Int. Fin. Markets, Inst. and Money 40 (2016) 1–13

Contents lists available at ScienceDirect

Journal of International Financial Markets, Institutions & Money

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

Intraday volatility interaction between the crude oil and equity markets

Dinh Hoang Bach Phan, Susan Sunila Sharma, Paresh Kumar Narayan*

Centre for Economics and Financial Econometrics Research, Deakin University, Australia

ARTICLE INFO

Article history:
Received 29 January 2015
Accepted 27 July 2015
Available online 7 August 2015

Keywords:
Volatility
Trading volume
Bid–ask spread
Cross-market
Predictability
Forecasting

ABSTRACT

This paper investigates the price volatility interaction between the crude oil and equity markets in the US using 5-min data over the period 2009–2012. Our main findings can be summarised as follows. First, we find strong evidence to demonstrate that the integration of the bid–ask spread and trading volume factors leads to a better performance in predicting price volatility. Second, trading information, such as bid–ask spread, trading volume, and the price volatility from cross-markets, improves the price volatility predictability for both in-sample and out-of-sample analyses. Third, the trading strategy based on the predictive regression model that includes trading information from both markets provides significant utility gains to mean-variance investors.

1. Introduction

The relationship between the crude oil and equity markets has been a popular subject in financial economics. While several studies examine the relationship between oil price returns and aggregate stock market returns, there is inconclusive evidence of the role of the oil price on stock returns.[1] Some studies (see, for instance, Jones and Kaul, 1996; Driesprong et al., 2008; Park and Ratti, 2008; Miller and Ratti, 2009) discover a negative effect of crude oil price returns on stock returns; while others (see, for instance, Chen et al., 1986; Huang et al., 1996; Wei, 2003) document no statistically significant effect or positive effect (see Gjerde and Saettem, 1999; Narayan and Narayan, 2010; Narayan and Sharma, 2011; Park and Ratti, 2008). For sectors of the NYSE, a recent study by Narayan and Sharma (2011) finds mixed evidence in that returns of sectors related to oil, such as transport and energy, respond positively to oil price changes, while the rest of the sectors' returns respond negatively. Recently, Phan et al. (2015b) show that the oil price can predict the US stock returns in both in-sample and out-of-sample tests.

Another strand of the literature looks at the cross-market volatility transmission and finds a significant result between the crude oil price and the US stock market[2] (see, inter alia, Agren, 2006; Hammoudeh et al., 2004; Aloui and Jammazi, 2009; Malik and Ewing, 2009; Arouri et al., 2011a; Mensi et al., 2013; Guesmi and Fattoum, 2014; Soucek and Todorova, 2013,

* Corresponding author at: Centre for Economics and Financial Econometrics Research, Deakin University, 221 Burwood Highway, Burwood, Victoria 3125, Australia. Tel.: +61 3 92446180.
E-mail address: paresh.narayan@deakin.edu.au (P.K. Narayan).

[1] See Jones and Kaul (1996), Park and Ratti (2008), Driesprong et al. (2008), Arouri (2011), Arouri et al. (2011a,b), Aloui et al. (2012), Narayan and Sharma (2014) and Phan et al. (2015a).

[2] There are a number of studies that investigate the volatility interaction between the crude oil and equity markets in the GCC countries (Malik and Hammoudeh, 2007; Arouri et al., 2011b; Awartani and Maghyereh, 2013; Jouini and Harrathi, 2014). All these studies find evidence for a significant volatility interaction.

http://dx.doi.org/10.1016/j.intfin.2015.07.007
1042-4431/

2014). Agren (2006) employs an asymmetric BEKK-GARCH model using weekly data on the aggregate markets of the US, UK, Japan, Norway, and Sweden, and discovers significant interaction in all countries except Sweden. Hammoudeh et al. (2004) examine the volatility interaction among five S&P oil sector stock indices and five oil prices from the US market using univariate and multivariate GARCH models. In the multivariate GARCH model, they find bidirectional interactions between the return volatility of the oil spot/futures market and oil sector indices. Malik and Ewing (2009) conclude a significant volatility interaction between oil prices and five US equity sector indices by utilising a bivariate GARCH model. Arouri et al. (2011a) applied a VAR-GARCH model to investigate the extent of volatility interaction between the crude oil and equity markets in Europe and the US at the sector level, and report evidence of a significant volatility interaction. In Europe, the transmission of volatility is uni-directional from oil markets to stock markets, but there is bi-directional volatility transmission in the US market. Recently, Soucek and Todorova (2013, 2014) find a significant spillover effect between the realised volatility of S&P500 and WTI crude oil futures contracts.

The literature, however, is not immune to limitations and we extend the literature in several ways. First, although the significant relationship between trading volume, bid–ask spread, and price volatility has been widely documented by the literature,[3] the bid–ask spread and trading volume have not been incorporated in an empirical model consisting of crude oil and equity market volatility. Our study, therefore, is motivated by this research gap and presents the first empirical analysis that addresses the relative importance of information on trading volume and bid–ask spread in testing cross-market volatility interaction between the crude oil and equity markets. Our contribution is based on an empirical model that consists of 5-min data on three nearby futures contracts, namely, E-mini S&P500 index futures, E-mini NASDAQ index futures, and Light Sweet Crude Oil (WTI) futures for the period 2 January 2009–31 December 2012. These futures contracts are the most actively-traded equity index futures and crude oil futures. We find strong evidence that information from the trading volume and bid–ask spread offers more accurate forecasts of the volatility of both markets.

Second, while most of the previous studies in this field use low-frequency data, such as monthly data (Aloui and Jammazi, 2009), weekly data (Arouri et al., 2011a; Agren, 2006; Malik and Ewing, 2009; Awartani and Maghyereh, 2013), or daily data (Hammoudeh et al., 2004; Malik and Hammoudeh, 2007; Arouri et al., 2011b), we employ intraday data in testing the volatility interaction between the crude oil and equity markets. Crude oil and equity markets are heavily traded and studies based on low-frequency data, such as daily, weekly or monthly data, may fail to capture information contained in intraday price movements. As volatility is a key input for market risk evaluation and derivatives pricing, intraday volatility modelling and forecasting is important to market participants who are involved in intraday trading, such as day traders, high-frequency portfolio managers, and programme traders (Wang and Wang, 2010). In fact, the availability of high-frequency data is considered valuable in measuring, modelling, and forecasting volatility (Barndorff-Nielsen and Shephard, 2007). Hansen and Lunde (2010) state that volatility is highly persistent, so that a more accurate measure of current volatility, which intraday data provide, is valuable for forecasting future volatility. In addition, the economic value of using intraday data in forecasting volatility has been widely evidenced in the literature (Sévi, 2014).[4] Thus, this paper contributes to the literature that examines the cross-market volatility transmission between the crude oil and equity markets by using data at 5-min frequency.

Third, a common feature of previous studies (Agren, 2006; Hammoudeh et al., 2004; Aloui and Jammazi, 2009; Malik and Ewing, 2009; Arouri et al., 2011a) is that they employ a GARCH model with different specifications to examine the volatility interaction between the crude oil price and equity markets. The evidence of the volatility interaction is based simply on the sign and the statistical significance of the other market volatility variable's parameter from the variance equation. Such significant cross-market volatility is well understood and we do not test this, rather, we test whether the cross-market volatility interaction can improve volatility forecasts. The results, thus, are strengthened and more robust in several ways. First, we construct three specifications of the EGARCH(1,1) model with different levels of trading information to predict the price volatility of the crude oil and equity markets.[5] We then compare the in-sample predictive accuracy among models to assess whether the model that includes the cross-market trading information is superior to others in predicting the price volatility of the crude oil and equity markets. Furthermore, we examine the economic significance of cross-market volatility interaction that has not been done previously. All we understand so far is that cross-market volatility is statistically significant but nothing is known about how beneficial the cross-market volatility is for investors. Following Campbell and Thompson (2008), we test whether the out-of-sample volatility forecasting can provide any utility gains to the mean-variance investor

[3] The mixture of distributions hypothesis proposed by Clark (1973) suggests a positive contemporaneous effect of trading volume on price volatility. Furthermore, the sequential arrival of an information hypothesis introduced by Copeland (1976), implies that the forecast ability of volatility can be improved by using the knowledge of lagged trading volume. These hypotheses have been tested in several empirical studies (Morgan, 1976; Westerfield, 1977; Jones et al., 1994; Wang and Yau, 2000; Cornell, 1981; Foster, 1995; Tauchen and Pitts, 1983; Najand and Yung, 1991; Rahman et al., 2002; Darrat et al., 2003; Hussain, 2011). Moreover, the literature shows that the bid–ask spread positively impacts the price volatility (Wang et al., 1994; Wang and Yau, 2000; Rahman et al., 2002; Worthington and Higgs, 2009; Hussain, 2011).

