Research Paper

Modeling superior predictors for crude oil prices

Sjur Westgaard,[1] Petter Osmundsen,[2] Daniel Stenslet[1] and Jo Ringheim[1]

[1]Department of Industrial Economics and Technology Management,
Norwegian University of Science and Technology, NO-7041 Trondheim, Norway;
emails: sjur.westgaard@iot.ntnu.no, lostensl@stud.ntnu.no, jokr@stud.ntnu.no
[2]Department of Industrial Economics and Risk Management, University of Stavanger,
NO-4036 Stavanger, Norway; email: petter.osmundsen@uis.no

ABSTRACT

A common perception in the literature is that oil price dynamics are most adequately explained by fundamental supply-and-demand factors. We use a general-to-specific approach and find that financial indicators are even more significant at modeling and predicting oil prices. We demonstrate empirically that the futures spreads level, high-yield bond spreads and PHLX Oil Service Sector (OSX) index are the best predictors of oil prices in the period February 2000–June 2013. (The OSX index is designed to track the performance of a set of companies involved in the oil services sector.) The OSX index is particularly interesting, as no study has analyzed its predictive power prior to our analysis. The relationship is intuitively meaningful, as stock prices, which strongly depend on the oil price, are determined in a market with well-informed investors that have strong incentives to gather correct market information. Moreover, the share prices serve as strong proxies or price signals, as they reflect future oil price expectations at any point of time. Furthermore, we demonstrate through an out-of-sample analysis that our most parsimonious model is superior to relevant benchmarks at forecasting oil price changes (two benchmarks were used: (1) a random walk

---

and (2) ARIMA$(2, 0, 2)$, which was optimized in-sample by minimizing the Akaike information criterion). Our findings do not necessarily imply that the financial sector determines oil prices. On the contrary, we take the view that fundamental information is traceable from financial markets, and, hence, financial predictors serve as indicators for oil price fundamentals.

1 INTRODUCTION

Oil is one of the most liquid and traded assets in the world, yet its price is subject to vast fluctuations and has a considerable impact on the global economy in various ways. The price fell dramatically (70%) in only a few months during the financial crisis in 2008. In the following years, the price recovered before falling even deeper from a level of US$112 per barrel in July 2014 to a twelve-year low of US$26 in January 2016. These extreme price fluctuations have evidently distressed international relations and political developments. The volatile price has put serious pressure on oil companies; it has also affected the economic stability in oil-dependent regions and caused considerable movements in financial markets. A great amount of research has been devoted to understanding the drivers and characteristics behind the movements. However, experts do not always agree when explaining the price dynamics.

The purpose of our research is threefold. First, we aim to determine the most important predictors for crude oil prices with a prediction horizon of one month. We apply a general-to-specific model selection approach, starting very generally by collecting a large set of fundamental and nonfundamental indicators that are motivated by a comprehensive review of the literature, and by general perceptions in the financial markets. We categorize these indicators into supply factors, demand factors, financial factors, geopolitical proxies and price shocks. The first two cover fundamental oil variables that reflect supply-and-demand dynamics. Geopolitical proxies and price shocks reflect political tension and unexpected events, respectively. Financial factors cover prices and measures that are determined by or traded in financial markets; much attention has been devoted to these in recent research. Some papers argue that the changes in oil prices over the past decade have been caused by financial factors (Cifarelli and Paladino 2010; Kaufmann 2011; Lombardi and Robays 2011). Others argue the opposite (Alquist and Gervais 2013; Fattouh *et al* 2013; Morana 2013), and some even suggest that financial factors could have a stabilizing effect on oil prices (Knittel and Pindyck 2013). We construct three pure prediction models that consist solely of lagged variables. Our analysis demonstrates that financial indicators, specifically futures spreads, high-yield bond spreads and the PHLX Oil Service Sector

(OSX) equity index, are the most significant predictors of changes in oil prices over the period February 2000–June 2013. The OSX index is particularly interesting, as no paper (to our knowledge) has analyzed its predictive power prior to our study.

Second, we aim to examine the validity of the prediction models. An out-of-sample analysis is performed where the models are compared with two benchmark autoregressive integrated moving average (ARIMA) models: a random walk and an in-sample optimized model. The models are evaluated based on their out-of-sample squared- and absolute error metrics. We find that our leanest model, consisting of three independent variables, delivers the best out-of-sample results. Consequently, this model is superior to the benchmarks at predicting oil prices in the out-of-sample period.

Finally, we evaluate the models' predictive power. Baumeister and Kilian (2016) claim that the drop in oil prices from mid-2014 was, to some extent, predictable. Our analysis supports their view, as our most parsimonious model is able to conjecture the price development to a certain extent. However, Baumeister and Kilian predict the changes using fundamental factors. Our model consists solely of financial factors; hence, it explains these effects differently.

This paper is organized as follows. Section 2 provides an overview of the relevant literature, while Section 3 offers a description of our data. In Section 4, we introduce the methodology. The results are presented and discussed in Section 5 and the conclusion is provided in Section 6.

2 OIL MARKET INDICATORS

In this section, we provide a thorough overview of explanatory oil price indicators and their studied and concluded importance. Here, supply-and-demand factors are referred to as fundamentals; a manifold of these will be discussed briefly. Furthermore, several nonfundamental indicators have been examined in the literature. These are categorized below as financial factors, political proxies and price shocks. The predictors included in our analysis are motivated by the relationships and findings below.

2.1 Demand factors

Oil is perhaps the most important source of energy in the world, accounting for roughly one-third of the world's total energy consumption (International Energy Agency 2016). It is generally assumed that oil prices and global economic growth are closely linked. For instance, Kilian and Hicks (2013) claimed that most of the oil price surge between 2003 and 2008 could be explained by unexpectedly strong gross domestic product (GDP) growth in emerging economies. Hamilton (1996, 2008) and Radetzki (2006) found a significant positive relationship between global GDP and oil prices. However, some papers have argued that GDP offers poor forecasting accuracy (Kesicki 2010; Ye *et al* 2006); Hooker (1996) even found strong evidence that

oil prices and macroeconomic indicators are unrelated. Another examined proxy for oil demand is the Kilian index (Kilian 2009), which is based on dry cargo freight rates and designed to capture changes in global demand for industrial commodities. Changes in the index are academically recognized to be positively correlated with oil prices (Baumeister and Kilian 2016; Broadstock and Filis 2014; He *et al* 2010; Kaufmann 2011; Kilian and Murphy 2014). Coleman (2012) studied another measure, the Organization for Economic Co-operation and Development (OECD) import dependence, which is based on the intuition that the Organization of the Petroleum Exporting Countries' (OPEC's) pricing power increases when OECD import quantum surges, and found a significant positive relationship with oil prices. Unemployment, inflation and monetary policies are other macroeconomic measures discussed in relation to oil prices. Leduc and Sill (2004) studied oil prices, monetary policies and economic downturns. Their conclusion was that the way monetary policies are conducted could possibly both amplify and mitigate the recessionary effects caused by high oil prices. A great amount of research concerns the macroeconomic impact caused by changes in oil prices, where increases in oil prices have been assigned the responsibility for economic downturns, periods of extreme inflation, low productivity and slow economic growth (Cunado and de Gracia 2005; Gupta and Goyal 2015; LeBlanc and Chinn 2004). However, Barsky and Kilian (2004) questioned the causal relationship between oil prices and economic growth, arguing that oil prices can be viewed as exogenous in relation to the US economy. This view was supported by Askari and Krichene (2010), who also argued that increased government spending would stimulate the economy, and thereby increase the global demand for oil and lift the oil price.

