Modelling the Volatility of TOCOM Energy Futures: A Regime Switching Realised Volatility Approach

Amir H. Alizadeh [a]
(A.Alizadeh@city.ac.uk)

Chih-Yueh Huang [b *]
(Chih-Yueh.Huang.1@cass.city.ac.uk)

Ian W. Marsh [a]
(I.Marsh@city.ac.uk)

Affiliations

[a] Faculty of Finance, Cass Business School, City University London, 106 Bunhill Row, London EC1Y 8TZ, United Kingdom.

[b] Knowledge Transfer Partnership (KTP) Associate, Faculty of Finance, University of Kent, Canterbury, Kent, CT2 7NZ, United Kingdom.

Abstract

This paper combines the Heterogeneous Autoregressive Realised Volatility (HAR-RV) model and the Markov Regime Switching (MRS) approach to estimate and forecast volatility of energy futures contracts traded at the Tokyo Commodity Exchange (TOCOM). The proposed MRS-HAR-RV model allows the dynamics of the realised volatility to change as market conditions change. The dataset consists of intraday prices for gasoline, kerosene and crude oil futures. Estimation results suggest MRS-HAR-RV model can capture dynamics of price volatility of energy futures better than alternative models. However, out-of-sample forecast evaluation results show that MRS-HAR-RV can only produce better forecasts for more liquid contracts. Moreover, MRS-HAR-RV model seems to less over-predict and more under-predict the volatility compared to HAR-RV, HAR-RV-CJ, GARCH, and MRS-GARCH models.

Keywords: Regime-switch; TOCOM; Realised volatility; petroleum futures
JEL codes: C320; G320; Q470

---

[*] **Corresponding author**

1 Introduction

Modelling volatility of asset prices has always been a key issue in financial econometrics because accurate volatility estimates and forecasts are essential for risk management, derivatives pricing, trading strategies, as well as portfolio optimisation and asset allocation. Therefore, a large body of the literature in financial econometric is devoted to modelling price volatility and finding the most appropriate techniques to capture its dynamics to produce accurate estimated and forecast of price volatility.

Volatility of commodity prices in general, and energy prices in particular, have been commonly characterised as autoregressive processes, thus the General Autoregressive Conditional Heteroscedasticity (GARCH) type models have been widely used to estimate and forecast price volatility of energy commodities (see for example, Sadorski, 2006, Fan et al. 2008, Agnolucci, 2009, Wei et al., 2010, Klein and Walther, 2016, Herrera et al., 2018, amongst others). Although GARCH type models can capture the long memory of volatility of commodity and energy prices, the poor forecasting performance of GARCH models has been pointed out in several studies, including by Figlewski (1997), Poon and Granger (2003), Cabedo and Moya (2003) and Sadeghi and Shavvalpour (2006), amongst others. Consequently, researchers proposed and developed alternative models to capture the time-varying dynamics of volatility of energy and commodity prices, including Fractionally Integrated GARCH (FIGARCH), Heterogeneous Autoregressive Model of Realised Volatility (HAR-RV), and even the Heston model of stochastic volatility (e.g. Vo 2009). For instance, Wei et al. (2010), Liu and Wan (2012) and Wei (2012) provide evidence that autoregressive realised volatility models outperform GARCH models in forecasting volatility of energy prices.

Increase in the availability of intraday and high frequency financial data has led to the development of a new concept for estimation of volatility, namely realised volatility (RV_t). Realised volatility is defined and estimated as the sum of squared intraday price changes but does not directly take into account the long memory of volatility. To overcome this issue, Andersen, Bollerslev, Diebold, and Labys (2003) model the realized volatility using a fractionally integrated autoregressive process, while Corsi (2009) propose the Heterogeneous Autoregressive - Realised Volatility (HAR-RV) model to capture the long-memory property of realised volatility. The HAR-RV model can be considered as a three-factor stochastic volatility model where the factors are lagged one day, one week and one month realised volatilities. However, as noted in many studies the dynamics and persistence of volatility may change over time as market and trading conditions change (e.g. Alizadeh et. al. 2008, and Lee and Yoder, 2007, and Nomikos and Pouliasis, 2011). Thus, simple HAR-RV models may not be able to capture the dynamics of volatility when there are changes in the state of the market or regime shifts.

In the present study, we extend the simple HAR-RV approach to a Markov Regime Switching HAR-RV (MRS-HAR-RV) model to allow for a more flexible specification for dynamics of volatility under different market conditions. Using the MRS-HAR-RV model, we estimate the volatility of the energy futures contracts traded on Tokyo Commodity Exchange (TOCOM). Following the estimation of MRS-HAR-RV, we examine the accuracy of volatility forecasts produced by MRS-HAR-RV in comparison with alternative approaches, including simple and regime switching GARCH models, Corsi (2009) HAR-RV and HAR-RV-CJ models used by Sévi (2014).

This paper contributes to the literature in several aspects. First, we model and forecast the realised volatility of TOCOM energy futures prices, where contracts with different maturities and high frequency data are considered. Second, we introduce and employ a Markov Regime Switching Heterogeneous Autoregressive Realised Volatility (MRS-HAR-RV) model to investigate the dynamics of realised volatility under different regimes. Third, we assess and compare the dynamics of realised volatility across contract with different maturities and liquidity levels. Fourth, we document and compare the high- and low-volatility regimes of realised volatility of TOCOM energy futures based on the average level of realised volatility in each regime, as well as the persistence of volatility. Finally, we investigate the performance of the MRS-HAR-RV model in predicting volatility and Value at Risk measures against different RV models, GARCH models and Historical Simulation, using variety of metrics. The results suggest that the regime switching realised volatility approach can explain the dynamics of time-varying volatility of TOCOM energy futures prices better compared to alternative models in-sample. However, the out-of-sample prediction of price volatility is not significantly improved and while the MRS-HAR-RV model seem to reduce the over-prediction of volatility, under-prediction of volatility increases.

The rest of this paper is structured as follows. Section 2 presents the earlier studies and recent literature on modelling and forecasting volatility of energy commodities. Section 3 introduces the HAR-RV and MRS-HAR-RV models. Section 4 describes the characteristics of the high frequency data and sampling scheme. The empirical results and forecast evaluations are presented in section 5 and the final section summarises the findings and concludes.

2 Literature Review

A number of studies in the literature suggest that high-frequency data are useful for estimating and predicting asset price volatility as intraday movements in prices are less subject to measurement error than price observations at a lower frequency (Andersen and Bollerslev, 1998). In this approach, an unbiased estimator of volatility, known as realised volatility (RV), can be arrived at using the squared values of intraday returns. For instance, daily realised volatility ($\mathrm{RV}_t^{(d)}$) is defined as the sum of intraday squared returns (Andersen et al., 2001 and 2003), which has been argued to be a more efficient estimate of volatility than daily squared returns (McAleer and Medeiros, 2008). Under the assumption that returns are independent with a zero mean, RV is also an unbiased estimator of true variance. Barndorff-Nielsen and Shephard (2004), Andersen et al. (2007), and Barndorff-Nielsen and Shephard (2007) further argue the importance of accounting for jumps in the estimation of realised volatility. Andersen et al. (2006) and McAleer and Medeiros (2008) provide a thorough survey of studies on the estimation and application of realised volatility.

The first study to employ a realised volatility approach to estimate the volatility of energy commodity prices (sweet crude oil) was by Martens and Zein (2004). This was followed by Wang et al. (2008), investigating volatility of NYMEX crude oil and natural gas futures prices using RV type models. They suggest that RV is an appropriate measure of volatility in both the crude oil and natural gas markets, as well as the realised correlation (RC) between the futures prices of the two commodities. They also point out that crude oil and natural gas price volatility measures can be characterized by slowly mean-reverting fractionally integrated processes with an estimated degree of integration between 0.25 and 0.45. Wei (2012) compares the accuracy of different volatility models, including six GARCH type models, ARFIMA-RV (Autoregressive Fractionally Integrated Moving Average Realised Volatility model) and Stochastic Volatility, in forecasting the volatility of fuel oil futures on the Shanghai Futures Exchange.

In a recent study, Sévi (2014) employs intraday data to forecast the volatility of WTI crude oil futures using different models based on the decomposition of realized variance into its positive or negative part (semivariances) and its continuous or discontinuous part (jumps). Considering eleven heterogenous autoregressive (HAR) models proposed in the literature (Andersen et al. 2007, Corsi, 2009, Chen and Ghysels, 2010, and Patton and Sheppard, 2015), Sévi (2014) reports that the model with independent squared jump has best in-sample forecast but does not significantly improve the out-of-sample forecast. Haugom et al. (2014) also analyse the realised volatility of WTI crude oil futures, using Corsi (2009) augmented HAR-RV model, which incorporates implied volatility (the CBOE Crude Oil Volatility Index) as well as other market variables, including trading volume, open interest, daily returns, bid-ask spread, and the slope of the forward curve. They report that incorporating implied volatility can significantly improve short term (daily and weekly) volatility forecasts, while including other market variables improves long term (monthly) volatility forecasts. Two other studies by Tseng et al. (2009) and Ma et al. (2017) employ the HAR Realised Range-Based Volatility (HAR-RRV) and its variations to model and forecast crude oil futures prices. Both studies report that HAR-RRV with inclusion of jump and sign components perform better than simple HAR-RV in predicting volatility of crude futures. Most recently, Lyócsa and Molnár (2018) compare the forecasting performance of various HAR-RV models and find combining the forecasts from different HAR-RV specifications can improve performance of forecasts.

Another branch of literature focuses on changes in the dynamics of price volatility under different market conditions. The main approach proposed for taking into account market conditions when estimating the time-varying volatility of asset prices is the Markov Regime Switching (MRS) model proposed by Hamilton (1989). The MRS approach has been extended to GARCH models, namely MRS-GARCH, to incorporate the effect of regime changes on the dynamics of volatility in GARCH models under different regimes (market conditions). For example, Lee and Yoder (2007) apply the MRS-GARCH model to the corn and nickel futures markets and report higher, yet insignificant, variance reduction when compared to OLS and single regime GARCH hedging strategies; while Alizadeh, et al. (2008) analyse three sets of energy commodities data, crude oil, gasoline and heating oil, and also find that the use of a MRS-MGARCH model improves hedging performance. Nomikos and Pouliasis (2010) also use two-state regime switching GARCH specification to model and forecast volatility of energy futures traded in NYMEX and ICE. They report that Mix-GARCH and MRS-GARCH models better capture the persistence of volatility in sample and produce better out-of-sample forecast compared to simple GARCH models. More recently, Herrera et al. (2018) report that regime switching GARCH models tend to produce better forecast for volatility of spot oil prices over long horizons compared to simple GARCH and Exponential GARCH models.

More recently Lux et al. (2016) investigate the performance of Markov Switching Multifractal (MSM) model and different GARCH type models in forecasting oil price volatility. They also report that MSM as a regime switching model performs better than alternatives in predicting volatility and Value-at-Risk of oil price. Klein and Walters (2016) propose a Mixture Memory GARCH (MMGARCH) approach to model and forecast volatility of oil prices where two GARCH and FIGARCH specifications are combined probabilistically with a mixture proportion variable defined as a logistic function. The MMGARCH model allows the memory of the volatility process to change according to the value of the mixture proportion variable. Klein and Walters (2016) report evidence on superiority of MMGARCH model compared to other GARCH type models in forecasting oil price volatility.

Given the importance of market conditions and the behaviour of price volatility, as well as the benefits of using high frequency intraday data in estimations of volatility, we propose a regime switching model which allows for changes in the dynamics of RV according to market conditions. In particular, we utilise the MRS-HAR-RV model to examine if the realised volatility of TOCOM energy futures is regime-dependent and to determine whether the MRS-HAR-RV model produces better forecasts and VaR estimates compared to single regime HAR-RV and HAR-RV models with jumps.

3 Methodology

In this section, we discuss set up of different model for estimation and forecasting of realised volatility of energy futures prices using high frequency intraday data, including Corsi (2009) HAR-RV, Anderson et al. (2007) RV with Jumps, and regime switching extension of HAR-RV.

3.1 Estimating Volatility with High-frequency Data

We begin the discussion with a continuous-time Geometric Brownian Motion process

$$dp_t = \mu_t dt + \sigma_t dw_t \ , \tag{1}$$

where p_t is the logarithm of the instantaneous price, μ_t is the time-varying drift term, σ_t is the diffusion parameter, also known as the instantaneous volatility of p_t, and w_t is the standard Brownian Motion process. The integrated variance of the price can then be defined as the integral of instantaneous variance (σ_t^2). For instance, the one-day integrated variance, our primary variable of interest, can be expressed as

$$IV_t^{(d)} = \int_{t-1d}^{t} \sigma_s^2 dw_s. \tag{2}$$

However, the integrated variance, defined by equation (2), is by nature a latent variable, so we need to find an observable variable in order to estimate volatility.

With the availability of intraday data, the sum of intraday squared returns, known as realised variance, has been utilised as the most common approximation of integrated variance. Many studies, including Andersen et al. (2001a, 2001b, 2003) and Barndorff-Nielsen and Shephard (2002a, 2002b), have shown that the realised variance converges to the integrated variance in probability. Hence, in this paper we follow the same approach in estimating realised variance. For instance, the realised variance for a one-day window$[t-1d, t]$, divided by M Δ-frequency intervals, is estimated by

$$RV_t^{(d)} = \sum_{j=0}^{M-1} r_{t-j\times\Delta}^2 \tag{3}$$

Where $RV_t^{(d)}$ is one-day realised variance, $\Delta = 1d/M$ is the frequency of intervals, M is the number of intervals in one day, and $r_{t-j\times\Delta}$ is the intraday return at Δ-frequency interval. As mentioned earlier, the RV itself is extremely erratic and does not capture the long memory of the variance of financial assets. Therefore, different approaches have been proposed in the literature to modify the RV and capture the long memory of the process.

3.2 The Heterogeneous Autoregressive Realised Volatility Model

Corsi (2009) proposes a HAR-RV approach to estimate the realised volatility, which can capture the long memory in volatility and can be simply estimated using OLS. The HAR-RV model is specified as an autoregressive process with lagged RV at different time horizons, namely one-day, one-week and one-month RV, as determinant of future RV, in the following form

$$RV_{t+1d}^{(d)}=\beta_0+\beta^{(d)}RV_t^{(d)}+\beta^{(w)}RV_t^{(w)}+\beta^{(m)}RV_t^{(m)}+\varepsilon_{t+1d}^{(d)}, \tag{4}$$

$$\varepsilon_{t+1d}^{(d)}\sim N(0,\Sigma\ \ ),$$

where $RV_{t+1d}^{(d)}$ is one-day ahead predicted realised volatility, and $\varepsilon_{t+1d}^{(d)}$ is the stochastic error term following a normal distribution with zero mean and constant variance, Σ. Daily realised volatility ($RV_t^{(d)}$) is calculated according to equation (3), whilst weekly and monthly realised volatility ($RV_t^{(w)}$ and $RV_t^{(m)}$) are calculated as the average of the past daily realised volatility over a one week (5 days) and one month (22 days), respectively. However, our benchmark HAR-RV is slightly different from the Corsi (2009) in two aspects. First, following Sévi (2014), we do not let the realised volatilities over the three horizons to overlap. That is, weekly realised volatility is measured as the average of daily realised volatility between $t-1$ and $t-5$, and monthly realised volatility is measured as the average of daily realised volatility between $t-6$ and $t-22$, shown as in the equation (5).

$$RV_t^{(w)}=\frac{1}{5}\sum_{i=1}^{5}RV_{t-i}^{(d)} \qquad and \qquad RV_t^{(m)}=\frac{1}{17}\sum_{i=6}^{22}RV_{t-i}^{(d)} \tag{5}$$

Second, because our dataset is constructed as a continuous time series of energy futures prices by rolling the contract on the day before maturity and volatility of commodity futures price tends to increase as contract maturity approaches (Samuelson Effect), we include a dummy variable in the model which counts the days-to-rollover. Therefore, our final benchmark HAR-RV is specified as

$$RV_{t+1d}^{(d)}=\beta_0+\beta^{(d)}RV_t^{(d)}+\beta^{(w)}RV_t^{(w)}+\beta^{(m)}RV_t^{(m)}+\beta^{(DTR)}DTR_t+\varepsilon_{t+1d}^{(d)}, \tag{6}$$

$$\varepsilon_{t+1d}^{(d)}\sim N(0,\Sigma\ \ ),$$

where DTR_t is the number of days to the rollover date at time t. Based on the Samuelson Hypothesis the volatility of futures contract tends to increase as rollover date approaches, thus $\beta^{(DTR)}$ is expected to be negative and significant.

