1999. (Understanding, optimizing, using and forecasting) realized volatility and correlation, Working Paper.

Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2001b. The distribution of realized exchange rate volatility. J. Amer. Statist. Assoc. 96 (453), 42–55. http://dx.doi.org/10.1198/016214501750332965.

Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2003. Modeling and forecasting realized volatility. Econometrica 71 (2), 579–625. http://dx.doi.org/10.1111/1468-0262.00418.
Bandi, F.M., Russell, J.R., 2008. Microstructure Noise, Realized Variance, and Optimal Sampling. Rev. Econom. Stud. 75 (2), 339–369. http://dx.doi.org/10.1111/j.1467-937X.2008.00474.x.
Bollerslev, T., Law, T.H., Tauchen, G., 2008. Risk, jumps, and diversification. J. Econometrics 144 (1), 234–256.
Brownlees, C.T., Gallo, G.M., 2010. Comparison of volatility measures: a risk management perspective. J. Financ. Econom. 8 (1), 29–56.
Chambers, M.J., 2011. Cointegration and sampling frequency. Econom. J. 14 (2), 156–185, URL http://www.jstor.org/stable/23118074.
Chambers, M.J., McCrorie, J.R., Thornton, M.A., 2018. Continuous time modelling based on an exact discrete time representation. Contin. Time Model. Behav. Relat. Sci. (14), 317–357.
Chen, W., Ma, F., Wei, Y., Liu, J., 2020. Forecasting oil price volatility using high-frequency data: New evidence. Int. Rev. Econ. Finance 66, 1–12.
Christoffersen, P.F., 1998. Evaluating interval forecasts. Internat. Econom. Rev. 841–862.
Clements, M.P., Galvao, A.B., Kim, J.H., 2008. Quantile forecasts of daily exchange rate returns from forecasts of realized volatility. J. Empir. Financ. 15 (4), 729–750. http://dx.doi.org/10.1016/j.jempfin.2007.12.001, URL https://www.sciencedirect.com/science/article/pii/S0927539807001168.
Corsi, F., 2009. A simple approximate long-memory model of realized volatility. http://dx.doi.org/10.1093/jjfinec/nbp001,
Corsi, F., Pirino, D., Reno, R., 2010. Threshold bipower variation and the impact of jumps on volatility forecasting. J. Econometrics 159 (2), 276–288.
Dacorogna, M., Gençay, R., Muller, U., Pictet, O., Olsen, R., 2001. An Introduction to High-Frequency Finance. Academic Press, California, USA.
Engle, R.F., Manganelli, S., 2004. Conditional autoregressive value at risk by regression quantiles. J. Bus. Econom. Statist. 22 (4), 367–381.
Fama, E.F., 1970. Efficient capital markets: A review of theory and empirical work. J. Finance 25 (2), 383–417.
Ghysels, E., Plazzi, A., Valkanov, R., Rubia, A., Dossani, A., 2019. Direct versus iterated multiperiod volatility forecasts. Annu. Rev. Finan. Econ. 11 (1), 173–195. http://dx.doi.org/10.1146/annurev-financial-110217-022808.
Ghysels, E., Plazzie, A., Valkanov, R., 2016. Why invest in emerging markets? The role of conditional return asymmetry. J. Finance 71 (5), 2145–2191, URL http://www.jstor.org/stable/44155388.
Ghysels, E., Sinko, A., Valkanov, R., 2007. MIDAS regressions: Further results and new directions. Econometric Rev. 26 (1), 53–90. http://dx.doi.org/10.1080/07474930600972467.
Ghysels, E., Valkanov, R., 2012. Forecasting volatility with MIDAS. In: Handbook of Volatility Models and their Applications. Wiley Online Library, pp. 383–401.
Grigoletto, M., Lisi, F., 2009. Looking for skewness in financial time series. Econom. J. 12 (2), 310–323.
Hansen, P.R., Lunde, A., 2005. A realized variance for the whole day based on intermittent high-frequency data. J. Financ. Econom. 3 (4), 525–554. http://dx.doi.org/10.1093/jjfinec/nbi028, URL http://jfec.oxfordjournals.org/content/3/4/525.abstract.
Hansen, P.R., Lunde, A., 2006. Realized variance and market microstructure noise. J. Bus. Econom. Statist. 24 (2), 127–160.
Haugom, E., Ray, R., 2017. Heterogeneous traders, liquidity, and volatility in crude oil futures market. J. Commod. Mark. 5, 36–49. http://dx.doi.org/10.1016/j.jcomm.2017.01.001, URL http://www.sciencedirect.com/science/article/pii/S240585131630188X.
