Evaluating the Empirical Performance of Alternative Econometric Models for Oil Price Forecasting

Chiara Longo, Matteo Manera, Anil Markandya
and Elisa Scarpa

NOTA DI LAVORO 4.2007

JANAURY 2007

IEM - International Energy Markets

Chiara Longo, *Fondazione Eni Enrico Mattei*
Matteo Manera, *Department of Statistics, University of Milan-Bicocca and Fondazione Eni Enrico Mattei*
Anil Markandya, *Department of Economics and International Development, University of Bath and Fondazione Eni Enrico Mattei*
Elisa Scarpa, *Risk Management Department, Intesa-San Paolo*

Evaluating the Empirical Performance of Alternative Econometric Models for Oil Price Forecasting

Summary

The relevance of oil in the world economy explains why considerable effort has been devoted to the development of different types of econometric models for oil price forecasting. Several specifications have been proposed in the economic literature. Some are based on financial theory and concentrate on the relationship between spot and futures prices ("financial" models). Others assign a key role to variables explaining the characteristics of the physical oil market ("structural" models). The empirical literature is very far from any consensus about the appropriate model for oil price forecasting that should be implemented. Relative to the previous literature, this paper is novel in several respects. First of all, we test and systematically evaluate the ability of several alternative econometric specifications proposed in the literature to capture the dynamics of oil prices. Second, we analyse the effects of different data frequencies on the coefficient estimates and forecasts obtained using each selected econometric specification. Third, we compare different models at different data frequencies on a common sample and common data. Fourth, we evaluate the forecasting performance of each selected model using static and dynamic forecasts, as well as different measures of forecast errors. Finally, we propose a new class of models which combine the relevant aspects of the financial and structural specifications proposed in the literature ("mixed" models). Our empirical findings can be summarized as follows. Financial models in levels do not produce satisfactory forecasts for the WTI spot price. The financial error correction model yields accurate in-sample forecasts. Real and strategic variables alone are insufficient to capture the oil spot price dynamics in the forecasting sample. Our proposed mixed models are statistically adequate and exhibit accurate forecasts. Different data frequencies seem to affect the forecasting ability of the models under analysis.

Keywords: Oil Price, WTI Spot And Futures Prices, Forecasting, Econometric Models

JEL Classification: C52, C53, Q32, Q43

This paper has been 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 Carlo Carraro, Giliola Frey, 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.

Address for correspondence:

Matteo Manera
Department of Statistics
University of Milan Bicocca
Via Bicocca degli Arcimboldi, 8 Building U7
20126 Milan Italy
Phone: +390264485819
Fax: +390264485878
E-mail: Matteo.Manera@unimib.it

1 Introduction

The relevance of oil in the world economy is undisputable. According to Eni (2006), the world oil production in 2005 amounted to 82,268 thousand barrels per day (tbd). OPEC countries produced 33,979 tbd (41.3% of the world oil production) in 2005, while OECD countries and Europe (25 countries) were responsible of 20,317 tbd (24.7%) and 2,631 tbd (3.2%), respectively. At 1 January 2006 world oil stocks were estimated at 1,124,291 million barrels. If OPEC countries alone hold 80.2% of world oil reserves, OECD and European countries can directly count only on 7% and 0.8%, respectively. Moreover, world oil consumption in 2005 was measured in 83,292 tbd, 59.6% of which originates from the OECD countries. The impact of oil on the financial markets is at least equally important. The NYMEX average weekly open interest volume (OIV)[1] on oil futures and options contracts was equal to 999,228 contracts during the period 2002-2005, it increased to 1,653,135 contracts during 2006 and until mid September 2006, with an increment of 65.4% over the past four years, whereas it jumped to over 2 million contracts in the third week of September 2006 (source: Commodity Futures Trading Commission, 2006).

The peculiar nature of oil price dynamics has attracted the attention of many researchers in recent years. Figure 1 depicts the behaviour of the WTI spot price over the period January 1986 - December 2005. From an inspection of this graph, it is easy to verify that both level and volatility of WTI spot price are highly sensitive to specific economic and geo-political events. For instance, the small price fluctuations of the years 1986-1990 are the result of the OPEC's production quotas repeated adjustments. The 1990 sharp increase in WTI spot price is obviously due to the Gulf war. The remarkable price falls of the period 1997-1998 coincide with the pronounced slowdown of Asian economic growth. The reduction in OPEC's production quotas of 1999 has been followed immediately by a sharp price increase. Finally, if the price decreases in 2001 are related to terrorist attack of 11 September, the reduction of the WTI spot price levels recorded in the period 2002-2005 are again justified by falling OPEC production quotas and spare capacity.

The more recent evolution of the WTI spot price demonstrates how oil price forecasting is challenging. On 11 August 2005 oil price has risen to over US$ 60 per barrel (pb), while

[1] Open interest volume is measured as the sum of all long contracts (or, equivalently, as the sum of all short contracts) held by market participants at the end of a trading day. It is a proxy for the flow of money into the oil futures and options market.

one year later it has topped out at the record level of US$ 77.05 pb. Experts have again attributed the spike in oil price to a variety of economic and geo-political factors, including the North Korean crisis, the Israel-Lebanon conflict, the Iranian nuclear threat and the decline in US oil reserves. At the end of the summer 2006, the WTI oil price has begun to decrease and reached the level of US$ 56.82 pb on 20 October 2006. In the meantime, OPEC has announced production cuts to stop the sliding price. On 16 January 2007 prices have been even lower: US$ 51.21 pb for the WTI spot price and US$ 51.34 for the first position of the NYMEX oil futures contract.

Given the relevance of oil in the world economy and the peculiar characteristics of the oil price time series, it is hardly surprising that considerable effort has been devoted to the development of different types of econometric models for oil price forecasting.

Several specifications have been proposed in the economic literature. Some are based on financial theory and concentrate on the relationship between spot and futures prices ("financial" models). Others assign a key role to variables explaining the characteristics of the physical oil market ("structural" models). These two main groups of models have often been compared to standard time series models, such as the random walk and the pure first-order autoregressive model, which are simple and, differently from financial and structural models, do not rely on additional explanatory variables.

It should be noticed that most of the econometric models for oil price forecasting available in the literature are single-equation, linear reduced forms. Two recent noticeable exceptions are represented by Moshiri and Foroutan (2006) and Dees et al. (2007). The first study uses a single-equation, non-linear artificial neural network model to forecast daily crude oil futures prices over the period 4 April 1983 - 13 January 2003. The second contribution discusses a multiple-equation, linear model of the world oil market which specifies oil demand, oil supply for non-OPEC producers, as well as a price rule including market conditions and OPEC behaviour. The forecasting performance of this model is assessed on quarterly data over the period 1995-2000.

The empirical literature is very far from any consensus about the appropriate model for oil price forecasting that should be implemented. Findings vary across models, time periods and data frequencies. This paper provides fresh new evidence to bear on the following key question: does a best performing model for oil price forecasting really exist, or aren't accurate oil price forecasts anything more than a mere illusion?

Relative to the previous literature, the paper is novel in several respects.

First of all, in this paper we test and systematically evaluate the ability of several alternative econometric specifications proposed in the literature to capture the dynamics of oil prices. We have chosen to concentrate our investigation on single-equation, linear reduced forms, since models of this type are the most widely used in the literature and by the practitioners. In this respect, our study complements the empirical findings presented in Moshiri and Foroutan (2006), which are focused on the forecasting performance of a single non-linear model.

Second, this paper analyses the effects of different data frequencies (daily, weekly, monthly and quarterly) on the coefficient estimates and forecasts obtained using each selected econometric specification. The factors which potentially affect the goodness of fit and forecasting performance of an econometric model are numerous, the most important being sample period and data frequency. The fact that no unanimous conclusions could be drawn by previous studies on the forecasting performance of similar models may depend, among other things, upon the particular data frequency used in each investigation.

Third, in this paper we compare different models at different data frequencies on a common sample and common data. For this purpose, we have constructed specific data sets which enable us to evaluate different types of econometric specifications involving different explanatory variables on the same sample period. Within our composite data base, the WTI spot oil price as well as the majority of the explanatory variables are recorded at different frequencies.

Fourth, we evaluate the forecasting performance of each selected model using static and dynamic forecasts, as well as different measures of forecast errors. In contrast with previous studies, which generally employ only fixed estimation and forecasting sample periods, in this paper static and dynamic forecasts are calculated by means of fixed as well as rolling forecasting windows. The latter method is of particular importance for time series exhibiting numerous price swings, as in the case of the WTI spot price.