3.3 The HAR-RV with Jumps

Anderson et al. (2007) propose a version of HAR-RV model that conditions the RV on "continues RV" and "jump components", known as HAR-RV-CJ, to model volatility of S&P 500, currencies and T-Bonds. Tseng et al. (2009) and Sévi (2014) also apply the HAR-RV-CJ to model oil price volatility. To estimate a HAR-RV-CJ, first the realised volatility must be decomposed into continuous and jump components. The continuous component can be estimated by means of the jump-robust realised volatility estimation using two main approaches; that is, bi-power variation (BPV) and median realised variance (MedRV). The BPV method, proposed by Barndorff-Nielsen and Shephard (2004), is computed as the sum of the product of two consecutive absolute returns in equation (7)

$$BPV_t = \frac{\pi}{2}\sum_{j=0}^{M-2}\left|r_{t-j\times\Delta}\right|\left|r_{t-(j+1)\times\Delta}\right| \tag{7}$$

where M is the intraday sampling frequency. BPV can capture the continuous realised volatility when a jump and a non-jump return occur in two consecutive time intervals since the product of the two is reduced by the non-jump return. However, BPV has a major drawback because when consecutive jumps occur, BPV will be biased upward due to multiplication of two large returns. In the same way, a zero return causes the product of two adjacent intervals to be zero, which leads to a downward biased in BPV. To overcome this issue, Anderson et al. (2012) propose MedRV, which is a more appropriate method of estimating jump-robust realised volatility. MedRV is calculated by the sum of the square median absolute returns for three adjacent returns as follows

$$MedRV_t = \frac{\pi}{6-4\sqrt{3}+\pi}\left(\frac{M}{M-2}\right)\sum_{j=1}^{M-2} med\left(\left|r_{t-(j-1)\times\Delta}\right|, \left|r_{t-j\times\Delta}\right|, \left|r_{t-(j+1)\times\Delta}\right|\right)^2 \tag{8}$$

According to equation (8), even if there are two large jump returns, MedRV returns the square of the lower jump return, which reduces the upward bias. In addition, MedRV is zero only when at least two of three adjacent returns are zero, so the downward bias due to the occurrence of zero returns can be reduced. In this paper, we employ both techniques to measure the continuous components for the comparison reason.

Once the realised volatility is decomposed into continuous and jump components, $C_t^{(d)}$and $J_t^{(d)}$, the full Heterogeneous Autoregressive Realized Volatility with Jumps (HAR-RV-CJ) model can be defined as

$$RV_{t+1d}^{(d)} = \beta_0 + \beta_c^{(d)} C_t^{(d)} + \beta_c^{(w)} C_t^{(w)} + \beta_C^{(m)} C_t^{(m)} + \beta_J^{(d)} J_t^{(d)} \tag{9}$$

$$+\beta_J^{(w)} J_t^{(w)} + \beta_J^{(m)} J_t^{(m)} + \beta_{st}^{(DTR)} DTR_t + \varepsilon_{t+1d}^{(d)}.$$

where $C_t^{(d)}, C_t^{(w)}, C_t^{(m)}, J_t^{(d)}, J_t^{(w)}$ and $J_t^{(m)}$ are daily, weekly and monthly continuous jump components of realised volatility respectively, and weekly and monthly components are estimated by similar technique used in equation (6). See appendix A for details of jump detection process and calculation of where $C_t^{(d)}$ and $J_t^{(d)}$.

3.4 The Markov Regime Switching HAR-RV

To allow the dynamics of realised volatility of TOCOM energy futures to change depending the state of the market, we extend the simple HAR-RV to a two-state Markov Regime Switching HAR-RV. The Markov Regime Switching models developed by Hamilton (1989) and extensively used in the literature to model financial and commodity prices and volatilities. The advantage of regime-switching volatility models, is that they allow for the persistence and dynamics of volatility to change according to the market conditions.[1] Therefore, we define a MRS-HAR-RV model, in which the coefficient stating the persistence of realised volatility can be regime dependent as follows

$$RV_{t+1d}^{(d)} = \beta_{0,st} + \beta_{st}^{(d)} RV_t^{(d)} + \beta_{st}^{(w)} RV_t^{(w)} + \beta_{st}^{(m)} RV_t^{(m)} + \beta_{st}^{(DTR)} DTR_t + \varepsilon_{t+1d,st}^{(d)}, \tag{10}$$

$$\varepsilon_{t+1d,st}^{(d)} \sim N(0, \Sigma_{st}),$$

[1] This is more important in the case of commodity and energy markets because it is argued that GARCH models induce a high degree of persistence in shocks that falsely implies high predictability but, in fact reflects regime shifts or structural breaks in the volatility process (Lamoureux and Lastrapes, 1990, and Nomikos and Pouliasis, 2011). This means that a regime-switching volatility model may be more suitable for modelling of energy commodities where there are structural breaks.

where $\mathrm{st} = \{1,2\}$ as we introduce a two-state MRS model, and all parameters (α_{st}, $\beta_{st}^{(d)}$, $\beta_{st}^{(w)}$, $\beta_{st}^{(m)}$ and Σ_{st}) are now state-dependent. The switching process between regimes depends on the conditional transition probability matrix

$$\boldsymbol{P} = \begin{pmatrix} \Pr(st_t = 1|st_{t-1} = 1) = p_{11} & \Pr(st_t = 1|st_{t-1} = 2) = p_{21} \\ \Pr(st_t = 2|st_{t-1} = 1) = p_{12} & \Pr(st_t = 2|st_{t-1} = 2) = p_{22} \end{pmatrix} = \begin{pmatrix} 1 - p_{12} & p_{21} \\ p_{12} & 1 - p_{21} \end{pmatrix}. \tag{11}$$

where p_{12} measures the probability of being in state one and switching to state 2 in the next period, while p_{21} is the probability of being in state 2 and switching to state 1 in the next period. Based on the conditional transition probability we can calculate the unconditional regime probability as

$$\Pr(st_t = 1) = p_{1,t} = \frac{p_{21}}{p_{12}+p_{21}}; \ \Pr(st_t = 2) = p_{2,t} = \frac{p_{12}}{p_{12}+p_{21}}. \tag{12}$$

Moreover, we can specifically rewrite the MRS-HAR-RV from the equation (10) to the following

$$RV_{t+1d}^{(d)} = p_{1,t}\big(\beta_{0,1} + \beta_1^{(d)} RV_t^{(d)} + \beta_1^{(w)} RV_t^{(w)} + \beta_1^{(m)} RV_t^{(m)} + \beta_1^{(DTR)} DTR_t + \varepsilon_{t+1d,1}^{(d)}\big) + \tag{13}$$

$$(1 - p_{1,t})(\beta_{0,2} + \beta_2^{(d)} RV_t^{(d)} + \beta_2^{(w)} RV_t^{(w)} + \beta_2^{(m)} RV_t^{(m)} + \beta_2^{(DTR)} DTR_t + \varepsilon_{t+1d,2}^{(d)}),$$

$$\varepsilon_{t+1d,1}^{(d)} \sim N(0, \Sigma_1), \ \varepsilon_{t+1d,2}^{(d)} \sim N(0, \Sigma_2).$$

Finally, assuming the state-dependent residuals follow a normal distribution, with mean zero and constant volatility, Σ_1 and Σ_2 for the two states respectively, the likelihood function for the entire sample is formed as a mixture of the probability distribution of the state variable, where:

$$f\left(RV_t^{(d)}, \theta\right) = \frac{p_{1,t}}{\sqrt{2\pi}\Sigma_1} \exp\left(-\frac{{\varepsilon_{t,1}^{(d)}}^2}{2\Sigma_1^2}\right) + \frac{p_{2,t}}{\sqrt{2\pi}\Sigma_2} \exp\left(-\frac{{\varepsilon_{t,2}^{(d)}}^2}{2\Sigma_2^2}\right) \tag{14}$$

with the log-likelihood function as

$$L(\theta) = \sum_{t=1}^{T} log f\left(RV_t^{(d)}, \theta\right), \tag{15}$$

where θ is the vector of parameters to be estimated. The log-likelihood function $L(\theta)$ is maximised using the BFGS estimation method subject to the constraint that

$p_{1,t} + p_{2,t} = 1$ and $0 \leq p_{1,t}, p_{2,t} \leq 1$.

4 Description of Data

There are six energy futures contracts listed in the Tokyo Commodity Exchange[2], namely gasoline, kerosene, crude oil, gasoil, Chukyo-gasoline and Chukyo-kerosene. However, considering that the lack of liquidity for gasoil, Chukyo-gasoline and Chukyo-kerosene futures, we only use gasoline, kerosene and crude oil futures, which have been actively traded for relatively long period. Our dataset comprises of high-frequency intraday futures price series from 22 September 2010 to 30 October 2015, obtained from Thomson Reuter Tick History. The beginning of sample is chosen to coincide with the extension of night session trading on 21 September 2010, which increased trading by five hours and was intended to attract more foreign trades. To evaluate the forecasting performance of different models, the sample is divided into model estimation and out-of-sample forecasting periods. The estimation period is from 22 September 2010 to 22 October 2014 (about four years), while the last one year of the data (250 days) is reserved for out-of-sample forecasting from 23 October 2014 to 30 October 2015.

We consider contracts across maturity spectrum, namely 1- to 6-month ahead contracts for gasoline and kerosene futures and current-month to 5-month ahead contracts for crude oil futures[3], to compare models and forecasting performances across all maturities. This is important because the pattern of trading activity with respect to contract maturity in TOCOM is surprisingly different from other energy and commodity futures exchanges (e.g. NYMEX and ICE)[4]. In fact, the most liquid contracts in TOCOM are the ones with the longest maturity (5 and 6-month) and trading

---

[2] TOCOM was established in 1984 when three exchanges, namely, Tokyo Textile Exchange, Tokyo Rubber Exchange and Tokyo Gold Exchange, merged and it is now the second Asian exchange trading energy futures after Singapore Exchange (SGX).

[3] The current-month crude oil futures contract is in fact the contract for settlement at the end of the current month. For example, on 10 March 2015, the current-month crude oil contract settles at end of March 2015, so the maturity spans from 1 to 22 trading days.

[4] For instance, Alquist and Kilian (2010) study NYMEX crude oil futures contracts and show that trading volume declines as maturity increases and long-maturity contracts are less liquid than short-maturity contracts.

activity declines as maturity declines.

The descriptive statistics of daily trading volume of futures contracts with different maturities presented in Panel A of Table 1.[5] Clearly there is a positive relation between maturity and trading activity in TOCOM energy futures. This could be due to the unavailability of contracts with maturities beyond 6 months, which means that participants with longer-term trading and hedging objectives are restricted to use contracts with the longest maturity (6-months) and roll over their positions over time. This in turn, increases the trading activity in contracts with longer maturities compared to shorter ones.

The descriptive statistics of returns on futures contracts are also reported in Table 1 (Panel B). The standard deviation of returns on Gasoline and Kerosene contracts seem to increase with maturity but decline slightly in crude oil futures. The increase in unconditional volatility of Gasoline and Kerosene futures with maturity is also in contrast to what is observed in other energy futures markets. Estimated coefficients skewness of returns are mostly negative while coefficients of kurtosis are greater than 3, indicating that returns are generally negatively skewed with extreme movements are ever present in the return series.

4.1 Sampling scheme and preliminary analysis

Before estimating the realised volatility models, we need to decide on sampling scheme and the optimal sampling frequency. For this purpose, we use calendar time sampling scheme as opposed to transaction (business) sampling because it is widely use in the literature dealing with high frequency data. Thus, we divide each trading day into M equally sized intervals by calendar time. Although a high frequency sampling is required for the estimate of RV to converge asymptotically to its true value, the choice of sampling is also dependent on the liquidity of the underlying asset and the trading activity within the interval chosen. In addition, Andersen and Bollerslev (1997 and 1998) and Taylor and Xu (1997) also show that market microstructure may affect the RV at higher frequencies. Hence, in this paper we choose the 15-minute interval as the sampling frequency for two reasons. First, the signature plots (Figures 1 to 3) show that

[5] Trading volume in TOCOM energy futures is reported as number of contracts traded. For all the three commodities considered in this paper, each futures contract is for delivery/settlement of 50kl (approximately 13,210 US gallons or 314.5 barrels) of the underlying commodity.

the realised volatility of sampling frequency is quite unstable when the sampling frequency is lower than 15-minutes interval but stabilises afterward. Second, considering the relatively low liquidity of TOCOM energy futures, it is appropriate to lower the sampling frequency to include more transactions in the time interval. Liu et al. (2015) suggest that 15-minute to one-hour interval is more suitable for less liquid assets. Although the realised volatility using quote may be a solution to the low liquidity issue, Hansen and Lunde (2006) find that the realised volatility using mid-quote is very likely to underestimate the realised volatility due to the microstructure disturbance. Therefore, our high-frequency data is still sampled with the 15-minute interval.

Figure 4 and 5 present the daily returns and annualised realised volatilities of gasoline, kerosene and crude oil futures for contracts longest maturity in our dataset; namely 6-month for gasoline and kerosene, and 5-month for crude oil. It can be seen that fluctuations in returns changes over time in a form of clustering effect which point to the existence of severe heteroscedasticity as shown in Figure 5, where several noticeable spikes are also observed in realised volatility. The most distinct spike occurs on 6 May 2011 (from the night session on 5 May 2011), due to the intraday flash crash in the oil market, when the oil price dropped sharply by 10%. The other spike takes place on 7 May 2012, when the anti-austerity party won the legislative election in Greece. To take into account and control the effect of these two extreme events, we introduce two binary dummy variables in all models.

Descriptive statistics of RV of futures contracts for different commodities and across maturities are presented in Table 2. The level of autocorrelation of realised volatility and the significance of Ljung and Box (1978)'s Q statistics for the first 22 lags of the autocorrelation function indicate that realised volatility of all energy futures are highly auto-correlated, confirming the long-memory property of realised volatility. The estimated coefficients of skewness and kurtosis also indicate that realised volatility series are highly skewed and leptokurtic, while the ADF test results confirm that the realised volatilities are stationary.

5 Empirical Results

The estimation results of the proposed realised volatility models as well as the out-of-sample forecasts and evaluation metrics, are reported and discussed in the following subsections. The in-sample analysis is performed over the period of 22 September 2010 to 22 October 2014 and the out-of-sample period is from 23 October 2014 to 30 October 2015.[6]

5.1 In-sample Analysis

We start the in-sample analysis by estimating the benchmark HAR-RV model using the Ordinary Least Square (OLS) method for three TOCOM energy futures across six different maturities. Estimation results are reported in Table 3, 4 and 5 for Gasoline, Kerosene and Crude oil futures, respectively. Estimated coefficients of lagged one-day, one-week and one-month volatility ($\beta_1^{(d)}, \beta_1^{(w)}$ and $\beta_1^{(m)}$) in HAR-RV models are all significant suggesting high degree of persistence in realised volatility of all futures contracts and maturities. For most futures contracts, the impact of one-day and one-month lagged realised volatility are greater than that of one-week when looking at the magnitude of the estimated coefficients. For example, for 5-month crude oil futures, the coefficients of one-day and one-month lagged realised volatility are 0.2608 and 0.1547, which are noticeably larger than that of the one-week one, 0.1114. In addition, estimated coefficients of days-to-rollover ($\beta_1^{(DTR)}$) are not significant for gasoline and kerosene futures contracts, with the exceptions of 1-month futures. The coefficients of days-to-rollover for contracts with 1-month maturity are negative and significant, which is in line with the Samuelson Hypothesis. In the case of crude oil futures, the estimated coefficients of days-to-rollover are positive and significant, indicating that as the rollover date (or maturity) approaches realised volatility declines. Although this observation is not in line with the Samuelson effect, it is also observed by Chen et al. (1999) in Nikkei-225 index futures.