2.2 Supply factors

The behavior of OPEC is often utilized when oil supply characteristics are explained. OPEC accounts for more than 40% of daily production and has close to three-quarters of the global oil reserves (International Energy Agency 2016). Rig activity is associated with oil companies' investments and, hence, is assumed to indicate their long-term production levels. Ringlund *et al* (2008) found a clear long-term negative relation between oil prices and the activity of oil rigs in non-OPEC regions. Similarly, Olimb and Oedegaard (2010) investigated the relationship between changes in rig activity and oil prices, finding that global rig utilization had a significant negative impact on oil prices the following month. Changes in crude oil production have also been well examined, where an increase is associated with a negative pressure on prices (Baumeister and Kilian 2016; Hamilton 2003; He *et al* 2010). Baumeister and Kilian (2016) argued that the recent increase in US shale production has definitely put a downward pressure on oil prices. Kaufmann (2011) examined the supply allocation

between OPEC and non-OPEC, stating that changes in this measure will have an important effect on the price due to the view of OPEC being the marginal supplier. A similar view was provided by Chevillon and Rifflart (2009).

The role of inventories has been frequently discussed in the literature. The perception is generally that changes in inventories are negatively correlated with oil prices. That is, an increase in the inventory levels is assumed to be related to a fall in oil prices. According to Ye *et al* (2005), inventory levels measure the supply-and-demand balance or imbalance, and hence capture changing market conditions explicitly. Ye *et al* presented a model for forecasting short-term oil prices using only lagged OECD petroleum inventory levels as explanatory variables, providing a well-fitted forecast both in-sample and out-of-sample. Ryan and Whiting (2016) found that the OECD inventory level lagged by two months had a significant relationship with oil prices. Inventory levels are generally accepted as an important fundamental factor when explaining oil price dynamics (Hamilton 2008; He *et al* 2010; Kilian and Murphy 2014; Merino and Ortiz 2005). Another measure related to inventory levels, which was evaluated by Olimb and Oedegaard (2010), named days of stock covered. The measure corresponds to the number of days that OECD inventories could supply their demand, where a decrease was demonstrated to have a positive relation with oil prices. As crude oil needs to be refined to final products, researchers have considered the refinery activity, or utilization, as a potential predictor for oil prices. Dees *et al* (2008) presented a forecasting model based on refinery activity and OPEC spare capacity, which performed well relative to models based on futures prices. Coleman (2012) examined the OPEC market share, yet another proxy for supply. He argued that OPEC's oil production accounts for a large portion of the global oil supply, and, hence, an increase in the market share would further increase its market power. A significant negative relationship between OPEC market share and oil prices was found.

2.3 Financial factors

Several nonfundamental indicators have been found to significantly influence oil prices. We call these financial factors, since they cover a wide range of prices and measures that are determined by or traded in financial markets. Futures contracts, exchange rates, stock indexes, interest rates, bond spreads, market volatility and other commodity prices are all examples of such factors. The growing number of financial participants in the oil market and their potential influence on the price have received much attention in the literature recently. It is frequently assumed that an increase in the number of traded futures contracts could trigger an uplift of oil prices, or that futures prices themselves reflect price expectations more accurately and thus can drive oil prices. In either case, Coppola (2008) concluded that spot and futures prices have a co-integrated relationship. Some researchers argue that many of the price changes

over the past decade are the result of an increase of participants in nonphysical trad-ing contracts (Cifarelli and Paladino 2010; Kaufmann 2011; Lombardi and Robays 2011), while others argue the opposite (Fattouh *et al* 2013). The latter standpoint is supported by Knittel and Pindyck (2013), who suggested that financial speculation could have a stabilizing effect, and by Morana (2013), who found that oil prices have been driven mainly by macroeconomic shocks, not financial speculation. Moreover, Alquist and Gervais (2013) applied Granger causality tests and concluded that finan-cial speculation has little or no effect on oil price changes. Other financial indicators are futures spreads or ratios, which are the difference and ratio, respectively, between the price of a futures contract and the spot price. Dees *et al* (2008) constructed a well-performing one-month forecasting model and showed that the change in the difference between four-month and one-month futures contracts (ie, a one-month contract minus a four-month contract) had a significant positive relationship with oil prices. Olimb and Oedegaard (2010) constructed the same measure by using a six-month contract minus a one-month contract, and found a significant negative relationship with oil prices. This is effectively the same conclusion. Alquist and Kilian (2010) suggested that the negative relationship between futures spreads and oil prices could indicate that oil price fluctuations are driven by precautionary demand.

Market volatility is commonly used to explain the risk (or fear) level among investors. The Chicago Board Options Exchange's Volatility Index (VIX) measures the implied volatility of Standard and Poor's 500 (S&P 500) index options, and it is generally assumed that the stock market tends to fall when the VIX increases. For instance, Sarwar (2012) found a strong negative correlation between changes in VIX and US stock market returns. Ryan and Whiting (2016) found that the VIX, lagged with one month, was the most important indicator in terms of forecasting oil prices. A large amount of literature has examined the relationship between oil prices and the stock market (Bernanke 2016; Hammoudeh *et al* 2004; Kang *et al* 2015; Sadorsky 1999), generally explaining a significant positive correlation between the two. How-ever, we do not find any studies conducted on the stock market's ability to forecast oil prices. Changes in oil prices have also been explained by exchange rates. Since the oil price is traded in dollars, the rationale is that an appreciation of the dollar makes it more expensive for market participants in other currencies to buy oil and, hence, reduces demand outside of the United States. However, Alquist and Gervais (2013) found that there is no evidence of a systematic relationship between the US exchange rate and the price of oil. This conclusion was further discussed by Baumeister and Kilian (2016), who argued that an appreciation of the dollar also stimulates exports outside the United States and thus increases supply, finally reducing the exchange rate effect. However, Chen *et al* (2010) found that commodity currency exchange rates such as the US dollar have surprisingly well-performing predictive power for

global commodity prices. A similar conclusion was provided by Pindyck and Rotemberg (1988), who argued that an equally weighted index of the dollar value of British, German and Japanese currencies has a significant negative relationship with oil prices.

