Journal of Commodity Markets 24 (2021) 100169

Contents lists available at ScienceDirect

Journal of Commodity Markets

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

Forecasting the dynamic relationship between crude oil and stock prices since the 19th century

Kris Ivanovski [*], Abebe Hailemariam

Monash Business School, Monash University, VIC, Australia

ARTICLE INFO

JEL classification:
C3
E5

Keywords:
Forecasting
Oil price
Stock price
Volatility and correlation
Multivariate GAS model
DCC-GARCH model

ABSTRACT

In this paper, we model and forecast the volatility and correlation between oil prices and stock returns. Employing a recently innovated generalized autoregressive score (GAS) model based on the score function and using long historical data spanning from 1871 to 2020, we find a time-varying relationship between oil prices and stock returns. Specifically, the dynamic correlations between crude oil and stock returns tend to rise during turbulent events over the sample period significantly. Our results show that the GAS(1,1) model outperforms the DCC-GARCH model. Our results on the dependent patterns between oil price and stock returns provide useful information for investors, portfolio managers and market participants.

1. Introduction

The relationship between oil prices and stock market returns has been featured in the literature and remains a contested topic in the post-global financial crises (GFC) period. Theoretically, the relationship between the two variables is ambiguous; it can be positive or negative. Following the pioneering work of Jones and Kaul (1996), the cash flow hypothesis, which states that expected discounted cash flows play a key role in determining the value of assets, has been used as a mechanism that links oil prices and stock returns. According to the cash flow hypothesis, the negative relationship between oil and stock prices arises from the fact that oil is an essential input in production for many firms. Thus, an increase in oil price lowers stock returns, thereby raising the cost of production that reduces future cash flows or earnings and dividends. Another channel through which higher oil price may have a negative or positive impact on stock prices is via the influence of oil price volatility on the effect of the sensitivity of oil price changes on the risk premium component of the discount rate and, hence, on cash flow. Recent studies show that an increase in oil prices may lead to higher stock returns as investors associate the rise in oil prices with a booming global economy (see, e.g., Chen et al., 2017; Chen and Xu, 2019; Kilian and Park, 2009; Kollias et al., 2013).

The theoretical ambiguity in the oil price – stock returns relationship has been mirrored in the empirical literature. A large volume of empirical studies shows a negative relationship between the two variables (see, e.g., Basher et al., 2012; Filis, 2010). However, some studies provide evidence of a positive relationship (see, e.g., Narayan and Narayan, 2010; Zhu et al., 2011). Therefore, the existing literature on the oil price-stock returns nexus is far from being resolved.

An emerging and growing strand of literature on the response of macroeconomic variables to oil price shocks has focused on the non-linearities that may occur during episodes of recessions and booms. There is growing evidence that the response of macroeconomic

* Corresponding author.
E-mail addresses: kris.ivanovski@monash.edu (K. Ivanovski), abebe.hailemariam@monash.edu (A. Hailemariam).

https://doi.org/10.1016/j.jcomm.2021.100169
Received 19 May 2020; Received in revised form 2 December 2020; Accepted 6 January 2021
Available online 12 January 2021
2405-8513/

variables to oil price shocks is time-varying (see, e.g., Baumeister and Peersman, 2013; Hailemariam et al., 2019; Silvapulle et al., 2017). Specifically, several studies provide evidence that the relationship between oil prices and stock returns is time-varying and non-linear (see, e.g., Broadstock and Filis, 2014; Silvapulle et al., 2017). An excellent survey of the literature on the oil price-stock returns relationship can be found in Smyth and Narayan (2018).

This paper builds on the latter strand of literature. Common to the existing literature on the non-linearity of the oil price-stock return relationship is based on parameter-driven models that use mean and higher-order moments. In this paper, we employ an observation-driven model based on the score function proposed by Creal et al. (2013), also known as the generalized autoregressive score (GAS) model. The main advantage of the GAS model is that it exploits the complete density structure rather than means and higher moments only. Utilising the GAS framework, we model and forecast the volatility and correlation between oil prices and stock returns and pin down the risk dependence structure between them in a non-linear context. This is the first study that employs the multivariate GAS model to forecast the volatility and correlation of oil prices and stock returns to the best of our knowledge. The only study closer to ours is Avdulaj and Barunik (2015), which utilizes the GAS model to evaluate oil–stocks diversification benefits.

Accordingly, this research makes several contributions to the literature. First, our empirical analysis utilizes historical time-series data possible by modelling the West Texas Intermediate (WTI) crude oil and Standard and Poor's 500 (S&P 500) stock prices since the first globalisation boom in the 1870s. We use monthly observations from 1871 to 2020, which runs from the beginning of the modern era of the petroleum industry. Consequently, our investigation not only captures the first and second oil price shocks, as well as the Global Financial Crisis (GFC) of the post-1950 period but the various world wars, the Great Depression of the 1930s, and numerous oil and stock market events since 1870s. Given such waves of political and economic changes over a long period, the GAS model is suitable to capture the time-varying dependence in oil price and stock returns.

Second, we utilise the Generalized Autoregressive Score (GAS) model developed by Creal et al. (2013) which provides a unified and consistent framework for capturing time-varying parameters in a broad class of non-linear models. The advantage of this model over other data-driven competing frameworks (i.e., DCC-GARCH and autoregressive conditional duration and intensity (ACD, ACI) is the ability to model the time-varying parameters based on the score function of the predictive model density at time *t*. The score function is also a practical choice for introducing a driving mechanism for the time-varying parameters and where likelihood evaluation is straightforward. Given that the GAS model is based on the score, it exploits the complete density structure rather than means and higher moments only (Creal et al., 2013). In addition to utilising the GAS model, we also use the widespread dynamic conditional correlations (DCC) GARCH model of Engle (1982) to examine the crude oil-stock price relationship and compare our results with the GAS model in estimating the long-run relationship between our return series.

Third, our motivation is also based on the predictive capability of both models. While it is important to have a well-specified model that describes the data, we are also interested in utilising the model to make predictions and, more importantly, to compare the out-of-sample evaluations of the competing models. In this regard, we evaluate the forecasting performance by comparing the forecasting results of the time-varying volatilities and dynamic conditional correlations from the DCC-GARCH and GAS models. This is of particular importance since "the ultimate test of any predictive model is its out-of-sample performance" Campbell (2008, pg.3). Also, only two studies to date have investigated the predictability of the crude oil-stock price relationship using historical data. Narayan and Gupta (2015) and Gupta and Wohar (2017) examine the out-of-sample predictability using the generalized least squares estimator and Qualitative VAR (Qual VAR) models, respectively. To the best of our knowledge, no other study has utilised the GAS and DCC-GARCH models to forecast the volatility and correlations between crude oil-stock prices over historical periods.

It is not yet clear from the existing literature which model has the better ability of forecasting. Besides, only a few studies have utilised the GAS model. For instance, Tafakori et al. (2018) evaluates the accuracy and forecasts of the Australian electricity markets using the asymmetric exponential GAS model and find it outperforms all the other models. The authors also state that the GAS model offers considerable flexibility and performance improvement for forecasting models in electricity markets. Moreover, our work closely follows Chen and Xu (2019), who estimate and forecast crude oil and gold prices from 2003 to 2018 using both the DCC-GARCH and GAS models. Their results suggest that the forecasting power of volatility and correlation in the GAS model is better than those of the DCC-GARCH model.

