Journal of Banking & Finance, 47, pp.1-14, 2014
https://doi.org/10.1016/j.jbankfin.2014.05.026

Forecasting Volatility of the U.S. Oil Market

Peter Molnar[a,*], Erik Haugom[b], Henrik Langeland[a], Sjur Westgaard[a]

[a] *The Norwegian University of Science and Technology, NO-7491 Trondheim, Norway*
[b] *Lillehammer University College, NO-2624 Lillehammer, Norway, and the Norwegian University of Science and Technology, NO-7491 Trondheim, Norway*

Abstract

We examine the information content of the CBOE Crude Oil Volatility Index (OVX) when forecasting realized volatility in the WTI futures market. Additionally, we study whether other market variables, such as volume, open interest, daily returns, bid-ask spread and the slope of the futures curve, contains predictive power beyond what is embedded in the implied volatility. In out-of-sample forecasting we find that econometric models based on realized volatility can be improved by including implied volatility and other variables. Our results show that including implied volatility significantly improves daily and weekly volatility forecasts, while including other market variables significantly improves daily, weekly and monthly volatility forecasts.
Keywords: Oil prices, realized volatility, implied volatility, volatility forecasting,
JEL: G14, G13, Q47, L94.

*Corresponding author.
Email addresses: peto.molnar@gmail.com (Peter Molnar), erik.haugom@hil.no (Erik Haugom), hlangeland@gmail.com (Henrik Langeland), sjur.westgaard@iot.ntnu.no (Sjur Westgaard)

1. Introduction

Accurate volatility forecasts are crucial for portfolio optimization, options and derivatives pricing, value-at-risk modeling, and hedging. Forecasting volatility has traditionally been done using the generalized autoregressive conditional heteroscedasticity (GARCH) approach of Bollerslev (1986) and Engle (1982), also in energy commodity markets (see e.g. Marzo and Zagaglia (2010) and Wei et al. (2010)).

A breakthrough in volatility measuring was provided when Andersen and Bollerslev (1998) introduced realized volatility as the sum of squared intra-daily returns. This made volatility almost an observable variable which can be modeled straightforwardly with standard time-series techniques.

It has long been recognized that there are other sources of information about future volatility than realized volatility. A natural candidate is the market's expectation of future volatility, commonly referred to as implied volatility (IV). Some previous studies (e.g. Lamoureux and Lastrapes (1993); Jorion (1995); Agnolucci (2009)) argue that forecasts obtained from implied volatility are both biased and inefficient. Evidence that IV improves volatility forecasts has also been presented (e.g. Day and Lewis (1993); Szakmary et al. (2003); Doran and Ronn (2005); Agnolucci (2009)). According to Jorion (1995), a failure to unearth IV's predictive power can only be interpreted in two ways; inefficient information processing in options markets or misleading test procedures. In highly liquid and transparent markets such as the WTI futures market the former is unlikely. Left is the latter, and in particular the discussion about the bias of the Black-Scholes (BS) formula (see e.g. Doran and Ronn (2005)). A way to avoid this possible problem (and several others) is to use a volatility index which is based on the market price of variance. Such an index was introduced for the WTI futures market in 2008 and is one of the main units of analysis in this paper. Volatility has also been linked to several other market variables. For instance, the relationship between volume and volatility is widely documented (e.g. Clark (1973) and Gallant et al. (1992)). In addition to possibly improve volatility forecasts, including additional variables in the analysis can increase our understanding of the market.

Even though realized and implied volatility in equity markets has been extensively studied (see e.g. Bollerslev et al. (2013) and references within), much less work has been done in this field for commodity markets. This is particularly suprising for the oil market,considering the market's economic importance (Sadorsky, 2006). Wang et al. (2008) studied the realized correlation between oil and gas markets and found the use of RV in energy markets to be highly appropriate, especially in areas such as volatility forecasting. Martens and Zein (2004) compared forecasts obtained from a long-memory model of RV with options-implied volatility for the WTI futures market. They found that both RV and IV contain useful information in volatility forecasting. Little work has been done regarding the WTI IV index, due to its recent inception. An exception is Padungsaksawasdi and Daigler (2013) who studied the return-IV relation, and concluded that IV increases with negative returns.

In this paper we examine the role of both volatility implied from the OVX and observable market variables when forecasting volatility for the WTI futures market. We apply the simple HeterogenousAutoregRessive (HAR) model of Corsi (2009) on realized volatility itself. Additionally, two fundamentally different types of variables are used in the model; the forward looking IV index and other exogenous market variables including volume, open interest, daily returns and the slope of the futures curve. The main findings can be summarized as follows. First, we find that including information from the OVX significantly improves the day-ahead and week-ahead volatility forecasts. Second, the exogenous market variables improve volatility forecasts for daily, weekly and monthly horizons. Of the additional explanatory market variables, the daily returns is the most important factor to improve volatility forecasts.

The rest of the paper will unfold as follows. Section 2 provides a description of the WTI crude oil market and data (high-frequency data used for estimation of realized volatility, implied volatility index and other market variables). Chapter 3 presents the results and Chapter 4 concludes.

2. The data set and preliminary analyses

2.1. The data

The data used in this paper are obtained from the CME Group and contain all trades of WTI futures contracts made through the CME Globex electronic trading platform from May 16 2007 through May 15 2012, for a total of 173.4 million ticker observations. A single WTI futures contract represents 1000 bbl of oil for physical delivery in Cushing Oklahoma. Front month trading ends on the third business day prior to the 25th of the month prior to delivery. If the contract is held until expiry physical delivery must be undertaken during the following month according to the specifications in the contract.

Daily average trading volume is 900,000 futures and options contracts and the largest open interest for all contracts has reached 7.5 million lots.[1] Trading starts 17:00 and ends 16:15 eastern time (ET) on weekdays, meaning that every day there is a 45 minute break. The trading week starts Sunday evening and ends Friday afternoon.

Figure 4 (a) shows how average volume for the first position changes throughout the month. The volume traded on the second position is a mirror image because of traders rolling their positions. The same mechanics seem to drive open interest as shown in figure 4 (b). Rolling is set to happen on the business day prior to the 9th every month.

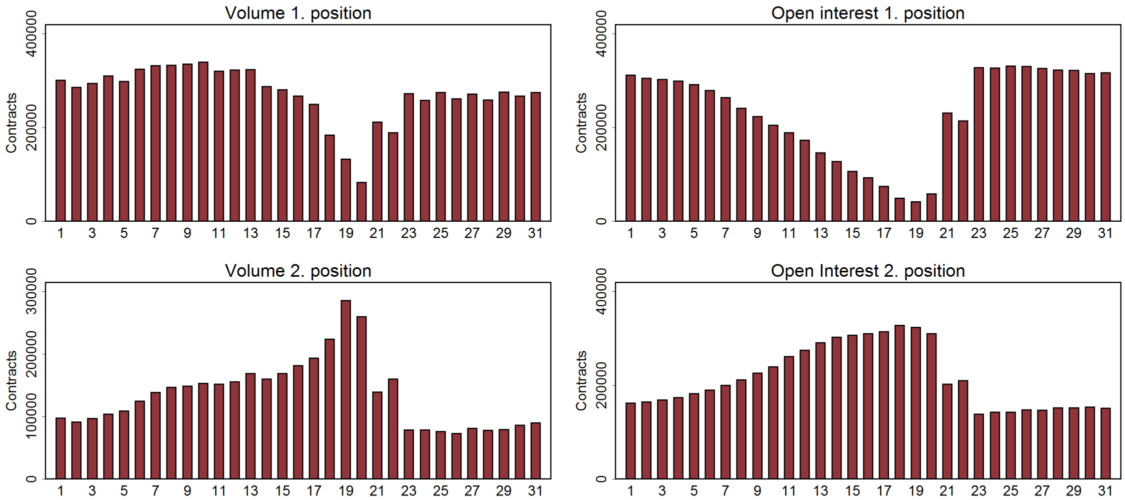
Trades happening more than 24 hours before closing and trading days with an early close (i.e before 16:15 ET) are also removed.[2] This leaves a total of 1246 trading days in the sample. In the next subsections we provide a detailed description and preliminary analyses of all the variables that will be used in the subsequent models.

2.2. Realized volatility

Andersen and Bollerslev (1998) proposed to use the realized variation as a proxy for the integrated variance. Realized variance can be written as the sum of squared intra-daily returns. Throughout this paper the measure realized volatility (RV) will

[1] Numbers for 2013. See www.cmegroup.com/trading/energy/

[2] Extended trading days can happen under special circumstances, usually around holidays.



(a) Volume **(b)** Open interest

Figure 1: Daily average values during each day of the month for volume and open interest during the period May 16 2007 to May 15 2012

be used and is defined as:

$$RV_t = \sqrt{RVar_t} = \sqrt{\sum_{j=1}^{M} r_{t,j}^2}, \qquad t = 1, \cdots, T. \tag{1}$$

where M is sampling frequency, i.e., the number of equally spaced intraday returns $r_{t,j}^2$ used in calculation of $RVar_t$ for a given day t.

To calculate the daily realized volatility, we first extract ticker prices prior to every minute over each trading day, i.e., the previous tick interpolation method, as recommended by Hansen and Lunde (2006) for theoretical reasons.

The choice of sampling frequency has a direct effect on the accuracy of the Realized Volatility (RV) measure. High sampling frequency is desirable and causes more information about the volatility of the underlying process to be captured. This can be seen in the last two columns of table 1; the realized measures based on higher frequencies display a higher autocorrelation.