[4] Using intraday data to forecast the volatility is beneficial for the portfolio choice (Fleming et al., 2003; Bandi et al., 2008), risk management activities (Giot and Laurent, 2004; Clements et al., 2008), option pricing (Duan, 1995; Heston and Nandi, 2000; Stentoft, 2008; Corsi et al., 2013), and density forecast amelioration (Geman, 2005; Hua and Zhang, 2008; Wong and Lo, 2009; Maheu and McCurdy, 2011). See Sévi (2014) for more discussion.

[5] Model 1 predicts the volatility of the crude oil or equity market based on its own lagged volatility only, while Model 2 is based on the information on volatility, bid–ask spread and the trading volume in its own-market. Model 3 contains lagged volatility, lagged bid–ask spread and lagged trading volume in its own-market and from the cross-market.

who allocates her portfolio between a risky asset and a risk-free bill. The evidence from in-sample analysis illustrates the integration of the bid–ask spread and trading volume factors from both markets improve the price volatility predicting. In addition, the improvement for the price volatility forecasting is also found in an out-of-sample analysis. Finally, we find that the improvement in price volatility forecasting is economically significant to the investors, as evidenced by the utility gain which, on average, is 12.37% per annum using a trading strategy based on the best forecasting model instead of a buy-and-hold trading strategy.

The remainder of this paper is organised as follows. The methodology and data are discussed in Section 2. Section 3 discusses the main findings, and the final section provides the summarised conclusions of this study.

2. Data and methodology

2.1. Data

The sample for this study is based on five-minute data frequency relating to three specific series, namely, E-mini S&P500 index futures and E-mini NASDAQ index futures that are traded on the Chicago Mercantile Exchange (CME), and Light Sweet Crude Oil (WTI) futures traded on the New York Mercantile Exchange (NYMEX). The trading time for WTI futures is from 6:00 p.m. to 5:15 p.m. of the next day according to New York time, from Sunday to Friday, with a 45-mine break each day beginning at 5:15 p.m. The trading time for E-mini S&P500 index futures and E-mini NASDAQ index futures are from 5:00 p.m. to 4:15 p.m. of the next day from Monday to Friday with a trading halt from 3:15 p.m. to 3:30 p.m. The trading data of the three contracts, including opening price, closing price, high price, low price, bid price, ask price, and trading volume, are downloaded from Thomson Reuter Tick History database for the period 2 January 2009 to 31 December 2012.

The bid–ask spread (*BAS*) is calculated as $BAS = (ASK - BID)/[(ASK + BID)/2]$ while the trading volume (*TV*) is measured as the natural log of trading volume in each 5-min interval. Given that the true volatility is unobservable, the empirical results may be sensitive to the chosen volatility measure. In this paper, the intraday volatility (*VO*) is calculated using three approaches, as below:

$$VO_t^{SQ} = \ln\left(CP_t/CP_{t-1}\right)^2$$
$$VO_t^{GK} = 0.5[\ln(HP_t) - \ln(LP_t)]^2 - [2\ln 2 - 1][\ln(CP_t) - \ln(OP_t)]^2$$
$$VO_t^{RS} = [\ln(HP_t) - \ln(OP_t)][\ln(HP_t) - \ln(CP_t)] + [\ln(LP_t) - \ln(OP_t)][\ln(LP_t) - \ln(CP_t)]$$

where VO_t^{SQ}, VO_t^{GK}, and VO_t^{RS} are the square return, volatility proposed by Garman and Klass (1980), and volatility proposed by Rogers and Satchell (1991) and Rogers et al. (1994), respectively. *HP*, *LP*, *CP*, and *OP* represent the high price, low price, closing price, and opening price, respectively.

The selected descriptive statistics of the variables are reported in Table 1. BAS_E and TV_E are the bid–ask spread and trading volume of the equity market, respectively, while BAS_O and TV_O are the bid–ask spread and trading volume of the crude oil market, respectively. VO_E^{SQ}, VO_E^{GK} and VO_E^{RS} are the three price volatility measures for the equity market, while VO_O^{SQ}, VO_O^{GK}, and VO_O^{RS} are the corresponding volatility measures for the crude oil market. The null hypothesis of normality based on the Jarque–Bera test is rejected at the 1% level of significance for all variables in both markets. A standard augmented Dickey–Fuller (ADF) test, which examines the null hypothesis of a unit root, suggests that the null hypothesis can be comfortably rejected at the 1% level of significance for all data series in both markets, which means that all variables are stationary. In the table, we also report the test of autoregressive conditional heteroskedasticity (ARCH) with the null hypothesis of "no ARCH" effect. We can reject the null hypothesis at the 1% level of significance and conclude that all variables (except the VO_E^{RS}) are heteroskedastic. Finally, the results of the Ljung–Box statistic also reject the null hypothesis of independence for each variable, which is suggestive of auto-correlation in the first lag and at least up to the 12th lag.

2.2. Methodology

2.2.1. Empirical model

This paper employs the EGARCH model[6] to remedy the presence of heteroskedasticity of variables noted in Table 1. Based on the Akaike Information Criterion and Schwarz Information Criterion, the lowest order EGARCH(1,1) model has

[6] According to Narayan and Narayan (2007), the EGARCH model has been demonstrated to be superior to the GARCH model in several ways. First, the EGARCH model does not restrict the parameters γ, α, and β in the variance equation as the GARCH model does. Second, the estimate of β allows us to evaluate the persistency of the shocks to the conditional variance. Third, the parameter γ measures the asymmetric or the leverage effect, so the EGARCH model allows us to test the volatility asymmetry. $\gamma > 0$ implies that positive shocks have a stronger effect than negative shocks, and vice versa. Fourth, the parameter α *shows* the magnitude of the conditional shock on the conditional variance.

Table 1
Selective descriptive statistics.

	Mean	SD	JB	ADF	ARCH(1)	ARCH(12)	LB(1)	LB(12)
Panel A: S&P 500								
BAS_E	0.000223	0.000057	0.00	0.00	0.00	0.00	0.00	0.00
TV_E	7.407827	1.898097	0.00	0.00	0.00	0.00	0.00	0.00
VO_E^{SQ}	0.006755	0.039188	0.00	0.00	0.00	0.00	0.00	0.00
VO_E^{GK}	0.006829	0.025104	0.00	0.00	0.00	0.00	0.00	0.00
VO_E^{RS}	0.007015	0.028670	0.00	0.00	0.00	0.00	0.00	0.00
Panel B: NASDAQ								
BAS_E	0.000189	0.000137	0.00	0.00	0.00	0.00	0.00	0.00
TV_E	5.039745	2.213027	0.00	0.00	0.00	0.00	0.00	0.00
VO_E^{SQ}	0.007237	0.041345	0.00	0.00	0.00	0.00	0.00	0.00
VO_E^{GK}	0.006551	0.037526	0.00	0.00	0.00	0.00	0.00	0.00
VO_E^{RS}	0.006704	0.055057	0.00	0.00	0.00	0.00	0.00	0.00
Panel C: crude oil								
BAS_E	0.000250	0.000351	0.00	0.00	0.00	0.00	0.00	0.00
TV_E	5.305737	1.900944	0.00	0.00	0.00	0.00	0.00	0.00
VO_E^{SQ}	0.020847	0.137236	0.00	0.00	0.00	0.00	0.00	0.00
VO_E^{GK}	0.019504	0.173764	0.00	0.00	0.00	0.00	0.00	0.00
VO_E^{RS}	0.020474	0.315063	0.00	0.00	0.75	1.00	0.00	0.00

Notes: BAS_E and TV_E are the bid–ask spread and trading volume of the equity market, respectively, while BAS_O and TV_O are the bid–ask spread and trading volume of the crude oil market, respectively. VO_E^{SQ}, VO_E^{GK} and VO_E^{RS} are the three price volatility measures, namely, square return, Garman and Klass (1980) volatility, and the volatility proposed by Rogers and Satchell (1991) and Rogers et al. (1994) for the equity market; while VO_E^{SQ}, VO_E^{GK}, and VO_E^{RS} are the corresponding volatility measures for the crude oil market. The magnitude of the mean and standard deviation of the volatility in three markets is multiplied by 10,000. In the fourth column of each panel, the table reports the *p*-value from the Jarque–Bera (JB) test, for which the null hypothesis is a joint hypothesis of the skewness and the excess kurtosis being zero. The *p*-values of the ADF test, which examines the null hypothesis of a unit root, are in the fifth column. The last four columns contain the *p*-values for the test of autoregressive conditional heteroskedasticity (ARCH) and the Ljung–Box (LB) test for the autocorrelation at lag 1 and lag 12.