Finally, we propose a new class of models which combine the relevant aspects of the financial and structural specifications proposed in the literature. Our "mixed" models generally produce forecasts which are more accurate than the predictions generated by the traditional financial and structural equations.

The paper is organized as follows. In Section 2 we briefly review the existing empirical literature related to oil price forecasting. Section 3 presents and describes the data collected for the empirical analysis. In Section 4 the empirical results obtained by forecasting oil prices with alternative econometric models are discussed. The performance of each model is analysed using different measures of forecasting ability and graphical evaluation "within" each class of models (i.e. financial, structural, time series and mixed models). Section 5 summarizes the forecasting performance of the alternative specifications, with particular emphasis on "between"-class analogies and differences. Some conclusions and directions for future research are presented in Section 6.

2 The existing literature on oil price forecasting

The literature on oil price forecasting has focused on two main classes of linear, single-equation, reduced-form econometric models. The first group ("financial" models) includes models which are directly inspired by financial economic theory and based on the market efficiency hypothesis (MEH), while models belonging to the second class ("structural" models) consider the effects of oil market agents and real variables on oil prices.[2] Both financial and structural models often use pure time series specifications for benchmarking.[3]

2.1 Financial and time series models

In general, financial models for oil price forecasting examine the relationship between the oil spot price at time t (S_t) and the oil futures price at time t with maturity T (F_t), analyzing, in particular, whether futures prices are unbiased and efficient predictors of spot prices. The reference model is:

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

[2] As pointed out in the Introduction and at the beginning of Section 2, the models analysed in this paper are linear, single-equation, reduced-forms. In this context, we use the term "structural model" to identify a specification whose explanatory variables capture the real and strategic (as opposed to financial) aspects of the oil market.

[3] Interesting exceptions are Pyndyck (1999) and Radchenko (2205), who propose alternative forecasting models in a pure time series framework. See Section 2.2 for details.

where the joint null hypothesis of unbiasedness (β_0=0 and β_1=1) should not be rejected, and no autocorrelation should be found in the error terms (efficiency). A rejection of the joint null hypothesis on the coefficients β_0 and β_1 is usually rationalised by the literature in terms of the presence of a time-varying risk premium.

A sub-group of models, which are also based on financial theory but have been less investigated, exploits the following spot-futures price arbitrage relationship:

$$F_t = S_t e^{(r+\omega-\delta)(T-t)} \tag{2}$$

where r is the interest rate, ω is the cost of storage and δ is the convenience yield.[4]

Samii (1992) attempts at unifying the two approaches described in equations (1) and (2) by introducing a model where the spot price is a function of the futures price and the interest rate. Using both daily (20 September 1991 - 15 July 1992) and monthly (January 1984 - June 1992) data on WTI spot price and futures prices with three- and six-month maturity, he concludes that the role played by the interest rate is unclear and that, although the correlation between spot and futures prices is very high, it is not possible to identify which is the driving variable.

An overall comparison of financial and time series models is offered by Zeng and Swanson (1998), who evaluate the in-sample and out-of-sample performance of several specifications. The authors use a daily dataset over the period 4 January 1990 - 31 October 1991 and specify a random walk, an autoregressive model and two alternative Error Correction models (ECM, see Engle and Granger, 1987), each with a different definition of long-run equilibrium. The deviation from the equilibrium level which characterizes the first ECM is equal to the difference between the futures price tomorrow and the futures price today, i.e. the so-called "price spread". In the second ECM, the error correction term recalls the relationship between spot and futures prices, which involves the cost of storage and the convenience yield, as reported in equation (2). The predictive performance of each model is evaluated using several formal and informal criteria. The empirical evidence shows that the

[4] The arbitrage relationship (2) means that the futures price must be equal to the cost of financing the purchase of the spot asset today and holding it until the futures maturity date (which includes the borrowing cost for the initial purchase, or interest rate, and any storage cost), once the continuous dividend yield paid out by the underlying asset (i.e. the convenience yield) has been taken into account. See, among others, Clewlow and Strickland (2000) and Geman (2005) for details on the arbitrage relationship (2) for energy commodities.

ECM specifications outperform the others. In particular, the ECM based on the cost-of-storage theory performs better than the ECM which specifies the error correction term as the spot-futures price spread.

Bopp and Lady (1991) investigate the performance of lagged futures and spot oil prices as explanatory variables in forecasting the oil spot price. Using monthly data on spot and futures prices for heating oil during the period December 1980 - October 1988, they find empirical support to the cost-of-storage theory.[5] The authors also compare a random walk against the reference financial model. In this case, the empirical evidence suggests that both models perform equally well.

Serletis (1991) analyses daily data on one-month futures price (as a proxy for the spot price) and two-month futures price (quoted at NYMEX) for heating oil, unleaded gasoline and crude oil, relative to the period 1 July 1983 - 31 August 1988 (the time series of gasoline starts on 14 March 1985). He argues that the presence of a time-varying premium worsens the forecasting ability of futures prices.

In the empirical literature on oil prices there is no unanimous consensus about the validity of MEH. For instance, Green and Mork (1991) offer evidence against the validity of unbiasedness and MEH, analysing monthly prices on Mideast Light and African Light/North Sea crude oils over the period 1978-1985. Nevertheless, the authors notice that, if the subsample 1981-1985 is considered, MEH is supported by the data, because of the different market conditions characterizing the two time periods.

The unreliability of unbiasedness and MEH is also pointed out by Moosa and Al-Loughani (1994), who analyse WTI monthly data covering the period January 1986 - July 1990. The authors exploit cointegration between the series on spot price and three-month and six-month futures contracts using an ECM, and show that futures prices are neither unbiased nor efficient. Moosa and Al-Loughani apply a GARCH-in-mean model to take into account the time-varying structure of the risk premium.

Gulen (1998) asserts the validity of MEH by introducing the posted oil price as an additional explanatory variable in the econometric specification. In particular, using monthly data on WTI (spot price and one-month, three-month and six-month futures prices) for the period March 1983 - October 1995, he verifies the explanatory power of the posted price by

[5] Two different spot prices are considered, namely the national average price reported by the Energy Information Administration (EIA) in the Monthly Energy Review, and the New York Harbor ex-shore price, while the futures contract is quoted at NYMEX.

using both futures and posted prices as independent variables. Empirical evidence from this study suggests that futures prices outperform the posted price, although the latter has some predictive content in the short horizon.

Morana's analysis (2001), based on daily data from 2 November 1982 to 21 January 1999, confirms that the Brent forward price can be an unbiased predictor of the future spot price, but in more than 50 percent of the cases the sign of the changes in oil price cannot be accurately predicted. He compares a financial model with a random walk specification and shows that, when considering a short horizon, both specifications are biased.

Chernenko et al. (2004) test the MEH by focusing on the price spread relationship:

$$S_{t+T} - S_t = \beta_0 + \beta_1 (F_t - S_t) + \varepsilon_t \tag{3}$$

Analysing monthly data on WTI for the period April 1989 - December 2003, the authors compare model (3) with a random walk specification and find that the empirical performance of the two models is very similar, confirming the validity of MEH.

The same model (3) is tested by Chin et al. (2005) with a monthly dataset on WTI spot price and three-month, six-month and twelve-month futures prices covering the period January 1999 - October 2004. The empirical findings are, in this case, supportive of unbiasedness and MEH.

Another interesting application of financial models to the oil spot-futures price relationship is proposed by Abosedra (2005), who compares the forecasting ability of the futures price in model (3) with a naïve forecast of the spot price. Specifically, assuming that the WTI spot price can be approximated by a random walk with no drift, he forecasts the daily one-month-ahead price using the previous trading day's spot price and constructs the naïve monthly predictor as a simple average of the daily forecasts. Using data for the period January 1991 - December 2001, he finds that both the futures price and the naïve forecast are unbiased and efficient predictors for the spot price. The investigation of the relationship between the forecast errors of the two predictors allows the author to conclude that the futures price is a semi-strongly efficient predictor, i.e. the forecast error of the futures price cannot be improved by any information embedded in the naïve forecast.

2.2 Structural and time series models

Structural models emphasise the importance of explanatory variables describing the peculiar characteristics of the oil market. Some examples are offered by variables which are strategic for the oil market (i.e. industrial and government oil inventory levels), “real” variables (e.g. oil consumption and production), and variables accounting for the role played by OPEC in the international oil market.