We build on Chen and Xu (2019) to study the dependence between oil and stock returns. The gold market and stock market have different features with respect to price determination and risk characteristics. Gold is considered a safe asset while the stock market is risky. As argued in Choudhry et al. (2015) and Tuysuz (2013), it is a stylized fact that the price of risky financial assets falls systematically in the entire market driven by contagion effects of losses. This results in 'flight to quality' that leads to an increase in prices of safer assets, such as gold. This indicates that the prices of risky assets (e.g., stock prices) and safe assets take different trajectories over the episodes of financial turbulence, including more recent times of COVID-19. Our study is motivated by the inherent risky nature of the stock market compared to the gold market.

Ultimately, we expect to provide empirical evidence for evaluating and forecasting the time-varying volatility and dynamic correlations between crude oil and stock prices using the well-established DCC-GARCH model and the recently developed GAS model. Our preliminary analysis in selecting the GAS specification is guided by likelihood ratio tests to explicitly choose the time-varying parameters of the crude oil and stock return series. The results indicate that both variance and correlation parameters are time-varying, and the t-distribution serves as the conditional distribution of the GAS model. The in-sample results suggest the GAS(1,1) performs better in capturing the volatility persistence and the non-linear interactions between crude oil and stock price returns. We also find the forecasting efficiency of the GAS(1,1) model performs better than the DCC-GARCH(1,1) model over the initial horizon forecasts based on the mean absolute error and the root mean squared error criterion. However, as the forecasting horizons increase, both models produce similar results.

The remainder of this paper is organised as follows: Section 2 presents the data and preliminary analysis in testing for multivariate normality. Section 3 presents the methodology relating to the DCC-GARCH(1,1) and GAS(1,1) models. Section 4 reports the in-sample empirical results. Section 5 conducts the forecasting exercise by comparing the forecasting results from both models to compare their accuracy. Section 6 concludes.

2. Data and preliminary analysis

2.1. Data

This paper aims to study the dynamic co-movements and forecast Standard and Poor's (S&P500) stock returns and West Texas Intermediate (WTI) crude oil returns over the monthly period 1871:01 to 2020:10. The unique feature of the data is that it virtually runs from the beginning of the first globalisation boom of the 1870s (Awaworyi Churchill et al., 2018; Jacks et al., 2011; Madsen and Ang, 2016). We choose monthly frequency since the use of daily and weekly data is not without its issues. For instance, daily frequency induces potential biases arising from non-synchronous trading days, the bid-ask effects, and the effects of illiquid asset prices (Arouri et al., 2011). Weekly frequency is not subject to volatility transmission mechanisms due to time aggregation and strong compensation effects for the positive and negative shocks (Antonakakis and Kizys, 2015; Sadorsky, 2014). In contrast, monthly frequency has adequate data points to analyse spillover effects over time (Lin and Li, 2015). Besides, since our variables span almost 150 years, the data is only available on a monthly basis. Both series are obtained from Global Financial Data (GFD).

The time series plots of the variables are illustrated in Fig. 1 (top panel). While stock prices have increased over the 150 years (albeit with peaks and troughs), crude oil prices seem to exhibit dramatic fluctuations driven by demand and supply-side factors. For instance, the oil-controls in the early 1880s, the gasoline power revolution in 1913, the Great Depression of 1929, the oil crises of the early 1970s, the various wars in the middle east during the early 1980s and 1990s, and the Global Financial Crisis of 2008–09 (Noguera, 2013). The returns series in Fig. 1 (lower panel) highlight that volatilities have changed across time, particularly during periods of uncertainty and recessions where volatility clustering is observed. This is noticeable in both return series. Accordingly, we seek to investigate the co-movement and forecasting properties between crude oil and stock price returns using time-varying volatility and correlation models over tranquil and uncertain periods from a historical perspective.

2.2. Preliminary analysis and test for multivariate normality

We calculate the returns series, r_t, of crude oil and stock prices as $r_t = \ln(P_t / P_{t-1}) \times 100$, where P_t is the monthly price at time t. We begin the analysis by providing descriptive statistics of both price and return series, illustrated in Table 1. Regarding the level price series (P_t), the mean values of stock prices are significantly larger than the mean value of crude oil prices, and the maximums are more than ten times larger relative to the minimum values. Both series are also positively skewed and report a kurtosis value greater than 2, suggesting the distribution of both level series is platykurtic.

In relation to the return series (r_t), the mean of stock price returns is quadruple than that of crude oil returns, with the latter showing

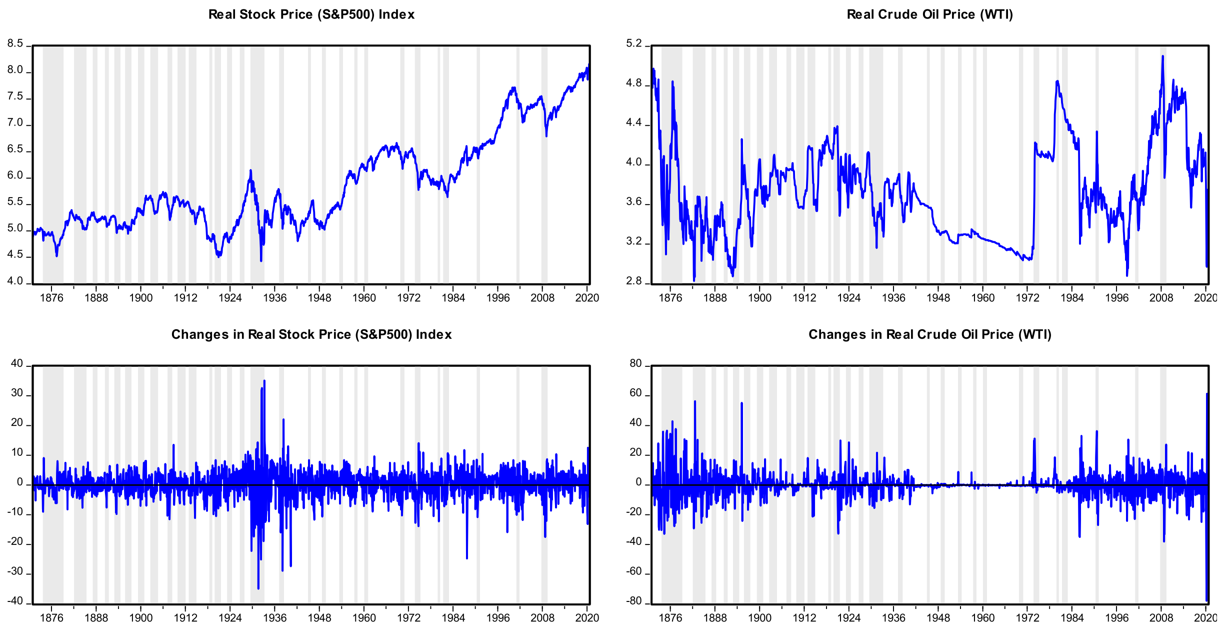

Fig. 1. Time series plots of stock prices (S&P500) and crude oil prices (WTI).
Notes: top panels are price series (in logs) and lower panels are return series (log-first difference multiplied by 100). Shaded regions denote US recessions as defined by the National Bureau of Economic Research (NBER).