Haugom, E., Ray, R., Ullrich, C.J., Veka, S., Westgaard, S., 2016. A parsimonious quantile regression model to forecast day-ahead value-at-risk. Finance Res. Lett. 16, 196–207. http://dx.doi.org/10.1016/j.frl.2015.12.006, URL http://www.sciencedirect.com/science/article/pii/S1544612315001385.
Haugom, E., Veka, S., Lien, G., Westgaard, S., 2014. Estimating and evaluating Value-at-Risk forecasts based on realized variance: empirical evidence from ICE Brent Crude oil futures. OPEC Energy Rev. 38 (4), 373–397. http://dx.doi.org/10.1111/opec.12024.
ICE, 2018. Brent crude futures - contract specifications. available at https://www.theice.com/products/219/brent-crude-futures.
Kambouroudis, D.S., McMillan, D.G., Tsakou, K., 2016. Forecasting stock return volatility: A comparison of GARCH, implied volatility, and realized volatility models. J. Futures Mark. 36 (12), 1127–1163. http://dx.doi.org/10.1002/fut.21783, URL https://onlinelibrary.wiley.com/doi/abs/10.1002/fut.21783.
Karatzas, I., Shreve, S.E., 1988. Brownian Motion and Stochastic Calculus. Springer-Verlag, New York.
Koenker, R., Bassett, G., 1978. Regression quantiles. Econometrica 46 (1), 33–50, URL http://www.jstor.org/stable/1913643.
Kupiec, P.H., 1995. Techniques for verifying the accuracy of risk measurement models. J. Deriv. 3, 73–84.
Laporta, A.G., Merlo, L., Petrella, L., 2018. Selection of value at risk models for energy commodities. Energy Econ. 74, 628–643. http://dx.doi.org/10.1016/j.eneco.2018.07.009, URL https://www.sciencedirect.com/science/article/pii/S0140988318302548.
Louzis, D.P., Xanthopoulos-Sisinis, S., Refenes, A.P., 2014. Realized volatility models and alternative value-at-risk prediction strategies. Econ. Model. 40, 101–116. http://dx.doi.org/10.1016/j.econmod.2014.03.025, URL https://www.sciencedirect.com/science/article/pii/S0264999314001229.
Mabrouk, S., 2009. Value-at-risk and expected shortfall estimations based on GARCH-type models: evidence from energy commodities. J. Energy Develop. 35 (1/2), 279–314.
McAleer, M., Medeiros, M.C., 2008. Realized volatility: A review. Econometric Rev. 27 (1–3), 10–45. http://dx.doi.org/10.1080/07474930701853509.
McCrorie, J.R., 2009. Estimating continuous-time models on the basis of discrete data via an exact discrete analog. Econom. Theory 25 (4), 1120–1137.
Nieto, M.R., Ruiz, E., 2016. Frontiers in VaR forecasting and backtesting. Int. J. Forecast. 32 (2), 475–501.
Patton, A.J., Sheppard, K., 2015. Good volatility, bad volatility: Signed jumps and the persistence of volatility. Rev. Econ. Stat. 97 (3), 683–697.
Poon, S.-H., Granger, C.W.J., 2003. Forecasting volatility in financial markets: A review. J. Econ. Lit. 41 (2), 478–539, URL http://www.jstor.org/stable/3216966.
Sadorsky, P., 2006. Modeling and forecasting petroleum futures volatility. Energy Econ. 28 (4), 467–488. http://dx.doi.org/10.1016/j.eneco.2006.04.005, URL https://www.sciencedirect.com/science/article/pii/S014098830600048X.
Shao, X.-D., Lian, Y.-J., Yin, L.-Q., 2009. Forecasting value-at-risk using high frequency data: The realized range model. Glob. Finance J. 20 (2), 128–136. http://dx.doi.org/10.1016/j.gfj.2008.11.003, URL https://www.sciencedirect.com/science/article/pii/S104402830900043X.
Thomakos, D.D., Wang, T., 2003. Realized volatility in the futures markets. J. Empir. Financ. 10 (3), 321–353. http://dx.doi.org/10.1016/S0927-5398(02)00052-X, URL https://www.sciencedirect.com/science/article/pii/S092753980200052X.
Wang, Y., Ma, F., Wei, Y., Wu, C., 2016. Forecasting realized volatility in a changing world: A dynamic model averaging approach. J. Bank. Financ. 64, 136–149.
Žikeš, F., Baruník, J., 2015. Semi-parametric conditional quantile models for financial returns and realized volatility. J. Financ. Econom. 14 (1), 185–226.