been chosen. We propose three specifications of the EGARCH(1,1) model that use different levels of trading information in predicting volatility of crude oil and equity markets. These three models are as follows:

$$\text{Model 1}: \begin{bmatrix} VO_t^E = \beta_0^E + \beta_1^E VO_{t-1}^E + \varepsilon_t \\ VO_t^O = \beta_0^O + \beta_1^O VO_{t-1}^O + \varepsilon_t \end{bmatrix} \\ \varepsilon_t \to N(0,\sigma_t^2) \\ \ln(0,\sigma_t^2) = \omega + \gamma\frac{\varepsilon_{t-1}}{\sigma_{t-1}} + \alpha\left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right| + \beta\ln(\sigma_{t-1}^2) \tag{1}$$

$$\text{Model 2}: \begin{bmatrix} VO_t^E = \beta_0^E + \beta_1^E BAS_{t-1}^E + \beta_2^E TV_{t-1}^E + \beta_3^E VO_{t-1}^E + \varepsilon_t \\ VO_t^O = \beta_0^O + \beta_1^O BAS_{t-1}^O + \beta_2^O TV_{t-1}^O + \beta_3^O VO_{t-1}^O + \varepsilon_t \end{bmatrix} \\ \varepsilon_t \to N(0,\sigma_t^2) \\ \ln(\sigma_t^2) = \omega + \gamma\frac{\varepsilon_{t-1}}{\sigma_{t-1}} + \alpha\left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right| + \beta\ln(\sigma_{t-1}^2) \tag{2}$$

$$\text{Model 3}: \begin{bmatrix} VO_t^E = \beta_O^E + \beta_1^E BAS_{t-1}^E + \beta_2^E TV_{t-1}^E + \beta_3^E VO_{t-1}^E + \beta_4^E BAS_{t-1}^O + \beta_5^E TV_{t-1}^O + \beta_6^E VO_{t-1}^O + \varepsilon_t \\ VO_t^O = \beta_0^O + \beta_1^O BAS_{t-1}^O + \beta_2^O TV_{t-1}^O + \beta_3^O VO_{t-1}^O + \beta_4^O BAS_{t-1}^E + \beta_5^O TV_{t-1}^E + \beta_6^O VO_{t-1}^E + \varepsilon_t \end{bmatrix} \\ \varepsilon_t \to N(0,\sigma_t^2) \\ \ln(\sigma_t^2) = \omega + \gamma\frac{\varepsilon_{t-1}}{\sigma_{t-1}} + \alpha\left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right| + \beta\ln(\sigma_{t-1}^2) \tag{3}$$

where BAS_t^E, TV_t^E, VO_t^E are the bid–ask spread, trading volume, and the price volatility of the equity market, respectively, while BAS_t^O, TV_t^O, VO_t^O are those of the crude oil market. ε_t is the residual from mean equation, and σ_t^2 is the conditional variance generated from the model.

Model 1 predicts the price volatility of the crude oil or equity market based on its own lagged volatility, while Model 2 is based on its own past trading information including volatility, bid–ask spread and trading volume. On the other hand, Model 3 predicts price volatility using lagged volatility, lagged bid–ask spread, and the lagged trading volume of its own-market and also from the cross-market. Our conjecture is that Model 2 would outperform Model 1 in predicting price volatility as

Model 2 utilises extra information, such as bid–ask spread and trading volume. Similarly, Model 3 is expected to be superior to Model 1 and Model 2 because of the additional information contained in the cross-market.

2.2.2. Forecasting evaluation

In order to compare the forecasting accuracy between models, we use the Mean Square Forecast Error (*MSFE*), which is considered to be amongst the most popular metrics for evaluating the forecasting accuracy. The *MSFE* of each model is calculated as follows:

$$MSFE = \frac{1}{T}\sum_{t=1}^{T}\left(VO_t - \widehat{VO}_t\right)^2 \tag{4}$$

where T is the number of the forecasted volatility, $\widehat{VO}_t$ is the forecasted volatility from each model, and VO_t is the actual volatility. To compare the competitor model to the benchmark model, we use Theil U statistics $= \frac{MSFE_1}{MSFE_0}$. If the competitor model's MSFE_1 is less than the $MSFE_0$ of the benchmark model (Theil $U < 1$), it indicates that the competitor model is more accurate in forecasting than the benchmark model, and vice versa.

Although Theil U can compare the *MSFE* of forecasting models, we need to use another test statistic to judge whether the difference is significant. We test the null hypothesis $H_0 : MSFE_0 \le MSFE_1$ against $MSFE_0 > MSFE_1$. The most popular method is Diebold and Mariano's (1995) and West's (1996) test statistic (*DMW*):

$$DMW = \sqrt{T}\frac{\bar{d}}{\sqrt{\hat{S}}} \tag{5}$$

where

$$\bar{d} = \frac{1}{T}\sum_{t=1}^{T}\hat{d}_t$$

$$\hat{d}_t = \left(VO_t - \widehat{VO}_{0t}\right)^2 - \left(VO_t - \widehat{VO}_{1t}\right)^2$$

$$\hat{S} = \frac{1}{T}\sum_{t=1}^{T}\left(\hat{d}_t - \bar{d}\right)^2$$

where VO_t, $\widehat{VO}_{0t}$, and $\widehat{VO}_{1t}$ are the actual volatility, forecasted volatility from the benchmark and competitor models, respectively. T is the number of observations for the out-of-sample period. The DMW statistic is equivalent to the t-statistic and has a standard normal asymptotic distribution when compared to non-nested models. However, Clark and McCracken (2001) and McCracken (2007) point out that this statistic has a nonstandard distribution when comparing forecasts from nested models. Clark and West (2007) propose a modified Diebold and Mariano (1995) and West (1996) test statistic which they refer to as the MSFE-adjusted statistic replacing $\hat{d}_t = \left(VO_t - \widehat{VO}_{0t}\right)^2 - \left(VO_t - \widehat{VO}_{1t}\right)^2$ with $\tilde{d}_{t+k} = \left(VO_t - \widehat{VO}_{0t}\right)^2 - \left[\left(VO_t - \widehat{VO}_{1t}\right)^2 - \left(VO_{0t} - \widehat{VO}_{1t}\right)^2\right]$. This test statistic is now widely used in the applied time series forecasting literature (for example, Rapach et al., 2010; Kong et al., 2011; and Neely et al., 2011).

2.2.3. Economic significance

In order to examine how beneficial the cross-market volatility is, we analyse the utility gains available for a mean-variance investor. Specifically, we compute the average utility for a mean–variance investor who allocates her portfolio between a risky asset and a risk-free asset with the aim of maximising her utility function, which has the following form:

$$Max\left[E(r_{t+1}|I_t) - \frac{\gamma}{2}Var(r_{t+1}|I_t)\right] \tag{6}$$

where γ is the relative risk aversion parameter; $E(r_{t+h})$ and $Var(r_{t+h})$ denote the expected mean and variance of index excess returns estimated by the forecast approaches. The return on a portfolio of risky asset and a risk-free asset is defined as:

$$r_{t+1}^{port} = r_{t+1}^{f} + \omega_t r_{t+1} \tag{7}$$

where r_{t+1}^{port}, r_{t+1}^{f}, r_{t+1} are the return of the portfolio, risk-free asset, and risky asset, respectively. ω_t denotes the proportion of the portfolio allocated to the risky asset. The risky asset weight, ω_t, is positively related to expected excess returns and negatively related to its variance. In other words, an investor will invest more in the risky asset if return is increasing, and

Table 2
Unconditional correlations.

	Square return		Garman and Klass volatility		Roger and Satchel volatility	
	Equity	Crude oil	Equity	Crude oil	Equity	Crude oil
Panel A: S&P500 index						
BAS_E	0.077	0.081	0.138	0.059	0.130	0.034
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
BAS_O	0.035	0.068	0.043	0.032	0.038	0.019
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
TV_E	0.171	0.112	0.275	0.096	0.243	0.056
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
TV_O	0.100	0.138	0.173	0.117	0.153	0.068
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
VO_E	1.000	0.144	1.000	0.081	1.000	0.041
	–	(0.00)	–	(0.00)	–	(0.00)
VO_O	0.144	1.000	0.081	1.000	0.041	1.000
	(0.00)	–	(0.00)	–	(0.00)	–
Panel B: NASDAQ index						
BAS_E	0.027	0.031	0.033	0.018	0.029	0.011
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
BAS_O	0.028	0.068	0.020	0.032	0.014	0.019
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
TV_E	0.175	0.108	0.194	0.095	0.134	0.055
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
TV_O	0.114	0.138	0.133	0.117	0.092	0.068
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
VO_E	1.000	0.114	1.000	0.053	1.000	0.021
	–	(0.00)	–	(0.00)	–	(0.00)
VO_O	0.114	1.000	0.053	1.000	0.021	1.000
	(0.00)	–	(0.00)	–	(0.00)	–

Notes: This table reports the correlations between bid–ask spread, trading volume, the price volatility of the equity market/crude oil market and each of three measures of price volatility (including square return, Garman and Klass (1980) volatility, and the volatility proposed by Rogers and Satchell (1991) and Rogers et al. (1994). BAS_E, TV_E, and VO_E are the bid–ask spread, trading volume, and the price volatility of the equity market, respectively, while BAS_O, TV_O, and VO_O are the corresponding variables for the crude oil market. Panel A reports the results when the equity market is proxied by the S&P500 index while the results when using the NASDAQ index are in Panel B.

will be equally discouraged from investing if its variance is rising over time. The optimal portfolio weight for risky asset return, therefore, is:

$$\omega_t^* = \frac{E_t(r_{t+1})}{\gamma Var_t(r_{t+1})} \quad (8)$$

Our approach is as follows. We follow the study of Campbell and Thompson (2008) and allow for borrowing only but not short-selling. This restricts the optimal portfolio weight, ω_t^*, for the risky asset to lie between 0 and 1.5. Following Narayan et al. (2013), the relative risk aversion parameter, γ, is set to six, which represents a medium level of risk position for an investor. We measure the utility gain as the difference between the utility of the models, and express the utility gain in the annualised percentage. In this way, the utility gain can be interpreted as the portfolio management fee that an investor would be willing to pay to have access to the additional information available in a competitor forecasting model.