Table 1
Descriptive statistics for WTI and S&P500.

	WTI		S&P500	
	P_t	r_t	P_t	r_t
Mean	46.66	−0.06	568.70	0.18
Std. dev	24.38	8.33	636.84	4.79
Min	16.90	−77.99	83.70	−34.96
Max	164.23	61.50	3506.60	35.24
Skewness	1.59	0.29	2.04	−0.40
Kurtosis	11.59	7.27	3.69	8.10
Jarque-Bera	1205.1***	10113***	2276.0***	4978.9***
ADF test	−5.29**	−11.89**	−0.502	−11.28**
Box-Pierce (10)	13718***	107.32***	16730***	54.46***
Box-Pierce^2 (10)	12279***	292.11***	15529***	836.24***
ARCH (10)	2067.59***	22.44***	23501.70***	44.30***
Observations	1798	1797	1798	1797

Notes: Jarque-Bera and Box-Pierce refer to the empirical statistics of the test for normality and autocorrelation, respectively. ADF test only considers the simplest case (unit root test with no constant, no trend).
ARCH (10) provides the statistics Engle (1982) test for conditional heteroscedasticity with 10 lags. *** and ** denote significance at 1% and 5% levels, respectively.

that average returns over the sample period are negative. In contrast, the unconditional volatility of oil price returns, as measured by the standard deviation, is nearly double the stock price returns. The standard deviations of both series are also larger than their respective averages in absolute levels, indicating the means of the return series are not significantly different from zero. This is commonly found in log-financial return time series that do not follow a random process. While oil price returns are positively skewed, stock prices are negatively skewed, and both report large kurtoses. These findings indicate that both return series are leptokurtic and heavy-tailed.

The Jarque-Bera tests for price and returns series reject the null hypothesis of normal distributions at the 1% level of significance, indicating a non-Gaussian distribution. Nevertheless, the ADF unit root tests confirm that the return series are stationary, although we find evidence of autocorrelation suggested by Box-Pierce testing. Using 10 lags, we find significant autocorrelation in the return and squared-return series. Also, the ARCH test (Engle, 1982) with 10 lags rejects the null hypothesis of homoscedasticity for the underlying return series.

Furthermore, while the results thus far suggest that the distribution of our return series are heavy-tailed, exhibit volatility clustering and heteroscedasticity, and non-linearities, we present further evidence for modelling and forecasting the variables using time-varying volatility models to capture the underlying stylized facts. Panel A of Table 2 shows a range of multivariate normality (MVN) tests for the full sample. The results overwhelmingly reject the null hypotheses of MVN. In Panel B, we further apply the Doornik and Hansen (2008) test for multivariate normality for different sub-samples.[1] The results suggest that the multivariate normal distribution of the return series is not easy to reject in the case of small sample size. In other words, the smaller the sample size is, the smaller the probability of normality is rejected. As a result, this exercise provides evidence of applying the GAS model with a given conditional distribution for crude oil and stock returns.

Table 2
Multivariate normality (MVN) tests.

Panel A: Full sample				
	Statistic	*p*-value	MVN	
Doornik-Hansen	1825.315	0.000	No	
Henze-Zirkler	170.525	0.000	No	
Royston	387.791	0.000	No	
Mardia Skewness	2002.169	0.000	No	
Mardia Kurtosis	35.400	0.000	No	
Panel B: Sub-sample (Doornik-Hansen)				
Sample size	20	50	75	100
Rejection rate	0.379	0.704	0.869	0.940
Total numbers	1778	1748	1723	1698

Notes: In Panel A, the MVN test indicates whether the series follows multivariate normality or not the 5% significance level. In panel B, the reject rate of the null hypotheses of multivariate normality is computed at the 5% significance level.

[1] To undertake this exercise, we apply the following technique as set out in Chen and Xu (2019). First, we divide repeatedly the return series into fixed sub-samples of size 20, 50, 75, and 100 in chronological order. Second, for each of the sub-samples at a given size, we apply the Doornik and Hansen (2008) technique to test multivariate normality at the 5% level of significance. Third, we record the number of rejection rates for each of the tests (where the rejection rate for each fixed sub-sample is calculated by the ratio of the rejection rate to the total number of the fixed sub-samples).

3. Methodology

For simplicity, let $r_t = (r_t^{Oil}, r_t^{Stock})$ be the return series of the two variables at time t, and by merging the univariate series, we create a bivariate return series with 1797 monthly observations. For the DCC-GARCH model, the conditional mean equations are fitted by the sample mean, while the univariate residual series are specified with a t-distribution. To generate forecasting results, we separate the data into two sub-samples. The first is an in-sample period to carry out estimations, and the other is an out-of-sample period for generating forecasting results. Based on the size of our dataset, the out-of-sample size is set to 75 horizon months. The empirical results, including the parameter and forecasting estimates, are obtained from the DCC-GARCH(1,1) and GAS(1,1) models discussed below.

3.1. The DCC-GARCH model

To examine the time-varying correlations, the typical DCC-GARCH(1,1) model proposed by Engle (2002) is defined as:

$$r_t = \mu_{t-1} + \varepsilon_t \tag{1}$$

$$\varepsilon_t = H_t^{1/2} u_t. \tag{2}$$

In Eq. (2), u_t is a $n \times 1$ zero mean vector and H_t is a $n \times n$ positive definite matrix of r_t. The conditional covariance matrix H_t can be decomposed as follows:

$$H_t = D_t R_t D_t \tag{3}$$

where $D_t = \mathrm{diag}(\sqrt{h_{11t}}, \ldots, \sqrt{h_{nnt}})$ is the $n \times n$ diagonal matrix of conditional standard deviations of the residuals and R_t denotes the matrix of time-varying conditional correlations decomposed as:

$$R_t = \mathrm{diag}(Q_t)^{-1/2} Q_t \mathrm{diag}(Q_t)^{-1/2} \tag{4}$$

where Q_t is an $n \times n$ symmetric positive definite matrix. Under the DCC model, $h_{ii,t}(i = 1, .., n)$ follows a univariate GARCH(1,1) process:

$$h_{iit} = \omega_i + \alpha_i \varepsilon_{it-1}^2 + \beta_i h_{it-1} \tag{5}$$

and Q_t is given by:

$$Q_t = (1 - a - b)\overline{Q} + a u_{t-1} u_{t-1}^{'} + b Q_{t-1} \tag{6}$$

where a and b are non-negative scalars satisfying the condition $a + b < 1$. $u_t = (u_{1t}, u_{2t}, .., u_{Nt})^{'}$ is the $N \times 1$ vector of standardised residuals, $\overline{Q}$ is the $N \times N$ unconditional variance matrix of u_t.

