Econometric Models for Oil Price Forecasting: A Critical Survey

Giliola Frey*
Matteo Manera**
Anil Markandya***
Elisa Scarpa****

Introduction

In the last two years the price of oil and its fluctuations have reached levels never recorded in the history of international oil markets. In 2007, the West Texas Intermediate (WTI) oil price, one of the most important benchmarks for crude oil prices, averaged around 72 $/b, while in 2008 the WTI price was around 100 $/b, with an increase of nearly 38 percent over the previous year. Within the past six months, WTI daily spot prices ranged from almost 150 $/b in early July to about 30 $/b towards the end of 2008.

The determinants of past, current, and future levels of the price of oil and its fluctuations have been the subject of analysis by academics and energy experts, given the relevance of crude oil in the worldwide economy. Although the share of liquid fuels in marketed world energy consumption is expected to decline from 37 percent in 2005 to 33 percent in 2030, and projected high oil prices will induce many consumers to switch from liquid fuels when feasible, oil will remain the most important source of energy, and liquid fuel consumption is expected to increase at an average annual rate of 1.2 percent from 2005 to 2030 (EIA 2008).

The crucial question of whether oil prices will rise in the future or will decline again is timely. According to EIA (2009), for example, under current economic and world crude oil supply assumptions, WTI prices are expected to average 43 $/b in 2009 and 55 $/b in 2010. The possibility of a milder recession or a faster economic recovery, lower non-OPEC production in response to current low oil prices and financial market constraints, and more aggressive action to lower production by OPEC countries could result in a faster and stronger recovery in oil prices. Consequently, it is extremely important for economists to provide accurate answers to the complex problem of forecasting oil prices.

This study aims at investigating the existing econometric literature on forecasting oil prices. In particular, we (i) develop a taxonomy of econometric models for oil price forecasting; (ii) provide a critical interpretation of the different methodologies; and (iii) offer a comprehensive interpretation and justification of the heterogeneous empirical findings in published oil price forecasts. The paper is structured as follows: we first introduce the historical framework which is necessary to understand oil price dynamics. The following section discusses and critically evaluates the different econometric models for oil price forecasting proposed in the literature. Finally we comment on alternative criteria for evaluating and comparing different forecasting models for oil prices.

International oil markets: A historical framework

The history of oil consumption and prices goes back to the second half of the 19th century. The introduction of oil distillation in 1853 gave rise to the use of kerosene for home lighting. Not until the end of the century did oil gain a much more relevant role, due to its use for the generation of electricity. At that time, the United States was the principal consumer and its North-Eastern region was the main source of

* Eurizon Capital, Milan.
** University of Milan-Bicocca, Milan.
*** BC Basque Centre for Climate Change, Bilbao.
**** Edison Trading, Milan.
A previous version of this paper was presented at the First FEEM Conference on the Economics of Sustainable Development held at the Fondazione Eni Enrico Mattei (FEEM), Milan, January 25–25, 2007. The authors would like to thank Marzio Galeotti, Alessandro Lanza, Michael McAleer and Yves Smeers for insightful discussion, as well as seminar participants at the University of Bath, FEEM and the University of Milan-Bicocca for helpful comments. The authors would also like to thank Linda Isola for excellent editorial assistance. This study does not necessarily reflect the views of Eurizon Capital and Edison Trading.

oil supply. The increasing consumption and the subsequent depletion of US North-Eastern reserves soon caused oil prices to rise, and Standard Oil, the oil company with a monopoly position at that time, was not able to control them. By the beginning of the 20th century, oil production was extended to Texas, generating over-supply and price reductions. In the meanwhile, oil consumption spread to Europe and oil reserves were also discovered in Iraq and Saudi Arabia, but the United States still remained the main consumer and maintained its dominance over the world oil market.

One of the major economic agents in the world oil market in that period was the Texas Railroad Commission (TRC) that was founded in 1891 as a regulatory agency aimed at preventing discrimination in railroad charges, later also controlled petroleum production, natural gas utilities as well as motor carriers. Given its dominant position in the US market, TRC was able to set oil prices by effectively fixing production quotas, at least until the formation of the Organization of Petroleum Exporting Countries (OPEC). The other major actors in the world oil markets were the so-called "seven sisters", five of which were American companies (Standard Oil of New Jersey (Esso), Standard Oil of California (Chevron), Standard Oil of New York (Mobil), Gulf Oil and TEXACO), together with Royal Dutch Shell and the Anglo Persian Oil Company (BP). The seven sisters started to operate after the break-up of Standard Oil by the US government. Their fairly complete monopoly and ability to work as a cartel allowed them to take control over oil prices for about fifty years.

World War II definitely marked the predominance of oil as an energy source. The excess of oil due to the cooperation between the United States and Saudi Arabia offered America and its allies a privileged access to this crucial resource. During the 1950s, new oil reserves were discovered in the Middle East, and new producers entered the market, making it difficult to limit oil production for the sake of controlling oil prices. In 1960 the Middle Eastern countries formed the OPEC, a cartel meant to avoid competition among its members and to prevent unsought price reductions. In 1970, for the first time, the growing US economy was not able to feed its increasing need of oil from domestic sources and became an importing country. The effects of this dependency became visible very soon after the Yom Kippur War in 1973, when the United States and many other Western countries supported Israel, catalyzing the reaction of the Arab exporting countries which declared an embargo. As a result, within six months the price of oil increased by 400 percent. Since 1973, the stability of oil prices has vanished, starting a period of large price fluctuations.

A second phase of uncertainty affected world oil prices in 1979 and 1980, when the Iranian Revolution and the Iraq-Iran War pushed oil prices to double. This period also revealed the inability of OPEC to act as a cartel. Saudi Arabia's warning that high prices would reduce consumption remained unheeded and prices kept on rising, while oil demand decreased. Furthermore, non-OPEC countries, attracted by the possibility of large gains at the high price level, increased their oil production and, consequently, helped match oil supply and demand. Later, between 1982 and 1985, OPEC policy was devoted to stabilize prices by setting production quotas below their previous levels. Unfortunately, this strategy was often hampered by the behaviour of some members, that kept on producing above their quotas. During this period, Saudi Arabia played the "swing producer" role, adjusting its production to demand in order to prevent price falls until 1986. Yet, burdened by this role, this country changed its strategy thereafter and increased its oil production, causing an abrupt price decrease.

Prices kept on falling until the Gulf War of 1990. The invasion of Kuwait in this year created a sudden price reversal, which was only normalized after 1993, when Kuwaiti exports outran their pre-war levels. In the early 1990s oil consumption started to rise again, aided by the growth of the Asian economies. The increasing rate of production by OPEC to meet the demand was then the origin of the drastic price reduction that occurred between 1997 and 1998, when the Asian growth slowed due to the financial and economic crises, and OPEC was faced by a massive oversupply at the same time. In 1999 the prices rose again, supported by the OPEC's strategy of reducing quotas, which was successful in spite of the increase in non-OPEC production, at least until the terrorist attack of September 11, 2001. During the years between 2002 and 2005, the majority of oil producer countries continued to adopt the policy of fixing low production quotas. This strategy, together with the inadequate response of non-OPEC countries to the increase in the oil demand, led to an increase in oil prices, which have kept on rising until the second half of 2008, when the monthly average price of WTI fell from 133 $/b in July 2008 to 41 $/b in December 2008 and January 2009.