There are, however, upper boundaries for sampling frequency. As frequency increases, microstructure noise increases. The optimal sampling frequency depends ultimately on the market liquidity. Andersen et al. (2001) argue that for a liquid market (which WTI futures market clearly is) 5-minute sampling interval is optimal. The same sampling frequency was chosen by e.g., Bandi and Russell (2006), Andersen et al. (2007) and Patton (2011). We follow these authors and and use a sampling frequency of five mintues in this study.

Figure 2 shows the slowly decaying autocorrelation for the 5-minute realized volatility. This indicates that the realized volatility exhibits long memory. The RV time series over the sample period is also stationary according to both the Dickey Fuller test and the Phillips-Perron test.[3]

Barndorff-Nielsen et al. (2008) found for that a decomposition of the RV into two separate measures; one based on positive and and another one on negative intra-daily returns, is useful as their results suggested a diferent autocorrelation structure for these two measures. On the basis of this finding in the equity market, we examined whether this for the WTI futures market. Our analysis showed that this effect was not present. Throughout the paper we therefore use the standard RV measure, as

[3]When splitting the sample into five subsamples, however, the same tests can not reject the existence of a unit root for all parts of the time series. Therefore, for the purpose of forecasting (when it is common to use only subsample of the data for model estimation) the time series appear as nonstationary. In other words, even though strictly speaking this time series does not have a unit root, it is relatively close to being an I(1) process. Therefore, we later estimate a model for both levels of volatility and percentage changes in volatility. Results from the Dickey Fuller test and the Phillips-Perron tests for various subsamples are available upon request.

Table 1: *RV calculated for 1, 3, 5, 10, 15, and 30 minute intervals. DF is the Dickey Fuller test and PP is the Phillips-Perron test for unit root. Values below -3.430 means rejection at the 1% significans level for both series.AC(1) and AC(10) are the autocorrelation for 1 and 10 lags respectively.*

Descriptive statistics of realized volatility measures

	Mean	Min	Max	Kurt	Skew	SD	DF	PP	AC_1	AC_{10}
RV_{1min}	0.373	0.129	1.326	7.121	2.015	0.184	-7.379	-5.245	0.916	0.820
RV_{3min}	0.363	0.110	1.313	7.233	1.994	0.180	-8.360	-5.470	0.894	0.791
RV_{5min}	0.359	0.107	1.131	6.603	1.906	0.176	-8.500	-5.201	0.891	0.791
RV_{10min}	0.354	0.111	1.158	6.532	1.881	0.175	-9.160	-6.148	0.874	0.782
RV_{15min}	0.353	0.105	1.311	7.251	1.986	0.179	-10.179	-7.859	0.847	0.762
RV_{30min}	0.346	0.104	1.313	7.510	1.999	0.181	-11.920	-7.806	0.796	0.725

defined by equation 1.

Figure 4 shows that there is no apparent monthly or weekly cycle in the RV measure. Although we study the WTI futures market at a daily frequency in this paper, we still provide some additional information about the market inferred from the high frequency data. Figure 4 (b) shows that early morning and mid-day has higher activity and volatility than the hours from 17:00 to 01:00 ET. There is an increase in volume and volatility at around 02:00 ET, which could represent markets opening in Europe and as markets open in the US there is another, even more distinct, rise in volatility and market activity. The drop at around 11:00 ET matches well with both markets closing in Europe and lunchtime in the US. However, we do not investigate intra-daily variations further.

2.3. Implied volatility

CBOE began calculating the CBOE Crude Oil Volatility Index (OVX) in June 2008 (using data back to May 2007) according to the "VIX methodology". The index

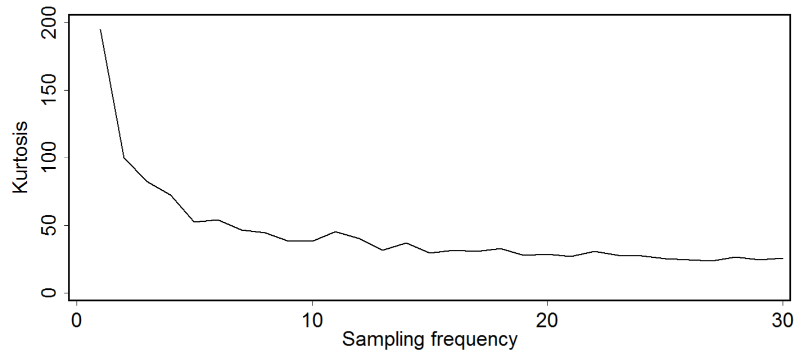


Figure 2: *Kurtosis of returns for sampling frequencies between 1 and 30 minutes.*

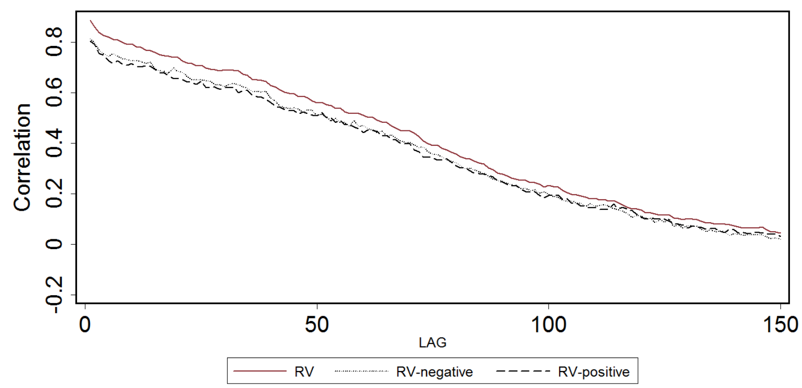
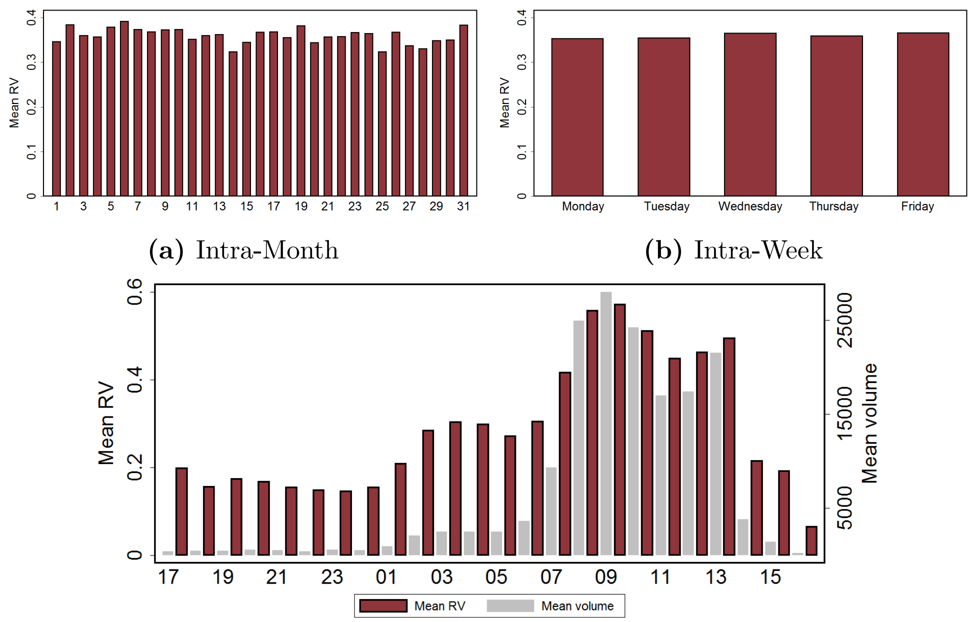


Figure 3: *Autocorrelation of RV based on 5-minute returns. Two additional lines show the autocorrelation for RV time series constructed using only squared negative returns and RV time series constructed using only squared positive returns.*



(a) Intra-Month

(b) Intra-Week

(c) Intra-Day

Figure 4: Seasonality of RV.

has become an important instrument for trading oil price volatility (Whaley, 2008).

The options underlying the OVX are options on the United States Oil fund (USO), an exchange traded fund (ETF) established to replicate the returns of the WTI benchmark price. This investment vehicle reduces transaction costs for investors seeking exposure to the oil price. Since USO needs a management and it is exposed to hedging risks it is unable to completely replicate the price of oil. This is obvious from Figure 5 which shows that the price of one share has been unable to keep up with the price of the front month contracts. It is also seen that the short term variations are very similar. The daily returns of the USO have a 88 % correlation to the first futures position and a 94 % correlation to the second futures position. The OVX represents expectations of the volatility of the USO which in practice it will means the expectations of the volatility of the contracts held by the USO.

The USO has no exposure to the spot price of oil and only holds futures contracts and other oil related derivatives. The main part of the fund's exposure to the oil market is contracts at the first and second position. The fund rolls these contracts during a 4–day window starting approximately 14 days before expiry of the first positions.[4] The rolling window for each month is publicly announced on the company website but further details about the rolling is not available to the market. It is therefore impossible to know the exact composition of the USO's contract holdings during this window other than it is likely to be shifting towards the second position. Over the course of one month the fund will therefore be holding mainly second position contracts about half of the days and first position contracts the other half.