Moving onto the result of MRS-HAR-RV reported in the right column of Table 3 to 5, it can be seen that significance of estimated coefficients in different confirm changes in the dynamics of the realised volatility of TOCOM energy futures under different regimes. In general, the coefficients of one-day lagged realised volatility are significant

---

[6] Due to space limitation, we only present and discuss the estimation results of HAR-RV and MRS-HAR-RV models for the three commodities. The estimation results of HAR-RV-CJ models (MedRV and BPV) as well as GARCH and MRS-GARCH models are not presented here and available from authors on request.

in both regime 1 and 2, while the coefficients of one-week and one-month realised volatility are not always significant in regime 1. This implies that realised volatility is less persistent in regime 1 than in regime 2. Moreover, we can look at the conditional transition probability and unconditional regime probability to further identify the two regimes and their existence. First, based on the estimated conditional transition probabilities (p_{12} and p_{21}) reported in Table 3 to 5, it can be seen that the probability of switching from regime 1 to regime 2 (p_{12}) is generally higher than the probability of switching from regime 2 to regime 1 (p_{21}). In other words, the probability of switching from a high volatility regime to a low volatility regime is higher than switching from low volatility regime to a high volatility regime. This indicates higher stability and persistence in the low volatility regime compared to the high volatility regime. In addition, Table 6 presents the number of days the probability of being a high volatility regime (when $p_{1,t} > 0.5$), is significantly lower than number of days of being in a low volatility regime (when $p_{2,t} > 0.5$), which is in line with the results of the condition transition probabilities (p_{12} and p_{21}).

To further confirm the identity of regimes, we compare the average level of realised volatility when unconditional probability of being in regime 1 is greater than 0.5 ($p_{1,t} > 0.5$) with that when unconditional probability of being in regime 2 ($p_{1,t} \leq 0.5$). According to results reported in Table 6, the average realised volatility in regime 1 ($p_{1,t} > 0.5$) is twice that of realised volatility of regime 2 ($p_{1,t} \leq 0.5$), which is consistent with the previous conjecture. Therefore, based on the estimated transition probabilities and realised volatilities, we can characterise the TOCOM energy futures market into two distinct regimes of high- and low-volatility levels.

Interestingly, the coefficients of days-to-rollover are also different under the two regimes, although their signs are mostly consistent with the results of a single regime. The magnitude of coefficients in high-volatility regimes is much higher than that in low-volatility regimes, which may be due to sharp price changes and steeper forward curve when the market is under the high-volatility regime. Finally, it can be observed that that in general MRS-HAR-RV models have a higher adjusted R-squared and SBIC than HAR-RV models, which suggests that MRS-HAR-RV can capture and explain the dynamic of realised volatility better than HAR-RV for all three energy futures.

To assess the relation between market micro-structure and regime-switching volatility, we explore the role of transactions initiated by buy and sell orders. Using the intraday data over the sample period, we sorted the transactions on each day into buy and sell initiated transactions[7]. We then defined the absolute difference between the number of transactions initiated by buy orders and number of transactions initiated by sell orders as "*order-trade imbalance*" (OTI). For instance, if in one day there are 50 transactions in total, where 30 of them have been initiated by buy orders and 20 have been initiated by a sell order, then OTI will be 10. The difference in sell or buy initiated transactions can be viewed as measure of buy or sell pressure in the market. Therefore, a relatively high OTI in a particular day leads to greater price movements and consequently higher volatility compared to a day when the OTI is relatively low[8]. Table 6 reports the average OTI in regime 1 and 2 for each commodity and maturity. It can be seen that higher average OTI is associated with high volatility regime and low average OTI is associated with low volatility regime.

5.2 Out-of-sample Forecast Evaluation

Although it is shown that MRS-HAR-RV captures changes in price volatility of the energy futures better than simple HAR-RV and HAR-RV-CJ in the in-sample, it is necessary to evaluate the performance of the models out-of-sample using different forecast evaluation metrics. Out-of-sample forecasts for all models are produced using recursive estimation approach where estimation period is extended after every forecast. Following the literature on comparison of volatility forecast, we use different loss functions that measure the difference between realised volatility and the forecast values of volatility. These include QLike, Mean Absolute Error (MAE), Mixed Mean Error for over-prediction (MMEO) and Mixed Mean Error for under-prediction (MMEU). The loss function used widely in most studies is the QLike statistic which is defined as

$$L(h_t, RV_t) = log(h_t) + \frac{RV_t}{h_t}, \tag{16}$$

where h_t is the forecasting value, and RV_t is the realised variance at time t. Patton (2011) proves that QLike loss function is robust even if realised volatility is an

[7] A sell-initiated transaction is identified as a transaction where the sell quote is closer to the transaction price, while a buy-initiated transaction is identified as a transaction where the buy quote is closer to the transaction price.

[8] This is in line with theoretical microstructure models and studies such as Kyle (1985), Admati and Pfleiderer (1988) and Huang and Stoll (1997) who suggest that net order flow causes the price movement.

imperfect proxy for true volatility. Similarly, MAE, MMO and MMEU are defined as

$$MAE = \frac{1}{M}\sum_{i=1}^{M} |RV_i \quad - h_i|, \tag{17}$$

$$MME(O) = \frac{1}{M}\left[\sum_{i=1}^{O} \sqrt{|RV_i - h_i|} + \sum_{i=1}^{U} |RV_i - h_i|\right] \tag{18}$$

$$MME(U) = \frac{1}{M}\left[\sum_{i=1}^{U} \sqrt{|RV_i - h_i|} + \sum_{i=1}^{O} |RV_i - h_i|\right] \tag{19}$$

where U is an indicator for under-prediction ($RV_t \quad - h_t < 0$) of realised variance, O is an indicator for overprediction ($RV_t \quad - h_t > 0$) of realised volatility variance, and M is the number of forecast points. According to the specification of the loss functions, it is obvious that MAE is a symmetric loss function, while MMEO more penalises over-prediction and MMEU more penalises under-prediction. Furthermore, the Diebold and Mariano (1995) test is used to statistically compare the forecast accuracy of different models. To implement the DM test, we calculate the difference in the loss function between two models as

$$d_t = L_{MRS-HAR-RV}(h_t, RV_t) - L_{alternative}(h_t, RV_t), \tag{20}$$

where $L_{MRS-HAR-RV}(h_t, RV_t)$ is the loss function for the MRS-HAR-RV model, and $L_{alternative}(h_t, RV_t)$ is loss function for the alternative model, including HAR-RV, HAR-RV-CJ, GARCH and MRS-GARCH. A negative d_t value indicates that MRS-HAR-RV has better forecasting ability than the alternative model, and vice versa. Therefore, the DM statistic is then defined as:

$$DM\ statistic = \frac{\bar{d}}{SE(\bar{d})} \tag{21}$$

where $\bar{d}$ is the average of the difference in Qlike loss function, and $SE(\bar{d})$ is standard error of d_t. Under the null $\mathrm{H}_0\colon \bar{d} \geq 0$ against the alternative $\mathrm{H}_1\colon \bar{d} < 0$, the DM statistic follows a standard normal distribution. Thus, a rejection of the null hypothesis means that MRS-HAR-RV has a superior predictive ability than the alternative.

The results of DM tests on the difference in QLike loss functions of MRS-HAR-RV against the alternative models including HAR-RV, HAR-RV-CJ (MedRV and BPV), GARCH and MRS-GARCH, are reported in Table 7**.** Comparison of QLike statistics and DM tests reveals that, in general the MRS-HAR-RV significantly outperform GARCH and MRS-GARCH models across all commodities and maturities, with the exception of MRS-GARCH model for Kerosene. However, comparison of QLike statistics of MRS-HAR-RV against HAR and HAR-RV-CJ indicates that MRS-HAR-RV forecasts do not significantly outperform HAR and HAR-RV-CJ models.

The results of comparison of MAE and DM tests are also reported in Panel B of Table 7. It appears that the results of MAE comparisons are similar to those of QLike between MRS-HAR-RV and simple HAR-RV but different between MRS-HAR-RV and jump models. MRS-HAR-RV outperforms HAR-RV for 1-, 4- and 5-month gasoline futures but not for current month crude oil futures. Regarding the comparison MAE of MRS-HAR-RV with those of HAR-RV-CJ, it seems that MRS-HAR-RV is still outperformed by both jump models for kerosene futures but produces better forecasts for gasoline and crude oil futures. Interestingly, although MedRV is considered as better measure for the continuous component than BPV, the forecasting performance of MedRV is worse than BPV in comparison to MRS-HAR-RV.

With respect to the results of MMEO and MMEU statistics, comparison of MAS-HAR-RV and other HAR-RV models seems to be quite different and not consistent with the result of QLike and MAE. For instance, looking at MMEO, we see that volatility forecasts by MRS-HAR-RV outperform HAR-RV for most gasoline futures, except 6-month ones. However, comparisons of MMEU statistics across different models reveal that MRS-HAR-RV does not outperform alternatives. In the case of kerosene futures, the results of comparisons of MMEU and MMEO of MRS-HAR-RV against alternatives reveal similar conclusions to the gasoline futures; that is, in general MRS-HAR-RV performs significantly better that other models when we look at MMEO statistics, but not so when MMEU statistics are compared. Finally, comparison of MMEO and MMEU across different volatility models for crude oil futures also suggests that MRS-HAR-RV significantly outperform HAR-RV, HAR-RV-CJ and GARCH type models when MMEO is considered. However, MRS-HAR-RV volatility forecasts do not outperform alternative models according to MMEU loss function. This leads to the conclusion that overall HAR-RV and HAR-RV-CJ tend to over-predict realised variance, whereas MRS-HAR-RV tends to under-predict the realised variance of

TOCOM energy futures.

Furthermore, it is also interesting to compare whether Markov regime-switching approach can also outperform the GARCH model. Based on the results of QLike and MAE reported in Table 7, it is clear that MRS-HAR-RV always outperforms simple GARCH model for all contracts while is outperformed by MRS-GARCH for three kerosene and one crude oil futures. Moreover, although GARCH type models are criticised for its overprediction, the result of MMEO shows that the overprediction level of MRS-GARCH is much lower than that of GARCH. Therefore, according to the forecast evaluation results presented, it seems that incorporating MRS technique also improves the forecasting performance of GARCH and reduce the over-predicted bias.

5.3 Value at Risk (VaR) Estimation

Forecasts of realised volatility can provide insights for market participants seeking to understand future market conditions, as well as being used to quantify market risk for risk management and trading decisions. One of the most popular approaches to quantifying market risk is VaR, which is the potential loss for a position (or a portfolio) given certain confidence level $(1 - u)$ over a fixed time horizon (k). Therefore, VaR can be defined as

$$\Pr\left(r_{t+1} < VaR_{u,t+1} \middle| \Omega_t\right) = u, \quad (22)$$

where Ω_t is the information given at time t. The VaR is estimated as the product of the u-percentile of assumed distribution of returns and the forecast of volatility (standard deviation), $VaR_{u,t+1} = F^{-1}(u)h_{t+1}$, where $F^{-1}(u)$ is the corresponding u-percentile (e.g. 0.5%, 1%, 5%) of assumed distribution, and h_{t+1} is the predicted volatility. To assess the practical implication of the results in terms of risk assessment and measurement, we compare the VaR estimates of the proposed models through an out-of-sample backtesting procedure proposed by Christoffersen (1998). These tests include unconditional coverage (UC) and conditional coverage (CC) log-likelihood ratio (LR) tests, which are presented in Appendix B. We also compare the model estimated VaRs with those from a non-parametric technique known as Historical Simulation (HS), where the VaR is given as the u percentile of historical returns.

We compare the 1% VaR (99% confidence level) of all models, namely HAR-RV, MRS-HAR-RV, HAR-RV-CJ as well as GARCH, MRS-GARCH. HS is estimated based on a sample 250 days (1 year). Using 250 observations for HS is believed to be appropriate because using longer samples tends smooth the percentile and underestimate the VaR. GARCH approaches are based on the forecasts from GARCH (1,1) and MRS-GARCH (1,1), and the standard normal distribution is used for the estimation of $F^{-1}(u)$. Instead of using standard normal percentile, we use Filtered Historical Simulation (FHS) to estimate the corresponding percentile for HAR-RV, MRS-HAR-RV and HAR-RV-CJ because realised volatility does not satisfy the normality assumption (see Table 2). The first step of FHS is to create a series of standardising returns by dividing the series of returns with historical standard deviation. Then, the u-percentile of this standardised returns series is employed as $F^{-1}(u)$ to calculate VaR for the RV approaches.

The VaR estimates and backtesting results for both long and short position[9] are reported in Figure 6 to 8 and Table 8 and 9. According to the figures, it seems that the VaR from the HS approach is smoother and underestimated compared with the other six approaches. Regarding the GARCH and MRS-GARCH approaches, the VaR estimates sometimes appear to be overestimated when comparing their distance from actual return, especially for gasoline and kerosene futures. Similarly, two HAR-RV-CJ models overestimate the VaR, since the VaR estimates are largely distant from actual returns. Moreover, it appears that the bias of BPV measure is higher than that of MedRV measure. Finally, the VaR estimates from HAR-RV and MRS-HAR-RV seems to be lower than HAR-RV-CJ models but around GARCH type models.

In order to more carefully examine the accuracy of VaR estimates, we need to further check the backtesting results in Table 8 and 9. For gasoline futures, all GARCH, MRS-GARCH and HS approaches fail to pass at least one of the backtesting tests, while the testing results of HAR-RV type models are similar. Specifically, all HAR-RV type models produce efficient VaR for all maturity contracts except for that HAR-RV-MedRV fails to pass the conditional coverage test for the 1-month futures. Interestingly, for 5-month gasoline futures, it seems that HAR-RV type models may overestimate the value of VaR since the PFs are all 0.00%. Moving to kerosene futures, HS and GARCH

[9] When investors take a long position, their payoff is positive when the price increases, so they only focus on the downside risk (negative VaR). Differently, the short position takers bear loss when the price decreases, so their concern is the upside risk (positive VaR).

approaches again fail to pass backtesting tests, while, surprisingly, HAR-RV and MRS-HAR-RV approaches do not pass the tests for five and four futures, respectively. Similar to MRS-HAR-RV, both HAR-RV-CJ models fail to pass the tests for four contracts despite for different maturities. For example, both MedRV and BPV approaches pass the backtesting tests for 6-month kerosene while MRS-HAR-RV does not. Finally, for crude oil futures, all HS, GARCH and HAR-RV-CJ approaches fail to pass backtesting tests for all contracts, while HAR-RV and MRS-HAR-RV performs almost evenly. The VaR estimated by HAR-RV and MRS-HAR-RV passes the backtesting tests for four and three futures respectively. The results of short-position VaR are consistent with those of long-position VaR but the performance is generally better. HS produce the least efficient results as for the long-position for all three TOCOM energy futures. GARCH type models pass the backtesting tests for gasoline futures but cannot for kerosene and crude oil. HAR-RV-CJ models performs evenly with HAR-RV and MRS-HAR-RV for gasoline and kerosene futures with all HAR-RV models producing efficient VaR, while are outperformed by both HAR-RV and MRS-HAR-RV for crude oil futures.