The DCC-GARCH(1,1) model is estimated using a two-step procedure. In the first step, the individual conditional variances are specified as univariate GARCH processes and in the second step, the standardised residuals from the first step are used to construct the conditional correlation matrix. Also, the model is estimated using Quasi-Maximum Likelihood (QML) estimator under a multivariate Student-t distribution with v degrees of freedom (see Fiorentini et al., 2003; Harvey et al., 1992). Thus, by combining this with the mean of Eq. (1), the conditional multivariate distribution of r_t is given by:

$$r_t | F_{t-1} \tilde{\ } t(r_t; \mu_{t-1}, H_t, v). \tag{7}$$

3.2. The GAS model

To minimise the notational burden, we follow Ardia et al. (2019) and Chen and Xu (2019) and specify the GAS(1,1) model by assuming r_t is a N-dimensional random vector at time t with conditional distribution as:

$$r_t | F_{t-1} \tilde{\ } p(r_t; \xi, \theta_t), \tag{8}$$

where F_{t-1} denotes the sigma-algebra of the past values of r_t up to time t and θ_t is a vector of time-varying parameters which fully characterises the function $p(\cdot)$ which depends only on F_{t-1} and a set of static additional parameters ξ.

The key feature of the GAS model is that the evolution in the time-varying parameter vector θ_t is driven by the score of the conditional distribution defined in Eq. (8), along with a first-order autoregressive component given by:

$$\theta_{t+1} = \kappa + A s_t + B \theta_t \tag{9}$$

where κ, A, and B, are vectors of coefficient matrices with proper dimensions collected in ξ, and s_t is a vector which is proportional to the score of Eq. (8) given by:

$$s_t = S_t \nabla_t. \tag{10}$$

The matrix S_t is a $N \times N$ is positive definite scaling matrix at time t and $\nabla_t = \frac{\partial \ln p(r_t;\theta_t)}{\partial \theta_t}$ is the score of Eq. (8) evaluated at θ_t. The pseudo-inverse square root matrix scaling is applied to the score. Creal et al. (2013) recommend setting the scaling matrix S_t to a power $\gamma > 0$ of the inverse of the information matrix of S_t to account for the variance of ∇_t. In particular, $S_t(\theta_t) = \Im_t(\theta_t)^{-\gamma}$ with $\Im_t(\theta_t) = E_{t-1}[\nabla_t \nabla_t^{\mathrm{T}}] = -E_{t-1}\left[\frac{\partial^2 \ln p(r_t;\theta_t)}{\partial \theta_t \partial \theta_t^{\mathrm{T}}}\right]$ where γ is fixed and typically takes the value in the set $\left\{0, \frac{1}{2}, 1\right\}$. The quantity, s_t, updates the time-varying parameters from θ_t to θ_{t+1} acting as a steepest ascent algorithm for improving the model's local fit given the current parameter position. In essence, this updating procedure resembles the well-known Newton-Raphson algorithm.[2]

Given that we assume a multivariate t-distribution, the conditional distribution parameters in the GAS(1,1) model are denoted as (μ_o, μ_s), (σ_o^2, σ_s^2), ρ_{os}, and ν_{os} which are the location, volatility, correlation, and shape parameters of the conditional t-distribution, respectively. However, while the volatilities of the conditional t-distribution in the GAS(1,1) specification are considered as time-varying parameters, the location, correlation, and shape parameters of the conditional t-distribution may not be time-varying. As a result, this needs to be tested using the Likelihood Ratio Testing (LRT) approach.

Moreover, to reduce the computational burden of the GAS(1,1) model, we consider the coefficient matrix A and B in Eq. (9) are a diagonal matrix where some of the components of the off-diagonal element that correspond to the fixed parameters are equal to zero:

$$A = \mathrm{diag}(0, 0, a_{\sigma_o}, a_{\sigma_s}, a_\rho, a_v)$$

$$B = \mathrm{diag}(0, 0, b_{\sigma_o}, b_{\sigma_s}, b_\rho, b_v)$$

and therefore, the unknown parameters in the GAS(1,1) specification are obtained by numerical maximisation of the log-likelihood function. Furthermore, and with regard to the mapping function, we use a non-linear exponential link function $\Lambda(\cdot)$ for the GAS model to map from $\widehat{\theta}_t$ to θ_t given by:

$$\theta_t = exp(\widehat{\theta}_t) + c$$

where $\widehat{\theta}_t$ has a linear dynamic specification. An advantage of the GAS approach is the automatic treatment of the link function to ensure that parameters remain in their appropriate domains. This feature of the GAS model is particularly important if θ_t is subject to some constraints within a domain. For example, if θ_t is a correlation parameter, the link function ensures that the correlation lies in the range of -1 and $+1$.

4. Empirical results

We first present the results for the LRT to determine if the location, correlation, and shape parameters are time-varying. We test the null hypothesis $H_0 : M = M_i$ versus the alternative of $H_A : M = M_{i+1}$, for $i = 1, 2, 3$, and where model M_i is a series of nested models of time-varying parameters. For instance, the model M_0 contains only the time-varying volatility parameters, while M_3 contains time-varying volatility, correlation, shape, and location. Under regular conditions, $M_0 \subset M_1 \subset M_2 \subset M_3$ and the LRT test statistic for a nested model will be chi-squared distributed, $\chi^2(k)$, with k-degrees of freedom that is equal to the difference in dimensionality between the model M_i and M_{i+1}.

Based on the results in Table 3, we select the M_1 model for the returns of crude oil and stock prices. Thus, for the GAS(1,1) model, the mean (μ_o, μ_s) and shape (ν_{os}) parameters are not considered as time-varying, while the variance (σ_o^2, σ_s^2) and correlation (ρ_{os}) parameters of the conditional t-distribution are time-varying.

The estimated parameters from the GAS(1,1) model, reported in Table 4, are all statistically significant at the 1% level (except for κ_ρ which is significant at the 5% level). In particular, the time-varying estimates are updated by the scaled score function and their first-order lagged parameters (and associated coefficients a_{σ_o}, a_{σ_s} a_ρ, b_{σ_o}, b_{σ_s}, b_ρ) are statistically significant at the 1% level. Since the estimated coefficients of b_{σ_o}, b_{σ_s}, and b_ρ are quantitatively large enough, this suggests that the GAS(1,1) performs well in capturing the strong volatility persistence and the non-linear interactions between crude oil and stock price returns. Following Ardia et al. (2019), we obtain the unconditional parameters as follows: $\mu_o = -0.158$; $\mu_s = 0.391$; $\kappa_o = 5.362$; $\kappa_s = 3.296$; $\kappa_\rho = 0.053$; and $\nu = 4.000$.

Table 5 presents the results of the DCC-GARCH(1,1) with the assumption that the univariate conditional return series follows a t-distribution. Once again, all of the estimated parameters except κ_ρ, are statistically significant at the 1% level and all of them are statistically significant at 5% level. In particular, we find both $\alpha_o + \beta_o$ and $\alpha_s + \beta_s$ sum up close to unity, which indicates high volatility persistence in crude oil and stock markets over historical periods. The dynamic correlation parameters (a_{os} and b_{os}) are positive and statistically significant at the 5% and 1% levels, respectively, and sum up to less than unity, suggesting that the time-varying correlations are mean-reverting. Also, the shape parameter (γ_{os}) which is equivalent to the degrees of freedom in the t-distribution, is above 4 and contains heavier tails, and thus is adequate for the conditional distribution of our return series.