Kaufmann (1995) models the real import price of oil using as structural explanatory variables the world oil demand, the level of OECD oil stocks, OPEC productive capacity, as well as OPEC and US capacity utilisation (defined as the ratio between oil production and productive capacity). The author also accounts for the strategic behaviour of OPEC and the 1974 oil shock with specific dummy variables. His analysis exploits an annual dataset for the period 1954-1989. Regression results show that his specification is successful in capturing oil price variations between 1956 and 1989, that is the coefficients of the structural variables are significant and the model explains a high percentage of the oil price changes within the sample period.

More recently, Kaufmann (2004) and Dees et al. (2007) specify a different forecasting model on a quarterly dataset. In particular, the first paper refers to the period 1986-2000, while the second contribution considers the sample 1984-2002. In these studies the authors pay particular attention to OPEC behaviour, using as structural regressors the OPEC quota (defined as the quantity of oil to be produced by OPEC members), OPEC overproduction (i.e. the quantity of oil produced which exceeds the OPEC quota), capacity utilisation and the ratio between OECD oil stocks and OECD oil demand. Using an ECM , the authors show that OPEC is able to influence real oil prices, while their econometric specification is able to produce accurate in-sample static and dynamic forecasts.

A number of authors introduce the role of the relative oil inventory level (defined as the deviation of oil inventories from their normal level) as an additional determinant of oil prices, for this variable is supposed to summarize the link between oil demand and production. In general, two kinds of oil stocks can be considered, namely industrial and governmental. The relative level of industrial oil stocks (*RIS*) is calculated as the difference between the actual level (*IS*) and the normal level of industrial oil stocks (*IS**), the latter corresponding to the industrial oil inventories de-seasonalised and de-trended. Since the government oil stocks tend to be constant in the short-run, the relative level of government oil stocks (*RGS*) can be obtained by simply removing the trend component.

Ye et al. (2002), (2005) and (2007) develop three different models based on the oil relative inventory level to forecast the WTI spot price. In their 2002 paper, the authors build up a model on a monthly dataset for the period January 1992-February 2001, where oil prices are explained in terms of the relative industrial oil stocks level and of a variable describing an oil stock level lower than normal. Ye et al. (2005) present a basic monthly model of WTI spot prices which uses, as explanatory variables, three lags of the relative industrial oil stock level, the lagged dependent variable, a set of dummies accounting for the terrorist attack of 11 September 2001 (*D01*) and a "leverage" (i.e. step) dummy equal to one from 1999 onwards (*S99*) and zero before 1999, aimed at picking a structural change of the OPEC behaviour in the oil market. The authors compare this specification with: i) an autoregressive model which includes AR(1) and AR(12) terms and dummies *D01* and *S99*; ii) a structural model where the oil spot price is a function of the one-month lag of the industrial oil inventories, the deviation of industrial oil stocks from the previous year's level, the one-month lag of the oil spot price, as well as the dummy variables *D*01 and *S*99. Each model is estimated over the period 1992-2003. The basic model outperforms the other two specifications: in particular, the time series model is unable to capture oil price variability. The performance of each model is evaluated by calculating out-of-sample forecasts for the period 2000-2003. The forecasting accuracy of the two structural models depends on the presence of oil price troughs or peaks within the sample period. When considering three-month-ahead forecasts, the basic model exhibits a higher forecasting performance in presence of oil price peaks, while the second structural specification outperforms the basic model in presence of oil price troughs. On the basis of this last evidence, Ye et al. (2007), using the same dataset, take into account the asymmetric transmission of oil stock changes to oil prices. The authors define a low (*LIS*) and a high (*HIS*) relative industrial oil stock level as follows:

$$
\begin{aligned}
&\begin{cases} LIS_t = RIS_t + \sigma_{IS} & \text{if} \quad RIS_t < -\sigma_{IS} \\ LIS_t = 0 & \text{otherwise} \end{cases} \\
&\begin{cases} HIS_t = RIS_t - \sigma_{IS} & \text{if} \quad RIS_t < \sigma_{IS} \\ HIS_t = 0 & \text{otherwise} \end{cases}
\end{aligned}
\qquad (4)
$$

where σ_{IS} indicates the standard deviation of the industrial oil stock level.

The estimated model is:

$$S_t = \alpha_0 + \alpha_1 S_{t-1} + \sum_{j=0}^{5} \psi_j D01_{jt} + \lambda S99_t + \sum_{i=0}^{k} \beta_i RIS_{t-i} + \sum_{i=0}^{k} \left(\gamma_i LIS_{t-i} + \delta_i LIS_{t-i}^2 \right) + \\ + \sum_{i=0}^{k} \left(\phi_i HIS_{t-i} + \varphi_i HIS_{t-i}^2 \right) + \varepsilon_t \tag{5}$$

which shows a more accurate forecasting performance than the linear specification proposed by Ye et al. (2005).

Following Ye et al. (2002), Merino and Ortiz (2005) specify an ECM with the percentage of relative industrial oil stocks and “speculation” (defined as the log-run positions held by non-commercials of oil, gasoline and heating oil in the NYMEX futures market) as explanatory variables. Evidence from January 1992 to June 2004 demonstrates that speculation can significantly improve the inventory model proposed by Ye et al., especially in the last part of the sample.

Zamani (2004) proposes a forecasting model based on a quarterly dataset for the period 1988-2004 and specifies an ECM with the following independent variables: OPEC quota, OPEC overproduction, *RIS*, *RGS*, non-OECD oil demand and a dummy for the last two quarters of 1990, which accounts for the Iraq war. The accuracy of the in-sample dynamic forecasts is indicative of the model’s capability of capturing the oil price evolution.

In the pure time series framework, two models, which are particularly useful for forecasting oil prices in the long-run, are proposed by Pindyck (1999) and Radchenko (2005). The data used by the authors cover the period 1870-1996 and refer to nominal oil prices deflated by wholesale prices expressed in US dollars (base year is 1967). Pindyck (1999) specifies the following model:

$$\begin{cases} S_t = \rho S_{t-1} + (\beta_1 + \phi_{1t}) + (\beta_2 + \phi_{2t})t + \beta_3 t^2 + \varepsilon_t \\ \phi_{1t} = \alpha_1 \phi_{1,t-1} + \upsilon_{1t} \\ \phi_{2t} = \alpha_2 \phi_{2,t-1} + \upsilon_{2t} \end{cases} \tag{6}$$

where ϕ_{1t} and ϕ_{2t} are unobservable state variables. He estimates the model with a Kalman filter and confronts its forecasting ability with the following specification:

$$S_t = \rho S_{t-1} + \beta_1 + \beta_2 t + \beta_3 t^2 + \varepsilon_t \qquad (7)$$

on the full dataset and three sub-samples, namely 1870-1970, 1970-1980 and 1870-1981. Model (6) offers a better explanation of the fluctuations of oil prices, while specification (7) produces more accurate forecasts.

Radchenko (2005) extends Pindyck's model, allowing the error terms to follow an autoregressive process:

$$\begin{cases} S_t = \rho S_{t-1} + \beta_1 + \phi_{1t} + \phi_{2t} t + \varepsilon_t \\ \phi_{1t} = \alpha_1 \phi_{1,t-1} + \upsilon_{1t} \\ \phi_{2t} = \alpha_2 \phi_{2,t-1} + \upsilon_{2t} \\ \varepsilon_t = \varphi \varepsilon_{t-1} + u_t \end{cases} \qquad (8)$$

The forecasting horizons are 1986-2011, 1981-2011, 1976-2011 and 1971-2011. Overall, the empirical findings confirm Pindyck's results, although the model is unable to account for OPEC behaviour, leading to unreasonable price declines. Nevertheless, the author suggests that forecasting results can be improved significantly by combining specification (8) with a random walk and an autoregressive model, which can be considered a proxy for future OPEC behaviour.

3 The data

We have constructed four different datasets, with the following frequencies: daily, weekly, monthly and quarterly. Prices refer to WTI crude oil spot price (*S*) and WTI crude oil futures prices contracts with one-month, two-month, three-month and four-month maturity (*F*1-*F*4), as reported by EIA. Weekly, monthly and quarterly data have been obtained by aggregating daily observations with simple arithmetic means, taking into account that the futures contract rolls over on the third business day prior to the 25th calendar day of the month preceding the delivery month. The sample covers the period 2 January 1986 - 31 December 2005.