Econometric models for oil price forecasting

In the existing empirical literature on oil price forecasting one can distinguish among three categories of econometric models:

- time series models exploiting the statistical properties of the data, namely autocorrelation and non-stationarity;
- financial models based on the relationship between spot and future prices; and
- structural models describing how specific economic factors and the behaviour of economic agents affect the future values of oil prices.

The following subsections will illustrate the main features of each class of econometric models for oil price forecasting, as well as the most relevant contributions which can be classified according to our proposed taxonomy.

(a) Time series models

Time series models aim at predicting future oil prices by exploiting relevant characteristics of historical data. In this respect, a wide range of models have been proposed which can be divided into three main groups, depending on their assumptions about the data-generation process: martingale sequences, autoregressive models and mean-reverting specifications. Given their simplicity, time series models have often been used as a benchmark for the forecasting performance of financial and structural models. In particular, the random walk model (a particular case of martingale sequence) is generally used to assess whether more complex and expensive models are indeed justified by an improvement in their forecasting performance.

A martingale sequence for the oil spot price S is a stochastic process such that the expected value of S at time t+1 conditional on all available information I up to time t is equal to the actual value of the oil spot price at time t:

(1)

$$E(S_{t+1} \mid I(t)) = S_t$$

Its applications in finance go back to the introduction of the “efficient market hypothesis (EMH)”, often credited to Fama (1965), which states that, in the presence of complete information and a large number of rational agents, actual prices reflect all available information and expectations for the future. In other words, current prices are the best predictor of tomorrow’s prices. A widely used form of the martingale process is the random walk specification:

(2)

$$S_{t+1} = S_t + \varepsilon_t$$

where ε_t is an uncorrelated error term with zero mean and constant variance. According to this model, prices deviate from their current level only because of casual fluctuations. The random walk with drift represents a simple extension of this formula, which introduces a linear trend in the data generation process:

(3)

$$S_{t+1} = \delta + S_t + \varepsilon_t$$

In this case prices are assumed to constantly increase (decrease) from their previous level, except for stochastic deviations.

Oil prices can follow an autoregressive (AR) process:

(4)

$$S_t = \phi_1 S_{t-1} + \ldots + \phi_q S_{t-p} + \varepsilon_t = \phi_p(L) S_t + \varepsilon_t$$

where p is the order of the AR(p) process, $\phi_p(L)$ is the polynomial in the lag operator L of order p, and ε_t is a white noise error term. Notice that this process can either be explosive or stable depending on whether the roots of the characteristic equation associated with $\phi_p(z) = 0$ are outside or inside the unit circle. In the case of autoregressive processes, prices are not driven by random fluctuations, instead they are predictable from their history.

Oil prices can also be driven by a mean reverting process. This assumption comes from the evidence that prices in financial markets tend to go back to their average level after a shock. According to this approach, prices can neither be explained by the random walk assumption nor simply inferred from their past values. Given a long-run equilibrium level S^*_t of the oil spot price and a mean reversion rate α, mean reverting models can be described as:

(5)

$$S_{t+1} - S_t = \alpha(S_t^* - S_t) + \varepsilon_t$$

According to equation (5), future price variations depend on the disparity between actual and long-run price levels, where the latter can be specified to be a function of a set of exogenous variables.

More generally, error correction models (ECM) are designed to capture movements towards an equilibrium level. Given two variables, X and Y, and an equilibrium level between the two variables, $Y=\alpha X$, variable Y tends to adjust to deviations from this equilibrium according to the following scheme:

(6)

$$Y_t = \alpha + \lambda_1(Y_{t-1} - \alpha X_{t-1}) + \varepsilon_t$$

where $Y_t^* = \hat{\alpha} X_t$ is the estimated equilibrium value for Y (see e.g. Engle and Granger 1987; Stock and Watson 1993).

In the empirical literature on oil price modelling and forecasting, several contributions provide empirical evidence that is supportive of the EMH. For instance, Morana (2001) notices that, during the period between January 4, 1982 and January 21, 1999, oil prices appeared to be characterized by a stochastic trend and exhibited alternating periods of high and low volatility. Since these features can be a symptom of underlying dependencies in the behaviour of oil prices, Morana (2001) suggests to use a martingale process to describe oil price dynamics. The reliability of a random walk model is also assessed by Chernenko et al. (2004) with an application to the crude oil future market.

Abosedra (2005) observes that the behaviour of the WTI spot price, S, during the period from January 1991 to December 2001 can be approximated by a random walk process with no drift. Consequently, the author proposes to forecast the one-month-ahead price of crude oil for every day using the previous trading day's spot price and to use the monthly average of these daily forecasts to obtain a monthly predictor of the future oil price X. To assess the statistical properties of this univariate forecast, the author suggests estimating the following relationship:

(7)

$$S_t = \alpha + \beta X_{t-1} + \varepsilon_t$$

and to test the null hypothesis α = 0 and β = 1, that is to test for the unbiasedness of X. However, since cointegration between S and X can lead to biased estimates of α and β in equation (7), the author follows Phillips and Loretan (1991) and suggests a non-linear estimation of and α and β:

(8)

$$S_t = \alpha + \beta X_{t-1} + \sum_{i=1}^{n} \rho_i (S_{t-1} - \alpha - \beta X_{t-i-1}) + \sum_{j=-m}^{m} \phi_j \Delta X_{t-j} + \varepsilon_t$$

Both single and joint tests of the null hypotheses α = 0 and β = 1, suggest that X is an unbiased predictor for future oil prices. Furthermore, the absence of autocorrelation in the residuals confirms the efficiency of the proposed forecast method.

The empirical evidence on autoregressive specifications is much more controversial. Bopp and Lady (1991) use an autoregressive specification to describe monthly heating oil prices from the New York Mercantile Exchange (NYMEX). Their analysis covers the period between December 1980 and October 1988, and confirms the good performance of the autoregressive model. An autoregressive representation is used by Lalonde et al. (2003) to analyze the behaviour of WTI crude oil prices. The authors show that this model has a very poor forecasting ability. Ye et al. (2005) verify the performance of an autoregressive specification with seasonal effects in predicting monthly oil prices in the period from January 2000 to January 2003. Their model takes into account the consequences of the reduction of OPEC production from 1999, using a leverage variable and a dummy variable capturing the effects of the twin towers terrorist attack, of which impact is supposed to extend from October 2001 to March 2002:

(9)

$$S_t = \alpha + \beta_1 S_{t-1} + \beta_2 S_{t-12} + \sum_{j=0}^{5} \gamma_j D01_j + S99 + \varepsilon_t$$

A dynamic forecasting exercise shows the poor performance of this model, which is not able to capture oil price variations.