To sum up, comparing the accuracy of VaR estimates reveals that HAR-RV type models can produce more accurate estimates than HS and GARCH type approaches across commodities and maturities. Nonetheless, even though HAR-RV and MRS-HAR-RV approaches can both produce efficient VaR estimates for crude oil futures, their performance for gasoline and kerosene is consistent with both HAR-RV CJ models.

6 Conclusion

This paper investigates the dynamics of realised volatility for TOCOM gasoline, kerosene and crude oil futures. The results of HAR-RV indicate a high level of persistence in realised volatility with a slow decay. Nonetheless, when we extend the HAR-RV to a 2-state Markov regime-switching HAR-RV, the empirical evidence shows that both the average volatility level and the persistence of realised volatility are different under each regime. More precisely, under one regime, RV is dependent on all one-day, one-week and one-month lagged RVs, whereas under the second regime, the impact of one-week and one-month lagged RVs disappears. Moreover, the MRS-HAR-RV model captures the dynamics of the realised volatility of TOCOM energy futures better than HAR-RV.

In out-of-sample tests, MRS-HAR-RV outperforms HAR-RV-CJ (MedRV and BPV) for most TOCOM energy futures, while it provides better forecast than HAR-RV for longer maturity futures, but not for shorter maturity futures (using the MAE loss function). We conjecture that the lack of liquidity for shorter maturity contracts increases the impact of OTI on volatility and then also the probability of regime-switching. Hence, any unexpected increase in OTI may lower the precision of forecasts of the unconditional regime probability, leading to less accurate forecasts of realised volatility. We further compare the difference in MAE, MMEO and MMEU, and find MRS-HAR-RV tends to under-predict realised variance, while HAR-RV, HAR-RV-CJ, GARCH and MRS-GARCH tend to over-predict. Moreover, incorporating MRS approach helps reduce the overprediction of GARCH-style models.

In the application of VaR estimation and valuation, both HAR-RV and MRS-HAR-RV outperform HS and GARCH approaches for all three commodities and across six maturities. HAR-RV and MRS-HAR-RV perform evenly with HAR-RV-CJ models for gasoline and kerosene futures while outperforms HAR-RV-CJ for crude oil futures.

Appendix A: Jump Detection

In this paper, we employ the adjusted jump ratio statistic proposed by Huang and Tauchen (2005) for jump detection with BPV, which can be calculated as follows:

$$Z_{BPV,t} = \sqrt{M}\frac{(RV_t - BPV_t)/RV_t}{\sqrt{(\frac{\pi^2}{4} + \pi - 5) \times \max(\frac{1}{M}, \frac{\widehat{QQ}_t}{BPV_t^2})}} \sim N(0,1), \tag{A.23}$$

$$\widehat{QQ}_t = M\frac{\pi^2}{4}\sum_{j=0}^{M-4}|r_{t-j\times\Delta}||r_{t-(j+1)\times\Delta}|\,|r_{t-(j+2)\times\Delta}||r_{t-(j+3)\times\Delta}|.$$

where $\widehat{QQ}_t$ is the quad-power quarticity that is used in Barndorff-Nielsen and Shephard (2004) for the estimator of the integrated quarticity. The $Z_{BPV,t}$ statistic converges to a standard normal distribution when the frequency approaches to the 0. A Similar test statistic is developed by Andersen et al. (2012) for the jump detection with MedRV. The statistic is calculated by:

$$Z_{MedRV,t} = \sqrt{M}\frac{(RV_t - MedRV_t)/RV_t}{\sqrt{0.96 \times \max(1, \frac{\widehat{MedRQ}_t}{MedRV_t^2})}} \sim N(0,1), \tag{A.24}$$

$$MedRQ_t = \frac{3\pi}{9\pi + 72 - 52\sqrt{3}}\left(\frac{M}{M-2}\right)\sum_{j=1}^{M-2} med\left(|r_{t-(j-1)\times\Delta}|, |r_{t-j\times\Delta}|, |r_{t-(j+1)\times\Delta}|\right)^4,$$

where $MedRQ_t$ estimates the integrated quarticity by using similar technique for $MedRV_t$. $Z_{MedRV,t}$ also follows a standard normal distribution when the frequency is high enough.

After testing for detecting jumps, both continuous and jump components can now be estimated as follows:

$$C_t = BPV_t \times I\left(Z_{BPV,t} > Z_\alpha\right) + RV_t \times \left[1 - I\left(Z_{BPV,t} > Z_\alpha\right)\right] \tag{A.25}$$

$$J_t = (RV_t - BPV_t) \times I\left(Z_{BPV,t} > Z_\alpha\right)$$

C_t and J_t are continuous and jump components, and Z_α is the critical value under the confidence level α. BPV_t and $Z_{BPV,t}$ can be replaced with $MedRV_t$ and $Z_{MedRV,t}$ when one uses MedRV as the measure of jump-robust realised volatility. $I(Z_{BPV,t} > Z_\alpha)$ is the indicator for jump detection, which is 1 when jump is detected and 0 otherwise. In equation (A.5), if jump is detected, the difference between RV and the jump-robust RV is treated as the jump components, and the jump robust RV is viewed as the continuous component. However, if the test shows no occurrence of jumps, the jump component is calculated as zero and RV is treated as the continuous component.

Appendix B: Backtesting VaR

Christoffersen (1998) outlines three tests for evaluation of VaR estimates in backtesting framework. These are unconditional coverage (UC), conditional coverage (CC) and the log-likelihood ratio (LR) tests. The first step to perform the backtesting is to calculate the percentage proportion of failure (PF), which is the proportion of actual returns that exceed the estimated VaR or the so-called "hit ratio", which is defined as follows

$$\mathrm{I}_{t+1} = \begin{cases} 1, if\ r_{t+1} < VaR_{u,t+1}|\Omega_t \\ 0, if\ r_{t+1} \geq VaR_{u,t+1}|\Omega_t \end{cases}, \tag{26}$$

where I_{t+1} is the indicator function for violation of VaR. The sum of the indicator function over the out-of-sample is the total hit numbers, and the proportion of hit numbers to total number of out-of-sample observations represents PF. The VaR estimate is considered to be efficient if the following condition is satisfied

$$E[\mathrm{I}_{t+1}|\Omega_t] = u \tag{27}$$

which implies that on average (1-PF) should be equal to the nominal confidence level u. The unconditional coverage (UC) test developed by Kupiec (1995), and the independence (IND), and conditional coverage tests (CC) proposed by Christoffersen (1998) are designed to examine whether PF is indifferent to the tolerance level p. We first denote π as PF calculated as

$$\pi = \frac{N_1}{N_0+N_1}, \tag{28}$$

where N_0 is the total number of indicator being *0* (no violation of VaR), and N_1 is the total number of indicator being *1* (violation of VaR). Then, the null hypothesis for the unconditional coverage test can be expressed as

$$H_0\colon \pi = u, \tag{29}$$

and the LR statistic for the UC test ($LR_{(UC)}$) can be defined as

$$LR_{(UC)} = 2\{\ln[(1-\pi)^{N_0}\pi^{N_1}] - \ln[(1-u)^{N_0}u^{N_1}]\} \sim \chi_{1-\alpha,1}, \tag{30}$$

where $LR_{(UC)}$ follows a chi-square distribution with degree of freedom 1 under given significant level α. If $LR_{(UC)} > \chi_{\alpha,1}$, H_0 is rejected, this implies that the VaR estimate is not efficient. However, an unconditional coverage test only examines whether the total PF exceeds nominal tolerance level on average but does not consider the cluster of violation (consecutive violation). An independent test is designed to examine the dependence of consecutive violation, and the LR statistic for independent test $LR_{(IND)}$ can be expressed as

$$LR_{(IND)} = 2\{\ln[(1-\pi_{01})^{N_{00}}\pi_{01}{}^{N_{01}}(1-\pi_{11})^{N_{10}}\pi_{11}{}^{N_{11}}] - \ln[(1-u)^{N_0}u^{N_1}]\} \sim \chi_{1-\alpha,1}, \tag{31}$$

$$\pi_{01} = \frac{N_{01}}{N_{00}+N_{01}} = \frac{N_{01}}{N_0}, \pi_{11} = \frac{N_{11}}{N_{10}+N_{11}} = \frac{N_{11}}{N_1}$$

where N_{ij} for *i,j = 0,1* is the number of indicator being *i* followed by indictor being *j*, andπ_{ij} is the corresponding probability. If *i = j = 1*, it indicates the occurrence of consecutive violation. $LR_{(IND)}$ also follows a chi-square distribution with degree of freedom 1 under given significant level α. Finally, the LR statistic for conditional coverage test $LR_{(CC)}$ is defined as the sum of $LR_{(UC)}$ and $LR_{(IND)}$, shown as

$$LR_{(CC)} = LR_{(UC)} + LR_{(IND)} \sim \chi_{1-\alpha,2}. \tag{32}$$

Similarly, $LR_{(CC)}$ follows a chi-square distribution but with degree of freedom 2 under given significant level α.

Reference

Admati, A. R., & Pfleiderer, P. (1988). A theory of intraday patterns: Volume and price variability. Review of Financial studies, 1(1), 3-40.

Agnolucci, P., 2009. Volatility in crude oil futures: a comparison of the predictive ability of GARCH and implied volatility models. Energy Economics 31 (2), 316–321.

Alizadeh, A. H., Nomikos, N. K., & Pouliasis, P. K. (2008). A Markov regime switching approach for hedging energy commodities. Journal of Banking & Finance, 32(9), 1970-1983.

Alquist, R., & Kilian, L. (2010). What do we learn from the price of crude oil futures? Journal of Applied Econometrics, 25(4), 539-573.

Andersen, T. G., & Bollerslev, T. (1997). Intraday periodicity and volatility persistence in financial markets. Journal of empirical finance, 4(2-3), 115-158.

Andersen, T. G., & Bollerslev, T. (1998). Deutsche mark–dollar volatility: intraday activity patterns, macroeconomic announcements, and longer run dependencies. the Journal of Finance, 53(1), 219-265.

Andersen, T. G., Bollerslev, T., and Diebold, F. X. (2007). Roughing it up: Including jump components in the measurement, modelling, and forecasting of return volatility. The Review of Economics and Statistics, 89(4), 701–720.

Andersen, T. G., Bollerslev, T., Christoffersen, P. F., and Diebold, F. X. (2006). Volatility and correlation forecasting. Handbook of economic forecasting, 1, 777–878.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & 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., 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. (2003). Modelling and forecasting realized volatility. Econometrica, 71(2), 579–625.

Andersen, T. G., Dobrev, D., & Schaumburg, E. (2012). Jump-robust volatility estimation using nearest neighbor truncation. *Journal of Econometrics*, *169*(1), 75-93.

Bandi, F. M., & Russell, J. R. (2008). Microstructure noise, realized variance, and optimal sampling. The Review of Economic Studies, 75(2), 339-369.

Barndorff-Nielsen, O. E. (2002a). 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.

Barndorff-Nielsen, O. E. and Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps. Journal of financial econometrics, 2(1), 1–37.

Barndorff-Nielsen, O. E. and Shephard, N. (2007). Variation, jumps, market frictions and high frequency data in financial econometrics. In Advances in Economics and Econometrics. Theory and Applications, Ninth World Congress, volume 3, (pp. 328–372).

Barndorff-Nielsen, O. E., & Shephard, N. (2002b). Estimating quadratic variation using realized variance. Journal of Applied econometrics, 17(5), 457-477.

Chen, Y.J., Duan, J.C., Hung, M.W., 1999. Volatility and maturity effects in the Nikkei index futures. Journal of Futures Markets 19, 895–909.

Cabedo, J. D., & Moya, I. (2003). Estimating oil price 'Value at Risk' using the historical simulation approach. Energy Economics, 25(3), 239-253.

Chan, K., & Fong, W. M. (2000). Trade size, order imbalance, and the volatility–volume relation. Journal of Financial Economics, 57(2), 247-273.

Chen, X. and Ghysels, E. (2010). News—good or bad—and its impact on volatility predictions over multiple horizons. Review of Financial Studies, hhq071.

Chordia, T., Roll, R., & Subrahmanyam, A. (2002). Order imbalance, liquidity, and market returns. Journal of Financial economics, 65(1), 111-130.

Christoffersen, P. F. (1998). Evaluating interval forecasts. International economic review, 841-862.

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

Fan, Y., Zhang, Y.J., Tsaic, H.T., Wei, Y.M., 2008. Estimating 'value at risk' of crude oil price and its spillover effect using the GED-GARCH approach. Energy Economics 30 (6), 3156–3171.

Figlewski, S. (1997). Forecasting volatility. Financial markets, institutions and instruments, 6(1), 1–88.

Glosten, L. R., & Harris, L. E. (1988). Estimating the components of the bid/ask spread. Journal of financial Economics, 21(1), 123-142.

Hansen, P. R., & Lunde, A. (2006). Realized variance and market microstructure noise. Journal of Business & Economic Statistics, 24(2), 127-161.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica: Journal of the Econometric Society, 357-384.

Haugom, E., Langeland, H., Molnár, P., & Westgaard, S. (2014). Forecasting volatility of the US oil market. Journal of Banking & Finance, 47, 1-14.

Herrera, A. M., Hu, L. and Daniel Pastor, D. (2018). Forecasting crude oil price volatility, International Journal of Forecasting, 34, 622-635.

Huang, R. D., & Stoll, H. R. (1997). The components of the bid-ask spread: A general approach. Review of Financial Studies, 10(4), 995-1034.

Klein, T. and Walther, T. (2016). Oil price volatility forecast with mixture memory GARCH. Energy Economics, 58, 46-58.

Kupiec, P. (1995). Techniques for verifying the accuracy of risk measurement models.

Kyle, A. S. (1985). Continuous auctions and insider trading. Econometrica: Journal of the Econometric Society, 1315-1335.

Lee, H. T., & Yoder, J. (2007). Optimal hedging with a regime-switching time-varying correlation GARCH model. Journal of Futures Markets, 27(5), 495-516.

Liu, L. Y., Patton, A. J., & Sheppard, K. (2015). Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes. Journal of Econometrics, 187(1), 293-311.

Liu, L., & Wan, J. (2012). A study of Shanghai fuel oil futures price volatility based on high frequency data: Long-range dependence, modelling and forecasting. Economic Modelling, 29(6), 2245-2253.

Ljung, G. M., & Box, G. E. (1978). On a measure of lack of fit in time series models. Biometrika, 65(2), 297-303.

Lux, T, Segnon, M., and Gupta, R., (2016), Forecasting crude oil price volatility and value-at-risk: Evidence from historical and recent data, Energy Economics, 56, 117-133.

Lyócsa, Š., & Molnár, P. (2018). Exploiting dependence: Day-ahead volatility forecasting for crude oil and natural gas exchange-traded funds. Energy, 155, 462-473.

Ma, F., Liu J., Huang, D., and Chen, H., (2017) "Forecasting the oil futures price volatility: A new approach" Economic Modelling, 64, 560–566.

Madhavan, A., Richardson, M., & Roomans, M. (1997). Why do security prices Change? A transaction-level analysis of NYSE stocks. Review of Financial Studies, 10(4), 1035-1064.

Martens, M., & Zein, J. (2004). Predicting financial volatility: High-frequency time-series forecasts vis-à-vis implied volatility. Journal of Futures Markets, 24(11), 1005-1028.

McAleer, M. and Medeiros, M. C. (2008). Realized volatility: A review. Econometric Reviews, 27(1-3), 10–45.

Nomikos, N. K. and Pouliasis, P. K. (2011). Forecasting petroleum futures markets volatility: The role of regimes and market conditions. Energy Economics, 33, 321-337.

Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. Journal of Econometrics, 160(1), 246-256.