Due to the limited availability of structural variables at high frequencies, the daily and weekly datasets include observations on the WTI prices only. Therefore, we have concentrated our analysis on financial and time series models at daily and weekly frequencies, whereas we have estimated the structural specifications using monthly and quarterly data.

The monthly dataset includes observations over the period January 1988 - August 2005 for the following variables: OECD government (*GS*) and industrial (*IS*) crude oil stocks; oil consumption in the OECD countries (*OC*); the world crude oil production (*WP*); the non-OPEC crude oil production (*NOP*); the commodity price index (*PPI*), with June 1982 as basis. All variables are expressed in million barrels per day (mbd) and are obtained from EIA, with the single exception of *PPI*, which is from the Bureau of Labor Statistics.

The quarterly data range from the first quarter of 1993 to the third quarter of 2005 and refer to the following variables: *GS*; *IS*; OECD (*OD*) and non-OECD (*NOD*) oil demand (source: International Energy Agency); OPEC (*OP*) and non-OPEC (*NOP*) crude oil production (expressed in mbd and obtained from EIA); OPEC sustainable oil production capacity (*PC*) in mbd (source: Petroleum Intelligence Weekly); OPEC quota (*OQ*) in mbd (source: EIA); the short-term interest rate (*I*), obtained from the Federal Reserve Board of Governors.

Moreover, we have constructed the following variables: OPEC overproduction (*OV*), as the difference between OPEC oil production and OPEC quota; OPEC capacity utilization (*CU*), as 100 times the ratio between production and productive capacity; OPEC spare capacity (*SC*), given by the difference between *PC* and *OP*.

The complete list of the variables employed in the empirical analysis is summarized in Table 1. Table 2 reports some descriptive statistics, disaggregated by frequency. It is worth noticing that the annualized standard deviation for financial prices is highest for the daily frequency and decreases as frequencies decrease. Conversely, the coefficient of variation shows a homogeneous behaviour of the WTI prices for all frequencies. The large majority of the other variables seem to be less volatile when the quarterly frequency is considered.

Prior to estimation, we have checked for the presence of unit roots in the variables using standard Augmented Dickey-Fuller tests. All variables are integrated of order one, or $I(1)$, with the exception of industrial oil stocks (at monthly and quarterly frequencies) and of world (i.e. OPEC and non-OPEC) crude oil production, which turn out to be stationary, or $I(0)$ (see Table 3). Moreover, we have tested for bi-variate cointegration between the WTI spot price and each futures price using the Johansen test. The empirical results (see Tables 4-7) are always supportive of the presence of one cointegrating relationship between the spot price and each futures price.

4 Empirical results

We have evaluated the forecasting performance of different econometric models available in the existing literature, which can be reconducted to the two main classes described in Section 2, namely "financial" and "structural" models. We also propose a new class of models which combine the relevant aspects of financial and structural models ("mixed" models), and are based on the assumption that the interaction between financial and macroeconomic variables can improve the understanding of oil price behaviour. Financial, structural and mixed models are confronted with pure time series specifications, such as the random walk with drift and the first-order autoregressive model.

The estimation period for time series and financial models runs from January 1986 up to December 2003, while the interval from January 2004 to December 2005 is used for forecast evaluation. Structural and mixed models have been estimated on the sample January 1993 - December 2003, and forecasts have been produced for the period January 2004 - August 2005. For all models and frequencies, the estimation method is Ordinary Least Squares (OLS). Financial models have also been expressed in terms of ECM, in order to exploit the cointegrating relationship between oil spot and futures prices.

Four different types of forecasts have been produced: i) static forecasts with fixed estimation and forecasting samples;[6] ii) dynamic forecasts with fixed estimation and forecasting samples;[7] iii) static forecasts with a two-year-width rolling estimation and

---

[6] A static forecast for the oil spot price is defined as a one-step-ahead forecast for S_t. Assume that the reference model is: $S_t = \alpha S_{t-1}+\beta X_t+\varepsilon_t$, where X_t is a generic regressor and ε_t is a classical error term. The fixed estimation sample is $t=1,\ldots,T$, whereas the forecast sample is $t=T+1,\ldots,T+k$. The reference model is estimated on the fixed estimation sample to obtain OLS estimates of the parameters, i.e. $\hat{\alpha}$ and $\hat{\beta}$. Then, the sequence of static forecasts is calculated as: $\hat{S}_{T+1} = \hat{\alpha}S_T + \hat{\beta}X_{T+1}$; $\hat{S}_{T+2} = \hat{\alpha}S_{T+1} + \hat{\beta}X_{T+2}$, etc.

[7] A dynamic forecast for the oil spot price is a multi-step-ahead forecast for S_t. Assume that the reference model is identical to the model described in Footnote 6. As in the static forecast framework, the reference model is estimated on the fixed estimation sample to obtain OLS estimates of the parameters, i.e. $\hat{\alpha}$ and $\hat{\beta}$. Then, the sequence of dynamic forecasts is calculated as: $\hat{S}_{T+1} = \hat{\alpha}S_T + \hat{\beta}X_{T+1}$; $\hat{S}_{T+2} = \hat{\alpha}\hat{S}_{T+1} + \hat{\beta}X_{T+2}$, etc. It is evident that one-step-ahead static and dynamic forecasts are identical, while, for $n>2$, n-step-ahead static forecasts differ from the corresponding n-step-ahead dynamic forecasts, since the sequence of actual values S_{T+1},

forecasting window; iv) dynamic forecasts with a two-year-width rolling estimation and forecasting window.

The computation of rolling forecasts involves the following steps. First, the base equation is estimated on a rolling window, whose width has been chosen to be equal to two years. Second, dynamic and static forecasts are produced on a two-year width forecasting window using the estimated coefficients obtained in the first step, and different measures of forecasting performance are computed. Third, we iterate on steps one and two by rolling both the estimation and forecast window by one period, until the end of sample is met. A direct evaluation of the impact of the forecasting approach (i.e. fixed sample versus rolling window) on the forecasting performance of each estimated model is obtained by calculating, for each forecasting measure, the simple arithmetic mean of its values obtained at each iteration.

Four canonical measures have been used to evaluate the forecasting performance of each estimated model: Mean Absolute Error (MAE); Mean Absolute Percentage Error (MAPE); Theil Inequality Coefficient (Theil); Root Mean Squared Error (RMSE).[8]

4.1 Financial models

In Section 3 we have pointed out that, independently of the frequency considered, the WTI spot price and the four WTI futures prices involved in the empirical analysis are $I(1)$. Moreover, the WTI spot price and each WTI futures price are cointegrated, that is there exists a stationary, long-run equilibrium relationship between the WTI spot price and the WTI futures price at different maturities. Equivalently, the residuals of the relationship between spot and futures prices:

S_{T+2}, …, enters the expression of the static forecasts, while the dynamic forecasts depend on the sequence of predicted values $\hat{S}_{T+1}$, $\hat{S}_{T+2}$,….

[8] Suppose that the forecast sample is $t=T+1,\ldots,T+k$ and that, at time t, the actual and fitted values of the dependent variable are, respectively, S_t and $\hat{S}_t$. The forecasting evaluation measures can be defined as:

$$MAE = \frac{\sum_{t=T+1}^{T+k} \left|\hat{S}_t - S_t\right|}{k}, \quad RMSE = \sqrt{\frac{\sum_{t=T+1}^{T+k}\left(\hat{S}_t - S_t\right)^2}{k}}, \quad Theil = \frac{\sqrt{\frac{\sum_{t=T+1}^{T+k}\left(\hat{S}_t - S_t\right)^2}{k}}}{\sqrt{\frac{\sum_{t=T+1}^{T+k}\hat{S}_t^2}{k}} + \sqrt{\frac{\sum_{t=T+1}^{T+k}S_t^2}{k}}}, \quad MAPE = 100\frac{\sum_{t=T+1}^{T+k}\left(\frac{\hat{S}_t - S_t}{S_t}\right)}{k}.$$

While RMSE and MAE are scale-dependent and should be used to compare forecasts for the same variable across different models, MAPE and Theil are scale-invariant. Moreover, Theil ranges from 0 to 1, with zero indicating perfect fit.

$$S_t = \alpha + \beta F_t + \varepsilon_t \qquad (9)$$

are $I(0)$, for T equal to one month, two months, three months and four months, respectively. The presence of cointegration between S_t and F_t can be exploited via the following ECM representation:

$$\Delta S_t = \alpha + \beta \Delta F_t + \gamma ECT_{t-1} + \eta_t \qquad (10)$$

where the error correction term (*ECT*) is given by the residuals of model (9).