3. Empirical results

In this section, we report the main findings as follows. We begin with the contemporaneous relationship between bid–ask spread, trading volume, and price volatility variables across the equity and crude oil markets. We then examine whether including additional information from the bid–ask spread, trading volume, and the price volatility from own-market and cross-market is important in predicting price volatility. Next, we compare the forecasting accuracy of our predictive regression models with three different out-of-sample periods. Finally, we investigate the economic significance of trading strategy based on the forecasting models.

3.1. Contemporaneous effect

In this sub-section, we investigate the contemporaneous relationship between bid–ask spread and trading volume to price volatility across the equity and crude oil markets by considering their correlations, which are reported below in Table 2. As explained earlier, three measures of price volatility are used. Panel A reports the results when the equity market is proxied by the S&P500 index, and the results when using the NASDAQ index are reported in Panel B.

Focusing on the relationship between bid–ask spread and price volatility, the correlation coefficients are significantly positive at the 1% level of significance in both the equity and crude oil markets and for all three measures of price volatility.

Table 3
Information criterion.

		Square return			Garman and Klass volatility			Roger and Satchel volatility		
		Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Panel A: S&P500 index										
Equity	AIC	−22.2130	−22.2223	−22.2247	−23.6944	−23.7219	−23.7243	−23.1389	−23.1624	−23.1635
	SIC	−22.2128	−22.2219	−22.2242	−23.6942	−23.7216	−23.7238	−23.1386	−23.1621	−23.1630
	ADJ *R*-squared	1.95%	4.41%	4.62%	29.77%	31.10%	31.22%	17.35%	19.26%	19.34%
Crude oil	AIC	−19.6075	−19.6531	−19.6579	−19.7144	−19.1575	−19.1105	−17.9552	−17.8822	−17.8835
	SIC	−19.6073	−19.6528	−19.6575	−19.7141	−19.1571	−19.1100	−17.9549	−17.8819	−17.8831
	ADJ *R*-squared	2.04%	4.55%	5.16%	−4.40%	−5.26%	2.27%	−10.76%	0.48%	0.67%
Panel B: NASDAQ index										
Equity	AIC	−22.0559	−22.0739	−22.0759	−22.8477	−22.8645	−22.8722	−22.1030	−21.7201	−21.7344
	SIC	−22.0557	−22.0735	−22.0754	−22.8475	−22.8641	−22.8717	−22.1027	−21.7198	−21.7340
	ADJ *R*-squared	1.93%	4.47%	4.66%	13.98%	15.38%	15.40%	−2.86%	4.69%	4.77%
Crude oil	AIC	−19.6075	−19.6531	−19.6569	−19.7144	−19.1575	−19.1083	−17.9552	−17.8822	−17.8828
	SIC	−19.6073	−19.6528	−19.6564	−19.7141	−19.1571	−19.1078	−17.9549	−17.8819	−17.8824
	ADJ *R*-squared	2.04%	4.55%	4.97%	−4.40%	−5.26%	1.93%	−10.76%	0.48%	0.56%

Notes: This table reports the Akaike Information Criterion, Schwarz Information Criterion, and the adjusted *R*-square of three EGARCH(1,1) models predicting volatility in the crude oil and equity markets. The predictive regression models are presented as Eqs. (1)–(3) in the main text. Three price volatility measures are used, namely, square return, Garman and Klass (1980) volatility, and the volatility proposed by Rogers and Satchell (1991) and Rogers et al. (1994). Panel A reports the results when using the S&P500 index as the equity market, while the results when using the NASDAQ index are reported in Panel B.

Interestingly, the price volatility in both markets is positively correlated with not only their own bid–ask spread but also the bid–ask spread from the cross-market. This relationship is similar but relatively weaker when using the NASDAQ index compared to the S&P500 index.

Turning to the relationship between price volatility and trading volume, we also observe that the correlation coefficients between these two variables are consistently positive and statistically significant at the 1% level across all volatility measures. This finding is consistent with the mixture of distributions hypothesis of Clark (1973), which suggests a positive contemporaneous relationship between trading volume and price volatility. The price volatility in the equity and crude oil markets is positively correlated with their own trading volume as well as the trading volume from the cross-market. The correlation coefficients vary and are in the range of 0.056–0.275 in the case of the S&P500 index, while the range is 0.055–0.175 when using the NASDAQ index. Similarly, the results suggest a positive and significant contemporaneous relationship among the three volatility measures in both the crude oil and equity markets.

Overall, the results represented in Table 2 confirm the positive contemporaneous relationships between bid–ask spread, trading volume and price volatility. It is also interesting that the relationships between these three variables are not only significant in their own-market but also in the cross-market. The significant correlation of the bid–ask spread and trading volume on price volatility across markets motivates us to empirically test whether including additional information from the bid–ask spread and trading volume can improve the predictability of price volatility predictability. This is the subject of the next section.

3.2. In-sample analysis

In order to empirically assess whether information from the bid–ask spread, trading volume, and price volatility contains any useful information for forecasting price volatility in the crude oil and equity markets, we implement three specifications of the EGARCH(1,1) model, as specified in Eqs. (1)–(3). Briefly, Model 1 predicts the price volatility of the crude oil or equity market based on its own lagged volatility, while Model 2 contains the bid–ask spread and trading volume from its own-market in the predictive regression model. On the other hand, Model 3 contains lagged volatility, lagged bid–ask spread, and lagged trading volume of its own-market as well as from the cross-market. In this setup, Model 1 is the weakest set-up in terms of information content, while Model 3 is the richest; Model 2, by comparison, falls somewhere in between.

Table 3 presents statistics on the empirical fit of all three models. In particular, we report the Akaike Information Criterion (AIC), Schwarz Information Criterion (SIC), and the adjusted *R*-squared (ARS). There are two main findings. First, the information from the bid–ask spread and trading volume improves the fit of the price volatility predictability model. To illustrate, Model 2 is found to be superior to Model 1, as an improvement in the AIC, SIC and ARS statistics is observed. In particular, in eight out of 12 cases the AIC and SIC from Model 2 are less than those from Model 1, while the ARS from Model 2 is greater than that of Model 1 in 10 out of 12 cases, suggesting that Model 2 has a better fit than Model 1 in predicting price volatility.

Second, the trading information from the cross-market also improves the fit of the price volatility predictive regression model. Out of 12 regressions across markets and volatility measures, there are eight times when the AIC and SIC of Model 3 are less than those obtained from Model 1. Similarly, Model 3 has the smallest AIC and SIC statistics compared to Model 2. Furthermore, the ARS from Model 3 is higher than those from Models 1 and 2 in all cases, suggesting strong evidence in favour of the value of trading information from the cross-market in predicting price volatility.

Table 4
Lagged effect.

	Square return		Garman and Klass volatility		Roger and Satchel volatility	
	Equity	Oil	Equity	Oil	Equity	Oil
Panel A: S&P500 index						
C	−0.0275	−0.0804	−0.0135	−0.0721	−0.0185	−0.0901
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$BAS_{E,t-1}$	0.0062	0.0141	0.0029	0.0153	0.0037	0.0173
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$BAS_{O,t-1}$	0.0004	0.0043	0.0001	0.0017	0.0002	0.0019
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$TV_{E,t-1}$	0.0023	0.0045	0.0011	0.0027	0.0017	0.0035
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.09)
$TV_{O,t-1}$	0.0002	0.0044	0.0003	0.0056	0.0003	0.0075
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$VO_{E,t-1}$	0.1108	0.0751	0.5058	0.3173	0.3748	0.2415
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$VO_{O,t-1}$	0.0069	0.1162	0.0035	0.0576	0.0013	0.0180
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Panel B: NASDAQ index						
C	−0.0146	−0.0446	−0.0083	−0.0407	−0.0115	−0.0547
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$BAS_{E,t-1}$	0.0026	0.0041	0.0017	0.0042	0.0021	0.0047
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$BAS_{O,t-1}$	0.0004	0.0046	0.0001	0.0021	0.0002	0.0023
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$TV_{E,t-1}$	0.0026	0.0041	0.0018	0.0026	0.0022	0.0038
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.04)
$TV_{O,t-1}$	0.0003	0.0045	-0.0001	0.0057	0.0002	0.0073
	(0.00)	(0.00)	(0.35)	(0.00)	(0.29)	(0.00)
$VO_{E,t-1}$	0.1089	0.0812	0.3470	0.1369	0.1714	0.0634
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$VO_{O,t-1}$	0.0075	0.1193	0.0044	0.0608	0.0017	0.0191
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Notes: BAS_E, TV_E, and VO_E are the bid–ask spread, trading volume, and price volatility of the equity market, respectively, while BAS_O, TV_O, and VO_O are the corresponding variables for the crude oil market. Three price volatility measures are used, namely, square return, Garman and Klass (1980) volatility, and the volatility proposed by Rogers and Satchell (1991) and Rogers et al. (1994). The specification of the model underlying the results is presented by Eq. (3) in the main text. The *p*-value of the coefficient for each variable is in parentheses. The coefficients of the constants, $TV_{E,t-1}$, and $TV_{O,t-1}$, in both the S&P500 and NASDAQ indices are multiplied by 10,000.