Pindyck (1999) analyzes the stochastic dynamics of crude oil, coal and natural gas prices using a large data set covering 127 years, and tries to assess whether time series models are helpful in forecasting long horizons. The analysis ranges from 1870 to 1996, considering nominal oil prices deflated by wholesale prices *(p)* (expressed in 1967 USD). The author proposes a model which accounts for fluctuations in both the level and the slope of a deterministic time trend:

(10)

$$\begin{cases} p_t = \rho p_{t-1} + \beta_1 + \beta_2 t + \beta_3 t^2 + \phi_{1t} + \phi_{2t} t + \varepsilon_t \\ \phi_{1t} = \alpha_1 \phi_{1,t-1} + \upsilon_{1t} \\ \phi_{2t} = \alpha_2 \phi_{2,t-1} + \upsilon_{2t} \end{cases}$$

where ϕ_{1t} and ϕ_{2t} are unobservable state variables. Assuming normally distributed and uncorrelated error terms, Pindyck computes a Kalman filter to estimate model (10). This procedure is a recursive estimate that calculates parameters via Maximum Likelihood, along with optimal estimates of the state variables. The initial values are usually estimated using OLS and assuming that the state variables are constant parameters. The author concentrates on three sub-samples (1870–1970, 1970–1980, 1870–1981) and the full dataset to compare the forecasting ability of the proposed model with respect to a model with mean reversion to a deterministic linear trend:

(11)

$$p_t = \rho p_{t-1} + \beta_1 + \beta_2 t + \varepsilon_t$$

Results show that the deterministic trend model performs better in forecasting oil prices. Nevertheless equation (10) provides a more accurate explanation of oil prices fluctuations.

Radchenko (2005) proposes a univariate shifting-trends model for the long-term forecasting of energy prices:

(12)

$$\begin{cases} p_t = \alpha p_{t-1} + \beta_1 + \phi_{1,t} + \phi_{2,t} t + \varepsilon_t \\ \phi_{1t} = \gamma_1 \phi_{1,t-1} + \mu_{1,t} \\ \phi_{2t} = \gamma_2 \phi_{2,t-1} + \mu_{2,t} \\ \varepsilon_t = \psi \varepsilon_{t-1} + \upsilon_t \end{cases}$$

which is a modified version of Pindyck (1999), where the error term ε is assumed to be an autocorrelated process, rather than a simple white noise. In particular, the author exploits the same dataset used by Pindyck (1999) and considers four different forecasting horizons: 1986–2011, 1981–2011, 1976–2011, 1971–2011. Radchenko (2005) suggests embedding equation (12) into a Bayesian framework and obtains results similar to Pindyck (1999), except for the autoregressive parameters α, γ_1 and γ_2 which appear less persistent. However, the author notices that forecasts from shifting-trend models cannot account for OPEC cooperation, thus predicting unreasonable oil price declines. As a solution, he suggests combining model (12) with an autoregressive model and a random walk model, which can be considered a proxy for future cooperation. Results confirm that forecasts can be improved by a combination of different models.

A comprehensive comparison of the different time-series models proposed is offered by Zeng and Swanson (1998), who analyze four futures markets – gold, crude oil, Treasury bonds and S&P500. The authors compare the performance of a random walk specification with an autoregressive model and an error correction model, where the deviation from the equilibrium level *(ECT)* is assumed to be equal to the difference between the future price for tomorrow and the futures for today's price, which is generally called the price spread:

(13)

$$F_t = \alpha + \beta_1 ECT_{t-1} + \sum_{i=1}^{n} \rho_i (F_{t-i}) + \varepsilon_t$$

Daily data from April 1, 1990 to October 31, 1995, with a rolling out-of-sample forecast over the period between April 1, 1991 and October 31, 1995, shows that ECM are preferable when short forecast horizons are considered.

Prices may revert to a non-constant and uncertain value, which can evolve stochastically through time. Factor models are the direct translation of this assumption, as they are meant to infer from the data the nature of the stochastic unobservable factors that drive a given phenomenon. Schwartz and Smith (2000) provide an interesting example of a factor model, where the spot price of a general commodity is decomposed into two factors, one capturing the equilibrium value (χ_t), the other the short-run departures from equilibrium (ξ_t). The short-run component ξ_t is assumed to follow an Ornstein-Uhlenbeck process reverting to a zero mean:

(14)

$$d\xi_t = -k\xi_t d_t + \sigma_\xi dz_\xi$$

while the long-run level χ_t is modelled according to a Brownian motion:

(15)

$$d\chi_t = \mu_\chi d_t + \sigma_\chi dz_\chi$$

with dz_ξ and dz_χ indicating the correlated increments of standard Brownian motion processes. Clearly, the Ornstein-Uhlenbeck process and the Brownian motion represent the extension in continuous time of the mean reverting process and the random walk process, respectively. Model shown in equations (14) and (15) can be generalized by including another stochastic factor, as the three factors model proposed by Schwartz (1997), where a stochastic interest rate is added as the determinant of spot prices and it is modelled as a mean-reverting process.

(b) Financial models

The relationship between spot (S) and futures (F) prices can be represented as:

(16)

$$F(t,T) = S(t)e^{r(T-t)}$$

where *F(t,T)* is the futures price at time *t* for maturity *T*, *r* is the interest rate, *S(t)* is the asset price at time *t*. The underlying assumption is that it is possible to replicate the payoff from a forward sale of an asset by borrowing money, purchasing the asset, "carrying" the asset until maturity and then selling the asset. This kind of arbitrage is known as the "cost-of-carry arbitrage". Referring to commodities (e.g. oil), relationship shown in equation (16) is no longer valid, unless it is modified to include the costs of storage *(w)*:

(17)

$$F(t,T) = S(t)e^{(r+w)(T-t)}$$

However, the activity of storing oil can provide some benefits, which are generally indicated with the term "convenience yield" (δ). Consequently, in the commodities market, the future-spot relationship becomes:

(18)

$$F(t,T) = S(t)e^{(r+w-\delta)(T-t)}$$

From equation (18) the market can be either in *contango* (future price exceeds spot price) or in backwardation (spot price exceeds future price), according to the relative size of *w* and δ.

Financial econometric models generally assume that futures and forward prices can be unbiased predictors for the future values of the spot price:

(19)

$$F_t = E(S_{t+1})$$

In order to test for unbiasedness, the following model can be specified:

(20)

$$S_{t+1} = \beta_0 + \beta_1 F_t + \varepsilon_{t+1}$$

In equation (20), F_t is an unbiased predictor of S_{t+1} if the joint hypothesis $\beta_0 = 0$ and $\beta_1 = 1$ is not rejected *(unbiasedness hypothesis)*, and it is also an efficient predictor if no autocorrelation is found in the error terms *(efficiency hypothesis)*. It is worth noticing that a violation of the unbiasedness hypothesis is generally interpreted as the presence of a risk premium.