Changes in interest rates have also been found to influence oil prices. Hamilton and Kim (2000) discussed various factors that influence GDP, finding that both bond spreads and changes in shorter-rate factors are good predictors. Although the relationship between interest rates and macroeconomic measures has been widely discussed and evaluated in academia, few analyses have been directly concerned with the relationship between interest rates and oil prices. Coleman (2012) measured the difference between Moody's Baa and Aaa corporate bond yield indexes. Although no significant relationship was found, it is generally argued that high-yield bond spreads reflect the strength of the economy by representing investors' willingness to take on risk (Gertler and Lown 1999). Other commodities can explain oil prices as well, and Baumeister and Kilian (2016) discussed the potential relationships between oil prices and industrial commodities. Saghaian (2010) found significant correlations between oil prices and other commodities, with no clear indication of a Granger causal relationship. Baumeister *et al* (2013) presented yet another approach to the forecasting of oil prices, using changes in crack spreads. This approach was based on the assumption that refiners choose to reduce production when refined products are traded at low prices relative to the actual crude price. Murat and Tokat (2009) constructed a forecasting model for oil prices using only crack spread futures, finding a causal impact of crack spread futures on crude oil prices in both the short run and the long run.

2.4 Political proxies and price shocks

Hamilton (2003) examined the impact of several exogenous disruptions in global oil supply. He found that wars taking place in oil-producing regions caused supply to decrease and oil prices to rise. Coleman (2012) incorporated numerous proxies for geopolitical imbalance, with the objective of capturing political tension. Specifically, he included the number of deaths due to terrorist attacks and the number of US military troops deployed in the Middle East. He found that both measures had a significant impact. Alhajji and Huettner (2000) also investigated political factors in the oil market in order to determine whether a dominant producer exists. The conclusion was that neither OPEC nor other core producers could be characterized as dominant producers. Price shocks are another nonfundamental factor. These are infrequent and unexpected events that affect oil prices. The literature acknowledges their existence and importance in understanding price changes, but shocks bring limited value to forecasting models. Frequently used shock variables are natural disasters, wars, terrorist attacks and political events (Coleman 2012; Hamilton 2008; Kilian and Lee 2014; Olimb and Oedegaard 2010; Ye *et al* 2006).

Our review of the literature shows that there exist numerous proxies for supply and demand, and that financial factors have received more attention and have been increasingly used in research during the last decade. Thus, it will be of great interest to evaluate the predictive power of both fundamental and nonfundamental variables using a general-to-specific framework.

3 DATA

We utilize a data set consisting of thirty-two time series for further analysis, including the Brent crude spot price. The data set is chosen based on the general-to-specific modeling approach presented by Campos *et al* (2005), which involves a comprehensive yet well-justified set of data being collected. For the sake of comprehensiveness, we have chosen to include a large set of determinants that reflect all significant relationships presented in the literature overview (see Section 2). The data set is listed in Table 1 and consists of time series of monthly arithmetic changes and absolute values. The sample period is February 2000–June 2016, corresponding to a sample size of 197 observations.[1]

4 METHODOLOGY

We construct three models: these are all pure prediction models and, hence, limited to lagged variables. The models are selected using a stepwise forward model selection method, each with a unique termination criterion. Furthermore, an out-of-sample testing framework is formulated in order to evaluate the models' validity and predictive power.

We use the following model to forecast the Brent crude oil spot price:

$$
\begin{aligned}
y_t &= \beta_0 + \gamma_{1,1}\beta_{1,1}x_{1,t} + \cdots + \gamma_{1,v}\beta_{1,v}x_{1,t-v} + \gamma_{2,1}\beta_{2,1}x_{2,t} \\
&\quad + \cdots + \gamma_{n,v}\beta_{n,v}x_{n,t-v} + u_t \\
&= \beta_0 + \sum_{k=1}^{n}\sum_{p=1}^{v}\gamma_{k,p}\beta_{k,p}x_{k,t-p} + u_t,
\end{aligned}
\tag{4.1}
$$

where y_t is the crude oil price at time t, and $x_{k,t-p}$ are the variables based on k different time series and p different lags. n and v are the number of variables and lags, respectively; $\gamma_{k,p} \in \{0, 1\}$ are the binary weights determining whether variable $x_{k,t-p}$ is included in the model; and $u_t = N$.

---

[1] The sample period is limited by data availability. Certain time series from the Environmental Investigation Agency were only available from the year 2000 onward. In a trade-off between fewer time series with larger samples versus more time series with smaller samples, we choose the latter in order to follow the general-to-specific approach.

TABLE 1 List of variables and literature hypothesis.

Variable	Description	Hypothesis
Oil	Europe Brent spot price	
GDP	Weighted geometric mean of world GDP	↑
Kilian	Kilian index of global economic activity	↑
Unem	US unemployment	↑
US_Imp	US net imports of petroleum products	↓
MS	US money supply	↑
Infl	US inflation	↑
G_Rig	Global rigs in operation	↓
US_Rig	US rigs in operation	↓
OECD_Inv	OECD petroleum inventories	↓
US_Inv	US petroleum inventories	↓
D_Stock	OECD inventories divided by OECD daily consumption	↕
Ref	US refinery utilization	↕
OPEC_SC	OPEC spare production capacity	↓
G_Prod	Global oil production	↓
N_OPEC_Prod	Non-OPEC oil production	↓
OPEC_Prod	OPEC oil production	↓
Saudi_Prod	Saudi oil production	↓
OPEC_MS	OPEC oil production market share	↕
F_Spread	Futures spread NYMEX six million versus one million oil futures spread ratio: six million/one million are used due to their liquidity	↕
VIX	Volatility Index	↓
Exch	Dollar exchange rate index	↕
R	Three-month US Treasury bill rate	↓
BB_Spread	Yield of the BofA Merrill Lynch US Corporate BB Index	↓
Metals	Metal commodity index	↑
Crack_Spread	Crack spread	↑
SP500	S&P 500 equity index	↑
Gold	Gold price	↑
OSX	Equity index with oil companies	↑
N_Gas	Natural gas price	↑
Agricul	Agriculture commodity index	↑
Terror	Monthly number of deaths due to terror attacks in the Middle East	↑

Data descriptions and sources are shown in Table A.6 (online). The rightmost column illustrates the expected change in oil prices in response to a positive change in the respective variable. NYMEX denotes New York Mercantile Exchange and BofA denotes Bank of America.

4.1 Model selection

The regression estimators are not valid if a time series is nonstationary. Hence, we employ an augmented Dickey–Fuller test at 5% significance for each time series. If a time series turns out to be nonstationary, the time series of arithmetic changes will be replaced by a time series of logarithmic changes, and a new test will be carried out. If this transformation turns out to be unsuccessful, the time series is excluded from further analysis (we avoid first differentiating).

Our model incorporates six lags of each variable, corresponding to a lag horizon of six months. The selected period is no larger than this in order to avoid any further reduction of the degrees of freedom and, more importantly, to reduce the risk of including false significant variables in the model. This risk arises when relationships are not backed by economic fundamentals, but rather are a consequence of a data set with many variables. Thus, our model includes thirty-one independent time series with six lags each, yielding 186 potential explanatory variables.