[4]See www.unitedstatesoilfund.com/uso-rolldates.php

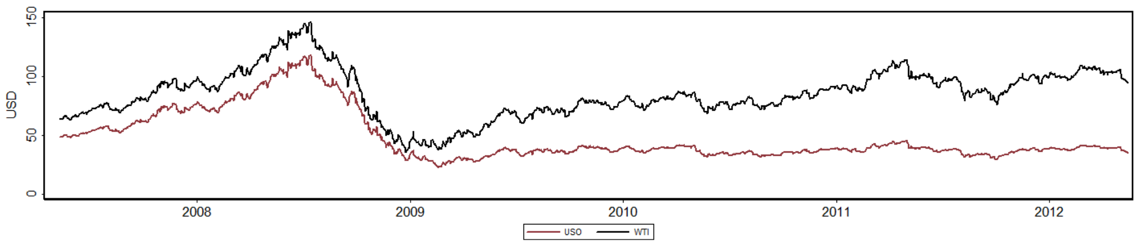


Figure 5: *The price of one share of the USO and the price of the WTI futures contracts rolling from 1. to 2. position the first business day prior to the 10th of each month for the period May 16 2007 to May 15 2012*

Because both the underlying and the options are traded side by side in a very liquid market with minimal transaction costs, prices are likely to be highly synchronized. It is therefore unlikely that the market will suffer from asymmetric information dissemination, which according to Jorion (1995) is a frequent source of measurement errors when investigating implied volatility.

2.4. Volume

Clark (1973) states that both traded volume and volatility are driven by the same underlying "news"-variable and will therefore be positively correlated. According to this theory, it is primarily the *number* of trades that will capture the content of this news variable. This relationship has been widely documented for stock markets for instance by Gallant et al. (1992) and Andersen (1996).

The contract specifications in the WTI futures market disturb the information content of the volume variable because of the previously mentioned monthly cycle. Volume on first position will start to decline as the contract nears expiry and volume on second position contract will increase as it closes in on first position as shown in figure 1. Rolling the contracts further away from expiry such as USO does, does little in removing this cycle. Therefore the volume traded is driven mostly by the structure of the market.

An alternative description of the volume-volatility relation, based on the microstructure theory described by O'Hara (1995), is grounded in different types of traders with asymmetric information. The theory states that informed traders will prefer to trade larger quantities, therefore the trade size has information content for prices (Kim and Verrecchia, 1991). However, this information can potentially be "hidden" by breaking a large trade into smaller parts, which according to (Chakravarty, 2001), is a widely used tactics.

Either way, it is possible that number of trades and the size of trades provide different information and are therefore linked differently to the volatility in the WTI futures market. We therefore decompose volume into these two components. Figure 6 plots number of trades and average size of trades calculated for each day of the month. The figure shows that the cyclic variation in volume is mainly evident in the number of trades component. We removed the cyclical component by running regression with dummy variables for each day of the month.

2.5. Open interest

Open interest is a variable which shows how many bets market participants already have taken, indicating depth and size of the market. High open interest means that a lot of subjects already have taken bets, and low open interest means that many subjects are not in the market yet. This market characteristic might be relevant for volatility forecasting. Girma and Mougoue (2002) found that lagged values of open interest affect volatility and can be used for short term predictions of price movements

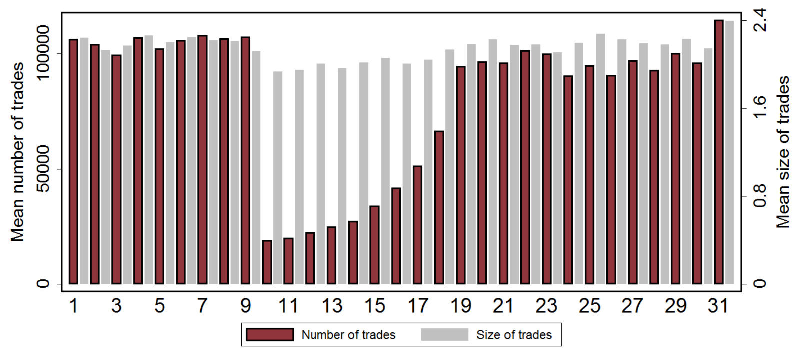


Figure 6: *The average size of trades and the average number of trades for each day of the month when rolling to second position the first business day prior to the 10th every month. Numbers are for the period May 16 2007 to May 15 2012*

in petroleum futures markets. Figlewski (1981) found that open interest is positively correlated with volatility in futures markets using a monthly average of open interest.

Since taking physical delivery of the oil is certainly not the goal of many of the subjects trading oil futures, open interest is a highly cyclical variable depending on the days to maturity of the contract, as described in the previous section. We remove the cyclical component in the same way as for number of trades.

2.6. Daily returns

Doran and Ronn (2005) found a positive correlation between price returns and volatility in energy markets, contrary to the "leverage effect" found in equity markets. Their explanation is it that higher commodity prices represents a threat to economic activity for energy importing countries. For oil markets this "inverse leverage effect" has been confirmed by Wei et al. (2010) who found significant indications of its existence in the *ICE Brent* market. However, together with Wang et al. (2008), Agnolucci (2009) and Padungsaksawasdi and Daigler (2013) found no evidence of such effects in the WTI futures market. Padungsaksawasdi and Daigler (2013) also conclude that although behavioral theories proved weak explanations for return-volatility relations in commodity markets they make more sense than the leverage hypothesis.

At a more fundamental level Kilian (2009) argues that changes in the oil price are primarily driven by the demand side and makes the distinction between precautionary demand and aggregate demand, pointing out their different effects on the US economy. Precautionary demand with negative effects, and aggregate demand with positive effects. The author concludes that the price build up since late 1990s is mainly the result of aggregate demand while during instances of political instability prices are driven by precautionary demand.

We do not have strong expectations about the sign of the leverage effect. However, we expect that positive and negative returns might have different impact and we allow for this in our estimation.

2.7. Bid-ask spread

The bid-ask spread (BAS) has been found to have a positive correlation to price volatility (e.g. Bollerslev and Melvin (1994); Roll (1984) and Wang and Yau (2000)). The BAS is commonly divided into three components which it must cover. The processing cost of orders, the cost for market makers of holding the futures (i.e. the cost of hedging) and the cost of adverse information. The cost of processing orders through the Globex electronic trading platform is minimal, and the instruments available to market makers make hedging their portfolio efficient. The variation in cost of hedging will therefore be driven by the cost of variance which is reflected in the OVX. One would therefore suspect some correlation between the BAS and the OVX. Our data set does not contain bid and ask quotes. therefore we use the effective spread (spread calculated from actual trades). Schultz (2000) and Huang and Stoll (1996) showed that the Roll-estimator (Roll, 1984) is an appropriate measure of spreads when applied to intra-daily data for liquid markets. Roll-estimator is based on the recognition that if trades fluctuate between spreads, returns will be negatively autocorrelated. The bid-ask spread is calculaed as follows:

$$BAS = 2\sqrt{\frac{-\sum_{t=1}^{t} \Delta P_t \Delta P_{t-1}}{T-1}}, \tag{2}$$

where ΔP is the price difference between two consecutive trades and T is the number of trading pairs during the day.

2.8. Futures curve slope

According to Litzenberger and Rabinowitz (1995) the crude oil market is expected to exhibit backwardation. The cause is that owning an extractive resource is equal to owning a call option with a pay-off equal to the spot price and strike price equal to the extraction cost. The producer will therefore evaluate the price of this option against having oil out of the ground. Without backwardation, this option would not be exercised (hence, no production), just as an option on a stock without dividend would not be exercised before expiration. In times of high volatility this option becomes more valuable which in turn requires stronger backwardation for the option to be exercised. The volatility should therefore be positively correlated with the degree of backwardation (Litzenberger and Rabinowitz, 1995).

Kogan et al. (2009) expand this theory based on a production economy framework. By observing that since capital investments for oil production are irreversible, the supply of oil would also be inelastic and futures prices volatile. Since spot prices are affected by the degree of optimality of the production capital stock, the absolute value of the slope will be larger when there is a large deviation from this optimality. This leads to a V-shaped relationship between volatility and the shape of the futures curve. In other words, an increasing degree of contango will also be positively correlated to volatility. The measure of the slope of the futures curve used by Kogan et al. (2009) is the following one:

$$SL_t = \ln\Big(\frac{P_{t,6}}{P_{t,3}}\Big), \tag{3}$$

where $P_{t,6}$ is the latest price tick observable at day t for the 6th position and $P_{t,3}$ for the 3rd position. In order to allow for the V-shaped relationship to be captured we will split the SL_t variable into SL_t^+ and SL_t^- after demeaning it:

$$SL_t^+ = max(SL_t, 0) \tag{4}$$

$$SL_t^- = min(SL_t, 0) \tag{5}$$

2.9. Data summary

Figure 7 displays selected variables[5] and the closing price of the front month contract from May 16 2007 to May 15 2012. It is apparent from this figure that this period was highly turbulent, as was the case in most markets during this period. In 2008 the price of oil reached record levels, only to be followed by a unprecedented decline and then a long recovery. These price fluctuations reveals an incoherent RV-price relation. At the beginning of the period they appear positively correlated but during the decline in price, volatility rises even more steeply. Then, RV seems to be negatively correlated to the price, but in a decreasing matter.

Comparing the graphs of the implied volatility measured by the OVX and the RV based on 5 minute intervals, it is obvious that they are highly correlated. Avereage implied volatility of 0.41 is higher than the average RV of 0.36. The difference in values could represent what Chernov (2007) described as the risk premium on variance. The RV measure seems to be noisier than the implied volatility measure. This should be expected since the OVX measure is the 30-day expected volatility while the RV is the instantaneous daily volatility.

From the panel showing the slope of the futures curve it is apparent that over the period, the market was mainly in contango. Since the values have been demeaned, this graph has been shifted down and the dotted line shows the original zero-line. The variable appear to be well correlated to the RV and IV measures. It also indicates a positive correlation between the value of the slope and volatility, as was suggested by Kogan et al. (2009). During and after the large price decline in the second part

[5]For descriptive statistics see Table 12 in Appendix.

of 2009 the steep slope of the futures curve would indicate that there was a large discrepancy between the actual capital stock and the optimal capital stock after the large decline in price.