Fama and French (1987) propose a detailed comparison between storage costs and risk premia applied to commodity markets. Although their study does not include crude oil prices, it clearly shows that empirical evidence in favour of storage costs is easier to detect than the existence of risk premia. Following this seminal paper, a significant part of the empirical literature has focused on risk premium models, although the findings on the existence of a risk premium are mixed. An attempt to model the cost of storage relationship has been proposed by Bopp and Lady (1991), who include in the regression a proxy which measures the number of months until expiration of the contracts corresponding to the futures price. Using monthly data on NYMEX heating oil from December 1980 to

October 1988, they confirm the statistical adequacy of this relationship. However, they also propose a simple random walk specification and a regression model of spot prices on futures prices, which seem to perform equally well. Samii (1992) estimates the WTI futures oil price (three and six months) as a function of the WTI spot price and an interest rate, using daily data for the years 1991–1992 and monthly data over the period 1984–1992. In particular, the author shows that oil storage should influence spot prices in the intermediate run, while in the long run prices should be led by a premium. Unfortunately, Samii (1992) does not find any robust evidence for either of the two hypotheses of cost storage and risk premium. The conclusion is that the interest rate does not play a relevant role, whereas spot and futures prices are highly correlated, although it is not possible to identify the causal direction of the relationship between spot and futures prices.

Gulen (1998) extends model shown in equation (20) by incorporating the effects of posted price *(C)*, i.e. the price at which oil is actually bought or sold by an oil company. The author proposes posted prices as an alternative predictor to futures prices and states that, if futures prices are the best predictor, then posted prices should have no explanatory power in the following regression model:

(21)

$$S_{t+1} = \beta_0 + \beta_1 F_t + \beta_2 C_t + \varepsilon_{t+1}$$

Gulen (1998) analyzes monthly data of WTI spot and futures prices for one-, three- and six-month ahead, computed as a simple mean of daily data and covering the period between March 1983 and October 1995. He shows that futures prices outperform the posted price and that futures prices are an efficient predictor of spot prices. However, the posted price seems to have a predictive content, although limited to the short run.

Zeng and Swanson (1998) use an ECM to forecast oil prices over the period 1991–1995. The specification of the long-run equilibrium refers to the cost-of-storage approach specified in equation (18), as the ECT is defined as:

(22)

$$ECT_{t-1} = F_{t-1} - e^{(r+\omega-\delta)cl} S_{t-1}$$

where cl denotes the number of days for the delivery cycle. As described in the previous section, Zeng and Swanson (1998) estimate also a random walk, an autoregressive model and an ECM, where the ECT is given by the price spread. The empirical evidence is supportive of the ECM. Chernenko et al. (2004) focus on the spreads between spot price and futures as well as forward prices by estimating the following modification of model (20):

(23)

$$S_t - S_{T-t} = \beta_0 + \beta_1 (F_{t|T-t} - S_{T-t}) + \varepsilon_t$$

In particular, the authors' strategy is to test for the absence of risk premia and, if the null is rejected, to investigate whether risk premia are time-varying or constant by testing for $\beta_1 = 1$. Results show that futures and forward prices do not generally outperform the random walk model and cannot be considered as rational expectations for the spot price. Furthermore, when the oil market is analyzed, risk premium does not seem to be a relevant factor, while the empirical performance of futures prices is very close to the random walk specification.

Chin et al. (2005) examine how accurate futures prices are in forecasting spot prices. They analyze the relationship between three-, six- and twelve-month ahead futures prices and the current spot price for crude oil (WTI), gasoline (Gulf Coast), heating oil (No.2 Gulf Coast) and natural gas (Henry Hub). Assuming that the spot price follows a random walk with drift and rational expectations, the authors estimate a logarithmic version of equation (23) with OLS and robust standard errors. For the period from January 1999 to October 2004, the authors show that futures prices at different maturities are unbiased predictors of spot oil prices, and they find empirical evidence in favour of the efficient market hypothesis.

The two hypotheses of storage costs and risk premium are tested by Green and Mork (1991) for the oil market during the period 1978–1985. They concentrate on Mideast Light and African Light/North Sea monthly prices using Generalized Method of Moments (GMM) estimates. The most interesting result is that in the years 1978–1985 there is no evidence of unbiasedness/efficiency, while the subperiod 1981–1985 seems to support the hypothesis of efficiency in the oil financial market. Serletis (1991) analyzes daily spot and futures prices of NYMEX heating oil and crude oil over the period between July 1, 1983 and August 31, 1988, as well as daily spot and futures prices of unleaded gasoline over the period between March 14, 1985 and August 31, 1988. The aim of his contribution is to measure the forecast information contained in futures prices and the time-varying risk premium. The empirical findings suggest that variations in the premium worsen the forecasting performance of futures prices.

Moosa and Al-Loughani (1994) use monthly data from January 1986 to July 1990 on WTI spot, three- and six-month futures prices to test unbiasedness and efficiency. Given the presence of cointegration between spot and futures prices, they extend equation (20) in an error correction form:

(24)

$$\Delta S_{t+1} = \alpha_0 + \alpha_1(S_t - \beta_1 F_{t-1}) + \alpha_2 \Delta F_{t-1} + \sum_{i=1}^{n} \gamma_i \Delta S_{t-i} + \sum_{i=1}^{n} \delta_i \Delta F_{t-j} + \varepsilon_t$$

In this case, unbiasedness corresponds to the null hypothesis $\alpha_0 = 0$, $\alpha_1 = -1$, $\alpha_2 = 1$, $\gamma_i = \delta_i = 0$, $\forall i$. Results show that futures prices are neither unbiased nor efficient. Assuming rational expectations and using a GARCH-in-mean specification to take into account non-constant volatility, the authors analyze the structure of the risk premium, which turns out to be time-varying.

Morana (2001) shows that one-month ahead forward prices are a poor predictor of futures spot prices, since in more than 50 percent of the cases they fail to predict the sign of oil price changes. The author compares the forecasting ability of the Brent forward price with the accuracy of a simple random walk model, using daily data from November 2, 1982 to January 21, 1999 and considering a long forecasting horizon (May 2, 1985–January 21, 1999) and a short forecasting period (November 21, 1988–January 21, 1999). The decomposition of the mean squared forecast error (MSFE) and the sign tests show that forecasting with forward prices or with a random walk does not yield significantly different results. Specifically, over a short time horizon both methods are biased, while, when a longer time period is considered, they do produce unbiased forecasts, although their performance resembles that of a random guess. Nevertheless, Morana (2001) points out that an appropriate use of forward prices can be promising, as they are reliable predictors when oil price volatility is small. Following Barone-Adesi et al. (1998) and Efron (1979), the author uses bootstrap methods to approximate the oil price density function, which is characterized by time-varying volatility. The resulting confidence intervals for oil price forecasts confirm that forecasting with forward prices future values of the price of oil is less reliable, as the confidence intervals tend to widen as volatility increases. Cortazar and Schwartz (2003) use a three factor model to explain the relationship between spot and futures prices. Daily data from the NYMEX over the period 1991–2001 confirm the accuracy of the model. The authors propose a minimization procedure as an alternative to the standard Kalman filter approach, which seems to produce more reliable results.