Table 4, on the other hand, reports the coefficients and the *p*-value of variables in the price volatility predictive regression model (Model 3). We find three main features of the results. First, the bid–ask spread coefficients are positive and statistically significant at the 1% level in all cases, indicating strong evidence of predictability of price volatility resulting from the bid–ask spread. This implies that an increase in liquidity (i.e., narrowing bid–ask spread) reduces the price volatility in both the equity and crude oil markets. Focusing on the results reported in Panel A for the S&P500 index, the magnitude of the bid–ask coefficients are in the range 0.0002–0.0153. The price volatility in the crude oil market as well as in the equity market is positively and significantly predicted not only by its own lagged bid–ask spread but also in the cross-market. The results are robust to the use of the equity market as similar results are found when using the NASDAQ index.

Second, we also find strong evidence of price volatility predictability resulting from trading volume. As can be seen from Panel A, the trading volume coefficients are positive and statistically significant at the 1% level across different price volatility measures, implying that an increase in the number of trades in the crude oil/equity market leads to an increase in price volatility. Our finding is in line with the sequential arrival of Copeland's (1976) information hypothesis and other empirical studies (such as Foster, 1995; Wang and Yau, 2000; Rahman et al., 2002; Darrat et al., 2003; Hussain, 2011). In addition, the trading volume not only significantly predicts its own-market price volatility but also the price volatility of the cross-market. Turning to Panel B, we observe that the results on trading volume coefficients when using the NASDAQ index are similar to those when using the S&P500 index. One exception is that the crude oil market's trading volume does not affect the Garman and Klass volatility and Roger and Satchel volatility of equity market, although the result is still significant for the square return volatility measure.

Our final result of interest relates to the price volatility coefficients, a manifestation of the volatility interaction between the crude oil and equity markets. When using the S&P500 index, the price volatility of the equity market is positively and significantly affected by the crude oil price volatility and vice versa. The results are robust to the price volatility measures. For example, when using the square return measure, the coefficient of the lagged crude oil price volatility is 0.0069, which means that a 1% increase in the price volatility of the crude oil market can lead to a rise of 0.0069% in the price volatility of the equity market. Reciprocally, the coefficient of lagged equity price volatility is 0.0751, suggesting that a 1% increase in the price volatility of the equity market can lead to a rise of 0.0751% in the price volatility of the crude oil market. Taken

Table 5
Statistics of the in-sample performance.

	(2) versus (1)		(3) versus (1)		(3) versus (2)	
	Theil *U*	*p*-Value	Theil *U*	*p*-Value	Theil *U*	*p*-Value
Panel A: S&P500 index						
VO_E^{SQ}	0.9749	(0.00)	0.9727	(0.00)	0.9978	(0.00)
VO_O^{SQ}	0.9203	(0.00)	0.9145	(0.00)	0.9936	(0.00)
VO_E^{GK}	0.9810	(0.00)	0.9793	(0.00)	0.9983	(0.00)
VO_O^{GK}	1.0179	(0.00)	0.9361	(0.00)	0.9196	(0.11)
VO_E^{RS}	0.9769	(0.00)	0.9760	(0.00)	0.9991	(0.00)
VO_O^{RS}	0.8985	(0.12)	0.8968	(0.12)	0.9981	(0.00)
Panel B: NASDAQ index						
VO_E^{SQ}	0.9741	(0.00)	0.9722	(0.00)	0.9980	(0.00)
VO_O^{SQ}	0.9203	(0.00)	0.9163	(0.00)	0.9956	(0.00)
VO_E^{GK}	0.9838	(0.00)	0.9835	(0.00)	0.9997	(0.00)
VO_O^{GK}	1.0179	(0.00)	0.9393	(0.00)	0.9228	(0.11)
VO_E^{RS}	0.9265	(0.16)	0.9259	(0.16)	0.9992	(0.00)
VO_O^{RS}	0.8985	(0.12)	0.8978	(0.12)	0.9992	(0.00)

Notes: VO_E^{SQ}, VO_E^{GK}, and VO_E^{RS} are the three price volatility measures, namely, square return, Garman and Klass (1980) volatility, and the volatility proposed by Rogers and Satchell (1991) and Rogers et al. (1994) of the equity market; while VO_O^{SQ}, VO_O^{GK}, and VO_O^{RS} are the corresponding volatility measures for the crude oil market. The table reports the predicting performance comparison between three EGARCH(1,1) predictive regression models, which are given as Eqs. (1)–(3) in the main text. The Theil *U* statistics = $MSFE_1/MSFE_0$, where $MSFE_1$ and $MSFE_0$ are the Mean Square Forecast Errors from the competitor and benchmark models, respectively. (2) versus (1) means that Model (2) is the competitor model and Model (1) is the benchmark model and similarly for (3) versus (1) and (3) versus (2). The *p*-value of Clark and West (2007) adjusted-*MSFE* which tests the null hypothesis $H_0: MSFE_0 \leq MSFE_1$ against $MSFE_0 > MSFE_1$ is in parentheses. Panel A reports the results when using the S&P500 index as the equity market, while the results when using the NASDAQ index are reported in Panel B.

together, it is worthy to note that the volatility interaction between the crude oil and equity markets is bi-directional. In the other words, the information on crude oil price volatility can be used to predict the equity price volatility, and vice versa. The results are robust when we use the NASDAQ index as the equity market.

So far the results obtained from Tables 3 and 4 together strengthen the evidence that the information from the bid–ask spread, trading volume, and price volatility in its own-market as well as from the cross-market successfully predict price volatility. To delve deeper into this, we subsequently compare the in-sample predicting performance among three EGARCH(1,1) models in predicting the volatility of the crude oil and equity markets using evaluation statistics. The Theil *U* statistics and the *p*-value for the *MSFE*-adjusted test are reported in Table 5. Comparing Model 2 and Model 1, we find that all Theil *U* statistics for both S&P500 and NASDAQ indices (reported in Panels A and B, respectively) are less than 1, except for the Theil *U* obtained from the crude oil market's Garman and Klass volatility measure, VO_O^{GK}. The Theil *U* statistics are in a range 0.8985–0.9838 and are statistically significant at the 1% level of significance in most cases, concluding that Model 2 outperforms Model 1. This result again confirms the predictability of market price volatility from the bid–ask spread and trading volume information.

Turning to the comparison between Model 3 and Model 1, it is clearly depicted that the introduction of the bid–ask spread, trading volume, and price volatility of both markets in the volatility predictive regression model reduces the gap between the model-estimated volatility and the actual volatility. Among 12 cases across three different volatility measures and two equity indices, the Theil *U* statistics are all less than 1, and nine of those 12 cases are statistically significant at the 1% level. Considering the results of the comparison of Models 3 and 2, which are reported in the last two columns of Table 5, we find that all Theil *U* statistics are less than 1 for both the S&P500 and NASDAQ indices. Except when we use the crude oil market's Garman and Klass volatility measure, VO_O^{GK}, the outperformance of Model 3 over Model 2 is consistently significant at the 1% level, regardless of the volatility measures and the markets in which they are conducted. The results imply that the price volatility can be predicted more accurately by utilising the trading transaction information from both the equity and crude oil market together (Model 3), rather than just using information from that market alone (Models 1 and 2).

In summary, the in-sample evidence so far demonstrates that the integration of the bid–ask spread and trading volume factors leads to a better performance than the use of lag volatility alone. Also, the trading information, such as bid–ask spread, trading volume, and the price volatility from the cross-market, improves the price volatility predictability. The above conclusions are robust across different types of analysis including the regression fitness, the significance of variables in the predictive regression model, and the predicting evaluation.

3.3. Out-of-sample forecasting results

To get an additional perspective on price volatility forecasting, the out-of-sample forecasting performance of previous predictive regression models is examined in this sub-section. We choose three out-of-sample periods for our analysis—2 January 2011–31 December 2012 (two years), 3 January 2012–31 December 2012 (one year), and 1 July 2012–31 December 2012 (six months). Table 6 presents the results based on the out-of-sample forecasting performance between three models

Table 6
Statistics of the out-of-sample performance using the S&P500 index.