The estimation results and the forecasting performance of model (10) for different frequencies and futures price contracts are reported in Tables 8-11. For each data frequency and futures price, the constant term α is not significant, while the coefficient β is significantly different from zero and close to one. These findings support the hypothesis that futures prices are unbiased predictors of spot prices. The coefficient of adjustment γ is always significant and negative; its absolute value decreases as futures maturity increases, indicating that convergence to the long-run equilibrium is faster for one-month than for our-month futures contracts.

For all data frequencies, with the exception of weekly data, the goodness of fit of the estimated model, summarized by the adjusted-R^2, decreases with the maturity of WTI futures prices. Moreover, models with the most satisfactory explanatory ability are all estimated on monthly and quarterly data.

Residual autocorrelation has been investigated with the Breusch-Godfrey Lagrange multipliers (LM) test.[9] Results highlight the presence of high-order serial correlation in the residuals for all models, except for the specification estimated on monthly data.

The presence of heteroskedasticity in the residuals has been checked with the White LM test.[10] The null hypothesis of homoskedasticity is rejected by the models estimated on daily,

---

[9] The null hypothesis of the Breusch-Godfrey (BG) test is no residual autocorrelation of order *p*. Under the null hypothesis, the BG test statistic has an asymptotic χ^2 distribution, with *p* degrees of freedom.

[10] The null hypothesis of the White (W) test for heteroskedasticity is that the squared regressors and regressors cross-products do not contribute to the explanation of the model squared residuals. Under the null hypothesis of homoskedasticity, the W test statistic is asymptotically χ^2 -distributed with *q* degrees of freedom, *q* being the number of squared regressors and regressors cross-products.

weekly and monthly data (with the exception of the model involving the one-month WTI futures contract).

The evaluation measures calculated on the static forecasts with fixed estimation and forecasting samples show that an increase of the maturity of futures prices worsens the forecasting capability of the estimated models. A noticeable exception is represented by weekly data, where model (10) involving the two-month futures prices produces the most accurate forecasts. When dynamic forecasts with fixed estimation and forecasting samples are considered, the values of MAE, RMSE, Theil and MAPE suggest that the models estimated on daily and weekly data generate inaccurate forecast irrespective of the maturity of the futures contract, while the forecasts obtained by model (10) on monthly and quarterly data are more satisfactory.

The empirical performance of static and dynamic forecasts with a two-year-width rolling estimation and forecasting window can be summarized as follows. All the forecasting evaluation criteria associated with the static forecasts point out that the model performance is a decreasing function of the maturity associated with the futures prices involved in the econometric specification. The results in terms of dynamic forecasting are very similar to the corresponding case with fixed estimation and forecasting samples.

The empirical results obtained by estimating the benchmark time series model (1) on monthly data are reported in Table 12. There are no significant differences in the coefficient estimates with respect to model (10), although the overall explanatory power of the regression is slightly lower. Each forecasting measure shows the reduced forecasting ability of model (1) with respect to the ECM specification (10). In particular, MAE and RMSE are much higher than the corresponding values obtained for model (10).

4.2 Structural and mixed models

Structural and mixed models have been estimated only for monthly and quarterly frequencies, due to the lack of data on the structural variables at higher frequencies.

For monthly data we propose two different specifications. In the basic mixed model the WTI spot price is regressed on the WTI futures price, OPEC consumption, the relative inventory industrial level of the previous month and a step dummy for 1999 (*S99*), which accounts for a structural change of the OPEC's behaviour in the international oil market:

$$S_t = \alpha + \beta F_t + \gamma OC_t + \delta RIS_{t-1} + \lambda S99_t + \varepsilon_t \tag{11}$$

The structural specification considers as explanatory variables the relative oil inventory level of the previous month as well as of the previous year, the world oil production of the previous month, the commodity price index, the step dummy *S99* and a set of dummy variables capturing the effects of 11 September 2001 (*D*01):

$$S_t = \alpha + \beta RIS_{t-1} + \gamma WP_{t-1} + \delta RIS_{t-12} + \phi PPI_t + \lambda S99_t + \sum_{j=0}^{5} \psi_j D01_{jt} + \varepsilon_t \tag{12}$$

The empirical findings show that the mixed model (11) has a much higher explanatory ability than the structural model (12). Moreover, the residuals of the mixed model (11) are less affected by serial correlation and heteroskedasticity (see Table 13).

The lagged relative industrial oil stock levels negatively affects the oil spot price, irrespective of the number of lags, although the estimated coefficients are very small in both specifications. An increase in the world oil production causes a reduction in the WTI spot price, as expected. On the contrary, in model (11) the rise of OECD oil consumption leads to an increase of the WTI oil spot price. There is also evidence that the commodity price index and the oil spot price move in the same direction. Finally, from inspection of Table 13, the forecasting evaluation measures indicate the dominance of the mixed model (11) over the structural model (12).

On quarterly data we estimate the following four different types of models:

$$\ln(S_t) = \alpha + \beta \ln F_t + \gamma \Delta \ln F_t + \delta \Delta \ln(NOD_t + OD_t) + \phi \ln IS_t + \vartheta \Delta \ln IS_t + \varepsilon_t \tag{13}$$

$$\ln(S_t) = \alpha + \delta \ln(NOD_t + OD_t) + \phi \ln IS_t + \vartheta \Delta \ln IS_t + \varphi \ln S_{t-1} + \varepsilon_t \tag{14}$$

$$S_t = \alpha + \beta RIS_{t-1} + \delta RIS_{t-4} + \phi_1 OP_{t-1} + \phi_2 (NOD_t + OD_t) + \lambda S99_t + \sum_{j=0}^{5} \psi_j D01_{jt} + \varepsilon_t \tag{15}$$

$$S_t = \alpha + \beta RIS_{t-1} + \delta RIS_{t-4} + \phi_1 OP_{t-1} + \phi_2 \left(NOD_t + OD_t\right) + \varphi S_{t-1} + \lambda S99_t + \sum_{j=0}^{5} \psi_j D01_{jt} + \varepsilon_t \qquad (16)$$

Specification (13) is a mixed model, model (15) is purely structural, while models (14) and (16) are structural specifications, where the lagged dependent variable is introduced among the regressors to solve for residual autocorrelation.

The mixed model (13) outperforms the other three specifications in term of explanatory power: the adjusted-R^2 is 0.99 for the mixed model (13), close to 0.94 for models (15) and (16), and equal to 0.83 for model (14). The diagnostic tests for serial correlation and heteroskedasticity do not highlight any problem in the residuals of each model (see Table 14).

As already discussed for monthly data, OPEC oil production and the industrial oil stocks variable, irrespective of the way it enters the specification (i.e. level or relative level), have a negative impact on the oil spot price. The world oil demand appears in three of four specifications, with a positive influence on the oil spot price. The forecasting performance of the mixed model (13) is clearly superior for both fixed and rolling forecasts. A comparison among the other three models shows that specification (14) provides the most favourable values for each forecasting indicator, although its explanatory power is the lowest.

In Tables 15-17 short-run and long-run marginal effects, as well as short-run and long-run elasticities are reported. With monthly data (see Table 15), the effects exerted by the relative oil inventories level (RIS) over the oil spot price are very small and exhibit a negative sign. In model (12) both short-run and long-run marginal effects are negative. In particular, in the long-run, as expected, the marginal effects of the relative oil inventories level over the oil spot price are larger, in absolute value, than in the short-run.

Larger short-run impacts are generated on the oil spot price by lagged world oil production (WP) and commodity prices index (PPI), being negative for the former and positive for the latter.

Estimation results of structural models (15) and (16) on quarterly data are reported in Table 17. The relative oil inventories level has still a negative marginal effect on oil spot prices, both in the short-run and in the long-run. OPEC oil production (OP) has a negative effect on the spot oil price. On the contrary, total oil demand (i.e. $NOD + OD$) positively affects the WTI spot price.

In Table 16 short-run and long-run elasticities of the oil spot price to different explanatory variables are presented. Within the mixed model (13), the response of the oil spot price to a contemporaneous change in industrial inventories (IS) is negative: the short-run elasticity is equal to -1.049, indicating that a variation in industrial oil stocks is associated with a decrease in the spot price of the same amount. When structural model (14) is estimated, the oil spot price is more reactive to industrial oil inventories. In this case, the short-run elasticity is equal to -2.101. The long-run elasticity, which represents the average response of the oil spot price to a change in industrial oil inventories within the estimation period, is equal to -6.256, showing a very high sensitivity of the oil spot price to oil inventories. Both short-run and long-run elasticities of the oil spot price to total oil demand are positive, being equal to 0.964 and 2.871, respectively, and indicate a strong reactivity of prices to quantities.