Another interesting evaluation of financial models is carried out by Abosedra (2005), who compares the performance of futures prices *(F)* with a simple univariate forecast *(X)*. As already mentioned, Abosedra (2005) assumes a random walk process with no drift for spot crude oil prices *(S)*, and suggests using the previous trading day spot price to forecast the one-month ahead price of crude oil for every trading day. The monthly forecast is set equal to the simple average of the daily forecasts. Using the approach described in the section related to time series models, the author establishes that the forward price and the simple univariate forecast are unbiased and efficient predictors for the future value of the spot price of oil. A more formal comparison of the two predictors is based on testing whether the forecast error related to each forecast can be improved by the information contained in the other forecast. This comparison corresponds to a test of the null hypothesis $\alpha_1 = 0$ and $\beta_1 = 0$, $i = 1,..., n$, in models:

(25)

$$S_T - F_{T-1} = \alpha_0 + \sum_{i=1}^{n} \alpha_i (S_{T-i} - X_{T-i-1}) + \varepsilon_t$$

(26)

$$S_T - X_{T-1} = \beta_0 + \sum_{i=1}^{n} \beta_i (S_{T-i} - F_{T-i-1}) + \varepsilon_t$$

Results show that futures prices can reduce the univariate forecast error, while the converse is not true.

These findings lead to conclude that futures prices are semi-strongly efficient.

Murat and Tokat (2009) analyze the relationship between crude oil prices and the crack spread futures. In the oil industry the crack spread is defined as the difference between the price of crude oil and the price of its products. In other words, the crack spread represents the profit margin that can be obtained from the oil refining process. An ECM is specified to assess the direction of the causal relationship between crude oil price and crack spread, as well as to predict the price of oil from the crack spread futures, using weekly data from the NYMEX over the period from January 2000 to February 2008. The empirical evidence suggests that the crack spread helps to predict oil prices. When its performance is compared with a random walk model and a regression of the spot price on futures oil prices, the authors find out that both crack spread and crude oil futures are preferable to the random walk specification, although futures prices are slightly more accurate than the crack spread futures.

(c) Structural models

Structural models relate the oil price behaviour to a set of fundamental economic variables. The variables that are typically used as the economic drivers of the spot price of oil can be grouped into two main categories: variables that describe the role played by OPEC in the international oil market, and variables that measure current and future physical oil availability. In this context researchers have generally considered measures of OPEC behaviour, such as production quotas, overproduction, capacity utilisation and spare capacity. It is well known that OPEC periodically establishes the quantity of oil to be produced by its members *(OQ)* in order to pursue oil market stability. It is also well acknowledged that, on several occasions, some OPEC countries have decided to produce more than their fixed production quotas. This overproduction *(OV)* is computed as the difference between OPEC production *(OP)* and quotas. Another relevant factor is production capacity. This variable is introduced in structural models in two different ways. On the one hand, some authors have used capacity utilization *(CU)*, computed as 100 times the ratio between production and productive capacity *(PC)*. On the other hand, some authors have proposed spare capacity *(SC)*, defined as the difference between production and productive capacity.

Besides the impact of OPEC, many authors have also recognized the importance of the current and future availability of physical oil. According to this view, the most crucial variable is represented by the level of inventories. Stocks are the link between oil demand and production and, consequently, they are a good measure of price variation. Most authors have considered two kinds of stocks, namely government *(GS)* and industrial *(IS)*. Due to their strategic nature, government inventories are not generated by a supply-demand mechanism and are generally constant in the short run. This explains the decision of many researchers to introduce in their models industrial stocks that vary in the short run and are able to account for oil price dynamics. When industrial inventories are considered, they are generally expressed in terms of the deviation from their normal level *(ISN)*, which is defined as the relative inventory level *(RIS)*. Operationally, *RIS* is calculated as:

(27)

$$RIS_t = IS_t - ISN_t$$

In equation (27), ISN_t indicates the de-seasonalized and de-trended industrial stock level, i.e.

(28)

$$ISN_t = \alpha_0 + \beta_1 t + \sum_{i=2}^{12} \beta_i D_i$$

where t is a linear trend and D_i is a set of monthly dummies, used to detect seasonal variations. Since government stocks are not subject to seasonality, their relative level *(RGS)* is specified as follows:

(29)

$$RGS_t = GS_t - GSN_t$$

being GSN_t the de-trended government stock level, defined as:

(30)

$$GSN_t = \alpha_0 + \beta_1 t$$

Zamani (2004) presents a short-term quarterly forecasting model of the real WTI price *(W)* that accounts for both the role of OPEC and the physical oil availability. Besides the significance of both kinds of relative inventory levels, the author includes in his model OPEC quotas, overproduction and non-OECD demand *(DN)* as explanatory variables. In particular, Zamani (2004) proposes an ECM, estimated using the two-step approach by Engle and Granger (1987), where the long-run equilibrium is specified as:

(31)

$$S_t = \alpha_1 + \alpha_2 OQ_t + \alpha_3 OV_t + \alpha_4 RIS_t + \alpha_5 RGS_t + \alpha_6 DN_t + \alpha_7 D90_t + \varepsilon_t$$

and the short-run dynamics is described by:

(32)

$$\Delta S_t = \beta_0 + \sum_{i=1}^{m} \beta_{1i} \Delta OQ_{t-i} + \sum_{i=1}^{m} \beta_{2i} \Delta OV_{t-i} + \sum_{i=1}^{m} \beta_{3i} \Delta RIS_{t-i} + \sum_{i=1}^{m} \beta_{4i} \Delta RGS_{t-i} + \sum_{i=1}^{m} \beta_{5i} \Delta DN_{t-i} + \beta_6 D90_t + \lambda \varepsilon_{t-1} + \mu_t$$

In equations (31) and (32), *D*90 is a dummy variable for the Iraqi War in the third and fourth quarter of 1990. Using data for the period 1988–2004, Zamani (2004) shows that an increase in all the explanatory variables generates a reduction of the price of oil, while the dummy variable and the non-OECD demand positively affect the real WTI price. It is worth noticing that the in-sample dynamic forecasts computed on the basis of this model are quite accurate, according to standard forecast evaluation criteria.

Ye et al. (2002, 2005 and 2006) use relative oil inventory levels to forecast oil prices. Ye et al. (2002) describe oil prices as a function of RIS and of a variable accounting for a lower-than-normal level of inventories. The specification is empirically tested using a monthly dataset which covers the period from January 1992 to February 2001. This model is generalized by Ye et al. (2005), who use monthly data from 1992 to 2003 to analyze the relationship between WTI spot price and oil stocks. Defining relative industrial inventories as described in equations (27) and (28), they suggest modeling the WTI spot price as:

(33)

$$S_t = \alpha_0 + \sum_{i=0}^{3} \beta_i RIS_{t-i} + \sum_{j=0}^{5} \gamma_j D01_j + S99 + \alpha_1 S_{t-1} + \varepsilon_t$$

where *D*01 is a dummy variable for the period between October 2001 and March 2002, which takes into consideration the consequences of the terrorist attack on 11 September 2001, and *S*99 is a leverage variable which captures the impact on the oil market of a structural change in the OPEC's behaviour. The evaluation of this model is conducted through a comparison with a pure time series model and the following regression:

(34)

$$S_t = \alpha_0 + \alpha_1 S_{t-1} + \sum_{j=0}^{5} \gamma_j D01_j + \beta_0 S99_t + \beta_1 IS_{t-1} + \beta_2 (IS_t - IS_{t-12}) + \varepsilon_t$$

where relative inventories are substituted by industrial inventories, which are assumed to affect oil prices with a one-month lag and to depend on the deviation from their previous year level. One-, two-, three- and six-month ahead forecasts over the period from January 2000 to January 2003 show that equation (33) outperforms the other two specifications. When considering the three-month ahead forecasts, equation (34) produces more satisfactory results in the presence of a price trough, while equation (33) is more accurate in the presence of price peaks. More recently, Ye et al. (2006) extend the work by Ye et al. (2005), allowing for asymmetric transmission of inventory changes to oil price. The authors claim that the response of the oil price should be different, depending on the level of the relative stocks:

(35)

$$\begin{cases} LIS_t = RIS_t + \sigma_{IS} & \text{if} \quad RIS_t < -\sigma_{IS} \\ LIS_t = 0 & \text{otherwise} \end{cases}$$

(36)

$$\begin{cases} HIS_t = RIS_t - \sigma_{IS} & \text{if} \quad RIS_t > \sigma_{IS} \\ HIS_t = 0 & \text{otherwise} \end{cases}$$

where *LIS* is the low inventory level, *HIS* is the high level of inventories, and σ_{IS} is the standard deviation of *IS* for the entire period. The specification proposed for the forecasting model introduces both linear and non-linear terms, according to the following scheme:

(37)

$$S_t = \alpha_0 + \alpha_1 S_{t-1} + \sum_{j=0}^{5} \gamma_j D01_j + S99 + \sum_{i=0}^{k} \beta_i RIS_{t-i} + \sum_{i=0}^{k} (\gamma_i LIS_{t-i} + \delta_i LIS_{t-i}^2) + \sum_{i=0}^{k} (\phi_i HIS_{t-i} + \psi_i HIS_{t-i}^2) + \varepsilon_t$$

Results show that the use of asymmetric behavior helps to predict oil prices and that the forecasting ability of equation (37) is stronger than the simple linear specification.

Kaufmann (1995) outlines a model for the world oil market that accounts for changes in the economic, geological and political environment. This model is divided into three blocks: demand, supply and real oil import prices *(PCO)*, analyzed over the period 1954–1989. Due to the presence of two dominant oil producers in the period under scrutiny, the author models oil prices as a function of the behaviour of both agents:

(38)

$$\frac{PCO_t - PCO_{t-1}}{PCO_{t-1}} = \alpha_0 CU_{TRC_t}^2 + \alpha_1 CU_t^2 + \alpha_2 \frac{PC_t - PC_{t-1}}{PC_{t-1}} CU_t \frac{OP_t}{WD_t} + \alpha_3 (DOPEC_t - DOPEC_{t-1}) + \alpha_4 S74_t + \alpha_5 PCO_{t-1} + \alpha_6 (SOECD_t \frac{OP_t}{WD_t}) + \varepsilon_t$$

where *WD* is the world oil demand, *DOPEC* is a dummy variable for the strategic behaviour of OPEC, *S*74 is a step dummy for the 1974 oil shock, and *SOECD* is the level of OECD stocks. Equation (38) appears to have a good explanatory power in detecting oil price variations. It is interesting to note that the key factor in OPEC's behaviour is OPEC capacity.

Focusing on the recent history of oil prices, Kaufmann et al. (2004 and 2006) modify equation (38) by excluding the role of the TRC. The new specification places much more emphasis on OPEC's behaviour, since it accounts for OPEC overproduction besides OPEC quota and capacity utilization. Furthermore, the modified model outlines the impact of a new variable – the number of days of forward consumption *(DAYS)* proxied by the ratio of OECD oil stocks to OECD oil demand. Their analysis is centered on the following equation:

(39)

$$PCO_t = \alpha_0 + \alpha_1 DAYS_t + \alpha_2 OQ_t + \alpha_3 OV_t + \alpha_4 CU_t + \sum_{i=1}^{3} \beta_i DS_i + \beta_4 D90_t + \varepsilon_t$$

where *DS* are seasonal dummies and *D*90 is a dummy variable for the Persian Gulf War in the third and fourth quarters of 1990. The two studies carried out based on quarterly data differ with respect to the time period considered, which is 1986–2000 in Kaufmann et al. (2004), while Kaufmann et al. (2006) refer to the time interval 1984–2000. An error correction representation of equation (39) is estimated via the Dynamic OLS (DOLS) approach proposed by Stock and Watson (1993) and using Full Information Maximum Likelihood (FIML). Results indicate that OPEC quotas, production and capacity utilization are important in affecting oil prices. In-sample dynamic forecasts from the first quarter of 1995 to the third quarter of 2000 suggest that the performance of the model depends on the considered time period, although the proposed specification is able to capture the consequences of various exogenous shocks on the oil price level.

Merino and Ortiz (2005), extending the various works of Ye et al. (2002, 2005 and 2006), investigate whether some explanatory variables can account for the fraction of oil price variations that is not explained by oil inventories. The authors acknowledge as possible sources of variation: the difference between spot and futures prices; speculation defined as the long-run positions held by non-commercials of oil, gasoline and heating oil in the NYMEX futures market; OPEC spare capacity along with the relative level of US commercial stocks; different long-run and short-run interest rates. Exploiting causality and cointegration tests, the authors identify the importance of the speculation variable which, among others, appears to add systematic information to the model. Given the presence of cointegration, the authors eventually propose an error correction model, where oil prices are function of the percentage of relative inventories on the total current level of inventories and of speculation *(SPEC)*:

(40)

$$\Delta W_t = \alpha_0 + \alpha_1 \Delta \frac{RIS_t}{IS_t} + \alpha_2 \frac{RIS_{t-1}}{IS_{t-1}} + \alpha_3 \Delta SPEC_t + \alpha_4 SPEC_{t-1} + \alpha_5 W_{t-1} + \varepsilon_t$$

Data from January 1992 to June 2004 show that speculation helps predicting prices throughout the whole sample, except for the period 2000–2001.

A different approach in forecasting oil prices is proposed by Lalonde et al. (2003), who test the impact of the world output gap and of the real US dollar effective exchange rate gap on WTI prices. A comparison with a random walk and with an AR(1) specification suggests that both variables play an important role in explaining oil price dynamics. In Dees et al. (2007) oil prices are driven by OPEC quotas and capacity utilization, which are shown to be statistically relevant over the period 1984–2002. Sanders et al. (2009) investigate the empirical performance of the EIA model for oil price forecasting at different time horizons. This model is a mixture of structural and time series specifications, which includes supply and demand as the main factors driving oil prices, and takes into account the impact of past forecasts. The authors find that EIA three-quarter ahead oil price forecasts are particularly accurate.

Evaluation and comparison of oil price forecast models

In this study we have described three broad classes of econometric models that have been proposed to forecast oil prices. We have also presented the different and often controversial empirical results in the relevant literature. Any attempt to compare alternative oil price forecasts should be based on a comprehensive evaluation of the underlying econometric approach and model specification.

There are a number of statistical issues which should be accounted for in the development of an econometric model. Heteroskedasticity (both unconditional and conditional) as well as autocorrelation in the errors of a regression model are common problems, which, if unsolved, lead to misleading statistical inference. Another issue that comes up frequently when dealing with financial data is non-stationarity, as it is acknowledged that prices are often integrated of order one, or even two. Granger and Newbold (1974) warn that spurious regressions may arise in the presence of non-stationary variables. However, when non-stationary prices are cointegrated, it is then possible to overcome the spurious regression problem and to embed in the forecasts the information provided by the existence of one (or more than one) long-run equilibrium.