Patton, A. J., & Sheppard, K. (2015). Good volatility, bad volatility: Signed jumps and the persistence of volatility. Review of Economics and Statistics, 97(3), 683-697.

Poon, S.-H. and Granger, C. W. (2003). Forecasting volatility in financial markets: A review. Journal of economic literature, 41(2), 478–539.

Sadeghi, M. and Shavvalpour, S. (2006). Energy risk management and value at risk modelling. Energy policy, 34(18), 3367–3373.

Sadorsky, P., 2006. Modelling and forecasting petroleum futures volatility. Energy Economics 28 (4), 467–488.

Samuelson, P. A. (1965). Proof that properly anticipated prices fluctuate randomly.
Sévi, B. (2014). Forecasting the volatility of crude oil futures using intraday data. European Journal of Operational Research, 235(3), 643–659.
Tseng, T.-C., Chung, H., & Huang, C.-S. (2009). Modelling jump and continuous components in the volatility of oil futures. Studies in Nonlinear Dynamics & Econometrics, 13(3), 5 (Article 5).
Taylor, S. J., & Xu, X. (1997). The incremental volatility information in one million foreign exchange quotations. *Journal of Empirical Finance*, *4*(4), 317-340.
Wang, T., Wu, J., and Yang, J. (2008). Realized volatility and correlation in energy futures markets. Journal of Futures Markets, 28(10), 993–1011.
Wei, Y. (2012). Forecasting volatility of fuel oil futures in china: GARCH-type, SV or Realized Volatility models? Physica A: Statistical Mechanics and its Applications, 391(22), 5546–5556.
Wei, Y., Wang, Y., and Huang, D. (2010). Forecasting crude oil market volatility: Further evidence using GARCH-class models. Energy Economics, 32(6), 1477–1484.

Table 1: Descriptive statistics of daily trading volume and returns of three TOCOM energy futures across maturities

	Mean	Std.	Skewness	Kurtosis	Q(22)	ADF
			Gasoline			
	Panel A: Trading volume					
1-month	265.86	260.39	4.58	36.96	299.65 [a]	-4.18 [a]
2-month	205.05	173.00	6.79	100.79	1048.85[a]	-4.50 [a]
3-month	258.76	192.30	3.49	26.51	1337.94 [a]	-4.70 [a]
4-month	405.63	299.63	4.39	44.50	1210.22 [a]	-4.86 [a]
5-month	1524.75	1328.78	2.67	12.70	1167.05 [a]	-4.48 [a]
6-month	6211.32	2697.35	1.17	5.59	2236.68 [a]	-4.33 [a]
	Panel B: Returns (Annualised)					
1-month	-1.13%	22.66%	-0.55	6.90	42.41 [a]	-6.49 [a]
2-month	-0.78%	22.84%	-0.45	6.28	45.56 [a]	-6.51 [a]
3-month	-0.58%	22.91%	-0.51	6.62	40.54 [a]	-6.86 [a]
4-month	-0.20%	22.88%	-0.49	6.52	33.31 [a]	-6.90 [a]
5-month	0.15%	22.98%	-0.45	6.50	27.10 [a]	-7.00 [a]
6-month	0.18%	23.07%	-0.41	6.43	26.50 [a]	-7.10 [a]
			Kerosene			
	Mean	Std.	Skewness	Kurtosis	Q(22)	ADF
	Panel A: Trading volume					
1-month	292.28	336.35	3.27	16.90	1252.78 [a]	-2.72
2-month	192.48	165.81	3.06	22.28	3059.87 [a]	-3.10
3-month	207.96	173.55	3.68	31.73	1881.69 [a]	-4.97 [a]
4-month	279.47	222.19	4.62	51.33	1384.75 [a]	-5.37 [a]
5-month	685.61	576.50	2.55	13.42	2106.88 [a]	-3.64 [b]
6-month	1510.01	857.98	2.64	22.07	4771.25 [a]	-3.70 [b]
	Panel B: Returns (Annualised)					
1-month	-1.81%	22.56%	-0.47	8.14	39.77 [a]	-6.80 [a]
2-month	-1.85%	22.71%	-0.56	8.03	42.81 [a]	-6.96 [a]
3-month	-2.03%	23.02%	-0.54	8.01	43.63 [a]	-7.11 [a]
4-month	-2.55%	23.22%	-0.51	8.00	45.01 [a]	-7.08 [a]
5-month	-2.92%	23.66%	-0.50	7.88	46.86 [a]	-7.06 [a]
6-month	-2.64%	24.06%	-0.47	8.15	46.58 [a]	-7.04 [a]
			Crude oil			
	Mean	Std.	Skewness	Kurtosis	Q(22)	ADF
	Panel A: Trading volume					
current-m.	61.09	83.38	5.59	54.02	161.95 [a]	-5.05 [a]
1-month	113.27	105.97	3.39	25.08	354.20 [a]	-5.24 [a]
2-month	164.19	149.64	2.83	16.31	590.87 [a]	-4.51 [a]
3-month	290.95	288.33	2.93	13.97	2040.02 [a]	-3.25
4-month	1340.78	1991.70	3.70	19.15	4818.85 [a]	-1.43
5-month	4059.61	2945.36	2.41	10.31	9499.13 [a]	-1.74
	Panel B: Returns (Annualised)					
current -m.	-3.02%	24.69%	-1.37	19.81	61.58 [a]	-6.50 [a]
1-month	-3.38%	29.99%	-0.42	8.81	56.13 [a]	-7.16 [a]
2-month	-3.07%	29.65%	-0.41	8.63	65.90 [a]	-7.13 [a]
3-month	-2.95%	29.29%	-0.45	8.63	58.68 [a]	-7.15 [a]
4-month	-2.77%	28.87%	-0.46	8.53	62.00 [a]	-7.10 [a]
5-month	-2.60%	28.55%	-0.43	8.33	60.75 [a]	-7.14 [a]

- The sample period is from 22 September 2010 to 30 October 2015.
- Trading volume is defined as the number of futures contracts traded per day. For all the three commodities considered in this paper, each futures contract is for delivery/settlement of 50kl (approximately 13,210 US gallons or 314.5 barrels).
- Q(22) is Q-statistic with 22 lags, and ADF is the augmented Dickey-Fuller test statistic.
- [a] indicates rejection of null hypothesis at the 1% significance level.

Table 2: Descriptive statistics for annualised realised variance of three TOCOM energy futures across maturities

	Mean	Std.	Skewness	Kurtosis	Q(22)	ADF
			Gasoline			
	Mean	Std.	Skewness	Kurtosis	Q(22)	ADF
1-month	0.059	0.089	4.92	37.89	413.72 [a]	-4.97 [a]
2-month	0.051	0.086	10.52	198.27	484.21 [a]	-5.22 [a]
3-month	0.050	0.084	11.24	217.39	521.17 [a]	-5.45 [a]
4-month	0.050	0.080	12.53	268.36	528.05 [a]	-5.12 [a]
5-month	0.051	0.079	13.64	307.25	699.83 [a]	-5.04 [a]
6-month	0.048	0.099	21.94	639.46	231.96 [a]	-5.41 [a]
			Kerosene			
	Mean	Std.	Skewness	Kurtosis	Q(22)	ADF
1-month	0.059	0.126	14.34	284.12	256.11 [a]	-5.45 [a]
2-month	0.049	0.095	12.46	254.88	321.32 [a]	-5.25 [a]
3-month	0.049	0.095	14.95	346.01	389.82 [a]	-5.34 [a]
4-month	0.049	0.089	15.94	393.06	450.42 [a]	-5.06 [a]
5-month	0.051	0.093	18.15	483.10	416.67 [a]	-4.88 [a]
6-month	0.053	0.092	16.02	394.41	596.05 [a]	-4.87 [a]
			Crude oil			
	Mean	Std.	Skewness	Kurtosis	Q(22)	ADF
Current-m.	0.064	0.199	11.27	186.25	137.53 [a]	-5.17 [a]
1-month	0.076	0.154	13.56	298.26	417.82 [a]	-4.78 [a]
2-month	0.077	0.163	12.75	250.04	487.82 [a]	-5.01 [a]
3-month	0.075	0.156	17.47	450.75	437.11 [a]	-4.89 [a]
4-month	0.077	0.142	16.61	416.14	623.27 [a]	-4.76 [a]
5-month	0.075	0.140	16.97	429.12	540.49 [a]	-4.91 [a]

- [a] and [b] indicate rejection at the 5% and 1% significance levels. Std. is the standard deviation. Q(22) is Q-statistic with 22 lags. ADF is the augmented Dickey-Fuller test statistic. The sample period is from 22 September 2010 to 30 October 2015.

Table 3: Estimation results of simple HAR-RV and MRS-HAR-RV for gasoline futures

HAR-RV $$RV_{t+1d}^{(d)} = \beta_{0,1} + \beta_1^{(d)} RV_t^{(d)} + \beta_1^{(w)} RV_t^{(w)} + \beta_1^{(m)} RV_t^{(m)} + \beta_1^{(DTR)} DTR_t + \beta_1^{(D1)} D1_t + \beta_1^{(D2)} D2_t + \varepsilon_{t+1d,1}^{(d)}, \quad \varepsilon_{t+1d,1}^{(d)} \sim N(0, \Sigma_1)$$

MRS-HAR-RV $$RV_{t+1d}^{(d)} = \beta_{0,st} + \beta_{st}^{(d)} RV_t^{(d)} + \beta_{st}^{(w)} RV_t^{(w)} + \beta_{st}^{(m)} RV_t^{(m)} + \beta_{st}^{(DTR)} DTR_t + \beta_{st}^{(D1)} D1_t + \beta_{st}^{(D2)} D2_t + \varepsilon_{t+1d,st}^{(d)}, \quad st = \{1,2\}, \varepsilon_{t+1d,st}^{(d)} \sim N(0, \Sigma_{st})$$

	1-month		2-month		3-month		4-month		5-month		6-month	
	HAR-RV	MRS-HAR	HAR-RV	MRS-HAR	HAR-RV	MRS-HAR	HAR-RV	MRS-HAR	HAR-RV	MRS-HAR	HAR-RV	MRS-HAR
$\beta_{0,1}$	0.0001***	0.0004***	4.28E-5**	0.0003***	5.21E-5***	0.0003***	5.04E-5***	0.0003***	5.74E-5***	0.0003***	7.84E-5***	0.0003***
	(2.56E-5)	(4.35E-5)	(1.67E-5)	(3.13E-5)	(1.51E-5)	(7.20E-5)	(1.33E-5)	(1.78E-5)	(1.09E-5)	(1.68E-5)	(1.03E-5)	(1.33E-5)
$\beta_1^{(d)}$	0.1687***	0.0753	0.2783***	0.1485***	0.3125***	0.3246**	0.2726***	0.2075***	0.2938***	0.2115***	0.1286***	0.0864***
	(0.0295)	(0.0893)	(0.0199)	(0.0272)	(0.0188)	(0.1590)	(0.0169)	(0.0223)	(0.0142)	(0.0196)	(0.0104)	(0.0135)
$\beta_1^{(w)}$	0.2157***	0.3692***	0.1196***	0.0439	0.1184***	-0.0374	0.1192***	-0.0460	0.1338***	0.0262	0.1515***	0.0623*
	(0.0509)	(0.1123)	(0.0315)	(0.0646)	(0.0292)	(0.0679)	(0.0275)	(0.0824)	(0.0227)	(0.0363)	(0.0183)	(0.0347)
$\beta_1^{(m)}$	0.2465***	-0.0142	0.1765***	0.0602	0.1428***	-0.0601	0.1851***	-0.0549	0.1743***	-0.0755	0.1643***	-0.1307***
	(0.0764)	(0.1371)	(0.0502)	(0.1197)	(0.0459)	(0.1271)	(0.0433)	(0.0407)	(0.0346)	(0.0734)	(0.0303)	(0.0355)
$\beta_1^{(DTR)}$	-2.2612**	6.1384***	1.5037**	3.0754**	0.7392	0.8735	0.8328	1.5072	0.1136	-0.5228	-0.0914	-2.1827
	(1.1052)	(1.9918)	(0.7488)	(1.4693)	(0.6845)	(3.9618)	(0.5877)	(1.2814)	(0.4882)	(0.9360)	(0.4815)	(1.4709)
$\beta_{0,2}$		6.48E-5***		4.03E-5***		3.51E-5***		3.42E-5***		3.00E-5***		2.24E-5***
		(2.65E-6)		(2.30E-6)		(1.46E-6)		(2.20E-6)		(1.67E-6)		(1.43E-7)
$\beta_2^{(d)}$		0.0494***		0.2480***		0.2269***		0.1725***		0.2732***		0.3108***
		(0.0063)		(0.0065)		(0.0038)		(0.0115)		(0.0100)		(0.0084)
$\beta_2^{(w)}$		0.1825***		0.0801***		0.1451***		0.1434***		0.1896***		0.1145***
		(0.0070)		(0.0080)		(0.0276)		(0.0112)		(0.0084)		(0.0234)
$\beta_2^{(m)}$		0.1554***		0.1074***		0.1156***		0.1476***		0.1137***		0.2163***
		(0.0095)		(0.0084)		(0.0225)		(0.0059)		(0.0075)		(0.0171)
$\beta_2^{(DTR)}$		-1.5114***		0.1832		0.2296		0.4337***		0.0502		0.0743
		(0.1437)		(0.1225)		(0.2412)		(0.1666)		(0.0921)		(0.1212)
Σ_1		0.0005***		0.0004***		0.0004***		0.0003***		0.0002***		0.0002***
		(9.51E-6)		(1.06E-5)		(2.42E-5)		(1.55E-5)		(5.87E-6)		(2.85E-5)
Σ_2		7.12E-5***		5.92E-5***		5.61E-5***		5.24E-5***		4.28E-5***		4.29E-5***
		(1.98E-6)		(1.77E-6)		(2.24E-6)		(1.86E-6)		(1.30E-6)		(3.38E-6)
p_{12}		0.5723		0.4564		0.4464		0.4329		0.3778		0.3564
p_{21}		0.1289		0.1049		0.0928		0.1005		0.1086		0.0889
$\bar{R}^2$	19.07%	45.84%	63.80%	76.46%	68.04%	78.93%	73.61%	81.56%	81.22%	84.38%	89.21%	91.43%
SBIC	6544.279	7221.092	6911.427	7506.899	6997.52	7587.82	7145.681	7648.697	7322.469	7772.782	7321.122	7833.976

- Sample period used for estimation is from 21 September 2010 to 22 October 2014. *, ** and *** denote the significance at 10%, 5% and 1% levels, respectively. The figure in parentheses is the standard error of coefficient. p_{12} and p_{21} are estimated switching probability from regime 1 to 2 and vice versa, respectively. SBIC is calculated as the log-likelihood value minus the penalty parameters.