We utilize a stepwise forward model selection approach, which involves adding the variable that maximizes the fit of the model at each step. We thus eliminate the issue of having a large processing time, which would arise if a best subset analysis was to be carried out. An advantage of using the stepwise selection is that it deals with the problem of multicollinearity. As correlated predictors capture much of the same movements in the response variable, they could show up with low joint significance in the model. When there is a large number of potential explanatory variables and observations, a too-relaxed termination criterion of the selection procedure could overestimate the amount of explained variance in the data, leading to overfitting of the model (see Campos *et al* 2005; Harrell 2001; Moutinho and Huarng 2013; Olejnik *et al* 2000). Consequently, the model will have a much better fit in-sample than out-of-sample. It is thus suggested that a stricter termination criterion be used in order to solve the issue. In the $Alpha_1$ model, we formulate a rather conservative termination criterion of a maximum significance level of 1% for all variables selected. This means that, at each time step, the procedure is terminated if one of the included explanatory variables has a significance level above 1%. For $Alpha_5$ and $Alpha_{10}$, the criteria are relaxed somewhat to 5% and 10%.

A Breusch–Godfrey test for autocorrelation is performed on the three models. Next, the White test is utilized in order to evaluate the heteroscedasticity of the models. Finally, we employ the Ramsey Regression Equation Specification Error Test (RESET) to test for model misspecification.

4.2 Model evaluation

We evaluate two aspects of our selected models: the degree of overfitting and the predictive power.

Empirical models will inevitably capture idiosyncratic noise. Increasing the complexity of a model will increase its ability to fine-tune itself to fit this idiosyncratic noise. A model that captures too much noise is said to be overfitted. The ability of the model to capture the real underlying oil price dynamics is evaluated by comparing the prediction error in-sample versus out-of-sample. Specifically, we compare the mean squared error (MSE) and mean absolute error (MAE) in-sample and out-of-sample.[2] See Appendix B.2 (available online) for their respective definitions.

Hyndman (2010) recommends two classes of benchmarks for prediction: one "naive", such as random walk, and one standardized, such as a well-specified ARIMA model.[3] We evaluate the prediction errors generated by our models with the errors generated by the random walk and the optimal $ARIMA(p, d, q)$ model. We employ MSE, MAE, error variance and frequency of correct direction predicted when evaluating the models.

Determining a proper out-of-sample window is a trade-off between model accuracy and validation period sample size. The in-sample period of our choice is from January 2000 to June 2013, corresponding to 161 observations. We take the view that June 2013–June 2016 was a turbulent time for oil price, featuring a steadily fluctuating price regime followed by two considerable drops, before a transition into a recovery phase. Thus, it serves as a sufficient out-of-sample window; accordingly, the out-of-sample period will consist of thirty-six observations.

5 RESULTS

In this section, we begin by presenting the stationarity test results. We continue by presenting the selected models, discussing misspecification test results and significant variables, and interpreting their economic meaning. Then, we present and discuss the out-of-sample results. Finally, we compare our models with that obtained by Baumeister and Kilian (2016) and comment on their differences.

5.1 Stationarity

The augmented Dickey–Fuller test was performed on all time series. At a 10% significance level, nonstationarity was found for changes in global GDP. Hence, an additional test on logarithmic changes was performed, arriving at the same conclusion. Consequently, global GDP was excluded from further analysis.

---

[2] It is common in the forecasting literature to also illustrate the prediction error by mean absolute percentage error (MAPE). However, since our model explains changes in oil prices, the MAPE values may yield results that are of little use.

[3] Hyndman (2010) suggests using the auto.arima() package in R in order to find a "well-specified" ARIMA model. The package finds the optimal $ARIMA(p, d, q)$ model by identifying the one with the lowest AIC value (see Appendix B.1 online).

TABLE 2 Regression output.

Variable, $x_{k,t-p}$	Alpha$_1$		Alpha$_5$		Alpha$_{10}$	
	$\beta_{k,p}$	$P > \lvert t \rvert$	$\beta_{k,p}$	$P > \lvert t \rvert$	$\beta_{k,p}$	$P > \lvert t \rvert$
$\triangle$Kilian_5					−11.025	0.055*
$\triangle$Kilian_6					14.337	0.018**
$\triangle$Unem_2					−0.521	0.006***
$\triangle$G_Rig_4			−0.269	0.011**	−0.440	0.000****
$\triangle$OECD_Inv_1			−1.374	0.037**	−1.270	0.044**
$\triangle$OECD_Inv_2			1.487	0.022**	1.477	0.017**
$\triangle$OECD_Inv_4			−1.319	0.040**	−1.470	0.015**
$\triangle$Ref_4					−0.331	0.019**
$\triangle$OPEC_MS_3					0.629	0.030**
$\triangle$F_Spread_1	−1.063	0.000****	−0.929	0.000****	−1.038	0.000****
$\triangle$VIX_5					−0.909	0.026**
$\triangle$BB_Spread_1	−0.034	0.000****	−0.029	0.000****	−0.034	0.000****
$\triangle$OSX_1	0.220	0.002***	0.249	0.000****	0.230	0.001****
$\triangle$Terror_5					−0.006	0.012**
R^2	0.391		0.478		0.575	
R^2_{adj}	0.379		0.453		0.533	
B-G (p-stat)	0.214		0.105		0.024**	
White (p-stat)	0.264		0.892		0.641	
RESET (p-stat)	0.314		0.320		0.083*	

Significance level: **** = 0.001, *** = 0.01, ** = 0.05, * = 0.1. The table shows coefficients and significance levels for each selected variable in the respective prediction model. Each model consists solely of lagged variables, and the number included in the variable name denotes the lag (in months) for the respective variable. Model properties are shown in the lower part of the table.

5.2 Model selection

The stepwise forward model selection was carried out with three different termination criteria. The coefficients, significances and model properties from this are summarized in Table 2. Not surprisingly, stricter termination criteria yielded models with fewer explanatory variables. The most important variables are changes in the futures spread level, bond spreads and the OSX index, all lagged with one month. A rise in the futures spread level is associated with a fall in oil prices the following month. A widening of the high-yield bond spread has the same directional relationship. A rise in the OSX index is associated with an oil price increase the consequent month. Other significant explanatory variables are global rig utilization and OECD inventories, which are selected in the Alpha$_5$ model.