The bid-ask spread appears to be correlated to the volatility variables from the beginning of the period until about the end of the first quarter of 2009 which can be explained by the high OVX levels, as an indication of market makers' cost of hedging. At that point the bid-ask spread settles at a low and narrow band. It is likely that this change is caused by a structural change in the market. For instance the entry of a more sophisticated market maker.

Another possible explanation is based on the observation that the reduction in bid-ask spreads happens at the beginning of the price recovery and when world markets displayed less turmoil. At this point it could be that market participants were less uncertain about the direction of the market and that the role of adverse information was reduced. At the end of the period is seems that there is some increase again in the BAS. At that point in time prices had again reached historically high levels.

3. Results

This chapter first explains the HAR-RV model of (Corsi, 2009). Then, the model is expanded by adding implied volatility and other explanatory variables. The resulting models are fitted to the sample values of daily, weekly and monthly RV using OLS regression. Each regression is performed for both the levels of RV and the first difference of RV. The models are then evaluated for out-of-sample predictions, again

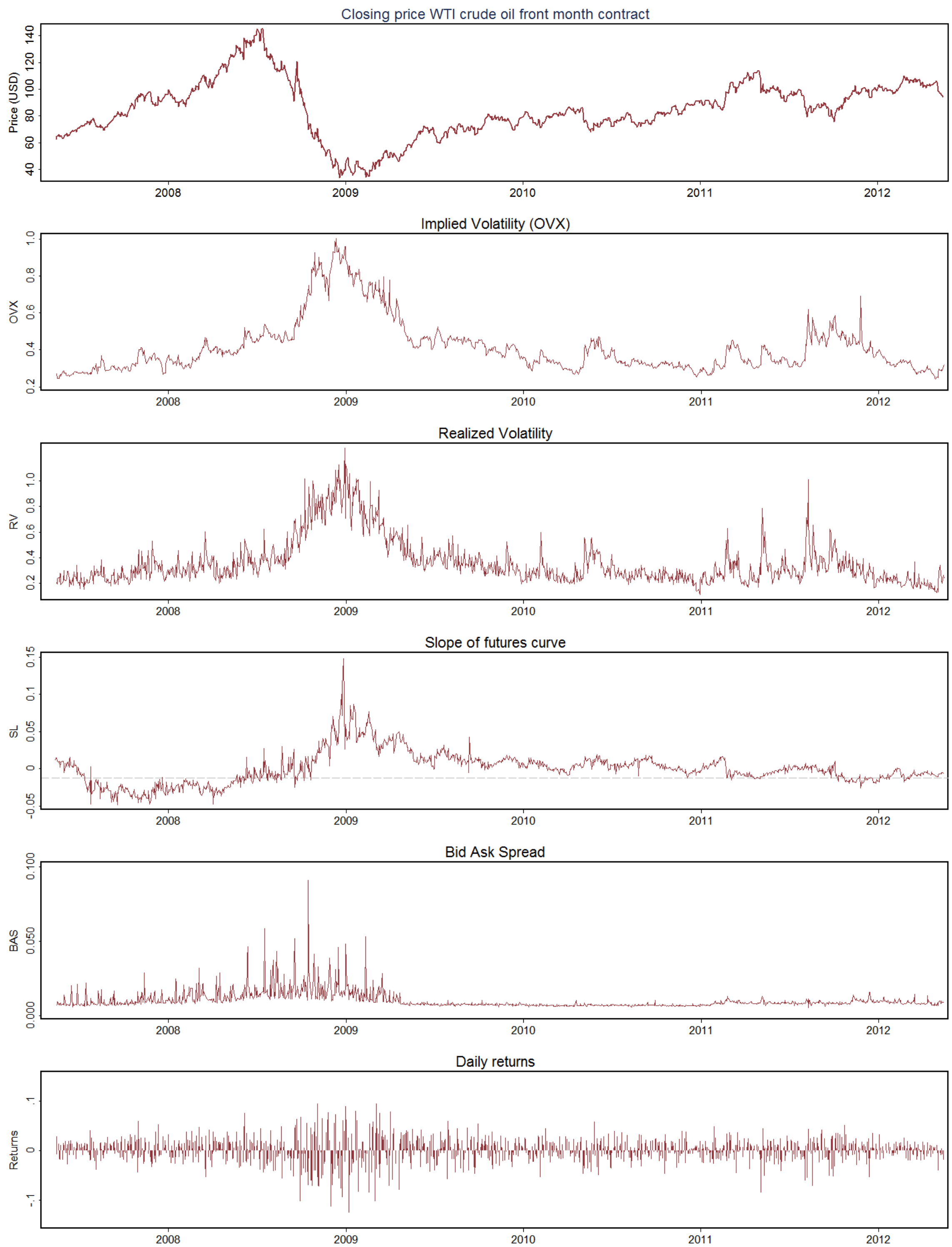


Figure 7: *Time series of the various variables for the period May 16 2007 to May 15 2012*

for the three different time-horizons and for both levels and first differences.

3.1. Model framework

The heterogeneous market hypothesis, postulated by Müller et al. (1997), claims that the asymmetric behavior of volatility is based on trader's different time-horizons. In brief, the short term trader will be influenced by both short term and long term volatility while a long term trader is not easily influenced by short term volatility. This gives rise to HAR-RV model, which is an approximate long-memory *cascading* model of realized volatility (Corsi, 2009). According to Andersen et al. (2007) the model has shown remarkably good forecasting performance comparable to the much more complicated long-memory ARFIMA model.

Realized volatilities over different time-horizons are defined as a simple average of the daily quantities. Weekly realized volatility (with 5 trading days per week) is defined as:

$$RV_t^{(w)} = \frac{1}{5}\Big(RV_t^{(d)} + RV_{t-1d}^{(d)} + \cdots + RV_{t-4d}^{(d)}\Big), \tag{6}$$

where RV is the realized volatility measure defined by equation 1. Monthly volatility is defined analogously. The partial volatility process at each level of the cascade is assumed to be a function of past realized volatility at the same timescale and the expectation of the next period values of the longer term partial volatilities (except for the monthly timescale which only has the AR(1) structure). With latent *partial* volatility defined as $\tilde{\sigma}_t^{(\cdot)}$ the model is shown below:

$$
\begin{aligned}
\tilde{\sigma}^{(m)}_{t+1m} =& c^{(m)} + \phi RV^{(m)}_t + \tilde{\omega}^{(m)}_{t+1m},\\
\tilde{\sigma}^{(w)}_{t+1w} =& c^{(w)} + \phi RV^{(w)}_t + \gamma^{(w)}\mathbb{E}_t[\tilde{\sigma}^{(m)}_{t+1m}] + \tilde{\omega}^{(w)}_{t+1w},\\
\tilde{\sigma}^{(d)}_{t+1d} =& c^{(d)} + \phi RV^{(d)}_t + \gamma^{(d)}\mathbb{E}_t[\tilde{\sigma}^{(w)}_{t+1w}] + \tilde{\omega}^{(d)}_{t+1d},
\end{aligned}
$$

where $RV^{(d)}_t, RV^{(w)}_t$ and $RV^{(m)}_t$ are the daily, weekly and monthly volatilities, respectively, as defined in equation 6.

By recursive substitution of the partial volatilities and setting $\tilde{\sigma}^{(d)}_t = \sigma^{(d)}_t$, the model can be written as follows:

$$
\sigma^{(d)}_{t+1d} = c + \beta^d RV^d_t + \beta^w RV^w_t + \beta_m RV^m_t + \tilde{\omega}^{(d)}_{t+1d} \tag{7}
$$

From this process of latent volatility the time series model of realized volatility becomes:

$$
RV^{(d)}_{t+1} = c + \beta^d RV^d_t + \beta^w RV^w_t + \beta_m RV^m_t + \omega_{t+1d}, \tag{8}
$$

where $\omega_{t+1d} = \tilde{\omega}^{(d)}_{t+1d} - \omega^{(d)}_{t+1d}$. Equation 8 is a a three factor stochastic volatility model labeled HAR(3)-RV with a simple autoregressive structure enabling treatment of volatilities realized over different intervals. A benefit of having such a simple model is that it can easily be extended by adding additional regressors.

3.2. Implementation

The HAR-RV model is extended by adding additional regressors, similarly in principle to what was done by Haugom et al. (2011) for the electricity market. Two main extensions are made; one using the IV measure, and one using the exogenous variables discussed in section 2. The specifications are shown in the following equations:

HAR-RV-IV:

$$RV_{t+1} = \beta_0 + \beta_1 RV_t + \beta_2 RV_t^w + \beta_3 RV_t^m + \beta_4 IV_t + \epsilon_{t+1} \tag{9}$$

HAR-RV-EX:

$$\begin{aligned} RV_{t+1} = &\beta_0 + \beta_1 RV_t + \beta_2 RV_t^w + \beta_3 RV_t^m + \beta_4 SIZE_t + \beta_5 NTR_t + \\ &\beta_6 OI_t + \beta_7 RTN_t^+ + \beta_8 RTN_t^- + \beta_9 BAS_t + \\ &\beta_{10} SL_t^+ + \beta_{11} SL_t^- + \epsilon_{t+1} \end{aligned} \tag{10}$$

HAR-RV-IV-EX:

$$\begin{aligned} RV_{t+1} = &\beta_0 + \beta_1 RV_t + \beta_2 RV_t^w + \beta_3 RV_t^m + \beta_4 IV_t + \beta_5 SIZE_t + \beta_6 NTR_t + \\ &\beta_7 OI_t + \beta_8 RTN_t^+ + \beta_9 RTN_t^- + \beta_{10} BAS_t + \\ &\beta_{11} SL_t^+ + \beta_{12} SL_t^- + \epsilon_{t+1} \end{aligned} \tag{11}$$

IV_t is the implied volatility measured by the OVX, $SIZE_t$ is the daily average size of trades, NTR_t is the average number of contracts traded during day t adjusted for monthly cycle, OI_t is in number of open interests at day t adjusted for monthly cycle, RTN_t^+ defined as $max(RTN_t, 0)$, and RTN_t^- as $min(RTN_t, 0)$ with RTN being the percentage change in price from market close at $t-1$ to market close at t, BAS_t is the bid-ask spread during day t estimated using the Roll estimator, SL_t^+ and SL_t^- represents the slope of the futures curve as specified by equations (4) and (5).