4.3 Time series models

When the model for the oil spot price is a random walk, the implicit assumption is that the best predictor for the oil price tomorrow is the oil price today. On the contrary, if we believe that the data generating process underlying the oil spot price is first-order autoregressive, we assume that the current value of the oil spot price does not embed the total amount of information needed for accurate forecasting. Instead, we are saying that the oil price today strictly influences the realization of the oil price tomorrow, the strength of this effect depending on how the autoregressive coefficient is close to zero or one.

Tables 18 and 19 summarize the estimation results, diagnostic tests and forecasting indicators associated with the random walk model and the autoregressive model. Specifically, we have estimated the following random walk with drift:

$$S_t = \alpha + S_{t-1} + \varepsilon_t \tag{17}$$

The drift α is not significant for all frequencies. The adjusted-R^2 is rather high: when daily and weekly data are considered, it is equal to 0.99 and 0.93, respectively, and it slightly decreases with the data frequencies. Serial correlation has been detected for all frequencies except the quarterly data, whereas heteroskedasticity affects model residuals for all data frequencies.

The proposed measures of forecasting evaluation, calculated on the static forecasts with fixed estimation and forecasting sample, suggest that the oil spot price today is a good predictor of the oil spot price tomorrow, but also that its forecasting ability decreases with data frequency. Similar conclusions emerge from the inspection of the values of MAE, RMSE, Theil and MAPE calculated for rolling static forecasts.

Conversely, both fixed and rolling dynamic forecasts exhibit an unexpected behaviour. In the first case, lagged oil spot price seems to be a better predictor for actual oil spot price when the model is estimated with daily and monthly data. In the second case, more accurate rolling dynamic forecasts are produced by the model estimated on weekly and monthly data.

When the data generating process for the oil spot price is supposed to be first-order autoregressive, the oil spot price is modelled as:

$$S_t = \alpha + \rho S_{t-1} + \varepsilon_t \tag{18}$$

Our empirical analysis shows that the constant term is statistically insignificant for daily and weekly data, while it becomes significant at 5 percent when model (18) is estimated on monthly and quarterly data. The autoregressive coefficient ρ is significant at all frequencies, and its value is generally very close to 1. The adjusted-R^2 ranges from 0.99 for daily data to 0.70 for quarterly data. Residual serial autocorrelation has not been detected for the quarterly model only, while only the weekly model has homoskedastic residuals.

The forecasting evaluation indicators outline that model (18) at daily frequency outperforms the same specification at the other frequencies, if static and dynamic forecasts with fixed sample and rolling static forecasts are considered. A different behaviour characterizes the rolling dynamic forecasts, where the best forecasting performance comes from the monthly model.

.

5 Overall comparison

In Section 4 the forecasting performance of financial, structural and mixed models is evaluated, in order to verify whether it is possible to identify, within each class, a best performing model. Simple time series specifications have been included in the evaluation procedure as benchmarks against which each model, “within” each class, can be compared.

This section aims at emphasizing the relevance of “between”-class comparisons for a thourough evaluation of the forecasting ability of each econometric specification.

Financial models generally exhibit, for all frequencies, a more satisfactory forecasting behaviour than pure time series specifications. While time series models seem to produce more accurate forecasts when fitting daily data, financial models are preferable with monthly and quarterly frequencies. It is interesting to notice that, within the class of financial models, monthly forecasts are the most accurate and outperform the forecasts obtained on quarterly data.

For all frequencies, the explanatory power of time series models is quite high when compared to more complex models, indicating that the inclusion of the lagged dependent variable captures most of the dynamics in the oil spot price. For forecasting purposes, however, we notice that pure time series models are less accurate than financial specifications for all frequencies, as all measures univocally indicate. Furthermore, the forecasting ability of pure time series models seems to be more sensitive to data frequency: the volatility of the values recorded by the majority of the indicators of forecasting performance is larger for time series models at different data frequencies.

The comparison between time series models and structural models suggests that the latter perform significantly better than the former at the estimation level for monthly and quarterly frequencies. However, this superiority dies away when the focus is on forecasting. On this respect, the only specifications which outperform the pure time series models are the mixed models, which include the oil futures price among the explanatory variables.

Within the class of mixed models, the most reliable forecasts are generated with monthly data, while for structural models the quarterly frequency produces better results. If, on the one hand, the quarterly dataset permits to propose several specifications for both structural and mixed models, on the other hand the use of monthly data allows us to estimate only two specifications, both affected by serial correlation and heteroskedasticity. One possible interpretation for the in-sample statistical inaccuracy of the models estimated on monthly data concentrates on data frequency: temporal aggregation of the data may help to eliminate error serial dependence and volatility clustering. However, we cannot exclude that the difficulty with monthly specifications is directly linked to the limited number of variables entering the estimated mixed and structural specifications.

Figures 2 and 3 graphically summarize the main empirical findings. Figure 2 illustrates, for the period January 2003 - December 2005, the different forecasting behaviour of financial models (1) and (10) at monthly frequency. As already noted, model (1) is unable to capture the future dynamics of the spot oil price, while model (10) produces a very accurate fit. Nevertheless, model (10) is of little use in a true out-of-sample forecasting framework. Actually, model (10) requires the prediction of the futures price, which shares the same difficulties as predicting the spot price.

The graphical comparison among financial model (10), mixed model (13), structural model (16) and the random walk (17) is reported in Figure 3. The quarterly financial and mixed models (10) and (13) perform fairly well, due to the presence of the futures contract among the explanatory variables. The random walk model (17) seems to capture the trend in the data, but fails to produce reliable forecasting values. Finally, the performance of structural model (16) is severely insufficient in capturing the oil price dynamics.

Although it is not possible to provide a rigorous ranking of the estimation and forecasting performance of the competing models, the empirical findings presented in this paper can be summarized as follows. First, financial models in levels do not produce satisfactory forecasts for the WTI spot price, since the forecasted price values generally “follow” the actual price values. Second, the financial ECM specification yields accurate in-sample forecasts. Financial ECM takes into account the short-run and long-run contemporaneous relationships between oil spot and futures prices, but it can hardly be employed for true out-of-sample forecasting, due to the presence of the oil futures price among the regressors. Third, real and strategic variables alone are insufficient to capture the oil spot price dynamics in the forecasting sample. This result explains the generally poor forecasting performance of the structural models, which are also heavily dependent on the correct specification of the forecasting mechanism for the exogenous variables. Fourth, our proposed mixed models, which exploit the combination of financial, real and strategic explanatory variables, are statistically adequate and exhibit accurate forecasts. Fifth, different data frequencies seem to affect both estimation and the forecasting ability of the models under analysis. In general, models estimated on low frequency data tend to generate more accurate forecasts. Finally, although pure time series models allow the researcher to compute true out-of-sample forecasts, their in-sample forecasting performance is far from being satisfactory.

6 Conclusions

The relevance of oil in the world economy as well as the specific characteristics of the oil price time series explain why considerable effort has been devoted to the development of different types of econometric models for oil price forecasting.

Several specifications have been proposed in the economic literature. Some are based on financial theory and concentrate on the relationship between spot and futures prices ("financial" models). Others assign a key role to variables explaining the characteristics of the physical oil market ("structural" models).

The empirical literature is very far from any consensus about the appropriate forecasting model that should be implemented. Findings vary across models, time periods and data frequencies.

Relative to the previous literature, the paper is novel in several respects.

First of all, we test and systematically evaluate the ability of several alternative econometric specifications proposed in the literature to capture the dynamics of oil prices. We have chosen to concentrate our investigation on single-equation, linear reduced forms, since models of this type are the most widely used in the literature and by the practitioners.

Second, we analyse the effects of different data frequencies (daily, weekly, monthly and quarterly) on the coefficient estimates and forecasts obtained using each selected econometric specification. The fact that no unanimous conclusions could be drawn by previous studies on the forecasting performance of similar models may depend, among other things, upon the particular data frequency used in each investigation.

Third, we compare different models at different data frequencies on a common sample and common data. We have constructed specific data sets which enable us to evaluate different types of econometric specifications involving different explanatory variables on the same sample period.