Table 4: Estimation results of simple HAR-RV and MRS-HAR-RV for kerosene futures

HAR-RV $RV_{t+1d}^{(d)} = \beta_{0,1} + \beta_1^{(d)} RV_t^{(d)} + \beta_1^{(w)} RV_t^{(w)} + \beta_1^{(m)} RV_t^{(m)} + \beta_1^{(DTR)} DTR_t + \beta_1^{(D1)} D1_t + \beta_1^{(D2)} D2_t + \varepsilon_{t+1d,1}^{(d)}, \quad \varepsilon_{t+1d,1}^{(d)} \sim N(0, \Sigma_1)$

MRS-HAR-RV $RV_{t+1d}^{(d)} = \beta_{0,st} + \beta_{st}^{(d)} RV_t^{(d)} + \beta_{st}^{(w)} RV_t^{(w)} + \beta_{st}^{(m)} RV_t^{(m)} + \beta_{st}^{(DTR)} DTR_t + \beta_{st}^{(D1)} D1_t + \beta_{st}^{(D2)} D2_t + \varepsilon_{t+1d,st}^{(d)}, \quad st = \{1,2\}, \varepsilon_{t+1d,st}^{(d)} \sim N(0, \Sigma_{st})$

	1-month		2-month		3-month		4-month		5-month		6-month	
	HAR-RV	MRS-HAR	HAR-RV	MRS-HAR	HAR-RV	MRS-HAR	HAR-RV	MRS-HAR	HAR-RV	MRS-HAR	HAR-RV	MRS-HAR-
$\beta_{0,1}$	0.0001***	0.0005***	6.68E-5***	0.0004***	6.04E-5***	0.0003***	5.97E-5***	0.0003***	6.65E-5***	0.0003***	6.15E-5***	0.0003***
	(1.69E-5)	(2.18E-5)	(1.77E-5)	(3.33E-5)	(1.36E-5)	(4.37E-5)	(1.18E-5)	(1.73E-5)	(1.15E-5)	(2.28E-5)	(1.21E-5)	(9.38E-5)
$\beta_1^{(d)}$	0.1606***	0.1156**	0.2374***	0.0626	0.3028***	0.3158***	0.2411***	0.0118	0.1980***	-0.0491	0.2384***	0.1963***
	(0.0172)	(0.0455)	(0.0191)	(0.0964)	(0.0148)	(0.0241)	(0.0135)	(0.0323)	(0.0123)	(0.0496)	(0.0135)	(0.0267)
$\beta_1^{(w)}$	0.1188***	0.0812	0.0921***	0.0770	0.0666***	-0.1401**	0.0800***	0.0642	0.0982***	0.0019	0.1336***	0.0561*
	(0.0302)	(0.0769)	(0.0315)	(0.0582)	(0.0240)	(0.0702)	(0.0225)	(0.0584)	(0.0211)	(0.0744)	(0.0224)	(0.0286)
$\beta_1^{(m)}$	0.2199***	-0.1545*	0.1275**	-0.0486	0.1122***	-0.2330***	0.1558***	0.0428	0.1837***	0.4151***	0.1763***	-0.0880*
	(0.0503)	(0.0830)	(0.0540)	(0.2401)	(0.0405)	(0.0214)	(0.0377)	(0.0805)	(0.0353)	(0.0994)	(0.0352)	(0.0518)
$\beta_1^{(DTR)}$	-3.2435***	-6.1435***	0.5564	0.7506	0.7478	3.3479***	0.7814	-0.3068	0.5476	1.7790	0.3686	0.9196
	(0.7467)	(1.3820)	(0.8152)	(4.6440)	(0.6300)	(0.3063)	(0.5361)	(0.9237)	(0.5201)	(1.1954)	(0.5527)	(5.0647)
$\beta_{0,2}$		5.83E-5***		4.54E-5***		2.84E-5***		3.62E-5***		3.93E-5***		3.21E-5***
		(2.20E-6)		(3.31E-6)		(5.13E-6)		(1.51E-6)		(1.87E-6)		(2.67E-6)
$\beta_2^{(d)}$		0.0289***		0.2368***		0.0428***		0.2572***		0.2017***		0.1947***
		(0.0077)		(0.0052)		(0.0095)		(0.0080)		(0.0078)		(0.0383)
$\beta_2^{(w)}$		0.0904***		0.0143		0.1460***		0.0789***		0.1219***		0.1213***
		(0.0069)		(0.0140)		(0.0062)		(0.0091)		(0.0082)		(0.0090)
$\beta_2^{(m)}$		0.1580***		0.1079***		0.2600***		0.0782***		0.1184***		0.2536***
		(0.0094)		(0.0051)		(0.0262)		(0.0075)		(0.0064)		(0.0190)
$\beta_2^{(DTR)}$		-0.7074***		0.2713***		0.2644		0.2626***		0.6054***		-0.1407
		(0.1242)		(0.0340)		(0.1739)		(0.0977)		(0.0975)		(0.1217)
Σ_1		0.0003***		0.0005***		0.0003***		0.0002***		0.0003***		0.0003***
		(6.84E-6)		(8.74E-5)		(1.64E-5)		(4.27E-6)		(6.37E-6)		(3.21E-5)
Σ_2		5.43E-5***		5.65E-5***		4.90E-5***		4.38E-5***		5.12E-5***		4.16E-5***
		(1.86E-6)		(4.97E-6)		(1.15E-6)		(1.23E-6)		(1.41E-6)		(3.26E-6)
p_{12}		0.4206		0.5254		0.4546		0.4826		0.3781		0.3201
p_{21}		0.1435		0.0856		0.1155		0.1460		0.0749		0.0943
$\bar{R}^2$	70.09%	80.34%	65.41%	74.01%	79.06%	86.89%	82.45%	89.52%	85.21%	88.27%	82.68%	85.33%
SBIC	6891.60	7385.36	6825.26	7573.75	7074.29	7645.14	7230.00	7684.98	7255.82	7747.40	7199.88	7762.27

- Sample period used for estimation is from 21 September 2010 to 22 October 2014. *, ** and *** denote the significance at 10%, 5% and 1% levels, respectively. The figure in parentheses is the standard error of coefficient. p_{12} and p_{21} are estimated switching probability from regime 1 to 2 and vice versa, respectively. SBIC is calculated as the log-likelihood value minus the penalty parameters.

Table 5: Estimation results of simple HAR-RV and MRS-HAR-RV for crude oil futures

HAR-RV	$RV_{t+1d}^{(d)} = \beta_{0,1} + \beta_1^{(d)} RV_t^{(d)} + \beta_1^{(w)} RV_t^{(w)} + \beta_1^{(m)} RV_t^{(m)} + \beta_1^{(DTR)} DTR_t + \beta_1^{(D1)} D1_t + \beta_1^{(D2)} D2_t + \varepsilon_{t+1d,1}^{(d)}, \quad \varepsilon_{t+1d,1}^{(d)} \sim N(0, \Sigma_1)$
MRS-HAR-RV	$RV_{t+1d}^{(d)} = \beta_{0,st} + \beta_{st}^{(d)} RV_t^{(d)} + \beta_{st}^{(w)} RV_t^{(w)} + \beta_{st}^{(m)} RV_t^{(m)} + \beta_{st}^{(DTR)} DTR_t + \beta_{st}^{(D1)} D1_t + \beta_{st}^{(D2)} D2_t + \varepsilon_{t+1d,st}^{(d)}, \quad st = \{1,2\}, \varepsilon_{t+1d,st}^{(d)} \sim N(0, \Sigma_{st})$

	Current-month		1-month		2-month		3-month		4-month		5-month	
	HAR-RV	MRS-HAR	HAR-RV	MRS-HAR	HAR-RV	MRS-HAR	HAR-RV	MRS-HAR	HAR-RV	MRS-HAR	HAR-RV	MRS-HAR
$\beta_{0,1}$	-0.0001***	5.41E-5***	7.21E-5***	0.0004***	7.21E-5***	0.0005***	7.21E-5***	0.0004***	6.26E-5***	0.0004***	6.94E-5***	0.0003***
	(3.39E-5)	(1.59E-5)	(2.25E-5)	(7.69E-5)	(2.18E-5)	(0.0002)	(1.71E-5)	(2.44E-5)	(1.51E-5)	(7.76E-5)	(1.59E-5)	(2.39E-5)
$\beta_1^{(d)}$	0.0724***	0.0826**	0.1433***	0.1282***	0.2032***	0.0708	0.2210***	0.2076***	0.2606***	0.1113**	0.2608***	0.2255***
	(0.0207)	(0.0334)	(0.0156)	(0.0112)	(0.0156)	(0.0959)	(0.0116)	(0.0365)	(0.0113)	(0.0455)	(0.0118)	(0.0125)
$\beta_1^{(w)}$	0.0885**	0.0852	0.1172***	0.1105**	0.0854***	-0.0168	0.0756***	0.0049	0.1129***	0.0641*	0.1114***	0.0528*
	(0.0373)	(0.0597)	(0.0273)	(0.0529)	(0.0264)	(0.0561)	(0.0197)	(0.0468)	(0.0184)	(0.0334)	(0.0194)	(0.0272)
$\beta_1^{(m)}$	0.1950***	0.2298*	0.1850***	0.0475	0.1776***	-0.0565	0.1465***	-0.1467**	0.1541***	-0.0790	0.1547***	-0.1007
	(0.0666)	(0.1193)	(0.0472)	(0.2689)	(0.0455)	(0.1066)	(0.0341)	(0.0697)	(0.0302)	(0.1539)	(0.0319)	(0.0707)
$\beta_1^{(DTR)}$	14.0343***	21.0914***	2.3447**	-0.4638	1.9100*	1.0779	2.1797***	3.9491***	1.9675***	1.8648	1.5408**	1.9333*
	(1.6612)	(1.5830)	(1.0208)	(4.6667)	(1.0082)	(3.6245)	(0.8014)	(1.2789)	(0.6978)	(1.2547)	(0.7347)	(0.9925)
$\beta_{0,2}$		-4.07E-6***		4.78E-5***		3.52E-5***		1.90E-5***		1.41E-5***		3.54E-5***
		(3.69E-7)		(7.96E-6)		(1.30E-5)		(2.61E-6)		(5.47E-6)		(2.08E-6)
$\beta_2^{(d)}$		0.0005		0.0246		0.2035***		0.0127		0.2656***		0.2240***
		(0.0024)		(0.0156)		(0.0066)		(0.0081)		(0.0009)		(0.0095)
$\beta_2^{(w)}$		0.0176***		0.0502***		0.0945***		0.0980***		0.2680***		0.0894***
		(0.0034)		(0.0106)		(0.0205)		(0.0046)		(0.0096)		(0.0073)
$\beta_2^{(m)}$		0.0265***		0.0804*		0.1020***		0.2992***		0.0947***		0.2733***
		(0.0062)		(0.0481)		(0.0392)		(0.0089)		(0.0264)		(0.0079)
$\beta_2^{(DTR)}$		2.7709***		1.3677***		1.4789***		1.7384***		1.4438***		0.1631
		(0.2395)		(0.3796)		(0.4014)		(0.1394)		(0.2889)		(0.1092)
Σ_1		0.0008***		0.0004***		0.0005***		0.0003***		0.0003***		0.0004***
		(9.54E-5)		(3.90E-5)		(0.0002)		(5.09E-6)		(1.66E-5)		(6.03E-6)
Σ_2		3.88E-5***		6.26E-5***		8.23E-5***		6.55E-5***		6.20E-5***		5.26E-5***
		(4.02E-6)		(6.17E-6)		(1.37E-5)		(2.04E-6)		(3.79E-6)		(1.55E-6)
p_{12}		0.5964		0.4300		0.4294		0.3334		0.2897		0.2868
p_{21}		0.1911		0.1832		0.1003		0.1111		0.0849		0.0923
$\bar{R}^2$	59.55%	71.17%	76.72%	83.81%	76.89%	81.29%	86.96%	91.39%	88.01%	90.04%	86.75%	89.21%
SBIC	6157.396	7386.73	6605.319	7119.792	6615.692	7178.596	6838.313	7324.865	6975.095	7446.656	6923.484	7503.817

- Sample period used for estimation is from 21 September 2010 to 22 October 2014. *, ** and *** denote the significance at 10%, 5% and 1% levels, respectively. The figure in parentheses is the standard error of coefficient. p_{12} and p_{21} are estimated switching probability from regime 1 to 2 and vice versa, respectively. SBIC is calculated as the log-likelihood value minus the penalty parameters.

Table 6: The average level of volatility in high- and low-volatility regimes in in-sample

	1-month	2-month	3-month	4-month	5-month	6-month
Gasoline						
$\overline{RV_1}$	40.84%	33.70%	35.60%	33.06%	30.56%	30.29%
$\overline{RV_2}$	17.72%	16.95%	17.15%	17.41%	18.07%	18.05%
N_1	157	193	142	168	182	145
N_2	824	788	839	813	799	836
$OTI._1$	279.91	131.50	154.01	215.28	576.93	1825.01
$OTI._2$	162.60	109.74	120.00	161.19	458.60	1410.39
Kerosene						
	1-month	2-month	3-month	4-month	5-month	6-month
$\overline{RV_1}$	35.25%	37.42%	32.16%	31.19%	31.53%	31.26%
$\overline{RV_2}$	16.93%	16.78%	16.69%	17.21%	18.21%	18.25%
N_1	192	108	173	159	141	168
N_2	789	873	808	822	840	813
$OTI._1$	298.86	138.33	134.28	169.77	267.86	535.99
$OTI._2$	188.43	108.55	98.74	119.01	250.49	448.85
Crude oil						
	Current-m.	1-month	2-month	3-month	4-month	5-month
$\overline{RV_1}$	40.89%	37.51%	39.07%	36.82%	38.55%	35.05%
$\overline{RV_2}$	13.34%	18.68%	20.09%	19.86%	20.35%	21.46%
N_1	222	226	148	189	176	183
N_2	759	755	833	792	805	798
$OTI._1$	50.35	68.15	80.41	128.31	380.91	1004.69
$OTI._2$	40.70	55.81	73.17	102.49	300.77	774.34

- Regime 1 and 2 are defined as unconditional regime probability greater than 0.5 and less than or equal to 0.5 respectively. $\overline{RV_1}$ and $\overline{RV_2}$ are the average realised volatility in regime 1 and 2 respectively. N_1 and N_2 are the number of days of being in regime 1 and 2 respectively. $OTI._1$ and $OTI._2$ are the average order-trade imbalance in regime 1 and regime 2 respectively.

Table 7: Comparison of forecasting performance of different volatility models against MRS-HAR-RV with Diebold-Mariano test