The negative sign on the futures spread coefficient supports the conclusions of Alquist and Kilian (2010), Dees *et al* (2008) and Olimb and Oedegaard (2010) that,

despite their common use as forecasts for the spot price, futures contracts tend to deliver inaccurate predictions. The negative coefficient implies that a decrease in the spot price relative to the futures price, corresponding to an increased spread, is generally followed by a decrease in the spot price the following month. Alquist and Kilian (2010) even demonstrate that futures prices tend to be less accurate at predicting oil prices than no-change forecasts; they further suggest that the negative relationship could indicate that oil price fluctuations are driven by precautionary demand. The predictive power offered by the bond spread is more intuitive and can be related to macroeconomics. A negative change in the bond spread indicates that the spread is tightening, hinting at a stronger economy that demands less risk premium. According to general demand perceptions (Hamilton 2008; Kilian and Murphy 2014), a stronger economy will consequently imply an increase in the demand for oil, driving oil prices upward. Similarly, a wider bond spread reflects a larger risk premium; this is often associated with a weaker economic outlook and, thus, less demand for oil. The strong positive correlation between the lagged OSX index and the oil price is interesting. A considerable amount of research has been devoted to explaining the relationship between the stock market and oil prices (Bernanke 2016; Hammoudeh *et al* 2004; Sadorsky 1999). However, none of these studies use stock market returns to predict oil prices. The OSX index is a price-weighted index composed of the common stocks of fifteen companies that provide oil drilling and production services, oil field equipment, support services and geophysical/reservoir services. The value of companies in this category is usually highly exposed to short-term oil price fluctuations, since their project portfolios typically have a fairly short time horizon and a high share of fixed cost, and the gearing is high. Thus, the lagged OSX indicator is intuitively meaningful, as stock prices reflect short-term oil price expectations. Oil service company investors have strong incentives, extensive expertise and access to leading sources of information, and can thus be better at evaluating forthcoming oil market dynamics.

The model with the largest set of variables has the best fit in-sample with an R^2_{adj} of 0.533. The Breusch–Godfrey test does not reveal significant autocorrelation for Alpha_1 and Alpha_5. It does, however, indicate that Alpha_{10} is autocorrelated at a 5% significance level. White's test does not provide evidence of heteroscedasticity in any model. The Ramsey RESET does not detect functional misspecification in Alpha_1 and Alpha_5, but it suggests that Alpha_{10} may be misspecified at a 10% significance level.

5.3 Out-of-sample evaluation

5.3.1 Overfitting

Model validity metrics are displayed in Table 3. Ex ante and ex post represent the out-of-sample and in-sample periods, respectively. We present normalized out-of-sample

TABLE 3 Model validation.

	Alpha$_1$		Alpha$_5$		Alpha$_{10}$	
	Ex ante	**Ex post**	**Ex ante**	**Ex post**	**Ex ante**	**Ex post**
MSE	0.0075	0.0049	0.0077	0.0043	0.0087	0.0036
MAE	0.0649	0.0574	0.0688	0.0518	0.0729	0.0485

Ex ante and ex post represent the out-of-sample and in-sample periods, respectively. MSE and MAE denote mean squared error and mean absolute error.

TABLE 4 Normalized model validation.

	Alpha$_1$		Alpha$_5$		Alpha$_{10}$	
	Ex ante	**Ex post**	**Ex ante**	**Ex post**	**Ex ante**	**Ex post**
MSE	0.0059	0.0049	0.0061	0.0043	0.0069	0.0036
MAE	0.0577	0.0574	0.0611	0.0518	0.0648	0.0485

This table shows values from Table 3, where the ex ante and ex post values are divided by the standard deviation of the oil price during their respective sample periods.

metrics in Table 4 in order to evaluate the validity of the models more accurately.[4] The ratios of the ex ante to ex post MSE and MAE illustrate the degree of overfitting. The ratios are considerably higher for Alpha$_5$ and Alpha$_{10}$ than for Alpha$_1$, indicating that the degree of overfitting is increasing as the termination criterion is relaxed. Alpha$_{10}$ is the most complex model and has the best fit in-sample, corresponding to the lowest ex post MSE and MAE. The same model has the highest ex ante MSE and MAE, which is likely a consequence of being erroneously tuned to idiosyncratic noise during the in-sample calibration. As anticipated, the results also show that the Alpha$_5$ ratio lies between the other two models, both for MSE and MAE.

A parsimonious model is preferred relative to a complex one if it offers equal or better predictive power. Our data suggests that Alpha$_5$ and Alpha$_{10}$ are more overfitted than Alpha$_1$; hence, the latter is favorable, as it offers better predictive power out-of-sample.

5.3.2 Predictive power

The predictive power was measured by comparing the models with two benchmark ARIMA models – a random walk and an ARIMA$(2, 0, 2)$, optimized in-sample by minimizing the AIC – on four different tests. The results displayed in Table 5 show that Alpha$_1$ performs best on three out of four tests, namely MAE, MSE and error variance.

---

[4] MSE is normalized based on variance, while MAE is normalized based on standard deviation.

TABLE 5 Model prediction evaluation.

	Alpha$_1$	Alpha$_5$	Alpha$_{10}$	ARIMA(0,1,0)	ARIMA(2,0,2)
MSE	0.0075	0.0077	0.0086	0.0136	0.00882
MAE	0.0649	0.0688	0.0730	0.0744	0.0750
Error variance	0.0075	0.0082	0.0091	0.0140	0.0089
%CD	0.5833	0.5556	0.5556	0.6111	0.6389

The selected models are compared with two benchmark models: ARIMA(1,0,1) (random walk) and ARIMA(2,0,2) (optimized in-sample). MSE and MAE denote mean squared error and mean absolute error. Error variance is the variance of errors between the respective model and the actual oil price. %CD represents the amount of time the models and the actual price move in the same direction.

Alpha$_5$ scores second best on the same three tests. The ARIMA$(2, 0, 2)$ model has the best frequency of correct direction predicted (% CD). Relative to Alpha$_1$, the optimized ARIMA model performs marginally better when predicting the direction of the price change, but somewhat weaker when explaining the magnitude of the movements. Alpha$_{10}$ does not perform as well as the two stricter Alpha models, but it is still better than random walk on three out of four tests. The random walk scores second best on % CD, but it makes large errors, which are reflected by the other indicators.

Figure 1 compares the predicted changes and the accumulated one-month forecasts for each of the models. Figure 2 illustrates both the stepwise and the cumulative sum of squared errors for every model. Alpha$_1$ and Alpha$_5$ appear to differ only marginally in terms of direction and magnitude predicted. Their predicted price developments are quite similar, and their monthly squared errors usually peak simultaneously. We observe that Alpha$_1$ generally displays the smallest error peaks, specifically after the end of 2014. The random walk follows the oil price with one-month lag, and it is therefore heavily punished during the turbulent oil price regime of early 2015, making a vast jump on the cumulative squared error plot. Roughly half of its cumulative error occurs at this point in time. Among all models, Alpha$_1$ has the lowest MSE, MAE and error variance, and thus appears to have the best predicting power.

Alpha$_1$ was able to conjecture the price decline from US\$112 in July 2014 to US\$85 in March 2016, and, after a slight recovery in mid-2015, further down to US\$74 in February 2016. Our results thus support what Baumeister and Kilian (2016) stated earlier: the vast decline in oil price could be partly explained by real-time information. They attribute the price decline starting in July 2014 to a negative shock in storage demand, which was either caused by a more positive outlook on oil production, a gloomier outlook on the global economy, or both. Furthermore, they assert that the even-steeper price decline occurring in December 2014 is consistent with negative demand shocks, which are reflected by an unexpectedly slowing global economy.