	(2) versus (1)		(3) versus (1)		(3) versus (2)	
	Theil *U*	*p*-Value	Theil *U*	*p*-Value	Theil *U*	*p*-Value
Panel A: two-year out-of-sample period						
VO_E^{SQ}	0.9723	(0.00)	0.9739	(0.00)	1.0017	(0.21)
VO_O^{SQ}	1.0206	(0.00)	0.9995	(0.00)	0.9793	(0.00)
VO_E^{GK}	0.9485	(0.03)	0.9659	(0.00)	1.0183	(0.69)
VO_O^{GK}	0.7192	(0.00)	0.6342	(0.00)	0.8818	(0.00)
VO_E^{RS}	0.9736	(0.00)	0.9731	(0.00)	0.9996	(0.00)
VO_O^{RS}	0.8903	(0.00)	0.8174	(0.00)	0.9181	(0.00)
Panel B: one-year out-of-sample period						
VO_E^{SQ}	0.9506	(0.00)	0.9625	(0.00)	1.0125	(0.05)
VO_O^{SQ}	1.1173	(0.00)	1.0619	(0.00)	0.9504	(0.00)
VO_E^{GK}	0.9330	(0.09)	0.9518	(0.08)	1.0201	(0.84)
VO_O^{GK}	0.6018	(0.00)	0.5039	(0.00)	0.8373	(0.00)
VO_E^{RS}	0.8922	(0.14)	0.8919	(0.14)	0.9997	(0.11)
VO_O^{RS}	0.8281	(0.00)	0.7582	(0.00)	0.9157	(0.00)
Panel C: six-month out-of-sample period						
VO_E^{SQ}	0.9258	(0.00)	0.9247	(0.00)	0.9988	(0.00)
VO_O^{SQ}	1.0052	(0.00)	0.9847	(0.00)	0.9796	(0.00)
VO_E^{GK}	0.9319	(0.12)	0.9476	(0.11)	1.0168	(0.83)
VO_O^{GK}	0.3952	(0.00)	0.4021	(0.00)	1.0175	(0.00)
VO_E^{RS}	0.8909	(0.15)	0.8903	(0.15)	0.9993	(0.13)
VO_O^{RS}	0.8057	(0.00)	0.8205	(0.00)	1.0184	(0.00)

Notes: VO_E^{SQ}, VO_E^{GK} and VO_E^{RS} are the three price volatility measures, namely, square return, Garman and Klass (1980) volatility, and the volatility proposed by Rogers and Satchell (1991) and Rogers et al. (1994) for the equity market; while VO_O^{SQ}, VO_O^{GK}, and VO_O^{RS} are the corresponding volatility measures for the crude oil market. (2) versus (1) means that Model (2) is the competitor model and Model (1) is the benchmark model and is similar for (3) versus (1) and (3) versus (2). The three EGARCH(1,1) predictive regression models are presented by Eqs. (1)–(3) in the main text. Three out-of-sample periods include 2 January 2011–31 December 2012 (two years), 3 January 2012–31 December 2012 (one year), and 1 July 2012–31 December 2012 (six months). The Theil U statistics = $MSFE_1/MSFE_0$, where $MSFE_1$ and $MSFE_0$ are the Mean Square Forecast Errors from the competitor and benchmark models, respectively. The *p*-value of Clark and West (2007) adjusted-*MSFE* which tests the null hypothesis $H_0 : MSFE_0 \leq MSFE_1$ against $MSFE_0 > MSFE_1$ is in parentheses.

in terms of the Theil *U* statistic and *MSFE*-adjusted test *p*-value when using the S&P500 index proxied for the equity market. The results based on the NASDAQ index are reported in Table 7.

When we use a two-year out-of-sample period, Model 3 appears to outperform Model 1 in all six cases (see Table 6). The Theil *U* statistics are less than one and are statistically significant at the 1% level. Likewise, Model 1 underperforms Model 2 consistently across three different measures of volatility and in both markets, except in the case when we use the square return conducted in the crude oil market. Meanwhile, the superiority of Model 3 over Model 2 is evidenced in four out of six cases. When we consider a one-year out-of-sample period, Model 3 performs better than Model 2 only in the crude oil market, while the outperformance of Model 2 and Model 3 over Model 1 appears most in all cases. The results based on a six-month out-of-sample period (see Panel C) are mixed. There are mixed cases of significant Theil *U* statistics of less-than-one, more-than-one, and also insignificant results.

In short, two main findings arise from these results. First, we find evidence that using information on bid–ask spread and trading volume can improve the price volatility forecast ability in an out-of-sample analysis. This is illustrated by the majority of the Theil *U* statistics reported in Table 6 which are significantly less than 1, supporting the superiority of Model 2 over Model 1. Similarly, the trading information from the cross-market is also helpful in forecasting price volatility, which is evidenced in the outperformance of Model 3 over Models 2 and 1. Second, the results are not very consistent across the three out-of-sample periods. The empirical evidence is strongest for the out-of-sample period of two-years, and weaker when the length of period declines, as expressed in the percentages of statistically significant less-than-one Theil *U* statistics over the total number of Theil *U* statistics decreasing from Panel A, Panel B to Panel C (83%, 67%, and 50%, respectively).

Focusing on the results reported in Table 7 which uses the NASDAQ index instead of the S&P500 index as the equity market, the first findings from Table 6 still hold. However, the results are consistent across all three out-of-sample periods, where the percentage of significant less-than-one Theil *U* statistics over the total number of Theil *U* is similar across Panel A, Panel B to Panel C (78%, 78%, and 83%, respectively).

3.4. Economic significance

The economic significance of the price volatility forecasting outperformance is considered in this sub-section. Table 8 reports the annualised utility gains of the trading strategy based on Model 3 compared to: (1) a buy-and-hold trading strategy; (2) a trading strategy based on forecasting using Model 1; and (3) a trading strategy based on forecasting using Model 2. The table reports the utility gains for each measure's price volatility forecasting from the first row to the sixth row

Table 7
Statistics of the out-of-sample performance using the NASDAQ index.

	(2) versus (1)		(3) versus (1)		(3) versus (2)	
	Theil *U*	*p*-Value	Theil *U*	*p*-Value	Theil *U*	*p*-Value
Panel A: two-year out-of-sample period						
VO_E^{SQ}	0.9754	(0.00)	0.9757	(0.00)	1.0002	(0.01)
VO_O^{SQ}	1.0206	(0.00)	1.0140	(0.00)	0.9936	(0.00)
VO_E^{GK}	0.9372	(0.00)	0.9374	(0.00)	1.0002	(0.00)
VO_O^{GK}	0.7192	(0.00)	0.6850	(0.00)	0.9525	(0.00)
VO_E^{RS}	0.9408	(0.00)	0.9280	(0.00)	0.9864	(0.00)
VO_O^{RS}	0.8903	(0.00)	0.8543	(0.00)	0.9596	(0.00)
Panel B: one-year out-of-sample period						
VO_E^{SQ}	0.9699	(0.00)	0.9748	(0.00)	1.0050	(0.00)
VO_O^{SQ}	1.1173	(0.00)	1.0908	(0.00)	0.9763	(0.00)
VO_E^{GK}	0.9369	(0.00)	0.9379	(0.00)	1.0011	(0.00)
VO_O^{GK}	0.6018	(0.00)	0.5494	(0.00)	0.9128	(0.00)
VO_E^{RS}	0.9792	(0.00)	0.9402	(0.00)	0.9601	(0.00)
VO_O^{RS}	0.8281	(0.00)	0.7924	(0.00)	0.9569	(0.00)
Panel B: six-month out-of-sample period						
VO_E^{SQ}	0.9661	(0.00)	0.9638	(0.00)	0.9976	(0.00)
VO_O^{SQ}	1.0052	(0.00)	0.9978	(0.00)	0.9927	(0.00)
VO_E^{GK}	0.9391	(0.00)	0.9127	(0.00)	0.9719	(0.00)
VO_O^{GK}	0.3952	(0.00)	0.3883	(0.00)	0.9825	(0.00)
VO_E^{RS}	0.9358	(0.00)	1.0666	(0.00)	1.1397	(1.00)
VO_O^{RS}	0.8057	(0.00)	0.7913	(0.00)	0.9822	(0.00)

Notes: VO_E^{SQ}, VO_E^{GK}, and VO_E^{RS} are three price volatility measures, namely, square return, Garman and Klass (1980) volatility, and the volatility proposed by Rogers and Satchell (1991) and Rogers et al. (1994) for the equity market; while VO_O^{SQ}, VO_O^{GK}, and VO_O^{RS} are the corresponding volatility measures for the crude oil market. (2) versus (1) means that Model (2) is the competitor model and Model (1) is the benchmark model and is similar for (3) versus (1) and (3) versus (2). The three EGARCH(1,1) predictive regression models are presented as Eqs. (1)–(3) in the main text. Three out-of-sample periods include 2 January 2011–31 December 2012 (two years), 3 January 2012–31 December 2012 (one year), and 1 July 2012–31 December 2012 (six months). The Theil *U* statistics = $MSFE_1/MSFE_0$, where $MSFE_1$ and $MSFE_0$ are the Mean Square Forecast Errors from the competitor and benchmark models, respectively. The *p*-value of Clark and West (2007) adjusted-*MSFE* which tests the null hypothesis $H_0: MSFE_0 \le MSFE_1$ against $MSFE_0 > MSFE_1$ is in parentheses.