Out of the 26 papers we have reviewed, 20 provide a test for autocorrelation, 15 for heteroskedasticity and 20 account for non-stationarity and cointegration (see Table 1). Needless to say, the absence of explicit references to the use of heteroskedasticity and error autocorrelation tests as well as to a systematic check for the presence of unit roots in the analyzed series does not imply that those issues have not been accounted for, and, above all, it cannot be interpreted as evidence for the presence of heteroskedasticity, autocorrelation or non-stationarities in the analyzed data. Rather, it denotes that some authors consider it unimportant to test the statistical adequacy of their models.

The frequency of the data influences the statistical characteristics of the series, as low frequencies tend to smooth volatility. As a consequence, the choice of the data frequency can produce significant effects on the performance of a forecasting model. In general, if daily data are more volatile than their weekly, monthly and yearly averages, low-frequency oil prices can be more easily predicted than their high-frequency counterparts. The data frequencies used by the contributions reviewed in our survey are not homogeneous. Yet monthly data are most widely employed by each of the three classes of models, while weekly data are used just twice.

In addition, the literature surveyed in our paper can help to answer another question: what is the gain, if any, from using a large set of control variables in a forecasting model? In other words, why don't we simply follow the idea that all relevant information to forecast the oil price is embedded in the price itself? Random walks, martingale processes and simple autoregressive models root their justification on this idea. In this respect, random walk and martin-

Table 1

Diagnostic checking and time series properties of the data

Year	Authors	Serial correlation	Heteroskedasticity	Non stationarity and cointegration
1991	Bopp and Lady	X		
1991	Green and Mork	X	X	X
1991	Serletis	X	X	X
1992	Samii	X		
1994	Moosa and Al-Loughani	X	X	X
1995	Kaufmann	X	X	
1998	Gulen			X
1999	Pindyck	X	X	X
2000	Schwartz and Smith	X		X
2001	Morana	X	X	X
2002	Ye et al.	X		
2002	Zeng and Swanson	X		X
2003	Cortazar and Schwartz	X	X	
2003	Lalonde et al.		X	X
2004	Chernenko et al.	X	X	
2004	Zamani			X
2005	Abosedra	X		X
2005	Chin et al.	X	X	X
2005	Kaufmann et al.	X	X	X
2005	Merino and Ortiz			X
2005	Radchenko	X	X	X
2005	Ye et al.	X	X	X
2006	Kaufmann et al.	X	X	X
2006	Ye et al.	X	X	X
2007	Dees et al			X
2009	Murat and Tokat			X

Notes: X indicates the the authors have checked for serial correlation and/or heteroskedasticity and/or nonstationarity and cointegration.

gale models exploit the actual value of the price to forecast its future values, while autoregressive specifications evaluate also the lagged price values. These models have been used in many papers as benchmarks to check the forecasting performance of more complex specifications. Specifically, 9 papers out of 26 use the random walk model as a benchmark, while 4 papers compare the forecasting results of their econometric models with simple autoregressive specifications. It is important to notice that the random walk and the autoregressive model never outperform the more general specifications.

Structural models are generally considered to be an extension of autoregressive specifications that integrate the information embedded in the price history using proxies for particular relevant aspects of the oil market and the world economy. Among the surveyed papers belonging to this category, two (Lalonde et al. 2003; Ye et al. 2005) use a benchmark model as a comparison. Of these two contributions, only Ye et al. (2005) show that structural models outperform time series specifications. Financial models are based on different assumptions, as they arise either from the arbitrage theory or from the REH. Out of 13 papers in this group, 6 formally compare their models with a benchmark, either a random walk or an autoregressive specification.

The comparison with specifications which could differ from the standard benchmark models is systematically used in the papers we have reviewed as a general strategy to assess the accuracy of oil price forecasts. In Tables 2 to 4 we report the criteria proposed by the reviewed literature to evaluate the forecasting accuracy of a model, and also demonstrate that model comparison is common practice for virtually all of the structural, financial and time series models considered in this survey. Some authors (e.g. Radchenko 2005) suggest that, rather than selecting among different forecasts produced by different models, a good strategy is to combine the forecasting performance of different specifications. By combining the forecasted values obtained from an autoregressive, a random walk and a shifting trend model, it is possible to obtain significant increases in the accuracy of the forecasts.

The type of econometric model used in forecasting the price of oil seems to affect the type of forecasts that is produced. As Tables 2 to 4 clearly show, the majority of time series and structural specifications mainly use dynamic forecasts to assess the perfor-

Table 2

Criteria for comparing in-sample and out-of-sample forecasts: time series models

Year	Authors	Type of forecast		Graphical evaluation	Model comparison		Forecast evaluation				
		Static	Dynamic		Formal	Informal	RMSE	MAPE	MAE	Theil	Others
In-sample forecasts											
2005	Abosedra	X			X						X
2005	Ye et al.		X	X	X		X	X	X	X	X
Out-of-sample forecasts											
1991	Bopp and Lady		X		X		X	X	X		X
1999	Pindyck		X	X	X						
2000	Schwartz and Smith		X	X	X						X
2001	Morana		X	X	X		X			X	X
2002	Zeng and Swanson		X		X		X	X	X		X
2003	Lalonde et al.		X	X	X		X				X
2004	Chernenko et al.	X			X				X		
2005	Ye et al.		X	X	X		X	X	X	X	X
2005	Radchenko		X	X	X		X				

Notes: X indicates the presence of a specific criterium; RMSE = root mean squared error; MAPE = mean absolute percentage error; MAE = mean absolute error

mance of the analyzed model, while in the class of financial models static and dynamic forecasts have been equally employed. Given the well-known difference between static and dynamic forecasts, the latter seem to be more reasonable in the present context. Graphical evaluation of the forecasting performance of a given econometric specification has been widely used for structural models and, though in a limited number of cases, for time series models as well. Conversely, graphical methods are rarely considered in financial models. Finally, it is worthy to note that the measures of forecast errors commonly used by the surveyed articles are the root mean squared error (RMSE), the mean absolute percentage error (MAPE), the mean average error (MAE) and the Theil inequality coefficient (Theil) (see also Tables 2 to 4). Those criteria have been taken into account mainly by time series as well as structural models, and only in few cases by financial models. Despite the relatively large number of criteria, which are available to evaluate the forecasting performance of each proposed model, it is not possible to identify which class of models outperforms the others in terms of forecasting accuracy.