Additionally, the models are used to predict weekly and monthly RV. When doing so the definition in equation 6 is used. In other words the models are used to predict next week's average volatility and next month's average volatility.

As was mentioned in section 2.2 the time series of realized volatility will partially appear like an integrated process. All calculations done for the level of RV are therefore also done for the percentage changes of RV. For the daily horizon we simply use the percentage change in RV from t to t+1. For the weekly horizon it means the percentage change from the last week's average volatility written $RV_t^{(w)}$ and the average volatility for the next week written $RV_{t+6}^{(w)}$. When having change over a monthly horizon on the left hand side of the equation it means the percentage change between last month's average RV, written $RV_t^{(m)}$, and next month's average RV, written $RV_{t+21}^{(m)}$.

3.3. In-sample modeling

The models described in the previous section, together with the original HAR-RV model from equation 8, were estimated for the levels of RV and the percentage changes of RV. This was done for three different horizons, next day, next week and next month. The results are shown in Table 2, Table 3 and Table 4.

Table 2 shows the results for the day-ahead models. Firstly, it is clear that all three coefficients in the HAR-RV (1) model are highly significant. More weight is put on the variables with a shorter time horizon and it appears that one-day volatility has the strongest influence on the next day's volatility level in the WTI futures market. This would be expected according to the model's underlying assumption that short term traders are mainly concerned about short term volatility.

When adding the IV variable in model (2), it becomes the main explanatory variable both in terms of statistical significance and in coefficient size. A substantial reduction is observed in all three RV-coefficients, but by far the largest reduction is seen for the monthly measure. This variable goes from being highly statistically significant to losing all explanatory power. Since the OVX measures the market expectation of 30-day volatility the long horizon part of the RV measures should be embedded. In other words it is a representation of what the long term traders think about volatility.

Adding the exogenous (EX) variables in model (3) moderately decreases the estimated coefficient of the short term component. The information content of the EX variables is therefore mainly overlapping the information in the daily RV measure. One can also observe an increase in the R-squared values for both models (2) and (3) which indicates that adding the variables improves the original model.

Model (3) shows that there are variations in the contributions of the different EX-variables. The two return-variables show the highest statistical significance of the EX variables. The two variables show a substantial difference in the size of their coefficients with 1.8 times larger effect when the return is negative than positive. This means that large returns (both positive and negative) increase volatility, but the increase is larger for negative returns, indicating a leverage effect.

The BAS measure also has a significant impact on volatility. However, as we can see from figure 7 there seems to be a structural change in the BAS from May 1 2009. We therefore split the sample into two sub-samples.[6] We find that the coefficient is significant in the first sub-sample, but not in the second. This is not surprising since the behavior of the variable changed substantially at that time. This supports our assumption that the market structure changed at May 1 2009. One possible explanation would be the entry of a new market maker.

When combining all the variables in model (4) the R-squared value indicates that *both* the IV *and* the EX measures contain additional information to that provided by the RV measures. By comparing the R-squared value for (2) and (3) it seems evident that the EX variables contain more information than the IV about next day volatility. Additionally, the increase in R-squared values induced by the IV variable is similar regardless of whether the model contains EX variables or not, implying that the IV variable contains separate information from the EX variables.

In general, the same effects from adding the IV and EX variables are seen in the RV coefficients when the models are fitted to the percentage change of RV. Slightly different effects from adding the IV and EX variables can be inferred from the R-

[6]Tables equivalent to Table 2, Table 3 and Table 4, just estimated for two sub-samples, are not included in the paper, but are available upon request

squared values. They suggest that the information content of the IV variable is larger when predicting change than levels. The same is seen for the EX variables when looking at the R-squared values in model (7).

The slope of the futures curve is unable to explain changes in volatility in model (7) and (8), but is a highly significant explanatory variable for levels of volatility seen in models (3) and (4).

Fitting the same four models to average RV over the next five days is presented in table 3. It is clear from (1) that the weight of the coefficients and the statistical significance is shifted towards the weekly measure of past volatility compared to when the model was fitted to daily values. When adding the IV variable, seen in model (2) the reduction of the RV coefficients is largest for the monthly measure but also substantial for the weekly and daily measure. In contrary to the daily horizon the effect of the monthly RV variable remains highly significant. It should also be noted that the IV measure is again the most important explanatory variable with the highest coefficient and significance.

In the case of estimating the model for weekly averages of RV, adding the EX variables still reduces all the RV coefficient but to a lesser extent than the IV variable. From the R-squared values it is seen that the increase from (1) to either (2) or (3) is almost equal. This could indicate that the IV variable performs better relative to the EX variables when it comes to next week average levels of RV. The significance of the effects from the EX variables are largely similar to what was found for the daily horizon, with returns, BAS and slope of futures curve having statistically significant coefficients.

When fitting the models to the first difference of weekly RV, model (5) in table 3 shows that the weight has been shifted towards the weekly measure of RV. Adding the IV measure induces the same effect as seen when fitted to RV values and the magnitude of all the RV-coefficients are reduced. Further, the effects from the EX variables, particularly in the OI variables, are slightly stronger in a statistical sense.

The in-sample fittings of the models to monthly RV measures are presented in table 4. For the monthly time-horizon the effect of the long term component of the time- series is the one with the highest significance and highest coefficient. This is evident for both levels and the first difference of the monthly RV.

Adding the IV variables, as seen in model (2), again reduces all coefficients and it becomes the main explanatory variable. The biggest reduction is now seen in the weekly RV measure but its statistical significance as well as the size of the coefficient is lower than when fitting the model to daily and weekly measures of RV.

The effect of adding the EX variables, as seen in model (3) of table 4, reduces the RV coefficients to a less extent than for the daily and weekly horizon. The negative returns and the BAS variable are still highly significant. Additionally, the effect from the number of trades becomes significant. Of the return variables only the negative returns now have any explanatory power. In addition, the number of trades has a significant impact when the models are estimated for the monthly measure of volatility.

The model combining both IV and EX variables (4) displays a higher R-squared value than models (1), (2) and (3). But the increase from adding the IV variable, seen from (3) to (4), is much less than what was observed for the daily and weekly horizon. This could indicate that when it comes to a monthly horizon the IV variable has less additional information to that provided by the HAR-RV-EX model.

For the first difference of monthly volatility, as seen in models (5)–(8) in table 4, the effect from IV is weaker when compared to the shorter horizons. Additionally, the effects from the EX variables are all, with the exception of open interest, stronger impacts on the dependent variable. This is again an indication that the IV variable performs better for shorter horizons.

Table 2: Daily horizon: *In-sample fitting of models for next day levels of RV, and the change in average RV from t to t+1, during May 16 2007 to May 15 2012. Models (2)–(4) and (6)–(8) are extensions of the model proposed by Corsi (2009). IV represents the implied volatility measure. EX represents exogenous variables including;* $SIZE_t$ *as daily average size of trades,* NTR_t *as average number of contracts traded per trade,* OI_t *as open interests,* RTN_t^+ *as price return for* $max(RTN_t, 0)$*, and* RTN_t^- *for* $min(RTN_t, 0)$*,* BAS_t *as measure of average bid-ask spread,* SL_t^+ *as slope of the futures curve if positive and* SL_t^- *the slope if negative.*

	$RV_{t+1}^{(d)}$				$\ln\left(\frac{RV_{t+1}^{(d)}}{RV_t^{(d)}}\right)$			
	HAR-RV (1)	HAR-RV-IV (2)	HAR-RV-EX (3)	HAR-RV-IV-EX (4)	HAR-RV (5)	HAR-RV-IV (6)	HAR-RV-EX (7)	HAR-RV-IV-EX (8)
$RV_t^{(d)}$	0.378***	0.311***	0.290***	0.223***	-1.374***	-1.518***	-1.494***	-1.646***
	(11.30)	(9.36)	(8.16)	(6.31)	(-16.13)	(-17.76)	(-16.02)	(-17.65)
$RV_t^{(w)}$	0.355***	0.227***	0.353***	0.229***	0.762***	0.485***	0.727***	0.446**
	(6.51)	(4.15)	(6.70)	(4.33)	(5.49)	(3.44)	(5.26)	(3.18)
$RV_t^{(m)}$	0.242***	0.037	0.208***	-0.007	0.551***	0.108	0.521***	0.033
	(5.54)	(0.78)	(4.39)	(-0.14)	(4.95)	(0.87)	(4.19)	(0.24)
OVX_t		0.464***		0.491***		1.007***		1.112***
		(8.98)		(8.92)		(7.57)		(7.64))
$SIZE_t$			0.013	0.027***			0.015	0.047**
			(1.93)	(4.06)			(0.83)	(2.64)
NTR_t			0.002	0.002*			0.000	0.001
			(1.87)	(2.16)			(0.17)	(0.37)
OI_t			-0.000	0.008			0.008	0.026
			(-0.01)	(1.44)			(0.51)	(1.75)
RTN_t^+			0.900***	0.765***			1.890***	1.583***
			(5.36)	(4.68)			(4.29)	(3.66)
RTN_t^-			-1.519***	-1.191***			-2.930***	-2.185***
			(-9.49)	(-7.47)			(-6.98)	(-5.18)
BAS_t			1.329**	1.209**			1.984	1.711
			(3.22)	(3.02)			(1.83)	(1.61)
SL_t^-			-0.007	0.115			0.277	0.553
			(-0.03)	(0.56)			(0.50)	(1.02)
SL_t^+			0.706**	0.854***			0.570	0.906
			(2.98)	(3.71)			(0.92)	(1.49)
β_0	0.009	-0.039***	-0.028	-0.137***	0.022	-0.083***	-0.035	-0.280***
	(1.84)	(-5.31)	(-1.27)	(-5.48)	(1.68)	(-4.40)	(-0.59)	(-4.25)
R^2	0.827	0.837	0.844	0.854	0.195	0.230	0.235	0.269