Table 8
Utility gain from out-of-sample forecasting.

	Six-month			One-year			Two-year		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Panel A: S&P500									
VO_E^{SQ}	−3.672	−2.569	−8.211	−3.274	3.018	−6.560	2.947	−0.510	−4.686
VO_O^{SQ}	6.898	4.041	4.996	24.081	−6.054	−3.798	24.279	−1.049	−1.831
VO_E^{GK}	4.027	1.449	−3.030	2.879	8.571	2.209	5.605	−0.570	−2.861
VO_O^{GK}	4.537	2.406	9.115	28.609	−4.396	−4.172	25.390	−4.056	0.018
VO_E^{RS}	5.066	2.061	−4.324	1.056	5.542	0.963	8.693	2.057	2.696
VO_O^{RS}	4.898	2.478	4.939	29.724	−3.185	−2.544	29.020	−0.319	−2.674
Average	3.626	1.644	0.581	13.846	0.583	−2.317	15.989	−0.741	−1.556
Panel B: NASDAQ									
VO_E^{SQ}	18.462	14.163	0.830	2.736	7.861	1.178	5.842	4.076	1.702
VO_O^{SQ}	−1.499	−4.357	−3.402	27.942	−2.192	0.063	21.283	−4.045	−4.827
VO_E^{GK}	19.683	8.954	−0.816	0.347	2.029	−2.504	5.167	0.158	3.548
VO_O^{GK}	2.296	0.165	6.874	22.582	−9.810	−9.418	28.720	−0.726	3.348
VO_E^{RS}	25.399	12.838	8.824	3.028	5.188	0.760	8.693	2.057	2.696
VO_O^{RS}	0.155	−2.265	0.196	24.569	−7.638	−7.048	29.020	−0.319	−2.674
Average	10.749	4.916	2.084	13.534	−0.760	−2.828	16.454	0.200	0.632

Notes: VO_E^{SQ}, VO_E^{GK}, and VO_E^{RS} are the three price volatility measures, namely square return, Garman and Klass (1980) volatility, and the volatility proposed by Rogers and Satchell (1991) and Rogers et al. (1994) of the equity market; while VO_O^{SQ}, VO_O^{GK}, and VO_O^{RS} are the corresponding volatility measures of the crude oil market. This table reports the annualised utility gains of trading strategy based on Model 3 to forecast the price volatility compare to: (1) Buy and hold trading strategy, (2) trading strategy based on the forecasting Model 1, (3) trading strategy based on the forecasting Model 2. Three out-of-sample periods include 2 January 2011–31 December 2012 (two years), 3 January 2012–31 December 2012 (one year), and 1 July 2012–31 December 2012 (six months). The three EGARCH(1,1) predictive regression models are presented as Eqs. (1)–(3) in the main text. The utility gain, in annualised percent, is the management fee the mean-variance investors are willing to pay for access to the forecasting model. The utility function $(r_{t+h}) - (\gamma/2)Var(r_{t+h})$. γ refers to the risk-aversion of investors and has value of six.

of each panel, and the average utility gains in the last row. The results based on the S&P500 index as the equity market are reported in Panel A, and the results based on the NASDAQ index are reported in Panel B of Table 8.

There are two important features worth highlighting from this analysis. First, the trading strategy based on the price volatility forecasting using Model 3 is superior to other strategies, as illustrated by the positive utility gains. Across the out-of-sample periods and equity markets, the average utility gains are positive in all cases when we compare with the buy-and-hold trading strategy. In addition, mean-variance investors are able to observe utility gains by using Model 3 instead of Model 1 to forecast price volatility, which is evidenced, on average, by four positive utility gains in the comparison (see column titled (2)). Turning to the comparison of Models 3 and 2, the average utility gains are positive in both the S&P500 and NASDAQ indexes in the six-month out-of-sample period. However, the results are mixed, with both positive and negative utility gains for the one-year and two-year out-of-sample periods.

The second noteworthy point is that the magnitude of the utility gains is sizeable. For example, across all out-of-sample periods and equity markets, the average utility gains when compared to a buy-and-hold strategy are in the range of 3.63% to 16.45% per annum. On average, the utility gain in this comparison is 12.37% per annum, which can be interpreted as the investors being willing to pay an extra 12.37% per annum to have access to the additional information available in the Model 3 forecasting approach. Comparing the trading strategies, t, based on comparisons between Models 3 and 1, the average utility gains are in the range of -0.74% to 4.92% per annum, while the range is -2.83% to 2.08% per annum when Model 3 is compared with Model 2.

4. Conclusion

This paper contributes to the existing literature by addressing the relative importance of information on trading volume and bid–ask spread using intraday data in predicting cross-market volatility in the crude oil and equity markets. This study uses three nearby futures contracts: E-mini S&P500 index futures, E-mini NASDAQ index futures, and Light Sweet Crude Oil (WTI) futures over the period 2 January 2009–31 December 2012.

In order to investigate the usefulness of the bid–ask spread and trading volume in predicting price volatility, we construct Model 1 which predicts the volatility of the crude oil or equity markets based on its own lagged volatility, while Model 2 is based on the information of the volatility, bid–ask spread, and trading volume of its own-market. On the other hand, we also examine the trading information from the cross-market in predicting price volatility by including the lagged volatility, lagged bid–ask spread, and lagged trading volume of its own-market and from the cross-market in the predictive regression model (see Model 3). For the purpose of comparing forecasting performance between predictive regression models, we use the Theil U statistics and the MSFE-adjusted test. Finally, we test the economic significance of the trading strategies based on the forecasting models.

Our findings are fourfold. First, we confirm the positive contemporaneous relationships between bid–ask spread, trading volume, and price volatility in which the relationships between the three variables are not only significant in their own-market but also in the cross-market set-up. Second, the evidence from in-sample analysis illustrates that the integration of the bid–ask spread and trading volume variables improve the price volatility predictability. Furthermore, the trading information, such as the bid–ask spread, trading volume, and price volatility from the cross-market, also significantly predicts price volatility. These findings are robust across different types of analysis including the regression fitness, the significance of variables in the predictive regression, and the forecasting evaluation.

Third, the bid–ask spread and trading volume from its own-market and cross-market can improve the price volatility forecast ability in an out-of-sample analysis. This is illustrated by the fact that the majority of Theil U statistics using different combinations of the models are reported at statistically significantly less than a value of 1. Finally, we find that the improvement in price volatility forecasting is economically significant to the investors. The trading strategy based on the best forecasting model has a utility gain, on average, of 12.37% per annum compared to using a buy-and-hold trading strategy.

References

Agren, M., 2006. Does Oil Price Uncertainty Transmit to Stock Markets? Working Paper. Uppsala University.
Aloui, C., Jammazi, R., 2009. The effects of crude oil shocks on stock market shifts behaviour: a regime switching approach. Energy Econ. 31, 789–799.
Aloui, C., Nguyen, D., Njeh, H., 2012. Assessing the impacts of oil price fluctuations on stock returns in emerging markets. Econ. Modell. 29, 2686–2695.
Arouri, M., 2011. Does crude oil move stockmarkets in Europe? A sector investigation. Econ. Modell. 28, 1716–1725.
Arouri, M., Jouini, J., Nguyen, D., 2011a. Volatility spillovers between oil prices and stock sector returns: implications for portfolio management. J. Int. Money Finance 30, 1387–1405.
Arouri, M., Lahiani, A., Nguyen, D., 2011b. Return and volatility transmission between world oil prices and stock markets of the GCC countries. Econ. Modell. 28, 1815–1825.
Awartani, B., Maghyereh, A., 2013. Dynamic spillovers between oil and stock markets in the Gulf Cooperation Council Countries. Energy Econ. 36, 28–42.
Bandi, F., Russell, J., Zhu, Y., 2008. Using high-frequency data in dynamic portfolio choice. Econ. Rev. 27, 163–198.
Barndorff-Nielsen, O., Shephard, N., 2007. Variation, jumps and high frequency data in financial econometrics. In: Blundell, R., Persson, T., Newey, W.K. (Eds.), 'Advances in Economics and Econometrics. Theory and Applications, Ninth World Congress', Econometric Society Monographs. Cambridge University Press, pp. 328–372.
Campbell, J., Thompson, S., 2008. Predicting excess stock returns out of sample: can anything beat the historical average? Rev. Finan. Stud. 24, 1509–1531.
Chen, N., Roll, R., Ross, S., 1986. Economic forces and the stock market. J. Business 59, 383–403.
Clark, P., 1973. A subordinated stochastic process model with finite variance for speculative prices. Econometrica 41, 135–155.
Clark, T., McCracken, M., 2001. Test of equal forecast accuracy and encompassing for nested models. J. Econ. 105, 85–110.