Table 3

Criteria for comparing in-sample and out-of-sample forecasts: financial models

Year	Authors	Type of forecast		Graphical evaluation	Model comparison		Forecast evaluation				
		Static	Dynamic		Formal	Informal	RMSE	MAPE	MAE	Theil	Others
In sample forecasts											
1992	Samii			X							
1998	Gulen					X					
2004	Chernenko et al.	X			X				X		
2005	Chin et al.		X	X	X		X		X		
1994	Moosa and Al-Loughani	X									X
2005	Abosedra	X			X						X
Out of sample forecasts											
1991	Bopp and Lady		X		X		X	X	X		X
2001	Morana		X	X	X		X			X	X
2002	Zeng and Swanson		X		X		X	X	X		X
2003	Contazar And Schwartz		X	X	X		X		X		X
2009	Murat and Tokat		X		X		X	X	X		X

Notes: X indicates the presence of a specific criterium; RMSE = root mean squared error; MAPE = mean absolute percentage error; MAE = mean absolute error

Table 4

Criteria for comparing in-sample and out-of-sample forecasts: structural models

Year	Authors	Type of forecast		Graphical evaluation	Model comparison		Forecast evaluation				
		Static	Dynamic		Formal	Informal	RMSE	MAPE	MAE	**Theil**	**Others**
In sample forecasts											
2002	Ye et al.		X								
2004	Zamani		X	X							
2005	Merino and Ortiz	X		X	X						
2005	Ye et al.		X	X	X		X	X	X	X	X
2007	Dees et al.	X	X	X	X		X				
2006	Ye et al.	X	X	X	X		X	X	X	X	X
2006	Kaufmann et al.	X	X	X			X				
Out of sample forecasts											
2003	Lalonde et al.		X	X	X		X				X
2005	Ye et al.		X	X	X		X	X	X	X	X
2006	Ye et al.	X	X	X	X		X	X	X	X	X

Notes: X indicates the presence of a specific criterium; RMSE = root mean squared error; MAPE = mean absolute percentage error; MAE = mean absolute error.

References

Abosedra, S. (2005), "Futures versus Univariate Forecast of Crude Oil Prices", *OPEC Review* 29, 231–241.

Barone-Adesi, G., F. Bourgoin and K. Giannopoulos (1998), "Don't Look Back", *Risk* 11, 100–104.

Bopp, A. E., and G. M. Lady (1991), "A Comparison of Petroleum Futures versus Spot Prices as Predictors of Prices in the Future", *Energy Economics* 13, 274–282.

Chernenko, S., K. Schwarz and J. H. Wright (2004), *The Information Content of Forward and Futures Prices: Market Expectations and the Price of Risk,* FRB International Finance Discussion Paper 808.

Chin, M. D., M. LeBlanch and O. Coibion (2005), *The Predictive Content of Energy Futures: An Update on Petroluem, Natural Gas, Heating Oil and Gasoline,* NBER Working Paper 11033.

Cortazar, G. and E. S. Schwartz (2005), "Implementing a Stochastic Model for Oil Futures Prices", *Energy Economics* 25, 215–238.

Dees, S., P. Karadeloglou, R. K. Kaufmann and M. Sanchez (2007), "Modelling the World Oil Market: Assessment of a Quarterly Econometric Model", *Energy Policy* 35, 178–191.

Efron, B. (1979), "Bootstrap Methods: Another Look at the Jackknife", *Annals of Statistics* 7, 1–26.

Energy Information Administration (EIA, 2008), *International Energy Outlook,* Washington DC.

Energy Information Administration (EIA, 2009), *Short-term Energy Outlook,* Washington DC.

Engle, R. F. and C. W. J. Granger (1987), "Co-integration and Error Correction: Representation, Estimation and Testing", *Econometrica* 55, 251–276.

Fama, E. F. (1965), "The Behaviour of Stock Market Prices", *Journal of Business* 38, 420–429.

Fama, E. F. and K. R. French (1987), "Commodity Future Prices: Some Evidence on Forecast Power, Premiums, and the Theory of Storage", *The Journal of Business* 60, 55–73.

Granger, C. W. J. and P. Newbold (1974), Spurious Regressions in Econometrics, *Journal of Econometrics* 2, 111–120.

Greenand, S. L. and K. A. Mork (1991), "Towards Efficiency in the Crude Oil Market", *Journal of Applied Econometrics* 6, 45–66.

Gulen, S. G. (1998), "Efficiency in the Crude Oil Futures Markets", *Journal of Energy Finance and Development* 3, 13–21.

Kaufmann, R. K. (1995), "A Model of the World Oil Market for Project LINK: Integrating Economics, Geology, and Politics", *Economic Modelling* 12, 165–178.

Kaufmann, R. K. (2004), Does OPEC Matter? An Econometric Analysis of Oil Prices", *The Energy Journal* 25, 67–91.

Lalonde, R., Z. Zhu and F. Demers (2003), *Forecasting and Analyzing World Commodity Prices,* Bank of Canada, Working Paper 2003–24.

Merino, A. and A. Ortiz (2005), "Explaining the So-called 'Price Premium' in Oil Markets", *OPEC Review* 29, 133–152.

Moosa, I. A., and N. E. Al-Loughani (1994), "Unbiasedness and Time Varying Risk Premia in the Crude Oil Futures Market", *Energy Economics* 16, 99–105.

Morana, C. (2001), "A Semiparametric Approach to Short-term Oil Price Forecasting", *Energy Economics* 23, 325–338.

Murat, A. and E. Tokat (2009), "Forecasting Oil Price Movements with Crack Spread Futures", *Energy Economics* 31, 85–90.

Phillips, P. C. B. and M. Loretan (1991), "Estimating Long-run Economic Equilibria", *Review of Economic Studies* 58, 407–436.

Pindyck, R. S. (1999), "The Long-run Evolution of Energy Prices", *The Energy Journal* 20, 1–27.

Radchenko, S. (2005), *The Long-run Forecasting of Energy Prices Using the Model of Shifting Trend,* University of North Carolina at Charlotte, Working Paper.

Samii, M. V. (1992), "Oil Futures and Spot Markets", *OPEC* Review 4, 409–417.

Sanders, D. R., M. R. Manfredo and K. Boris (2009), "Evaluating Information in Multiple Horizon Forecasts: The DOE's Energy Price Forecasts", *Energy Economics* 31, 189–196.

Schwartz, E. and J. E. Smith (2000); "Short-term Variations and Long-term Dynamics in Commodity Prices", *Management Science* 46, 893–911.

Schwartz, E. and J. E. Smith (1997), "The Stochastic Behaviour of Commodity Prices: Implications for Valuation and Hedging" *Journal of Finance* 51, 923–973.

Serletis A. (1991), "Rational Expectations, Risk and Efficiency in Energy Futures Markets", *Energy Economics* 13, 111–115.

Stock, H. J., and Watson M. W. (1993), "A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems", *Econometrica* 61, 783–820.

Ye, M., J. Zyren and J. Shore (2002), "Forecasting Crude Oil Spot Price Using OECD Petroleum Inventory Levels", *International Advances in Economic Research* 8, 324–334.

Ye, M., J. Zyren and J. Shore (2005), A Monthly Crude Oil Spot Price Forecasting Model Using Relative Inventories, *International Journal of Forecasting* 21, 491–501.

Ye, M., J. Zyren and J. Shore (2006), "Forecasting Short-run Crude Oil Price Using High and Low Inventory Variables", *Energy Policy* 34, 2736–2743.

Zamani, M. (2004), *An Econometrics Forecasting Model of Short Term Oil Spot Price,* Paper presented at the 6th IAEE European Conference, Zurich, 2–3 September.

Zeng, T. and N. R. Swanson (1998), "Predective Evaluation of Econometric Forecasting Models in Commodity Future Markets", *Studies in Nonlinear Dynamics and Econometrics* 2, 159–177.