	Gasoline					Kerosene					Crude oil				
	HAR-RV	HAR-RV-CJ MedJ	HAR-RV-CJ BPVJ	GARCH	MRS-GARCH	HAR-RV	HAR-RV-CJ MedJ	HAR-RV-CJ BPVJ	GARCH	MRS-GARCH	HAR-RV	HAR-RV-CJ MedJ	HAR-RV-CJ BPVJ	GARCH	MRS-GARCH
						Panel A: Improvement in QLike									
1-m.	-0.005	0.001	-0.004	-0.060*	-0.105***	-0.002	0.007	-0.002	-0.099**	-0.066	-0.611***	-12.79***	-16.59***	-0.450***	-0.627***
2-m.	0.008	0.013	-0.006	-0.088***	-0.019	0.013	0.013	0.012	-0.047*	0.034	0.033	0.023	0.068	-0.069**	-0.012
3-m.	0.003	-0.012	-0.052**	-0.073***	-0.091*	0.004	0.020	-0.004	-0.088***	-0.029	0.040	0.041	0.058	-0.023	0.062
4-m.	0.002	0.006	0.000	-0.086***	-0.044***	0.005	0.018	0.000	-0.059**	0.003	-0.002	0.022	0.038	-0.070**	-0.001
5-m.	-0.015**	0.003	-0.006	-0.085***	-0.060***	0.001	0.028	0.027	-0.063***	0.011	-0.011	0.021	0.014	-0.085***	-0.024*
6-m.	-0.046***	-0.051***	-0.001	-0.091***	-0.032**	-0.005	0.021	0.013	-0.079***	-0.013	-0.012	0.021	0.015	-0.103***	-0.046***
						Panel B: % Improvement in MAE									
1-m.	-3.84**	-3.52**	-3.57**	-24.56***	-12.08***	2.72	2.27	3.64	-16.87***	-1.79	3.73	-8.35**	-7.14**	-16.34***	-2.19
2-m.	1.46	0.15	4.57	-32.64***	-17.29***	1.40	-0.32	3.05	-30.84***	0.02	1.54	-0.49	-1.21	-40.09***	-25.2***
3-m.	1.98	-0.12	0.30	-30.45***	-13.63***	1.69	3.88	2.09	-34.67***	-9.13***	3.65	-2.58	-1.60	-31.59***	-11.26**
4-m.	-0.24	-1.38	2.22	-34.77***	-18.30***	0.56	2.92	7.92	-35.11***	-22.12	-0.99	-7.06**	-10.48***	-37.74***	-21.87***
5-m.	-4.51***	-3.45	-5.58	-36.53***	-24.09***	3.99	3.93	6.79	-36.37***	-4.23	-1.06	-3.24	-5.24	-42.69***	-27.98***
6-m.	-6.54***	-6.84***	-2.28	-40.39***	-21.69***	0.64	2.20	0.46	-40.85***	-10.18***	-2.70*	-4.01	-5.77	-46.82***	-28.86***
						Panel C: % Improvement in MMEU									
1-m.	0.01	-0.01	0.01	0.08	-0.02	0.04	0.04	0.04	0.24	-0.01	-0.25***	-0.36***	-0.38***	0.04	0.05
2-m.	0.05	0.05	0.03	0.17	0.11	0.01	0.03	-0.01	0.26	0.06	0.05	0.04	0.17	0.38	0.28
3-m.	0.02	0.02	-0.04***	0.20	0.08	0.01	0.05	0.01	0.25	0.07	0.07	0.12	0.17	0.35	0.23
4-m.	0.01	0.06	0.03	0.15	0.07	0.03	0.05	0.05	0.29	0.09	0.00	0.22	0.27	0.34	0.29
5-m.	0.00	0.12	0.12	0.14	0.03	-0.03**	0.09	0.01	0.27	0.05	-0.05***	0.21	0.23	0.23	0.12
6-m.	-0.04***	-0.05***	0.16	0.19	0.08	0.03	0.18	0.17	0.27	0.07	-0.01	0.23	0.22	0.25	0.15
						Panel C: % Improvement in MMEO									
1-m.	-0.04***	-0.04**	-0.04**	-0.31***	-0.07**	-0.03*	-0.04*	-0.02	-0.42***	0.01	0.40	0.40	0.42	-0.26***	-0.07*
2-m.	-0.03***	-0.06***	0.00	-0.48***	-0.27***	0.00	-0.03*	0.03	-0.59***	-0.07***	-0.05***	-0.06***	-0.23***	-0.97***	-0.61***
3-m.	-0.01	-0.01	0.05	-0.48***	-0.18***	0.01	-0.02	0.02	-0.61***	-0.14***	-0.05**	-0.18***	-0.22***	-0.84***	-0.38***
4-m.	-0.03**	-0.10**	-0.02*	-0.48***	-0.24***	-0.02**	-0.03**	0.01	-0.66***	-0.18***	-0.02	-0.32***	-0.40***	-0.85***	-0.53***
5-m.	-0.05***	-0.16***	-0.18***	-0.46***	-0.22***	0.06	-0.09***	0.02	-0.62***	-0.09***	0.03	-0.27***	-0.28***	-0.75***	-0.41***
6-m.	0.00	0.02	-0.20***	-0.55***	-0.26***	-0.02	-0.19***	-0.17***	-0.69***	-0.17***	-0.02	-0.30***	-0.31***	-0.82***	-0.44***

- The out-of-sample is from 23 October 2014 to 30 October 2015, a total of 250 forecasts. For QLike loss function comparisons, figures represent the decrease in QLike of MRS-HAR-RV compared to alternative models. For MAE, MMEO amd MMEU, figures represent the percentage change in loss functions. Negative value means MRS-HAR-RV outperforms the alternative model. The DM test is for significance of the difference between the loss function of the MRS-HAR-RV and the competing model. Finally, *, ** and *** denote the significance at 10%, 5% and 1% levels, respectively.

Table 8: Results of back-testing the VaR estimates of long positions for different energy futures and across maturity spectrum

		HAR	MRS-HAR	MedRV	BPV	GARCH	MRS-GARCH	HS
Panel A: Gasoline								
1-month	PF	1.60%	1.20%	1.60%	1.60%	4.80%	5.20%	4.80%
	LUC	0.77	0.09	0.77	0.77	19.02***	22.32***	19.02***
	LCC	0.93	0.19	4.91*	0.93	25.28***	24.41***	25.28***
2-month	PF	1.20%	1.60%	1.60%	1.60%	4.00%	3.60%	4.40%
	LUC	0.09	0.77	0.77	0.77	12.96***	10.23***	15.89***
	LCC	0.19	0.93	0.93	0.93	21.49***	14.92***	23.21***
3-month	PF	1.20%	1.20%	0.80%	0.80%	4.40%	5.20%	4.40%
	LUC	0.09	0.09	0.11	0.11	15.89***	22.32***	15.89***
	LCC	0.19	0.19	0.16	0.16	23.21***	22.57***	23.21***
4-month	PF	0.80%	0.80%	0.80%	0.80%	3.60%	5.20%	3.60%
	LUC	0.11	0.11	0.11	0.11	10.23***	22.32***	10.23***
	LCC	0.16	0.16	0.16	0.16	20.16***	27.66***	11.31***
5-month	PF	0.00%	0.00%	0.00%	0.00%	3.60%	4.40%	3.20%
	LUC	N/A	N/A	N/A	N/A	10.23***	15.89***	7.73***
	LCC	N/A	N/A	N/A	N/A	20.16***	16.45***	8.33**
6-month	PF	1.60%	0.80%	1.60%	0.80%	3.60%	3.20%	3.20%
	LUC	0.77	0.11	0.77	0.11	10.23***	7.73***	7.73***
	LCC	0.93	0.16	0.93	0.16	20.16***	9.18**	8.33**
Panel B: Kerosene								
1-month	PF	1.60%	1.60%	2.00%	1.60%	3.20%	4.40%	3.60%
	LUC	0.77	0.77	1.96	0.77	7.73***	15.89***	10.23***
	LCC	0.93	0.93	2.20	4.91*	26.74***	16.45***	20.16***
2-month	PF	3.20%	2.80%	3.20%	2.80%	3.20%	4.40%	2.80%
	LUC	7.73***	5.50**	7.73***	5.50**	7.73***	15.89***	5.50**
	LCC	9.18**	5.96*	9.18**	5.96*	19.31***	16.45***	12.29***
3-month	PF	2.80%	1.60%	2.80%	2.80%	4.40%	4.80%	4.00%
	LUC	5.50**	0.77	5.50**	5.50**	15.89***	19.02***	12.96***
	LCC	12.29***	0.93	5.96*	5.96*	23.21***	21.61***	21.49***
4-month	PF	3.20%	2.40%	2.80%	3.20%	4.00%	4.40%	4.40%
	LUC	7.73***	3.56*	5.50**	7.73***	12.96***	15.89***	15.89***
	LCC	13.38***	3.90	12.29***	9.18**	21.49***	16.45***	23.21***
5-month	PF	2.40%	2.40%	2.00%	2.00%	3.60%	4.80%	4.40%
	LUC	3.56*	3.56*	1.96	1.96	10.23***	19.02***	15.89***
	LCC	6.03**	3.90	5.15*	2.20	14.92***	20.33***	19.08***
6-month	PF	2.40%	2.80%	1.20%	2.00%	3.60%	5.20%	4.00%
	LUC	3.56*	5.50**	0.09	1.96	10.23***	22.32***	12.96***
	LCC	3.90	5.96*	0.19	2.20	14.92***	24.41***	16.84***
Panel C: Crude oil								
C-month	PF	5.20%	3.60%	4.40%	4.40%	4.80%	4.80%	3.20%
	LUC	22.32***	10.23***	15.89***	15.89***	19.02***	19.02***	7.73***
	LCC	24.41***	11.31***	23.21***	23.21***	30.09***	25.28***	9.18**
1-month	PF	2.80%	4.40%	4.80%	4.80%	7.60%	8.00%	3.60%
	LUC	5.50**	15.89***	19.02***	19.02***	45.19***	49.45***	10.23***
	LCC	5.96*	17.00***	19.40***	19.40***	52.24***	59.02***	10.98***
2-month	PF	2.00%	3.20%	5.60%	6.40%	7.20%	8.40%	4.00%
	LUC	1.96	7.73***	25.78***	33.15***	41.06***	53.80***	12.96***
	LCC	2.20	8.33**	25.96***	34.14***	45.75***	56.60***	13.74***
3-month	PF	2.00%	2.00%	4.80%	5.20%	7.20%	7.20%	4.00%
	LUC	1.96	1.96	19.02***	22.32***	41.06***	41.06***	12.96***
	LCC	2.20	2.20	19.40***	24.41***	45.75***	45.75***	13.74***
4-month	PF	2.00%	0.40%	4.00%	3.60%	6.80%	7.60%	3.60%
	LUC	1.96	1.18	12.96***	10.23***	37.04***	45.19***	10.23***
	LCC	2.20	1.19	13.74***	10.98***	42.52***	49.18***	10.98***
5-month	PF	1.60%	1.60%	5.60%	4.40%	6.80%	8.00%	4.00%
	LUC	0.77	0.77	25.78***	15.89***	37.04***	49.45***	12.96***
	LCC	0.93	0.93	27.44***	16.45***	42.52***	50.80***	13.74***

- VaR is calculated based on 250 observations in out-of-sample from 23 October 2014 to 30 October 2015.
- PF is the probability of failures (violations). The LR statistic in bold denotes rejection of each test. The critical value for the unconditional coverage and independent test ($LR_{(UC)}$ and $LR_{(UC)}$) is 2.7055, and that for the conditional coverage likelihood ratio test $LR_{(CC)}$ is 4.6052. *, ** and *** denote the significance at 10%, 5% and 1% levels, respectively.

Table 9: Results of back-testing the VaR estimates of short position for different energy futures and across maturity spectrum

		HAR	MRS-HAR	MedRV	BPV	GARCH	MRS-GARC	HS
Panel A: Gasoline								
1-month	PF	0.80%	0.40%	0.80%	0.80%	1.60%	2.00%	2.00%
	LUC	0.11	1.18	0.11	0.11	0.77	1.96	1.96
	LCC	0.16	1.19	0.16	0.16	0.93	2.20	2.20
2-month	PF	0.40%	0.80%	0.40%	0.80%	0.40%	1.60%	2.40%
	LUC	1.18	0.11	1.18	0.11	1.18	0.77	3.56*
	LCC	1.19	0.16	1.19	0.16	1.19	0.93	6.03**
3-month	PF	0.40%	0.40%	0.40%	0.80%	0.40%	1.60%	2.00%
	LUC	1.18	1.18	1.18	0.11	1.18	0.77	1.96
	LCC	1.19	1.19	1.19	0.16	1.19	0.93	5.15*
4-month	PF	0.40%	0.80%	0.40%	0.40%	0.80%	0.80%	1.60%
	LUC	1.18	0.11	1.18	1.18	0.11	0.11	0.77
	LCC	1.19	0.16	1.19	1.19	0.16	0.16	4.91*
5-month	PF	0.40%	0.40%	0.40%	0.40%	0.80%	1.60%	1.60%
	LUC	1.18	1.18	1.18	1.18	0.11	0.77	0.77
	LCC	1.19	1.19	1.19	1.19	0.16	0.93	4.91*
6-month	PF	0.80%	0.80%	0.80%	0.80%	0.80%	0.40%	2.40%
	LUC	0.11	0.11	0.11	0.11	0.11	1.18	3.56*
	LCC	0.16	0.16	0.16	0.16	0.16	1.19	6.03**
Panel B: Kerosene								
1-month	PF	0.80%	0.80%	0.80%	0.80%	0.80%	2.00%	1.60%
	LUC	0.11	0.11	0.11	0.11	0.11	1.96	0.77
	LCC	0.16	0.16	0.16	0.16	0.16	2.20	0.93
2-month	PF	0.80%	0.80%	0.80%	0.80%	0.80%	2.80%	2.40%
	LUC	0.11	0.11	0.11	0.11	0.11	5.50**	3.56*
	LCC	0.16	0.16	0.16	0.16	0.16	5.96*	6.03**
3-month	PF	0.80%	0.80%	0.80%	1.20%	0.80%	2.80%	2.40%
	LUC	0.11	0.11	0.11	0.09	0.11	5.50**	3.56*
	LCC	0.16	0.16	0.16	0.19	0.16	7.40**	6.03**
4-month	PF	0.80%	1.20%	0.80%	1.20%	0.40%	2.40%	2.80%
	LUC	0.11	0.09	0.11	0.09	1.18	3.56*	5.50**
	LCC	0.16	0.19	0.16	0.19	1.19	6.03**	7.40**
5-month	PF	0.80%	0.80%	0.80%	0.80%	0.40%	2.40%	3.60%
	LUC	0.11	0.11	0.11	0.11	1.18	3.56*	10.23***
	LCC	0.16	0.16	0.16	0.16	1.19	6.03**	11.31***
6-month	PF	1.20%	1.20%	0.80%	0.80%	0.80%	1.60%	3.60%
	LUC	0.09	0.09	0.11	0.11	0.11	0.77	10.23***
	LCC	0.19	0.19	0.16	0.16	0.16	0.93	11.31***
Panel C: Crude oil								
C-month	PF	3.60%	2.00%	2.40%	2.80%	2.80%	2.80%	0.80%
	LUC	10.23***	1.96	3.56*	5.50**	5.50**	5.50**	0.11
	LCC	11.31***	2.20	6.03**	7.40**	12.29***	7.40**	0.16
1-month	PF	0.80%	2.40%	5.20%	6.40%	5.60%	8.00%	4.00%
	LUC	0.11	3.56*	22.32***	33.15***	25.78***	49.45***	12.96***
	LCC	0.16	3.90	22.57***	33.28***	27.44***	49.72***	16.84***
2-month	PF	0.40%	1.20%	4.80%	6.00%	5.60%	6.80%	3.60%
	LUC	1.18	0.09	19.02***	29.40***	25.78***	37.04***	10.23***
	LCC	1.19	0.19	19.40***	29.53***	27.44***	37.77***	14.92***
3-month	PF	0.40%	0.80%	3.20%	3.60%	5.20%	7.20%	3.60%
	LUC	1.18	0.11	7.73***	10.23***	22.32***	41.06***	10.23***
	LCC	1.19	0.16	9.18**	11.31***	24.41***	41.59***	14.92***
4-month	PF	0.40%	0.40%	2.80%	2.40%	4.00%	7.20%	3.20%
	LUC	1.18	1.18	5.50**	3.56*	12.96***	41.06***	7.73***
	LCC	1.19	1.19	7.40**	6.03**	13.74***	41.59***	9.18**
5-month	PF	0.40%	0.40%	3.20%	2.40%	4.00%	5.60%	3.60%
	LUC	1.18	1.18	7.73***	3.56*	12.96***	25.78***	10.23***
	LCC	1.19	1.19	9.18**	6.03**	16.84***	25.96***	14.92***

- VaR is calculated based on 250 observations in out-of-sample from 23 October 2014 to 30 October 2015.
- PF is the probability of failures (violations). The LR statistic in bold denotes rejection of each test. The critical value for the unconditional coverage and independent test ($LR_{(UC)}$ and $LR_{(UC)}$) is 2.7055, and that for the conditional coverage likelihood ratio test $LR_{(CC)}$ is 4.6052. *, ** and *** denote the significance at 10%, 5% and 1% levels, respectively.