Fourth, we evaluate the forecasting performance of each selected model using static and dynamic forecasts, as well as different measures of forecast errors. In contrast with previous studies, in this paper static and dynamic forecasts are evaluated by means of fixed as well as rolling forecasting windows. The latter method is of particular importance for time series exhibiting numerous price swings, as in the case of the WTI spot price.

Finally, we propose a new class of models which combine the relevant aspects of the financial and structural specifications proposed in the literature. Our “mixed” models generally produce forecasts which are more accurate than the predictions generated by the traditional financial and structural equations.

Although it is not possible to provide a rigorous ranking of the estimation and forecasting performance of the competing models, the empirical findings presented in this paper can be summarized as follows. Financial models in levels do not produce satisfactory forecasts for the WTI spot price. The financial ECM specification yields accurate in-sample forecasts. Real and strategic variables alone are insufficient to capture the oil spot price dynamics in the forecasting sample. Our proposed mixed models, which exploit the combination of financial, real and strategic explanatory variables, are statistically adequate and exhibit accurate forecasts. Different data frequencies seem to affect both estimation and the forecasting ability of the models under analysis. Although pure time series models allow the researcher to compute true out-of-sample forecasts, their in-sample forecasting performance far from being satisfactory.

The empirical results presented in this paper point out that a best performing econometric model for oil price forecasts is still to appear in the literature. For this reason, we suggest two promising directions for future work in this area. First, it could be useful to develop more accurate economic models for key financial and structural driving variables. Examples are provided by models which combine physical oil reserves with economic and regulatory variables (e.g. Moroney and Berg, 1999), or which describe OPEC as well as non-OPEC behaviour (see, among others, Dees et al., 2007). Models of this type can be used as forecasting mechanisms for the driving variables, and are likely to improve the out-of-sample forecasting performance of the financial and structural models currently available in the literature. Second, it is crucial to identify a set of variables which accurately reflect changes in oil market expectations, such as the non-commercial long positions on oil futures markets used to proxy oil futures prices (Merino and Ortiz, 2005).

References

Abosedra S. (2005), “Futures versus univariate forecast of crude oil prices”, *OPEC Review*, **29**, 231-241.

Bopp A. E., and Lady G. M. (1991), “A comparison of petroleum futures versus spot prices as predictors of prices in the future”, *Energy Economics*, **13**, 274-282.

Chinn M. D., LeBlanch M., and Coibion O. (2005), “The predictive content of energy futures: an update on petroleum, natural gas, heating oil and gasoline”, *NBER Working Paper* n. 11033.

Clewlow L., and Strickland C. (2000), *Energy Derivatives: Pricing and Risk Management*, London, Lacima Publications.

Commodity Futures Trading Commission (2006), *The Commitments of Traders Report*, 5 December 2006 (www.cftc.org/cftc/cftccotreports.htm).

Dees S., Karadeloglou P., Kaufmann R.K. and Sanchez M. (2007), “Modelling the world oil market: assessment of a quarterly econometric model”, *Energy Policy*, **35**, 178-191.

Engle R. F., and Granger C.W.J. (1987), “Co-integration and error correction: representation, estimation and testing”, *Econometrica*, **55**, 251-276.

Eni (2006), *World Oil and Gas Review*, Rome, Eni SpA, Strategies & Development Department.

Geman H. (2005), *Commodities and Commodity Derivatives*, Chichester, Wiley.

Green S. L., and Mork K. A. (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.

MacKinnon J. G. (1991), “Critical values for cointegration tests”. In: R.F. Engle and C. Granger (eds.), *Long-run Economic Relationships*, Oxford, Oxford University Press, 267-276.

MacKinnon J. G. (1996), “Numerical distribution functions for unit root and cointegration tests”, *Journal of Applied Econometrics*, **11**, 601-618.

Merino A., and Ortiz A. (2005), “Explaining the So-called ‘Price Premium’ in Oil Markets”, *OPEC Review*, **29**, 133-152.

Moosa I. A., and Al-Loughani N. E. (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.

Moroney J.R. and Berg D. (1999), “An integrated model of oil production”, *The Energy Journal*, **20**, 105-124.

Moshiri S. and Foroutan F. (2006), “Forecasting nonlinear crude oil futures prices”, *The Energy Journal*, **27**, 81-95.

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.

Serletis A. (1991), “Rational expectations, risk and efficiency in energy futures markets”, *Energy Economics*, **13**, 111-115.

Ye M., Zyren J., and Shore J. (2002), “Forecasting crude oil spot price using OECD petroleum inventory levels”, *International Advances in Economic Research*, **8**, 324-334.

Ye M., Zyren J., and Shore J. (2005), “A monthly crude oil spot price forecasting model using relative inventories”, *International Journal of Forecasting*, **21**, 491-501.

Ye M., Zyren J., and Shore J. (2007), “Forecasting short-run crude oil price using high and low inventory variables”, *Energy Policy*, forthcoming.

Zamani M. (2004), “An econometrics forecasting model of short term oil spot price”, Paper presented at the 6th IAEE European Conference.

Zeng T., and Swanson N. R. (1998), “Predictive evaluation of econometric forecasting models in commodity future markets”, *Studies in Nonlinear Dynamics and Econometrics*, **2**, 159-177.

Table 1. Complete list of variables used in the empirical analysis

Variable	Sample	Frequency	Source	Acronym
WTI spot price	2/1/1986 31/12/2005	D, W, M, Q	EIA	*S*
WTI futures price contract 1	2/1/1986 31/12/2005	D, W, M, Q	EIA	*F1*
WTI futures price contract 2	2/1/1986 31/12/2005	D, W, M, Q	EIA	*F2*
WTI futures price contract 3	2/1/1986 31/12/2005	D, W, M, Q	EIA	*F3*
WTI futures price contract 4	2/1/1986 31/12/2005	D, W, M, Q	EIA	*F4*
OECD government oil stocks	1/1988-8/2005 Q1/1993-Q3/2005	M, Q	IEA	*GS*
OECD industrial oil stocks	1/1988-8/2005 Q1/1993-Q3/2005	M, Q	IEA	*IS*
Non OPEC countries oil production	1/1988-8/2005 Q1/1993-Q3/2005	M, Q	EIA	*NOP*
OECD oil consumption	1/1988-8/2005	M	EIA	*OC*
World oil production	1/1988-8/2005	M	EIA	*WP*
Commodity price index	1/1988-8/2005	M	BLS	*PPI*
OECD oil demand	Q1/1993-Q3/2005	Q	IEA	*OD*
Non-OECD countries oil demand	Q1/1993-Q3/2005	Q	IEA	*NOD*
OPEC oil production	Q1/1993-Q3/2005	Q	EIA	*OP*
OPEC sustainable oil production capacity	Q1/1993-Q3/2005	Q	PIW	*PC*
OPEC quota	Q1/1993-Q3/2005	Q	EIA	*OQ*
Short-term interest rate	Q1/1993-Q3/2005	Q	FRBG	*I*
OPEC overproduction	Q1/1993-Q3/2005	Q	Computed as: *OP-OQ*	*OV*
OPEC capacity utilization	Q1/1993-Q3/2005	Q	Computed as: (*OP*/*PC*)*100	*CU*
OPEC spare capacity	Q1/1993-Q3/2005	Q	Computed as: *PC-OP*	*SP*

Notes to Table 1. D = daily frequency; W = weekly frequency; M = monthly frequency; Q = quarterly frequency; Qi = ith quarter, i=1,2,3,4; EIA = Energy Information Administration; BLS = Bureau of Labor Statistics; IEA=International Energy Agency; PIW=Petroleum Intelligence Weekly; FRBG = Federal Reserve Board of Governors.