* p<0.05, ** p<0.01, *** p<0.001

Table 3: Weekly horizon: *In-sample fitting of models for next 5 day average levels of RV, and the percentage change in RV from recent 5 days to next 5 days, during May 16 2007 to May 15 2012. Models (2)–(4) and (6)–(8) are extensions of the model proposed by Corsi (2009). IV represents the implied volatility measure. EX represents exogenous variables including;* $SIZE_t$ *as daily average size of trades,* NTR_t *as average number of contracts traded per trade,* OI_t *as open interests,* RTN_t^+ *as price return for* $max(RTN_t, 0)$*, and* RTN_t^- *for* $min(RTN_t, 0)$*,* BAS_t *as measure of average bid-ask spread,* SL_t^+ *as slope of the futures curve if positive and* SL_t^- *the slope if negative.*

	$RV_{t+6}^{(w)}$				$\ln\left(\frac{RV_{t+6}^{(w)}}{RV_t^{(w)}}\right)$			
	HAR-RV (1)	HAR-RV-IV (2)	HAR-RV-EX (3)	HAR-RV-IV-EX (4)	HAR-RV (5)	HAR-RV-IV (6)	HAR-RV-EX (7)	HAR-RV-IV-EX (8)
$RV_t^{(d)}$	0.242***	0.188***	0.208***	0.150***	0.489***	0.378***	0.442***	0.315***
	(9.04)	(7.07)	(7.23)	(5.26)	(6.29)	(4.81)	(5.20)	(3.68)
$RV_t^{(w)}$	0.387***	0.283***	0.371***	0.263***	-1.473***	-1.685***	-1.522***	-1.756***
	(8.86)	(6.46)	(8.68)	(6.15)	(-11.62)	(-13.03)	(-12.09)	(-13.71)
$RV_t^{(m)}$	0.325***	0.159***	0.274***	0.086*	0.855***	0.517***	0.718***	0.313*
	(9.30)	(4.14)	(7.11)	(2.05)	(8.42)	(4.56)	(6.32)	(2.49)
OVX_t		0.378***		0.428***		0.768***		0.926***
		(9.13)		(9.64)		(6.28)		(6.96)
$SIZE_t$			0.002	0.015**			0.001	0.028
			(0.44)	(2.73)			(0.08)	(1.73)
NTR_t			-0.001	-0.001			-0.004	-0.003
			(-1.01)	(-0.80)			(-1.63)	(-1.46)
OI_t			0.006	0.013**			0.027*	0.043**
			(1.22)	(2.82)			(1.98)	(3.12)
RTN_t^+			0.470***	0.352**			0.807*	0.551
			(3.45)	(2.66)			(2.01)	(1.39)
RTN_t^-			-1.033***	-0.746***			-2.183***	-1.562***
			(-7.95)	(-5.79)			(-5.70)	(-4.05)
BAS_t			1.300***	1.195***			3.212**	2.982**
			(3.88)	(3.69)			(3.25)	(3.07)
SL_t^-			0.135	0.241			0.249	0.473
			(0.79)	(1.46)			(0.49)	(0.95)
SL_t^+			0.604**	0.733***			1.514**	1.789**
			(3.14)	(3.95)			(2.67)	(3.21)
β_0	0.016***	-0.023***	0.006	-0.088***	0.047***	-0.034	0.005	-0.200***
	(3.97)	(-3.93)	(0.34)	(-4.39)	(3.89)	(-1.94)	(0.10)	(-3.30)
R^2	0.875	0.883	0.885	0.893	0.112	0.139	0.158	0.190

* p<0.05, ** p<0.01, *** p<0.001

Table 4: Monthly horizon: *In-sample fitting of models for next 20 day average levels of RV, and the change in RV from recent 20 days to next 20 days, during May 16 2007 to May 15 2012. Model (2)–(4) and (6)–(8) are extensions of the model proposed by Corsi (2009). IV represents the implied volatility measure. EX represents exogenous variables including;* $SIZE_t$ *as daily average size of trades,* NTR_t *as average number of contracts traded per trade,* OI_t *as open interests,* RTN_t^+ *as price return for* $max(RTN_t, 0)$*, and* RTN_t^- *for* $min(RTN_t, 0)$*,* BAS_t *as measure of average bid-ask spread,* SL_t^+ *as slope of the futures curve if positive and* SL_t^- *the slope if negative.*

	$RV_{t+21}^{(m)}$				$\ln\left(\frac{RV_{t+21}^{(m)}}{RV_t^{(m)}}\right)$			
	HAR-RV (1)	HAR-RV-IV (2)	HAR-RV-EX (3)	HAR-RV-IV-EX (4)	HAR-RV (5)	HAR-RV-IV (6)	HAR-RV-EX (7)	HAR-RV-IV-EX (8)
$RV_t^{(d)}$	0.173***	0.131***	0.182***	0.136***	0.414***	0.366***	0.438***	0.357***
	(6.07)	(4.54)	(5.95)	(4.45)	(5.18)	(4.47)	(5.14)	(4.12)
$RV_t^{(w)}$	0.363***	0.281***	0.308***	0.224***	0.750***	0.657***	0.653***	0.503***
	(7.79)	(5.92)	(6.80)	(4.87)	(5.76)	(4.88)	(5.17)	(3.87)
$RV_t^{(m)}$	0.373***	0.243***	0.345***	0.201***	-1.397***	-1.543***	-1.627***	-1.883***
	(9.99)	(5.82)	(8.48)	(4.44)	(-13.38)	(-13.06)	(-14.25)	(-14.76)
OVX_t		0.296***		0.330***		0.334**		0.591***
		(6.60)		(6.90)		(2.61)		(4.37)
$SIZE_t$			-0.007	0.003			0.003	0.020
			(-1.17)	(0.49)			(0.21)	(1.20)
NTR_t			-0.004***	-0.004***			-0.010***	-0.009***
			(-5.30)	(-5.23)			(-4.43)	(-4.33)
OI_t			0.008	0.013**			0.062***	0.072***
			(1.54)	(2.67)			(4.44)	(5.14)
RTN_t^+			0.190	0.099			0.211	0.047
			(1.32)	(0.70)			(0.52)	(0.12)
RTN_t^-			-0.807***	-0.586***			-1.476***	-1.078**
			(-5.87)	(-4.23)			(-3.84)	(-2.75)
BAS_t			2.445***	2.364***			6.390***	6.237***
			(6.89)	(6.78)			(6.45)	(6.33)
SL_t^-			0.447*	0.529**			1.169*	1.303*
			(2.46)	(2.96)			(2.29)	(2.57)
SL_t^+			-0.035	0.065			2.119***	2.285***
			(-0.17)	(0.33)			(3.71)	(4.03)
β_0	0.033***	0.002	0.054**	-0.019	0.084***	0.048**	-0.008	-0.139*
	(7.47)	(0.30)	(2.80)	(-0.86)	(6.77)	(2.64)	(-0.14)	(-2.27)
R^2	0.846	0.852	0.861	0.866	0.158	0.162	0.242	0.254

* $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table 5: *Mean squared errors of predictions using rolling windows of different sizes. * indicates the best performing window.*

	RV_{t+1}		$\ln(RV_{t+1}/RV_t)$	
	HAR-RV	HAR-RV-IV-EX	HAR-RV	HAR-RV-IV-EX
w=125	0.00608	0.00638	0.0363	0.0408
w=250	0.00610	0.00556	0.0353	0.0333
w=500*	0.00430	0.00364	0.0362	0.0332
w=750	0.00494	0.00400	0.0382	0.0338

3.4. Out-of-sample forecasting

Out-of-sample predictions are used to make direct comparisons of the performance of the models relative to the actual values of the time series.

The choice of estimation-window for out-of-sample predictions will affect how the models are performing. On the one hand a larger window will make the model estimates more reliable as the underlying data sample increases and on the other hand a large window make the model less adaptable to changes in the market. Table 5 displays the mean squared error (MSE) for predictions made by two of the model specifications when using four different rolling windows. The numbers indicate that a two year window is preferable. Based on this crude comparison, a two year window rolling window is chosen for subsequent out-of-sample predictions.
In order to evaluate the relative performance of volatility models Mincer and Zarnowitz (1969) suggested running a regression as described by the following equation:

$$RV_{t+1} = \beta_0 + \beta_1 \hat{v}_{Model1,t} + \beta_2 \hat{v}_{Model2,t} + \epsilon_{t+1} \tag{12}$$

where RV_{t+1} is the observed realized volatility at $t+1$, $\hat{v}_{Model1,t}$ is the forecast obtained

Table 6: Mincer and Zarnowitz (1969) regressions for comparison of 1-day ahead forecasts of volatility from various models. Each model was estimated using a 500-day rolling window.