Figure 1: Realised volatility under different sampling frequency for gasoline futures

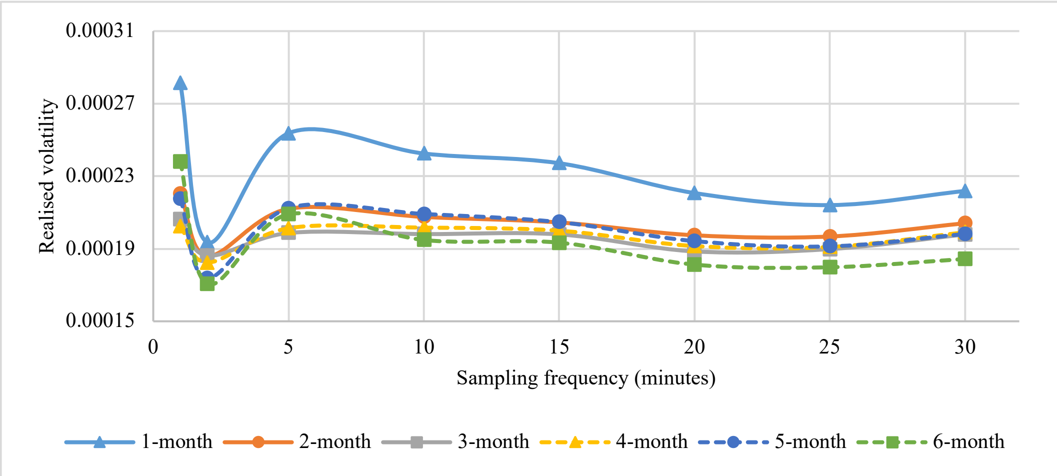

Figure 2: Realised volatility under different sampling frequency for kerosene futures

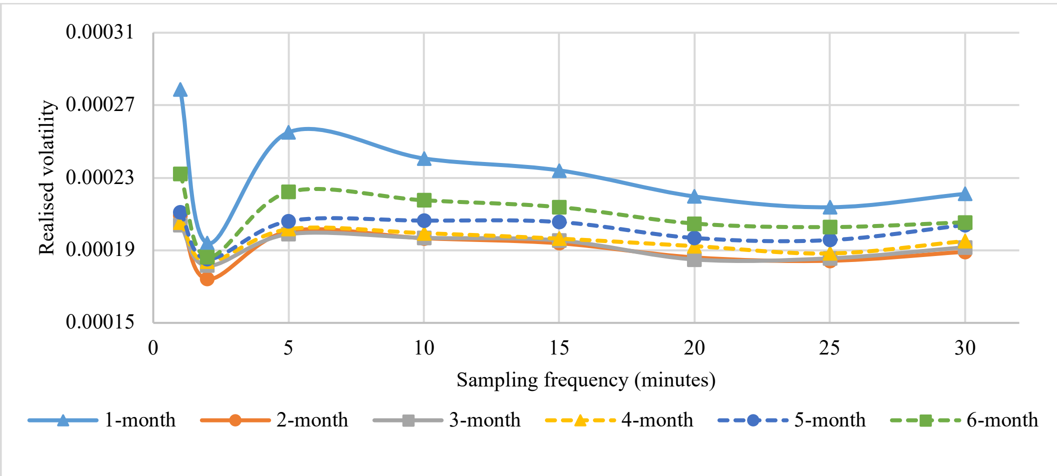

Figure 3: Realised volatility under different sampling frequency for crude oil futures

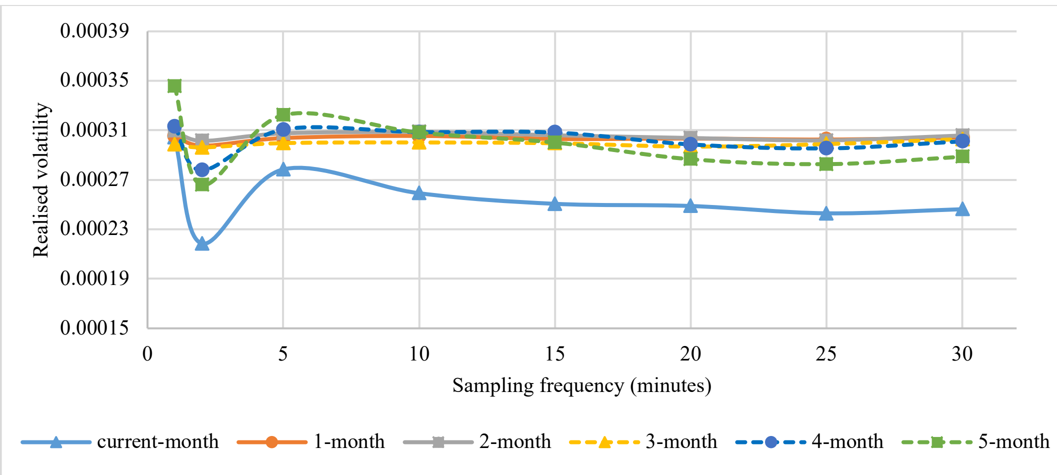

Figure 4: Daily log-return of futures contracts with longest maturity

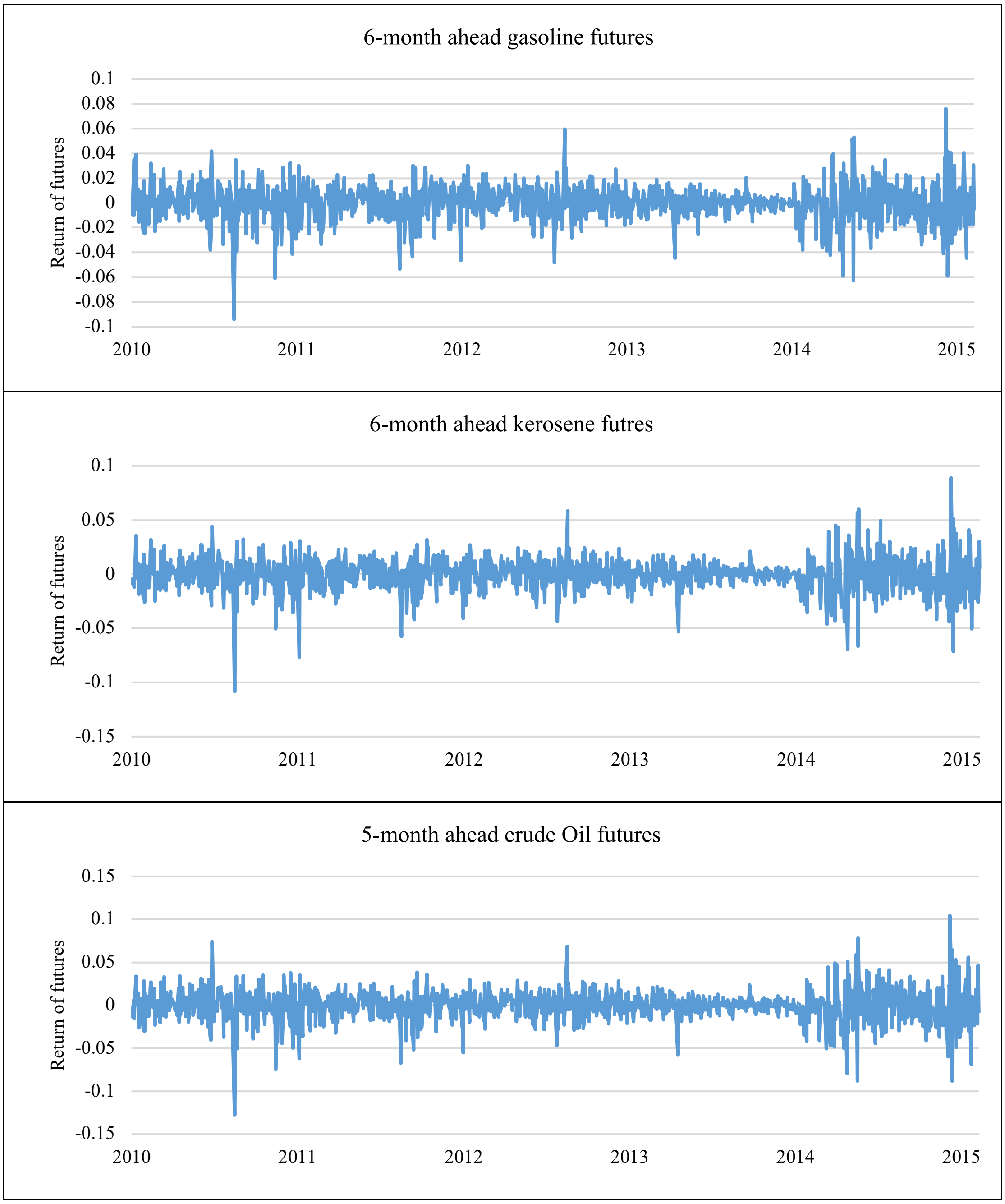

Figure 5: Daily realised volatility of futures contracts with longest maturity

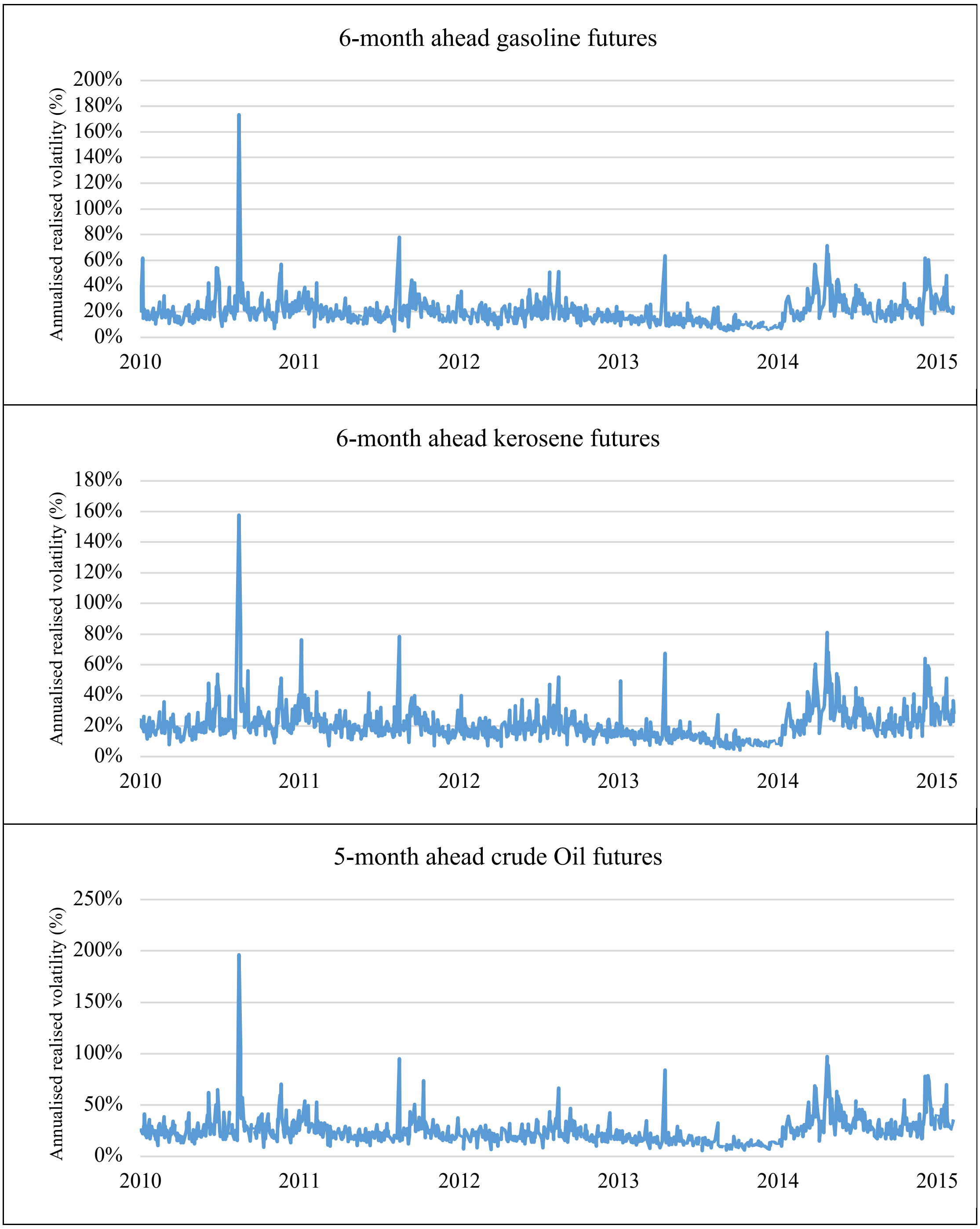

Figure 6: Value at Risk (1%) of 6-month gasoline futures

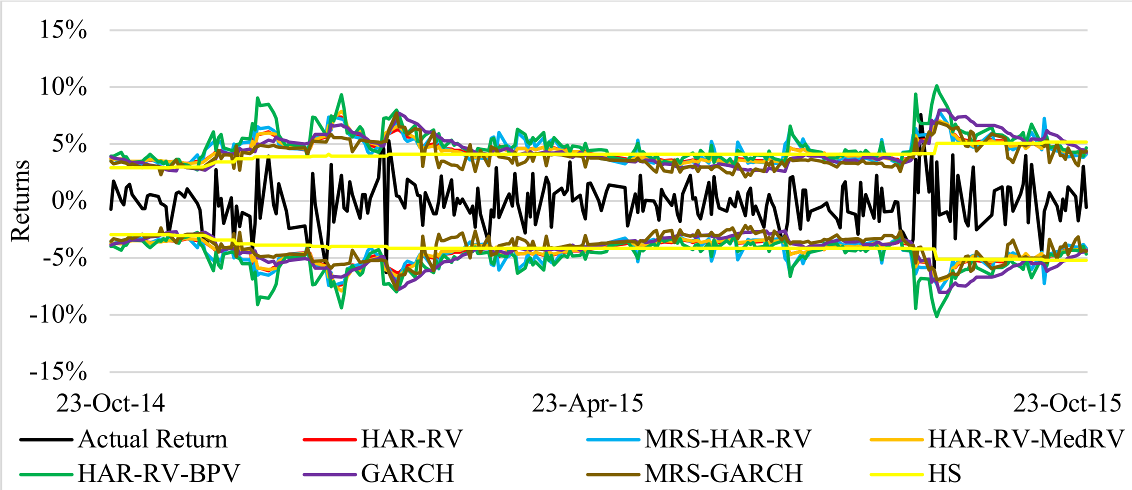

Figure 7: Value at Risk (1%) of 6-month kerosene futures

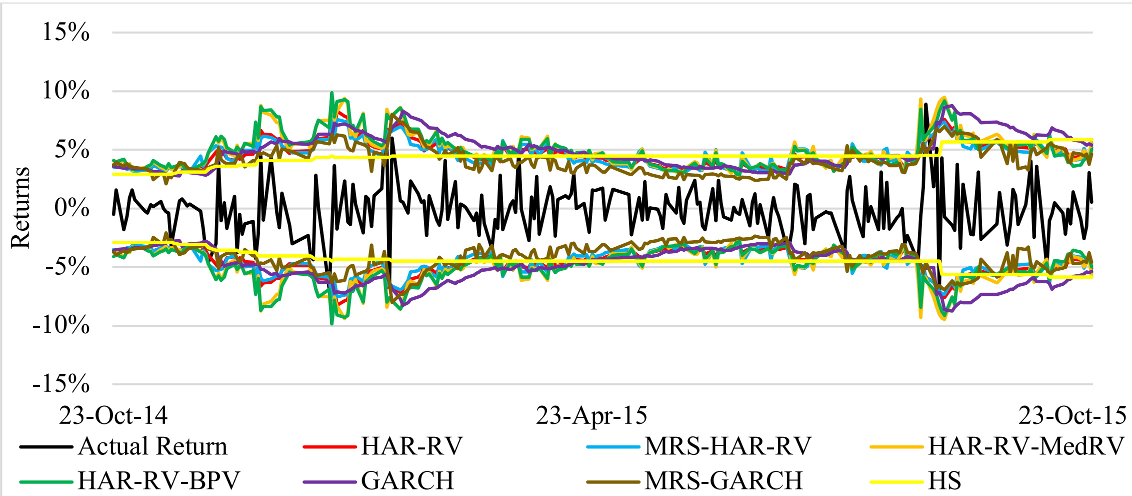

Figure 8: Value at Risk (1%) of 6-month crude oil futures