Figs. 2 and 3 illustrate the time-varying returns, volatilities, and correlation filtrations from the DCC-GARCH(1,1) and GAS(1,1)

[2] Note, in the actual calculation, the time-varying parameters of the variance, correlation, and degrees of freedom are restricted. However, the time-varying parameters of the of the dynamic evolution are a linear structure. As a result, we require a non-linear link function with a mapping from $\tilde{\theta}_t$ to θ_t, and where $\tilde{\theta}_t$ has a linear dynamic evolution similar to Eq. (9) as recommend by Creal et al. (2013).

Table 3
The likelihood ratio test (LRT) for the multivariate GAS(1,1) model.

Hypothesis	LRT	p-value	size
$M_0 = M_1$	13.910	0.000	1723
$M_1 = M_2$	−0.018	1.000	1723
$M_2 = M_3$	38.107	0.000	1723

Table 4
GAS(1,1) model estimates.

	Coef.	Std. err.	*p*-value
κ_{μ_o}	−0.158	0.019	0.000
κ_{μ_s}	0.390	0.087	0.000
κ_{φ_o}	0.013	0.005	0.003
κ_{φ_s}	0.050	0.017	0.001
κ_{ρ}	0.563	0.249	0.011
κ_{ν}	−13.061	0.000	0.000
a_{σ_o}	0.143	0.013	0.000
a_{σ_s}	0.077	0.012	0.000
a_{ρ}	0.093	0.029	0.001
b_{σ_o}	0.992	0.004	0.000
b_{σ_s}	0.958	0.013	0.000
b_{ρ}	0.629	0.162	0.000
AIC	19389.352		
BIC	19454.774		
np	12.000		
LogL	9682.676		

Table 5
DCC-GARCH(1,1) model estimates.

	Coef.	Std. err.	*p*-value
μ_o	−0.156	0.023	0.000
ω_o	0.014	0.006	0.022
α_o	0.202	0.031	0.000
β_o	0.797	0.034	0.000
γ_o	3.4634	0.168	0.000
μ_s	0.361	0.090	0.000
ω_s	0.954	0.263	0.000
α_s	0.135	0.022	0.000
β_s	0.821	0.028	0.000
γ_s	7.998	1.445	0.000
a_{os}	0.023	0.010	0.025
b_{os}	0.947	0.031	0.000
γ_{os}	4.696	0.234	0.000
Information criteria			
Akaike	11.325		
Bayes	11.370		
Shibata	11.325		
Hannan-Quinn	11.342		
LogL	−9742.888		

models. Initially, while the time-varying volatilities and correlations may adequately explain the return series variations, further inspection suggests this is not the case. For instance, the DCC-GARCH (1,1) model produces excessive volatility estimation for oil returns, especially during the early parts of the sample period for which the oil market was initially characterised by high price volatility due to supply-side factors. Similarly, we notice excessive volatility in stock price returns during the Great Depression of the late 1920s and early 1930s. In contrast, the GAS(1,1) model is more reasonable in capturing return and volatilities in the series since it produces less frequent fluctuations and smaller magnitudes. The reason is that the scaled score of the t-distribution causes the return and volatility dynamics to not respond to large variations of r_t. In other words, large fluctuations in r_t are due to the heavy-tailed distribution of the data and should not be entirely attributed to increases in the variance.

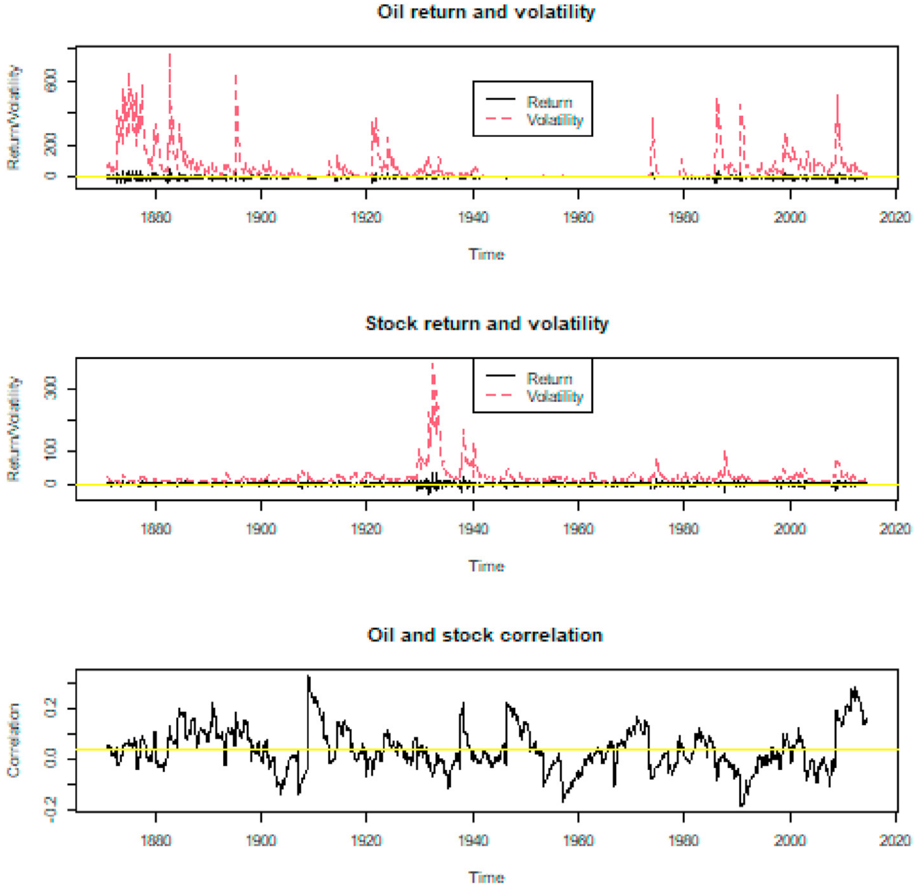

Fig. 2. Volatility and correlation from DCC-GARCH(1,1) model (in-sample).

5. Forecasting the oil-stock relationship

We now undertake the forecasting exercise to investigate the volatilities and dynamic conditional correlations between crude oil returns and stock price returns using the two models. We choose our out-of-sample forecasting horizon to be up to 75 months to capture the recent extreme events, including the oil price plunge of 2014–2016 and the COVID-19 pandemic. To assess the forecasting accuracy of the two models, we compute two measures of realized volatility and correlation widely used in the empirical literature. The first is the square of monthly returns as a measure for realized volatility. The second is realized correlations, calculated by multiplying the two returns sequences simultaneously between crude oil and stock prices (see, e.g., Ferland and Lalancette, 2006; Lopez, 2001).

To compare the forecasting capability of the DCC-GARCH(1,1) and GAS(1,1) models on the out-of-sample period, we follow Chen and Xu (2019) and calculate the mean absolute error (MAE) and the root mean square error (RMSE) as follows:

$$MAE_{\sigma^2} = \frac{\sum_{t=1}^{N} \left| \widehat{\sigma}_t^2 - \sigma_t^2 \right|}{N}, MAE_{\rho} = \frac{\sum_{t=1}^{N} \left| \widehat{\rho}_t - \rho_t \right|}{N},$$

$$RMSE_{\sigma^2} = \sqrt{\frac{\sum_{t=1}^{N} \left(\widehat{\sigma}_t^2 - \sigma_t^2 \right)^2}{N}}, RMSE_{\rho} = \sqrt{\frac{\sum_{t=1}^{N} \left(\widehat{\rho}_t - \rho_t \right)^2}{N}},$$

where N is the consecutive out-of-sample observations. $\widehat{\sigma}_t^2$, $\widehat{\rho}_t$ are the rolling forecasts of the volatility and correlations of time t from the two models, while σ_t^2 and ρ_t are the realized volatilities and correlations at time t. Determining the forecasting accuracy of the DCC-GARCH(1,1) and GAS(1,1) models is straightforward – the lower the metric of MAE or RMSE, the better the forecasting ability of the model (i.e., the evaluation criterion).