FIGURE 1   Monthly individual and accumulated out-of-sample forecasts, July 2013–June 2016.
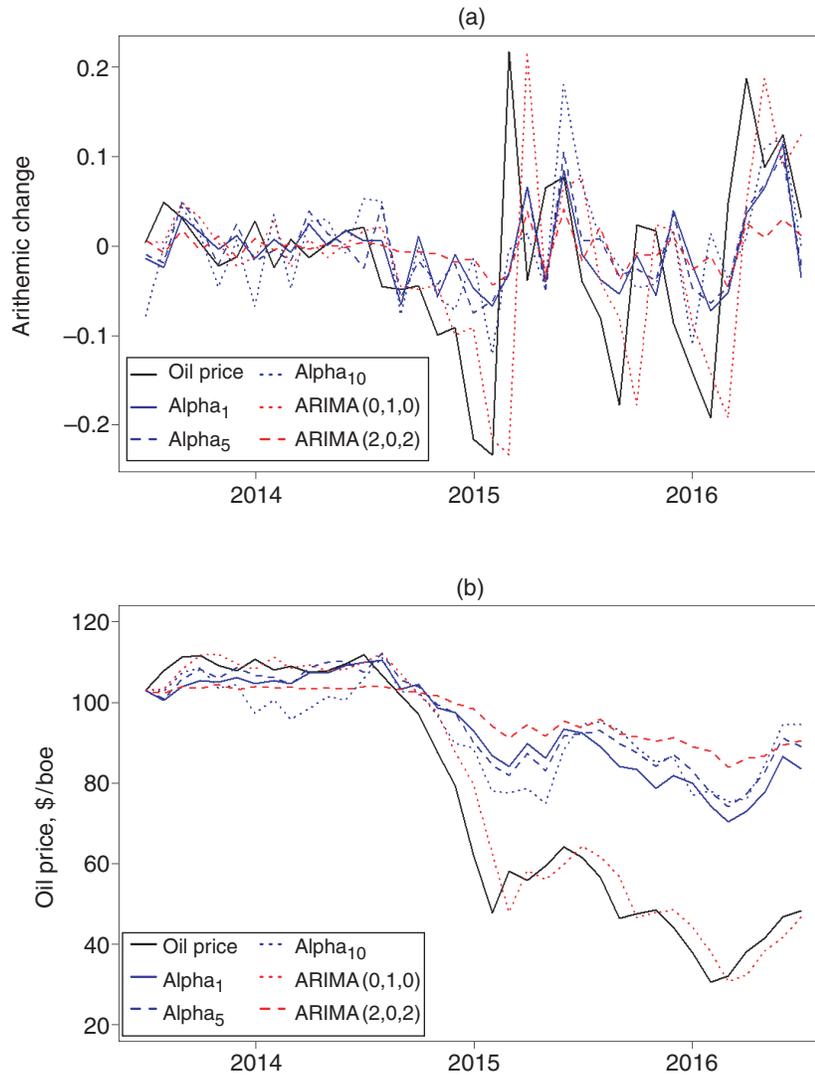
(a) Monthly out-of-sample forecasts. (b) Accumulated monthly out-of-sample forecasts.

Our most parsimonious model explains these effects differently: it is a function of futures spreads, bond spreads and the OSX index, which are variables that do not directly reflect oil supply and demand. Why we end up with different explanatory

FIGURE 2   Monthly individual and accumulated squared forecasting errors, July 2013–
June 2016.
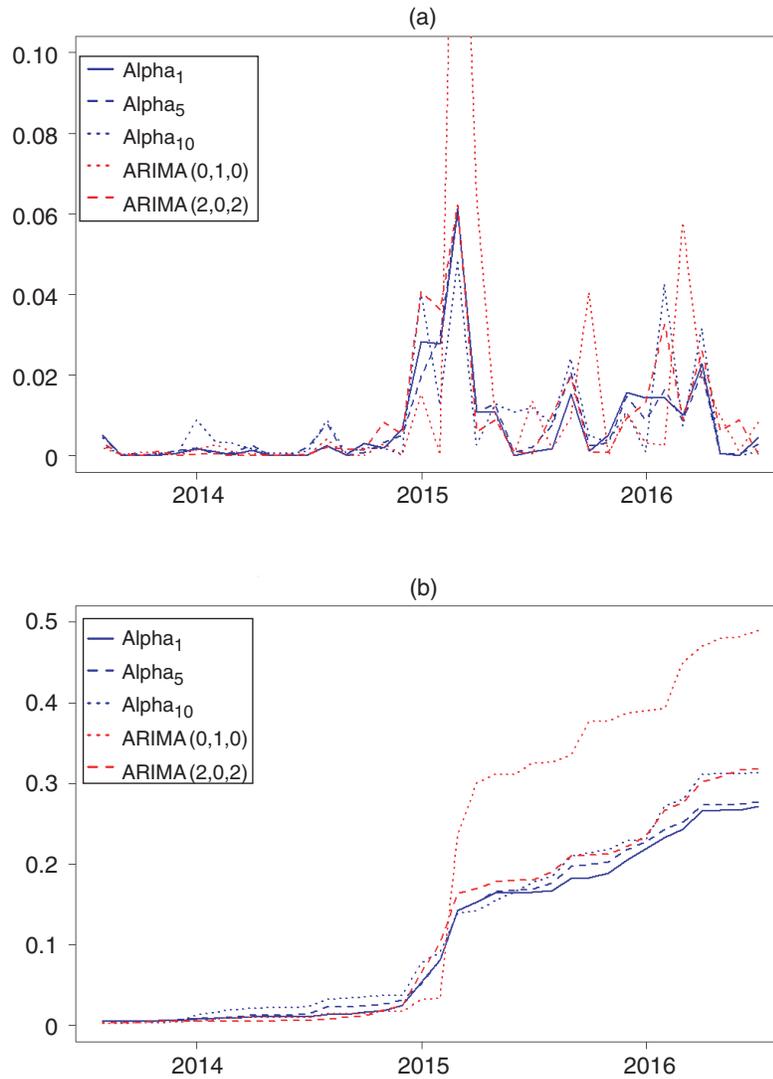
(a) Monthly squared forecasting errors. (b) Accumulated monthly squared forecasting errors.

variables may have several explanations. First, it is known that oil price factors and
their individual strengths vary in time. Since the model selection process is affected
by the in-sample window used to calibrate the model, some of the difference might

be attributed to the in-sample window alone. Second, and more importantly, determining significant explanatory variables depends on the variables fed into the model and the restriction on the number of lags. Baumeister and Kilian (2016) used four explanatory variables, with twenty-four lags each, while we used thirty-one variables, each with six lags. In any case, two unique models will never completely overlap, and our model may capture effects Baumeister and Kilian did not pick up, and vice versa. This is supported by an observation from Figure 2. Here, $Alpha_1$ performed well during summer and fall 2014 but was unable to detect the abrupt price drop occurring during December 2015, which Baumeister and Kilian attributed to a negative demand shock. Since $Alpha_1$ does not contain a pure demand variable, this effect may have been overlooked by the model.[5] Third, our stepwise model selection approach does not search all possible model combinations; hence, a better model could possibly be constructed by our data set. In any case, it is reasonable to believe that $Alpha_1$, apparently not containing any isolated supply or demand variable, is still able to provide certain pieces of supply-and-demand-related information as well as additional information. At the very least, that would explain why the relatively parsimonious $Alpha_1$ delivers a predictive power similar to Baumeister and Kilian.