Clark, T., West, K., 2007. Approximately normal tests for equal predictive accuracy in nested models. J. Econ. 138, 291–311.
Clements, M., Galvao, A., Kim, J., 2008. Quantile forecasts of daily exchange rate returns from forecasts of realized volatility. J. Empir. Finance 15, 729–750.
Copeland, T., 1976. A model of asset trading under the assumption of sequential information arrival. J. Finance 31, 1149–1168.
Cornell, B., 1981. The relationship between volume and price variability in futures markets. J. Futures Markets 1, 303–316.
Corsi, F., Fusari, N., La Vecchia, D., 2013. Realizing smiles: options pricing with realized volatility. J. Finan. Econ. 107, 284–304.
Darrat, F., Rahman, S., Zhong, M., 2003. Intraday trading volume and return volatility of the DJIA stocks: a note. J. Bank. Finance 27, 2035–2043.
Diebold, F., Mariano, R., 1995. Comparing predictive accuracy. J. Business Econ. Stat. 13, 253–263.
Driesprong, G., Jacobsen, B., Maat, B., 2008. Striking oil: another puzzle? J. Finan. Econ. 89, 307–327.
Duan, J., 1995. The GARCH option pricing model. Math. Finance 5, 13–32.
Fleming, J., Kirby, C., Ostdiek, B., 2003. The economic value of volatility timing using "realized" volatility. J. Finan. Econ. 67, 473–509.
Foster, A., 1995. Volume–volatility relationships for crude oil futures. J. Futures Markets 15, 929–951.
Garman, M., Klass, M., 1980. On the estimation of security price volatilities from historical data. J. Business 53, 67–78.
Geman, H., 2005. Commodity and Commodity Derivatives: Modelling and Pricing for Agriculturals, Metals and Energy. John Wiley and Sons, Edition.
Giot, P., Laurent, S., 2004. Modelling daily value-at-risk using realized volatility and ARCH type models. J. Empir. Finance 11, 379–398.
Gjerde, O., Saettem, F., 1999. Causal relations among stock returns and macroeconomic variables in a small, open economy. J. Int. Finan. Markets Inst. Money 9, 61–74.
Guesmi, K., Fattoum, S., 2014. Return and volatility transmission between oil prices and oil-exporting and oil-importing countries. Econ. Modell. 38, 305–310.
Hammoudeh, S., Dibooglu, S., Aleisa, E., 2004. Relationships among US oil prices and oil industry equity indices. Int. Rev. Econ. Finance 13, 427–453.
Hansen, P., Lunde, A., 2010. Forecasting volatility using high frequency data. In: Clements, M.P., Hendry, D.F. (Eds.), Oxford Handbook of Economic Forecasting, vol. 19. Blackwell, Oxford, pp. 525–556.
Heston, S., Nandi, S., 2000. A closed-form GARCH option valuation model. Rev. Finan. Stud. 13, 585–625.
Hua, Z., Zhang, B., 2008. Improving density forecast by modelling asymmetric features: an application to S&P500 returns. Eur. J. Oper. Res. 185, 716–725.
Huang, R., Masulis, R., Stoll, H., 1996. Energy shocks and financial markets. J. Futures Markets 16, 1–27.
Hussain, S., 2011. The intraday behaviour of bid–ask spreads, trading volume and return volatility: evidence from DAX30. Int. J. Econ. Finance 3.
Jones, C., Kaul, G., 1996. Oil and the stock market. J. Finance 51, 463–491.
Jones, C., Kaul, G., Lipson, M., 1994. Transactions, volume and volatility. Rev. Finan. Stud. 7, 631–651.
Jouini, J., Harrathi, N., 2014. Revisiting the shock and volatility transmissions among GCC stock and oil markets: a further investigation. Econ. Modell. 38, 486–494.
Kong, A., Rapach, D., Strauss, J., Zhou, G., 2011. Predicting market components out of sample: asset allocation implications. J. Portf. Manage. 37, 29–41.
Maheu, J., McCurdy, T., 2011. Do high-frequency measures of volatility improve forecasts of return distributions? J. Econ. 160, 69–76.
Malik, F., Ewing, B., 2009. Volatility transmission between oil prices and equity sector returns. Int. Rev. Finan. Anal. 18, 95–100.
Malik, F., Hammoudeh, S., 2007. Shock and volatility transmission in the oil, US and Gulf equity markets. Int. Rev. Econ. Finance 16, 357–368.
McCracken, M., 2007. Asymptotics for out of sample tests of Granger causality. J. Econ. 140, 719–752.
Mensi, W., Beljid, M., Boubaker, A., Managi, S., 2013. Correlations and volatility spillovers across commodity and stock markets: linking energies, food, and gold. Econ. Modell. 32, 15–22.
Miller, J., Ratti, R., 2009. Crude oil and stock markets: stability, instability, and bubbles. Energy Econ. 31, 559–568.
Morgan, I., 1976. Stock prices and heteroskedasticity. J. Business 49, 496–508.
Najand, M., Yung, K., 1991. A GARCH estimation of the relationship between volume and price variability in futures market. J. Futures Markets 11, 613–621.
Narayan, K., Sharma, S., 2014. Firm return volatility and economic gains: the role of oil prices. Econ. Modell. 38, 142–151.
Narayan, P., Narayan, S., 2007. Modelling oil price volatility. Energy Policy 35, 6549–6553.
Narayan, P., Narayan, S., 2010. Modelling the impact of oil prices on Vietnam's stock prices. Appl. Energy 87, 356–361.
Narayan, P., Sharma, S., 2011. New evidence on oil price and firm returns. J. Bank. Finance 35, 3253–3262.
Narayan, P., Narayan, S., Sharma, S., 2013. An analysis of commodity markets: what gain for investors? J. Bank. Finance 37, 3878–3889.
Neely, C., Rapach, D., Tu, J., Zhou, G., 2011. Forecasting the Equity Risk Premium: The Role of Technical Indicators. Working Paper. Federal Reserve Bank of St. Louis.
Park, J., Ratti, R., 2008. Oil price shocks and stock markets in the U.S. and 13 European countries. Energy Econ. 30, 2587–2608.
Phan, D., Sharma, S., Narayan, P., 2015a. Oil price and stock returns of consumers and producers of crude oil. J. Int. Finan. Markets Inst. Money 34, 245–262.
Phan, D., Sharma, S., Narayan, P., 2015b. Stock return forecasting: some new evidence. J. Int. Rev. Finan. Anal. 40, 38–51.
Rahman, S., Lee, C., Ang, K., 2002. Intraday return volatility process: evidence from NASDAQ stocks. Rev. Quant. Finance Account. 19, 155–180.
Rapach, D., Strauss, J., Zhou, G., 2010. Out-of-sample equity premium prediction: combination forecast and links to the real economy. Rev. Finan. Stud. 23, 821–862.
Rogers, L., Satchell, S., 1991. Estimating variance from high, low, and closing prices. Anal. Appl. Prob. 1, 500–512.
Rogers, L., Satchell, S., Yoon, Y., 1994. Estimating the volatility of stock prices: a comparison of methods that use high and low prices. Appl. Finan. Econ. 4, 241–247.
Sévi, B., 2014. Forecasting the volatility of crude oil futures using intraday data. Eur. J. Oper. Res. 235, 643–659.
Soucek, M., Todorova, N., 2013. Realized volatility transmission between crude oil and equity futures markets: a multivariate HAR approach. Energy Econ. 40, 586–597.
Soucek, M., Todorova, N., 2014. Realized volatility transmission: the role of jumps and leverage effects. Econ. Lett. 122, 111–115.
Stentoft, L., 2008. Option Pricing Using Realized Volatility. Working Paper. HEC Montréal, CREATES, CREF, and CIRANO.
Tauchen, G., Pitts, M., 1983. The price variability volume relationship on speculative markets. Econometrica 51, 485–505.
Wang, G., Yau, J., 2000. Trading volume, bid–ask spread and price volatility in futures markets. J. Futures Markets 20, 943–970.
Wang, G., Michalski, R., Jordan, J., Moriarty, E., 1994. An intraday analysis of bid–ask spreads and price volatility in the S&P 500 index futures market. J. Futures Markets 14, 837–859.
Wang, Y., Wang, Y., 2010. Intraday volatility patterns in the Taiwan stock market and the impact on volatility forecasting. Asia-Pac. J. Finan. Stud. 39, 70–89.
Wei, C., 2003. Energy, the stock market, and the Putty–Clay investment model. Am. Econ. Rev. 93, 311–323.
West, K., 1996. Asymptotic inference about predictive ability. Econometrica 64, 1067–1084.
Westerfield, R., 1977. The distribution of common stock price changes: an application of transactions time and subordinated stochastic models. J. Finan. Quant. Anal. 12, 743–765.
Wong, H., Lo, Y., 2009. Option pricing with mean reversion and stochastic volatility. Eur. J. Oper. Res. 197, 179–187.
Worthington, A., Higgs, A., 2009. Modelling the intraday return volatility process in the Australian equity market: an examination of the role of information arrival in S&P/ASX 50 Stocks. Int. Business Res. 1, 8.