The MAE and RMSE estimates are reported in Table 6. The results suggest that the forecasting ability of volatility and dynamic correlations for the GAS(1,1) model is preferred to those of DCC-GARCH(1,1). However, while the GAS(1,1) model improves the forecasting efficiency for less than 20 horizon months, as the horizon months increase beyond 20, the forecasting performance of the GAS(1,1) model is close to those of DCC-GARCH(1,1) model. Interestingly, Fig. 4 shows that the MAE for stock returns from the

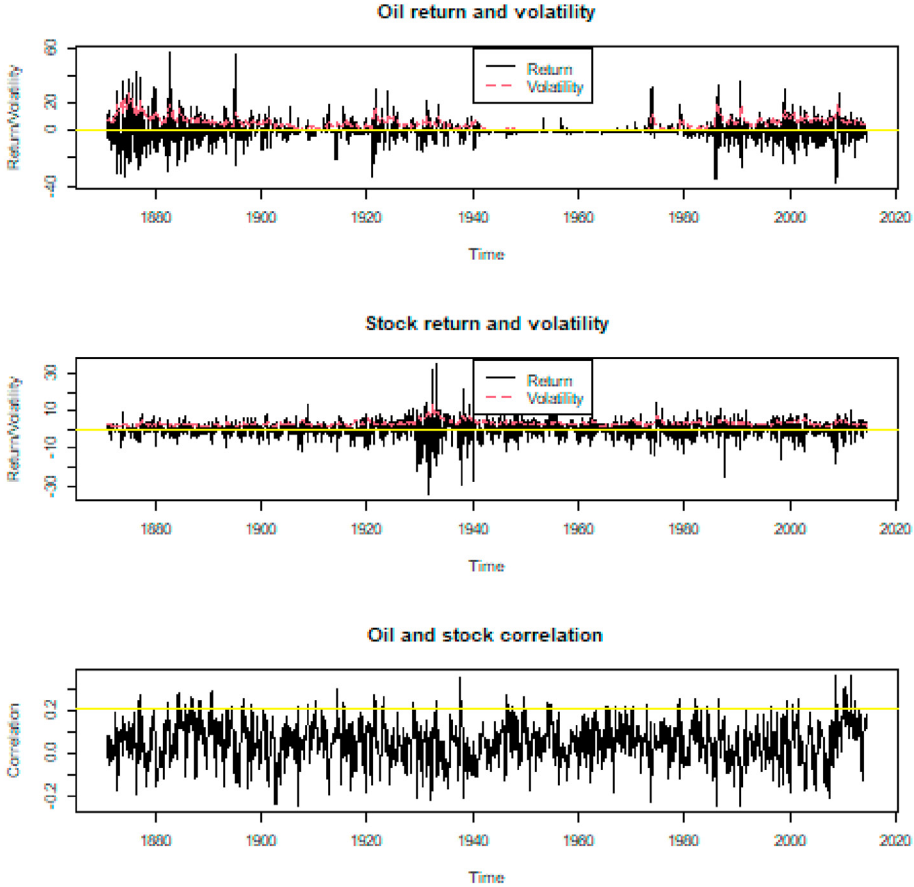

Fig. 3. Volatility and correlation from GAS(1,1) model (in-sample).

Table 6
MAE and RMSE results.

Month Ahead	MAE						RMSE					
	DCC-GARCH(1,1)			GAS(1,1)			DCC-GARCH(1,1)			GAS(1,1)		
	σ_O^2	σ_S^2	ρ_{OS}	σ_O^2	σ_S^2	ρ_{OS}	σ_O^2	σ_S^2	ρ_{OS}	σ_O^2	σ_S^2	ρ_{OS}
1	27.165	23.466	0.221	18.631	11.225	0.003	27.165	23.466	0.221	18.631	11.225	0.003
5	191.329	28.228	0.293	172.004	7.156	0.286	228.126	32.981	0.325	197.991	8.037	0.359
10	251.498	40.649	0.175	231.046	9.437	0.165	290.649	44.779	0.236	250.954	10.314	0.257
20	209.755	34.076	0.225	214.616	15.676	0.225	253.715	38.881	0.298	241.015	18.773	0.293
30	157.363	30.074	0.195	194.498	14.203	0.193	210.702	34.468	0.274	221.280	17.125	0.271
40	123.680	26.709	0.166	157.482	12.431	0.169	182.878	31.123	0.241	193.326	15.380	0.240
50	108.021	24.786	0.187	135.301	12.613	0.172	165.253	29.116	0.258	174.359	15.772	0.256

GAS(1,1) model is consistently lower than that of GARCH(1,1) for the entire forecast horizon. Thus, in terms of the criterion of MAE, the GAS(1,1) model provides more accurate forecasts compared to that of GARCH (1,1) at least at lower horizon months, and this result is also confirmed in Fig. 4.

We continue with our forecasting exercise and conduct additional comparisons between the two models. We use the estimated results from the previous section to forecast volatilities and dynamic correlations for the out-of-sample period. To preserve space, we only present the MAE measure. While Fig. 4 previously suggested that the GAS(1,1) model performs better at lower horizon months, Figs. 5 and 6 also confirm this finding. These figures illustrate that differences exist in the rolling forecast of volatility and correlations between the return series. Again, we observe the DCC-GARCH(1,1) model produces a larger forecasting error than the GAS(1,1), as seen in Fig. 5. For example, the lower panel of Fig. 5 shows that the GAS(1,1) estimates closely track the stock return square for much of the out-of-sample forecast period. In contrast, the GARCH(1,1) estimates are way higher than the observed return series until mid-2018 and have become consistently lower since then. More importantly, Fig. 5 shows a large spike in the oil and stock return series during the COVID-19 pandemic. In comparing the performance of the two models, the GAS(1,1) model captures the spike relatively better than the

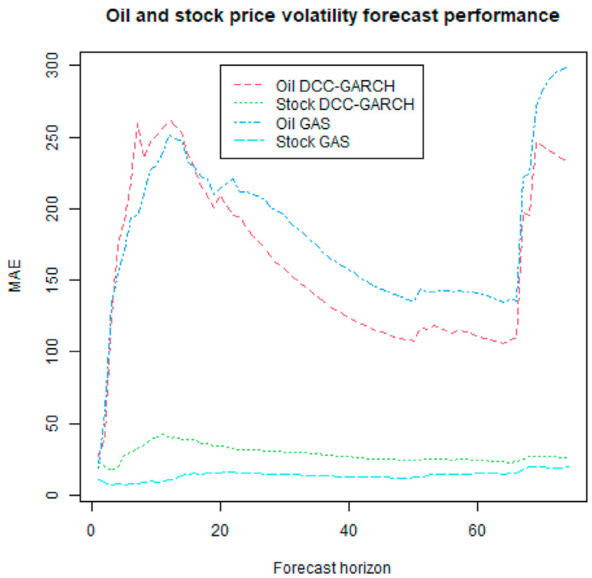

Fig. 4. Volatility forecast performance (out-of-sample).