In this section, we have shown that $Alpha_1$ is not overfitted in-sample and outperforms the relevant benchmarks in predicting oil prices the next month by evaluating the forecast accuracy out-of-sample. Thus, our model could advise a wide range of stakeholders and decision makers in the oil industry. In a scenario where the futures spread level is declining, the high-yield bond spread is tightening and the OSX index is rising, our model predicts that oil prices will rise for the next month, and vice versa. This may be useful information for a decision maker who is currently on the fence about storing or selling oil, or a refinery management who is seeking their optimal degree of utilization for future profits.

6 CONCLUSION

Motivated by a comprehensive study of the literature and some general perceptions about financial markets, we provide an analysis of a broad spectrum of fundamental and nonfundamental indicators for crude oil prices. We present three pure prediction models, each selected with a unique termination criterion. We find that the most significant oil price predictors in the period February 2000–June 2013 are futures

---

[5] $Alpha_{10}$ is a function of the Kilian index, ie, it contains a demand variable, and was somewhat more able to capture this price drop. However, $Alpha_{10}$ is also a function of fourteen variables, and the contribution from the Kilian variable might have been softened by the other variables.

spreads, high-yield bond spreads and the OSX equity index, all financial predictors and lagged with one month. The strong positive correlation between the OSX index and oil prices is particularly interesting, as no study, to our knowledge, has analyzed its predictive significance prior to this paper. The relationship is intuitively meaningful, as the OSX index is essentially the collective opinion of investors who are strongly exposed to the oil market. These stock prices, which very strongly depend on the oil price, are determined in a market with well-informed investors, who possess extensive market expertise and have strong incentives to gather correct market information.

Furthermore, we evaluate the models out-of-sample, where the models are compared with two benchmark ARIMA models. We demonstrate that our two most parsimonious models are superior to the benchmarks at predicting oil prices. Moreover, our leanest model, consisting of only three variables, provides the best out-of-sample results, according to our squared out-of-sample error evaluation. This demonstrates that the more complex models are overfitted to the idiosyncratic in-sample noise. The fact that the model with the best predictive power is also the most parsimonious leads us to the conclusion that this model predicts oil price dynamics more adequately than relevant benchmarks and more complex models. Baumeister and Kilian (2016) claim that the drop in oil prices from mid-2014 was, to some extent, predictable. Our analysis supports their view, as our most parsimonious model is able to predict the price decline from US$112 in July 2014 to US$85 in March 2015, and, after the slight recovery in mid-2015, further down to US$74 in February 2016. However, Baumeister and Kilian predict the changes using fundamental factors. Our model consists of solely financial factors, and, hence, it explains these effects differently.

Our findings do not necessarily imply that the financial sector determines oil prices. On the contrary, we take the view that fundamental information is traceable from financial markets and, hence, financial predictors serve as indicators for oil price fundamentals.

Our findings may provide suggestions for future research. As for oil service companies, the values of oil companies also vary with oil price. Their response to short-term price fluctuations, however, is considerably lower, due to their project portfolios being of longer duration and less gearing. It might be worthwhile, however, to examine whether various oil company indexes may serve as a leading indicator for the oil price in the medium term. If this proves to be the case, it would be interesting to know whether different categories of oil companies predict oil prices at different points in time in the future, eg, whether an index of the largest integrated international oil companies is a leading indicator of prices further into the future than an index of independent oil companies. One might also consider similar approaches for natural gas prices and electricity prices, applying indexes for natural gas companies and electricity companies as leading indicators, or, if available, indexes of their corresponding service industries.

DECLARATION OF INTEREST

REFERENCES

Alhajji, A. F., and Huettner, D. (2000). OPEC and world crude oil markets from 1973 to 1994: cartel, oligopoly, or competitive? *Energy Journal* **21**(3), 31–60.

Alquist, R., and Gervais, O. (2013). The role of financial speculation in driving the price of crude oil. *Energy Journal* **34**(3), 35–54.

Alquist, R., and Kilian, L. (2010). What do we learn from the price of crude oil futures? *Journal of Applied Econometrics* **25**(4), 539–573 (http://doi.org/dgkrtc).

Askari, H., and Krichene, N. (2010). An oil demand and supply model incorporating monetary policy. *Energy* **35**(5), 2013–2021 (http://doi.org/bq886k).

Barsky, R. B., and Kilian, L. (2004). Oil and the macroeconomy since the 1970s. *Journal of Economic Perspectives* **18**(4), 115–134 (http://doi.org/brpvw5).

Baumeister, C., and Kilian, L. (2016). Understanding the decline in the price of oil since June 2014. Working Paper 5755, CESifo. URL: https://ssrn.com/abstract=2747889.

Baumeister, C., Kilian, L., and Zhou, X. (2013). Are product spreads useful for forecasting? An empirical evaluation of the Verleger Hypothesis. Discussion Paper DP9572, Centre for Economic Policy Research.

Bernanke, B. (2016). The relationship between stocks and oil prices. Article, February 19, Brookings. URL: http://brook.gs/2qrUEIZ.

Broadstock, D. C., and Filis, G. (2014). Oil price shocks and stock market returns: new evidence from the United States and China. *Journal of International Financial Markets, Institutions and Money* **33**, 417–433 (http://doi.org/b67r).

Campos, J., Ericsson, N. R., and Hendry, D. F. (2005). General-to-specific modeling: an overview and selected bibliography. Discussion Paper 838, FRB International Finance (http://doi.org/fzptqm).

Chen, Y., Rogoff, K., and Rossi, B. (2010). Can exchange rates forecast commodity prices? *Quarterly Journal of Economics* **125**(3), 1145–1194.

Chevillon, G., and Rifflart, C. (2009). Physical market determinants of the price of crude oil and the market premium. *Energy Economics* **31**(4), 537–549 (http://doi.org/dfxn3k).

Cifarelli, G., and Paladino, G. (2010). Oil price dynamics and speculation: a multivariate financial approach. *Energy Economics* **32**(2), 363–372 (http://doi.org/fsbdw2).

Coleman, L. (2012). Explaining crude oil prices using fundamental measures. *Energy Policy* **40**, 318–324 (http://doi.org/ctvfnd).

Coppola, A. (2008). Forecasting oil price movements: exploiting the information in the futures market. *Journal of Futures Markets* **28**(1), 34–56. (http://doi.org/drxtqv).

Cunado, J., and de Gracia, F. P. (2005). Oil prices, economic activity and inflation: evidence for some Asian countries. *Quarterly Review of Economics and Finance* **45**(1), 65–83 (http://doi.org/bjp5b3).

Dees, S., Gasteuil, A., Kaufmann, R., and Mann, M. (2008). Assessing the factors behind oil price changes. Working Paper 855, January, European Central Bank.