	MZW1	MZW2	MZW3	MZW4	MZW5	MZW6	MZW7
HAR-RV-IV-EX	0.927				0.979	1.059	0.970
	(0.03)				(0.08)	(0.11)	(0.12)
HAR-RV		0.949			-0.060		
		(0.03)			(0.09)		
HAR-RV-IV			0.895			-0.137	
			(0.03)			(0.11)	
HAR-RV-EX				0.968			-0.047
				(0.03)			(0.13)
constant	0.024	0.014	0.029	0.012	0.027	0.027	0.026
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
R^2	0.607	0.533	0.556	0.574	0.607	0.608	0.607

from one of the models and $\hat{v}_{Model2,t}$ is the forecast from a second model. The main benefit with this procedure is that it will directly give an indication of the differences between the models. The method is also applied to evaluations of weekly and monthly volatility predictions by exchanging RV_{t+1} with $RV_{t+6}^{(w)}$ and $RV_{t+21}^{(m)}$ respectively. The results from these comparisons are displayed in table 9, 10, and 11, respectively.[7]

The regressions in table 9 show that when predicting RV one day ahead, adding both IV and EX variables significantly improves forecasting. When comparing the HAR-RV-IV model and the HAR-RV-EX the regression is inconclusive. This shows that the performance of the two models is comparable in making predictions. The coefficients indicate that the IV is slightly better at predicting the level of volatility, while the EX variables are slightly better at predicting the one-day change in volatility. When comparing the combined model to the other models the tests all show that the

[7] Additional comparison statistics such as the mean square error (MSE) and the mean absolute error (MAE) for RV_{t+1} and $\ln(\frac{RV_{t+1}}{RV_t})$ predictions can be found in the Appendix.

Table 7: Mincer and Zarnowitz (1969) regressions for comparison of 1-week ahead forecasts of volatility from various models. Each model was estimated using a 500-day rolling window.

	MZW1	MZW2	MZW3	MZW4	MZW5	MZW6	MZW7
HAR-RV-IV-EX	0.794				1.011	1.083	0.976
	(0.03)				(0.10)	(0.12)	(0.17)
HAR-RV		0.816			-0.253		
		(0.04)			(0.11)		
HAR-RV-IV			0.778			-0.311	
			(0.03)			(0.13)	
HAR-RV-EX				0.810			-0.194
				(0.03)			(0.17)
constant	0.065	0.053	0.064	0.060	0.078	0.074	0.069
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
R^2	0.465	0.399	0.412	0.441	0.468	0.469	0.466

Table 8: Mincer and Zarnowitz (1969) regressions for comparison of 1-month ahead forecasts of volatility from various models. Each model was estimated using a 500-day rolling window.

	MZM1	MZM2	MZM3	MZM4	MZM5	MZM6	MZM7
HAR-RV-IV-EX	0.672				0.807	0.834	0.906
	(0.03)				(0.08)	(0.09)	(0.23)
HAR-RV		0.650			-0.161		
		(0.04)			(0.08)		
HAR-RV-IV			0.631			-0.180	
			(0.03)			(0.09)	
HAR-RV-EX				0.675			-0.242
				(0.03)			(0.23)
constant	0.103	0.101	0.108	0.102	0.113	0.111	0.105
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
R^2	0.397	0.312	0.325	0.385	0.400	0.400	0.398

Table 9: Mincer and Zarnowitz (1969) regressions for comparison of 1-day ahead forecasts of percentage change in volatility from various models. Each model was estimated using a 500-day rolling window.

	MZDR1	MZDR2	MZDR3	MZDR4	MZDR5	MZDR6	MZDR7
HAR-RV-IV-EX	0.884				0.845	0.879	0.859
	(0.06)				(0.10)	(0.14)	(0.14)
HAR-RV		0.868			0.059		
		(0.08)			(0.12)		
HAR-RV-IV			0.839			0.006	
			(0.07)			(0.15)	
HAR-RV-EX				0.836			0.030
				(0.07)			(0.14)
constant	0.005	-0.003	-0.009	0.005	0.004	0.005	0.005
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
R^2	0.212	0.141	0.171	0.170	0.213	0.212	0.212

Table 10: Mincer and Zarnowitz (1969) regressions for comparison of 1-week ahead forecasts of percentage change in volatility from various models. Each model was estimated using a 500-day rolling window.

	MZWR1	MZWR2	MZWR3	MZWR4	MZWR5	MZWR6	MZWR7
HAR-RV-IV-EX	1.023				0.885	1.009	0.622
	(0.08)				(0.15)	(0.16)	(0.29)
HAR-RV		0.934			0.171		
		(0.09)			(0.15)		
HAR-RV-IV			0.897			0.016	
			(0.08)			(0.17)	
HAR-RV-EX				1.019			0.418
				(0.08)			(0.29)
constant	0.013	-0.006	-0.009	0.014	0.010	0.012	0.013
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
R^2	0.177	0.138	0.135	0.174	0.178	0.177	0.179

Table 11: Mincer and Zarnowitz (1969) regressions for comparison of 1-month ahead forecasts of percentage change in volatility from various models. Each model was estimated using a 500-day rolling window.

	MZMR1	MZMR2	MZMR3	MZMR4	MZMR5	MZMR6	MZMR7
HAR-RV-IV-EX	0.813				0.899	0.948	0.105
	(0.06)				(0.09)	(0.10)	(0.43)
HAR-RV		0.646			-0.113		
		(0.06)			(0.10)		
HAR-RV-IV			0.635			-0.173	
			(0.06)			(0.10)	
HAR-RV-EX				0.814			0.711
				(0.06)			(0.43)
constant	0.015	-0.016	-0.014	0.015	0.018	0.020	0.015
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
R^2	0.223	0.127	0.125	0.225	0.224	0.226	0.225

combined model outperforms the other models.

In the comparisons of the predictions of the various models for the one week ahead horizon, a slightly different dynamic is seen. 10 shows that including IV improves the forecasts significantly when focusing on the level of RV, but when making predictions of the change in volatility the IV variable no longer significantly improves the performance of the original HAR-RV model. Adding the EX variables clearly improves the predictions, and there is evidence that the HAR-RV-EX model outperforms the HAR-RV-IV model both when predicting the level and change of RV. When comparing the combined model to the rest of the models, it still performs better than the simpler HAR-RV-EX model indicating that the IV variable contain some information about the future week volatility level. When predicting the change in RV, adding the IV-measure only slightly improves the forecasts. The improvement is significant at the 5 % level.

When comparing the monthly predictions seen in Table 11 a similar pattern to the one in 10 is seen. The EX variables clearly contribute more to the precision of the predictions than the IV variables. When comparing the predictions of the HAR-RV-IV and the HAR-RV-EX it is evident that the latter is performing better for both levels and differences. Nevertheless, the IV variable still leads to an improvement to the HAR-RV-EX model when levels are predicted. When predicting changes the combined model does not perform any better than the simpler HAR-RV-EX model.

4. Conclusion

In this paper we combine realized volatility with implied volatility (IV) and other explanatory market variables (EX) in a forecasting model based on the HAR-RV model described by Corsi (2009). As a proxy for IV the oil volatility index published by the CBOE is used. Market variables added to the model are volume, open interest, daily returns, the bid-ask spread and the slope of the futures curve.

Our results show that the HAR-RV model fits the RV time series significantly better when both the IV and EX variables are added to the model. The effect of adding the IV variable is strongest when the model was fitted to next-day levels and weakest when the model was fitted to next-month percentage changes in RV. The effect of adding the EX variables is stronger for the longer horizons than the short. The daily return variables have highly statistically significant effects for both RV levels and differences for all time-horizons. The bid-ask spread and the slope of the futures curve also have significant effects when modeling realized volatility for WTI futures.

The results for the out-of-sample forecasting shows that the largest improvement is obtained by adding both the IV-measure and the EX variables. Implied volatility is the strongest driver for short-term predictions, whereas other market variables (particularly the bid-ask spread) are most relevant for long-term forecasts.

This work shows that including implied volatility and other market variables improves volatility forecasts for the WTI futures market. An additional finding is that leverage effects (the relationship between past returns and volatility) is very different from the one found in equity markets. It has a V-shape, meaning that large returns (both positive and negative) increase volatility in this market. We suggest the leverage effect for oil, as well as other commodities, to be investigated further.

References

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

Andersen, Torben G (1996), 'Return volatility and trading volume: An information flow interpretation of stochastic volatility', *The Journal of Finance* **51**(1), 169–204.

Andersen, Torben G. and Tim Bollerslev (1998), 'Deutsche mark-dollar volatility: Intraday activity patterns, macroeconomic announcements, and longer run dependencies', *The Journal of Finance* **53**(1), pp. 219–265.

Andersen, Torben G, Tim Bollerslev and Francis X Diebold (2007), 'Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility', *The Review of Economics and Statistics* **89**(4), 701–720.

Andersen, Torben G, Tim Bollerslev, Francis X Diebold and Paul Labys (2001), 'The distribution of realized exchange rate volatility', *Journal of the American statistical association* **96**(453), 42–55.

Bandi, Federico M and Jeffrey R Russell (2006), 'Separating microstructure noise from volatility', *Journal of Financial Economics* **79**(3), 655–692.

Barndorff-Nielsen, Ole, Silja Kinnebrock and Neil Shephard (2008), 'Measuring downside risk-realised semivariance', *CREATES Research Paper* (2008-42).