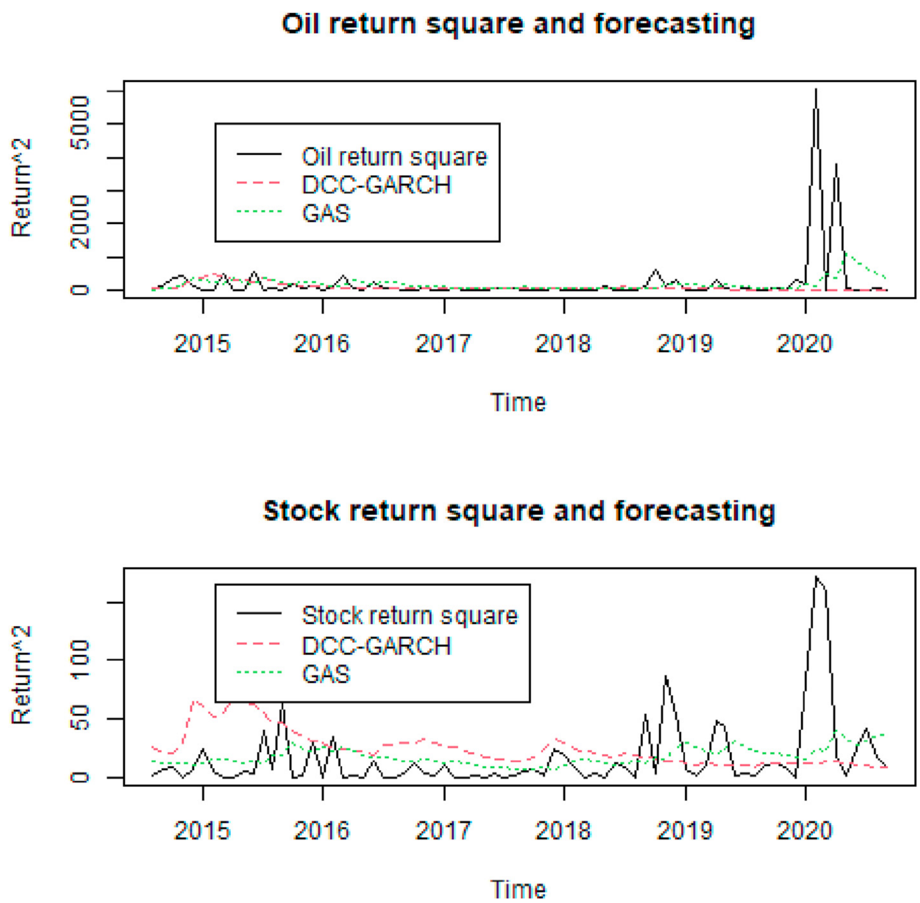

Fig. 5. DCC-GARCH and GAS volatility forecast comparison.

GARCH(1,1) model. Similar findings are illustrated in Fig. 6, where the MAE of the GAS(1,1) model relatively lower than that of the GARCH(1,1) model for most of the forecast horizon. However, both models provide similar forecasting performance beyond 60 horizon

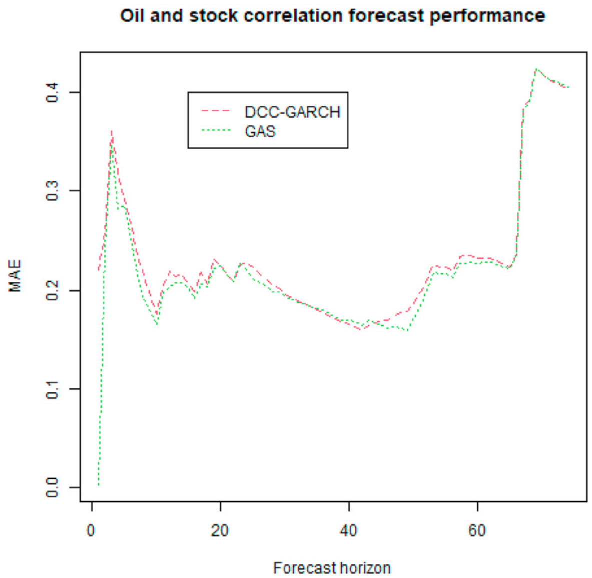

Fig. 6. Correlation forecast performance (out-of-sample).

months.

To sum up, we find evidence that the GAS model outperforms the DCC-GARCH model in various cases. Nonetheless, while there is no strict underlying theory to suggest that the GAS model is superior to the DCC-GARCH model, our empirical results indicate that the GAS(1,1) produces better forecasting results than the DCC-GARCH(1,1) model, at least for early horizon months. Our results support the findings from Chen and Xu (2019) who also find that the forecasting power of volatility and correlation in the multivariate GAS model is better than the DCC-GARCH model.

6. Conclusion

The role of oil price on stock returns has become a familiar subject of research in the financial and energy economics literature. However, little empirical work has focused on their interplay over historical periods and forecasting performance. Indeed, Narayan and Gupta (2015, p. 18) state that using both in-sample and out-of-sample analysis "is of paramount importance, since the existence of in-sample predictability does not necessarily ensure out-of-sample forecasting gains". Accordingly, we model and forecast the historical relationship between WTI oil and the S&P500 stock price index since the late 19th century. In doing so, we first apply the novel multivariate GAS(1,1) model to estimate and forecast the time-varying volatility and dynamic correlations and examine whether this model produces better forecasting predictivity than those of the standard DCC-GARCH(1,1) model.

Given this framework, our results can be summarised as follows. Initial testing suggests that the volatility and correlation parameters are time-varying, thus providing evidence of applying the GAS model with a conditional t-distribution. According to the in-sample estimation, we find the GAS(1,1) performs better in capturing the strong volatility persistence and the non-linear interactions between crude oil and stock price returns. We find the DCC-GARCH(1,1) model produces excessive volatility estimation for oil returns, particularly during the early parts of the sample period characterised by supply-side factors and the Great Depression of the late 1920s. Also, the dynamic conditional correlations are larger and changing more frequently in the GAS(1,1) model as opposed to the DCC-GARCH(1,1) model, which is also confirmed by the estimated model parameters.

It is obvious to see the dynamic correlations between crude oil and stock returns to rise during turbulent events over the sample period significantly. Such events caused by financial crises, oil price shocks, the recent COVID-19 pandemic or external events (i.e., geopolitical tension and wars) alter the behaviour of the relationship over time. These phenomena may point to a dependant pattern between the two series over time and provide useful information for investors, portfolio managers, and market participants (Smyth and Narayan, 2018).

Using the out-of-sample period to evaluate both models' forecasting performance, we utilise the mean absolute error (MAE) and the root mean squared error (RMSE) criterion to compare the forecasting ability. The results once again are in favour of the GAS(1,1) model over the DCC-GARCH(1,1) in terms of forecasting the time-varying volatility and dynamic correlations. However, while the forecasting efficiency of the GAS(1,1) performs better, it does so over the early horizon months. As the forecasting months increase over the out-of-sample period, the forecasting performance of the GAS(1,1) model produces similar results to that of the DCC-GARCH(1,1) model.