Fattouh, B., Kilian, L., and Mahadeva, L. (2013). The role of speculation in oil markets: what have we learned so far? *Energy Journal* **34**(3), 7–33 (http://doi.org/b67s).

Gertler, M., and Lown, C. (1999). The information in the high-yield bond spread for the business cycle: evidence and some implications. *Oxford Review of Economic Policy* **15**(3), 132–150.

Gupta, P., and Goyal, A. (2015). Impact of oil price fluctuations on Indian economy. *OPEC Energy Review* **39**(2), 141–161 (http://doi.org/b67t).

Hamilton, J. D. (1996). This is what happened to the oil price–macroeconomy relationship. *Journal of Monetary Economics* **38**(2), 215–220 (http://doi.org/b94pzv).

Hamilton, J. D. (2003). What is an oil shock? *Journal of Econometrics* **113**(2), 363–398 (http://doi.org/dschfz).

Hamilton, J. D. (2008). Understanding crude oil prices. Working Paper 14492, November, National Bureau of Economic Research (http://doi.org/b4pfr2).

Hamilton, J. D., and Kim, D. H. (2000). A re-examination of the predictability of economic activity using the yield spread. Working Paper 7954, October, National Bureau of Economic Research (http://doi.org/cmx28r).

Hammoudeh, S., Dibooglu, S., and Aleisa, E. (2004). Relationships among US oil prices and oil industry equity indices. *International Review of Economics and Finance* **13**(4), 427–453 (http://doi.org/fh3s6f).

Harrell, F. (2001). *Regression Modeling Strategies*, 1st edn. Springer.

He, Y., Wang, S., and Lai, K. K. (2010). Global economic activity and crude oil prices: a cointegration analysis. *Energy Economics* **32**(4), 868–876 (http://doi.org/fw659t).

Hooker, M. A. (1996). What happened to the oil price–macroeconomy relationship? *Journal of Monetary Economics* **38**(2), 195–213 (http://doi.org/fhvs4t).

Hyndman, R. J. (2010). Benchmarks for forecasting. Blog Post, August 25, *Hindsight*. URL: http://bit.ly/2qt2Rq5.

International Energy Agency (2016). Key world energy statistics. Report, IEA. URL: http://bit.ly/2d3VuOs.

Kang, W., Ratti, R. A., and Yoon, K. H. (2015). Time-varying effect of oil market shocks on the stock market. *Journal of Banking and Finance* **61**(Supplement 2), S150–S163 (http://doi.org/b67m).

Kaufmann, R. K. (2011). The role of market fundamentals and speculation in recent price changes for crude oil. *Energy Policy* **39**(1), 105–115 (http://doi.org/dgdsq5).

Kesicki, F. (2010). The third oil price surge: what's different this time? *Energy Policy* **38**(3), 1596–1606 (http://doi.org/bwv7pq).

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

Kilian, L., and Hicks, B. (2013). Did unexpectedly strong economic growth cause the oil price shock of 2003–2008? *Journal of Forecasting* **32**(5), 385–394.

Kilian, L., and Lee, T. K. (2014). Quantifying the speculative component in the real price of oil: the role of global oil inventories. *Journal of International Money and Finance* **42**, 71–87 (http://doi.org/b67n).

Kilian, L., and Murphy, D. P. (2014). The role of inventories and speculative trading in the global market for crude oil. *Journal of Applied Econometrics* **29**(3), 454–478 (http://doi.org/f5v877).

Knittel, C. R., and Pindyck, R. S. (2013). The simple economics of commodity price speculation. Working Paper 18951, April, National Bureau of Economic Research (http://doi.org/b67p).

LeBlanc, M., and Chinn, M. D. (2004). Do high oil prices presage inflation? The evidence from G-5 countries. Working Paper 04-04, Santa Cruz Center for International Economics. URL: https://ssrn.com/abstract=509262.

Leduc, S., and Sill, K. (2004). A quantitative analysis of oil-price shocks, systematic monetary policy, and economic downturns. *Journal of Monetary Economics* **51**(4), 781–808 (http://doi.org/bfmpj4).

Lombardi, M. J., and Robays, I. V. (2011). Do financial investors destabilize the oil price? Working Paper 1346, European Central Bank. URL: http://ssrn.com/abstract=1847503.

Merino, A., and Ortiz, Á. (2005). Explaining the so-called "price premium" in oil markets. *OPEC Energy Review* **29**(2), 133–152 (http://doi.org/dsqhch).

Morana, C. (2013). Oil price dynamics, macro-finance interactions and the role of financial speculation. *Journal of Banking and Finance* **37**(1), 206–226 (http://doi.org/b67q).

Moutinho, L., and Huarng, K.-H. (2013). *Quantitative Modelling in Marketing and Management*. World Scientific, Singapore.

Murat, A., and Tokat, E. (2009). Forecasting oil price movements with crack spread futures. *Energy Economics* **31**(1), 85–90. (http://doi.org/bq3v5g).

Nasdaq (2017). PHLX oil service sector index methodology. Document, Nasdaq. URL: http://bit.ly/2rrjQqJ.

Olejnik, S., Mills, J., and Keselman, H. (2000). Using Wherry's adjusted R^2 and Mallow's C_p for model selection from all possible regressions. *Journal of Experimental Education* **68**(4), 365–380.

Olimb, M., and Oedegaard, T. M. (2010). Understanding the factors behind crude oil price changes. Master's Thesis, Norwegian University of Science and Technology.

Pindyck, R. S., and Rotemberg, J. J. (1988). The excess co-movement of commodity prices. Working Paper 2671, July, National Bureau of Economic Research (http://doi.org/d2hgdw).

Radetzki, M. (2006). The anatomy of three commodity booms. *Resources Policy* **31**(1), 56–64 (http://doi.org/bzdgbr).

Ringlund, G. B., Rosendahl, K. E., and Skjerpen, T. (2008). Does oilrig activity react to oil price changes? An empirical investigation. *Energy Economics* **30**(2), 371–396 (http://doi.org/c9kjcm).

Ryan, L., and Whiting, B. (2016). Multi-model forecasts of the West Texas Intermediate crude oil spot price. *Journal of Forecasting*, online preprint (http://doi.org/b67f).

Sadorsky, P. (1999). Oil price shocks and stock market activity. *Energy Economics* **21**(5), 449–469 (http://doi.org/dkhx3d).

Saghaian, S. (2010). The impact of the oil sector on commodity prices: correlation or causation? *Journal of Agricultural and Applied Economics* **42**(3), 477–485.

Sarwar, G. (2012). Is VIX an investor fear gauge in BRIC equity markets? *Journal of Multinational Financial Management* **22**(3), 55–65. (http://doi.org/b67g).

Ye, M., Zyren, J., and Shore, J. (2005). A monthly crude oil spot price forecasting model using relative inventories. *International Journal of Forecasting* **21**(3), 491–501 (http://doi .org/frwz8c).

Ye, M., Zyren, J., and Shore, J. (2006). Forecasting short-run crude oil price using high- and low-inventory variables. *Energy Policy* **34**(17), 2736–2743 (http://doi.org/cjjp74).