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

Bollerslev, Tim, Daniela Osterrieder, Natalia Sizova and George Tauchen (2013), 'Risk and return: Long-run relations, fractional cointegration, and return predictability', *Journal of Financial Economics* **108**(2), 409 – 424.

Bollerslev, Tim and Michael Melvin (1994), 'Bid—ask spreads and volatility in the foreign exchange market: An empirical analysis', *Journal of International Economics* **36**(3), 355–372.

Chakravarty, Sugato (2001), 'Stealth-trading: Which traders' trades move stock prices?', *Journal of Financial Economics* **61**(2), 289–307.

Chernov, Mikhail (2007), 'On the role of risk premia in volatility forecasting', *Journal of Business & Economic Statistics* **25**(4), 411–426.

Clark, Peter K (1973), 'A subordinated stochastic process model with finite variance for speculative prices', *Econometrica: Journal of the Econometric Society* pp. 135–155.

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

Day, Theodore E and Craig M Lewis (1993), 'Forecasting futures market volatility', *The Journal of Derivatives* **1**(2), 33–50.

Diebold, Francis X and Robert S Mariano (2002), 'Comparing predictive accuracy', *Journal of Business & economic statistics* **20**(1).

Doran, James S and Ehud I Ronn (2005), 'The bias in Black-Scholes/Black implied volatility: An analysis of equity and energy markets', *Review of Derivatives Research* **8**(3), 177–198.

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

Figlewski, Stephen (1981), 'Futures trading and volatility in the GNMA market', *The Journal of Finance* **36**(2), 445–456.

Gallant, A Ronald, Peter Eric Rossi and George Tauchen (1992), 'Stock prices and volume', *Review of Financial studies* **5**(2), 199–242.

Girma, Paul Berhanu and Mbodja Mougoue (2002), 'An empirical examination of the relation between futures spreads volatility, volume, and open interest', *Journal of Futures Markets* **22**(11), 1083–1102.

Hansen, Peter R and Asger Lunde (2006), 'Realized variance and market microstructure noise', *Journal of Business & Economic Statistics* **24**(2), 127–161.

Haugom, Erik, Sjur Westgaard, Per Bjarte Solibakke and Gudbrand Lien (2011), 'Realized volatility and the influence of market measures on predictability: Analysis of Nord Pool forward electricity data', *Energy Economics* **33**(6), 1206–1215.

Huang, Roger D and Hans R Stoll (1996), 'Dealer versus auction markets: A paired comparison of execution costs on NASDAQ and the NYSE', *Journal of Financial Economics* **41**(3), 313–357.

Jorion, Philippe (1995), 'Predicting volatility in the foreign exchange market', *The Journal of Finance* **50**(2), 507–528.

Kilian, Lutz (2009), 'Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market', *The American Economic Review* **99**(3), 1053–1069.

Kim, Oliver and Robert E Verrecchia (1991), 'Market reaction to anticipated announcements', *Journal of Financial Economics* **30**(2), 273–309.

Kogan, Leonid, Dmitry Livdan and Amir Yaron (2009), 'Oil futures prices in a production economy with investment constraints', *The Journal of Finance* **64**(3), 1345–1375.

Lamoureux, CG and WD Lastrapes (1993), 'Forecasting stock-return variance: toward an understanding of stochastic implied volatilities', *Review of Financial Studies* **6**(2), 293–326.

Litzenberger, Robert H and Nir Rabinowitz (1995), 'Backwardation in oil futures markets: Theory and empirical evidence', *The Journal of Finance* **50**(5), 1517–1545.

Martens, Martin and Jason Zein (2004), 'Predicting financial volatility: High-frequency time-series forecasts vis-à-vis implied volatility', *Journal of Futures Markets* **24**(11), 1005–1028.

Marzo, Massimiliano and Paolo Zagaglia (2010), 'Volatility forecasting for crude oil futures', *Applied Economics Letters* **17**(16), 1587–1599.

Mincer, Jacob A and Victor Zarnowitz (1969), The evaluation of economic forecasts, *in* 'Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance', NBER, pp. 1–46.

Müller, Ulrich A, Michel M Dacorogna, Rakhal D Davé, Richard B Olsen, Olivier V Pictet and Jacob E von Weizsäcker (1997), 'Volatilities of different time resolutions—analyzing the dynamics of market components', *Journal of Empirical Finance* **4**(2), 213–239.

O'Hara, Maureen (1995), *Market microstructure theory*, Vol. 108, Blackwell Cambridge.

Padungsaksawasdi, Chaiyuth and Robert T Daigler (2013), 'The return-implied volatility relation for commodity ETFs', *Journal of Futures Markets* .

Patton, Andrew J (2011), 'Volatility forecast comparison using imperfect volatility proxies', *Journal of Econometrics* **160**(1), 246–256.

Roll, Richard (1984), 'A simple implicit measure of the effective bid-ask spread in an efficient market', *The Journal of Finance* **39**(4), 1127–1139.

Sadorsky, Perry (2006), 'Modeling and forecasting petroleum futures volatility', *Energy Economics* **28**(4), 467–488.

Schultz, Paul (2000), 'Regulatory and legal pressures and the costs of Nasdaq trading', *Review of Financial Studies* **13**(4), 917–957.

Szakmary, Andrew, Evren Ors, Jin Kyoung Kim and Wallace N Davidson III (2003), 'The predictive power of implied volatility: Evidence from 35 futures markets', *Journal of Banking & Finance* **27**(11), 2151–2175.

Wang, George HK and Jot Yau (2000), 'Trading volume, bid–ask spread, and price volatility in futures markets', *Journal of Futures markets* **20**(10), 943–970.

Wang, Tao, Jingtao Wu and Jian Yang (2008), 'Realized volatility and correlation in energy futures markets', *Journal of Futures Markets* **28**(10), 993–1011.

Wei, Yu, Yudong Wang and Dengshi Huang (2010), 'Forecasting crude oil market volatility: Further evidence using GARCH-class models', *Energy Economics* **32**(6), 1477–1484.

Whaley, Robert (2008), 'Understanding VIX', *Available at SSRN 1296743* .

5. Appendix

Table 12: *Descriptive statistics for regression variables. NTR and OI are adjusted for monthly cycle and are divided by 10,000 and 100,000 respectively.*

Variable	Obs	Mean	Std. Dev.	Min	Max
$\ln(\frac{RV_{t+1}}{RV_t})$	1246	0.000	0.208	-0.712	0.757
RV_t	1246	0.360	0.176	0.115	1.256
RV_{t5}	1246	0.360	0.166	0.138	1.018
RV_{t20}	1246	0.360	0.160	0.173	0.960
IV_{OVX}	1246	0.413	0.144	0.243	1.004
NTR	1246	11.098	0.689	9.063	13.39
$SIZE$	1246	2.142	0.441	1.202	3.537
OI	1246	3	0.405	0.479	4.143
RTN^+	1246	0.0090	0.0141	0.0000	0.0947
RTN^-	1246	-0.0090	0.0156	-0.1267	0.0000
RG^*	1246	0.021	0.013	0.005	0.126
BAS	1246	0.0141	0.008	0.007	0.129
SL^-	1246	-0.0076	0.0109	-0.0488	0.0000
SL^+	1246	0.0077	0.0147	0.0000	0.1478

Table 13: Test statistics for out-of sample predictions of next day levels and differences. Mean squared errors (MAE) and mean absolute errors (MAE) compared to acutal values are shown for the four different model specifications. * indicates statistically significant difference in the MSE values of the marked model compared to the MSE values for the basic HAR-RV model according to the Diebold and Mariano (2002) test.

	HAR-RV	HAR-RV-IV	HAR-RV-EX	HAR-RV-IV-EX
RV_{t+1}				
MSE	0.004113	0.00416	0.003914*	0.003461*
MAE	0.0450	0.0451	0.04296	0.044192
$\ln(\frac{RV_{t+1}}{RV_t})$				
MSE	0.03617	0.03511	0.03512	0.0332*
MAE	0.1460	0.145	0.1435	0.1387

* p<0.05, ** p<0.01,

Table 14: Test statistics for out-of sample predictions of next week levels and differences. Mean squared errors (MAE) and mean absolute errors (MAE) compared to acutal values are shown for the four different model specifications. * indicates statistically significant difference in the MSE values of the marked model compared to the MSE values for the basic HAR-RV model according to the Diebold and Mariano (2002) test.

	HAR-RV	HAR-RV-IV	HAR-RV-EX	HAR-RV-IV-EX
RV_{t+1}				
MSE	0.00406	0.004064	0.0038177*	0.003707*
MAE	0.04268	0.04269	0.03987	0.03908
$\ln(\frac{RV_{t+1}}{RV_t})$				
MSE	0.03622	0.0364	0.0348	0.03468
MAE	0.1389	0.1389	0.1337	0.1328

* p<0.05, ** p<0.01,

Table 15: Test statistics for out-of sample predictions of next month levels and differences. Mean squared errors (MAE) and mean absolute errors (MAE) compared to acutal values are shown for the four different model specifications. * indicates statistically significant difference in the MSE values of the marked model compared to the MSE values for the basic HAR-RV model according to the Diebold and Mariano (2002) test.

	HAR-RV	HAR-RV-IV	HAR-RV-EX	HAR-RV-IV-EX
RV_{t+1}				
MSE	0.003521	0.003539	0.003189	0.003163
MAE	0.04227	0.04539	0.04040	0.04020
$\ln(\frac{RV_{t+1}}{RV_t})$				
MSE	0.03667	0.03679	0.03189	0.0320
MAE	0.1542	0.1541	0.1352	0.1352

* p<0.05, ** p<0.01,