The results of this study open potential avenues for future research. First, a natural extension is to investigate the asymmetries between crude oil and stock returns over historical periods. Indeed, Narayan and Gupta (2015) suggest both positive and negative oil price changes are important predictors of US stock returns. However, they do not consider the asymmetric effect on the time-varying components (i.e., on volatility and dynamic correlations) using both GAS and DCC-GARCH models. Second, while our study examines aggregate stock returns, it will be interesting to investigate the time-varying relationship between crude oil and sector stock returns over historical periods similar to Degiannakis et al. (2013). Sectoral decomposition may reveal important information regarding the time-varying relationship and forecasting accuracy, given that industries respond differently to oil price movements (or shocks). It may also provide valuable information for the purposes of hedging or portfolio allocation. However, the challenge here is to obtain sectoral stock return data back to the 19th century.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Antonakakis, N., Kizys, R., 2015. Dynamic spillovers between commodity and currency markets. Int. Rev. Financ. Anal. 41, 303–319.
Ardia, D., Boudt, K., Catania, L., 2019. Generalized autoregressive score models in R: the GAS package. J. Stat. Software 88 (6), 1–28.
Arouri, M.E.H., Jouini, J., Nguyen, D.K., 2011. Volatility spillovers between oil prices and stock sector returns: implications for portfolio management. J. Int. Money Finance 30 (7), 1387–1405.
Avdulaj, K., Barunik, J., 2015. Are benefits from oil–stocks diversification gone? New evidence from a dynamic copula and high frequency data. Energy Econ. 51, 31–44.
Awaworyi Churchill, S., Ivanovski, K., Smyth, R., 2018. The environmental Kuznets curve in the OECD: 1870–2014. Energy Econ. 75, 389–399.
Basher, S.A., Haug, A.A., Sadorsky, P., 2012. Oil prices, exchange rates and emerging stock markets. Energy Econ. 34 (1), 227–240.
Baumeister, C., Peersman, G., 2013. Time-varying effects of oil supply shocks on the US economy. Am. Econ. J. Macroecon. 5 (4), 1–28.
Broadstock, D.C., Filis, G., 2014. Oil price shocks and stock market returns: new evidence from the United States and China. J. Int. Financ. Mark. Inst. Money 33, 417–433.
Campbell, J.Y., 2008. Estimating the equity premium. Canadian Journal of Economics/Revue canadienne d'économique 41 (1), 1–21.
Chen, C.-D., Cheng, C.-M., Demirer, R., 2017. Oil and stock market momentum. Energy Econ. 68, 151–159.
Chen, R., Xu, J., 2019. Forecasting volatility and correlation between oil and gold prices using a novel multivariate GAS model. Energy Econ. 78, 379–391.
Choudhry, T., Hassan, S.S., Shabi, S., 2015. Relationship between gold and stock markets during the global financial crisis: evidence from nonlinear causality tests. Int. Rev. Financ. Anal. 41, 247–256.
Creal, D., Koopman, S.J., Lucas, A., 2013. Generalized autoregressive score models with applications. J. Appl. Econom. 28 (5), 777–795.
Degiannakis, S., Filis, G., Floros, C., 2013. Oil and stock returns: evidence from European industrial sector indices in a time-varying environment. J. Int. Financ. Mark. Inst. Money 26, 175–191.
Doornik, J.A., Hansen, H., 2008. An omnibus test for univariate and multivariate normality. Oxf. Bull. Econ. Stat. 70, 927–939.
Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica: Journal of the Econometric Society 987–1007.
Engle, R.F., 2002. Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models. J. Bus. Econ. Stat. 20 (3), 339–350.
Ferland, R., Lalancette, S., 2006. Dynamics of realized volatilities and correlations: an empirical study. J. Bank. Finance 30 (7), 2109–2130.
Filis, G., 2010. Macro economy, stock market and oil prices: do meaningful relationships exist among their cyclical fluctuations? Energy Econ. 32 (4), 877–886.
Fiorentini, G., Sentana, E., Calzolari, G., 2003. Maximum likelihood estimation and inference in multivariate conditionally heteroscedastic dynamic regression models with Student t innovations. J. Bus. Econ. Stat. 21 (4), 532–546.
Gupta, R., Wohar, M., 2017. Forecasting oil and stock returns with a Qual VAR using over 150 years off data. Energy Econ. 62, 181–186.
Hailemariam, A., Smyth, R., Zhang, X., 2019. Oil prices and economic policy uncertainty: evidence from a nonparametric panel data model. Energy Econ. 83, 40–51.
Harvey, A., Ruiz, E., Sentana, E., 1992. Unobserved component time series models with ARCH disturbances. J. Econom. 52, 129–157.
Jacks, D.S., Meissner, C.M., Novy, D., 2011. Trade booms, trade busts, and trade costs. J. Int. Econ. 83 (2), 185–201.
Jones, C.M., Kaul, G., 1996. Oil and the stock markets. J. Finance 51 (2), 463–491.
Kilian, L., Park, C., 2009. The impact of oil price shocks on the US stock market. Int. Econ. Rev. 50 (4), 1267–1287.
Kollias, C., Kyrtsou, C., Papadamou, S., 2013. The effects of terrorism and war on the oil price–stock index relationship. Energy Econ. 40, 743–752.
Lin, B., Li, J., 2015. The spillover effects across natural gas and oil markets: based on the VEC–MGARCH framework. Appl. Energy 155, 229–241.
Lopez, J.A., 2001. Evaluating the predictive accuracy of volatility models. J. Forecast. 20 (2), 87–109.
Madsen, J., Ang, J., 2016. Finance-led growth in the OECD since the nineteenth century: how does financial development transmit to growth? Rev. Econ. Stat. 98 (3), 552–572.
Narayan, P.K., Gupta, R., 2015. Has oil price predicted stock returns for over a century? Energy Econ. 48, 18–23.
Narayan, P.K., Narayan, S., 2010. Modelling the impact of oil prices on Vietnam's stock prices. Appl. Energy 87 (1), 356–361.
Noguera, J., 2013. Oil prices: breaks and trends. Energy Econ. 37, 60–67.
Sadorsky, P., 2014. Modeling volatility and correlations between emerging market stock prices and the prices of copper, oil and wheat. Energy Econ. 43, 72–81.
Silvapulle, P., Smyth, R., Zhang, X., Fenech, J.-P., 2017. Nonparametric panel data model for crude oil and stock market prices in net oil importing countries. Energy Econ. 67, 255–267.
Smyth, R., Narayan, P.K., 2018. What do we know about oil prices and stock returns? Int. Rev. Financ. Anal. 57, 148–156.
Tafakori, L., Pourkhanali, A., Fard, F.A., 2018. Forecasting spikes in electricity return innovations. Energy 150, 508–526.
Tuysuz, S., 2013. Conditional correlations between stock index, investment grade yield, high yield and commodities (gold and oil) during stable and crisis periods. Int. J. Econ. Finance 5 (9), 28.
Zhu, H.-M., Li, S.-F., Yu, K., 2011. Crude oil shocks and stock markets: a panel threshold cointegration approach. Energy Econ. 33 (5), 987–994.