Table 2. Descriptive statistics

Variable	Obs.	Mean	Median	Max.	Min.	Std. Dev.	Annualised Std. Dev.	CV (%)
Daily frequency								
S	5023	24.05	20.37	69.91	10.25	10.53	166.50	43.78
F1	5023	24.03	20.37	69.81	10.42	10.54	166.65	43.86
F2	5023	23.87	20.24	69.98	10.54	10.57	167.13	44.28
F3	5023	23.72	20.09	70.23	10.58	10.56	166.97	44.52
F4	5023	23.56	20.05	70.41	10.71	10.52	166.34	44.65
Weekly frequency								
S	1175	24.07	20.37	68.47	10.87	10.50	62.00	43.62
F1	1175	23.44	20.34	66.09	10.93	9.75	62.06	41.60
F2	1175	23.94	20.23	68.60	10.91	10.50	62.30	43.86
F3	1175	23.80	20.14	69.03	11.05	10.50	62.30	44.20
F4	1175	23.65	20.09	69.28	11.23	10.52	62.06	44.48
Monthly frequency								
S	240	24.04	20.40	66.12	11.32	10.48	36.30	43.59
F1	240	24.02	20.44	66.13	11.35	10.49	36.34	43.67
F2	240	23.86	20.25	66.63	11.73	10.53	36.48	44.13
F3	240	23.71	20.02	67.06	11.86	10.53	36.48	44.41
F4	240	23.55	19.96	67.42	11.96	10.49	36.34	44.54
GS	212	61.21	62.66	75.20	47.50	0.07	22.57	0.11
IS	212	129.32	129.32	139.82	117.77	0.04	14.15	0.03
OC	212	45.44	46.26	51.74	37.43	3.24	11.22	7.13
WP	212	64.45	64.15	73.96	56.94	4.41	15.28	6.84
NOP	212	38.23	37.96	42.86	34.65	2.12	6.93	5.55
PPI	212	125.90	125.15	157.60	104.60	11.19	41.57	8.89
Quarterly frequency								
S	80	23.84	20.54	62.57	12.68	10.02	20.26	42.04
F1	80	23.82	20.53	62.59	12.67	9.92	20.26	41.64
F2	80	23.66	20.34	63.35	12.77	9.95	20.34	42.04
F3	80	23.50	20.29	63.91	12.86	9.94	20.34	42.28
F4	80	23.35	20.32	64.28	12.96	9.89	20.26	42.35
GS	51	21.58	21.31	24.91	19.43	0.01	2.78	0.06
IS	51	43.02	43.06	46.24	39.51	0.03	2.85	0.07
OD	51	46.94	47.20	50.40	41.50	2.16	4.32	4.60
NOD	51	28.11	27.90	34.20	23.80	2.88	5.76	10.25
OP	51	27.66	27.52	31.43	24.65	1.79	3.58	6.47
NOP	51	38.66	38.42	42.70	34.76	2.32	4.64	6.00
PC	51	29.63	30.34	31.95	25.96	1.67	3.46	5.84
OQ	51	24.52	24.52	28.00	21.07	1.54	3.08	6.28
OV	51	3.14	3.57	8.09	0.63	1.62	3.24	51.59
CU	51	93.37	94.07	98.35	81.64	3.69	7.38	3.95
SC	51	1.98	1.65	5.75	0.53	1.17	2.34	59.09
I	51	3.73	4.41	6.02	0.92	1.64	3.28	43.97

Notes to Table 2. For names of variables see Table 1; Obs = number of observations; Mean = sample mean; Median = sample median; Min. = minimum value in the sample; Max. = maximum value in the sample; Annualised Std. Dev. = std. dev. multiplied by the square root of the number of periods in the year (i.e. 250 days, 35 weeks, 12 months and 4 quarters); CV = Coefficient of variation, calculated as std. dev. divided by the mean.

Table 3. Augmented Dickey-Fuller tests for financial and macroeconomic variables

Variables	Deterministic components	Number of lags (p)	ADF statistic
Daily Financial Variables			
S	$\alpha_0 \neq 0, \alpha_1 \neq 0$	0	-1.991
ΔS	$\alpha_0 = 0, \alpha_1 = 0$	0	-73.304**
F1	$\alpha_0 = 0, \alpha_1 \neq 0$	0	-1.800
$\Delta F1$	$\alpha_0 = 0, \alpha_1 = 0$	0	-71.763**
F2	$\alpha_0 = 0, \alpha_1 \neq 0$	0	-1.154
$\Delta F2$	$\alpha_0 = 0, \alpha_1 = 0$	0	-70.800**
F3	$\alpha_0 = 0, \alpha_1 \neq 0$	0	-0.755
$\Delta F3$	$\alpha_0 = 0, \alpha_1 \neq 0$	0	-70.685**
F4	$\alpha_0 = 0, \alpha_1 = 0$	0	-0.485
$\Delta F4$	$\alpha_0 = 0, \alpha_1 \neq 0$	0	-72.486**
Weekly Financial Variables			
S	$\alpha_0 = 0, \alpha_1 = 0$	2	1.348
ΔS	$\alpha_0 = 0, \alpha_1 = 0$	1	-29.583**
F1	$\alpha_0 = 0, \alpha_1 \neq 0$	1	-1.566
$\Delta F1$	$\alpha_0 = 0, \alpha_1 = 0$	0	-39.341**
F2	$\alpha_0 = 0, \alpha_1 = 0$	2	1.467
$\Delta F2$	$\alpha_0 = 0, \alpha_1 = 0$	1	-28.505**
F3	$\alpha_0 = 0, \alpha_1 = 0$	2	1.693
$\Delta F3$	$\alpha_0 = 0, \alpha_1 = 0$	1	-28.558**
F4	$\alpha_0 = 0, \alpha_1 = 0$	2	1.912
$\Delta F4$	$\alpha_0 = 0, \alpha_1 \neq 0$	1	-28.867**
Monthly Financial Variables			
S	$\alpha_0 = 0, \alpha_1 = 0$	2	1.667
ΔS	$\alpha_0 = 0, \alpha_1 = 0$	0	-12.559**
F1	$\alpha_0 = 0, \alpha_1 = 0$	2	1.699
$\Delta F1$	$\alpha_0 = 0, \alpha_1 = 0$	0	-12.626**
F2	$\alpha_0 = 0, \alpha_1 = 0$	2	1.012
$\Delta F2$	$\alpha_0 = 0, \alpha_1 = 0$	0	-12.310**
F3	$\alpha_0 = 0, \alpha_1 = 0$	2	2.162
$\Delta F3$	$\alpha_0 = 0, \alpha_1 = 0$	1	-11.234**
F4	$\alpha_0 = 0, \alpha_1 = 0$	2	2.364
$\Delta F4$	$\alpha_0 = 0, \alpha_1 \neq 0$	1	-11.582**

Table 3. Continued

Variables	Deterministic components	Number of lags (p)	ADF statistic
Quarterly Financial Variables			
S	$\alpha_0 = 0, \alpha_1 = 0$	0	1.878
ΔS	$\alpha_0 = 0, \alpha_1 = 0$	0	-8.080**
F1	$\alpha_0 = 0, \alpha_1 = 0$	0	1.909
$\Delta F1$	$\alpha_0 = 0, \alpha_1 \neq 0$	0	-8.046**
F2	$\alpha_0 = 0, \alpha_1 = 0$	0	2.108
$\Delta F2$	$\alpha_0 = 0, \alpha_1 = 0$	0	-7.890**
F3	$\alpha_0 = 0, \alpha_1 = 0$	0	2.324
$\Delta F3$	$\alpha_0 = 0, \alpha_1 = 0$	0	-7.639**
F4	$\alpha_0 = 0, \alpha_1 = 0$	0	2.531
$\Delta F4$	$\alpha_0 = 0, \alpha_1 = 0$	0	-7.359**
Monthly Macroeconomic Variables			
GS	$\alpha_0 = 0, \alpha_1 = 0$	1	4.574
ΔGS	$\alpha_0 \neq 0, \alpha_1 = 0$	0	-11.571**
IS	$\alpha_0 \neq 0, \alpha_1 = 0$	12	-4.579**
OC	$\alpha_0 = 0, \alpha_1 = 0$	14	9.780
ΔOC	$\alpha_0 \neq 0, \alpha_1 = 0$	13	-5.766**
WP	$\alpha_0 \neq 0, \alpha_1 \neq 0$	0	-4.017**
NOP	$\alpha_0 = 0, \alpha_1 \neq 0$	1	-1.578
ΔNOP	$\alpha_0 = 0, \alpha_1 = 0$	0	-19.925**
PPI	$\alpha_0 = 0, \alpha_1 = 0$	1	3.269
ΔPPI	$\alpha_0 \neq 0, \alpha_1 = 0$	0	-12.534**
Quarterly Macroeconomic Variables			
CU	$\alpha_0 \neq 0, \alpha_1 = 0$	1	-2.289
ΔCU	$\alpha_0 = 0, \alpha_1 = 0$	0	-6.266**
GS	$\alpha_0 = 0, \alpha_1 = 0$	0	4.379
ΔGS	$\alpha_0 \neq 0, \alpha_1 = 0$	0	-5.528**
IS	$\alpha_0 \neq 0, \alpha_1 = 0$	4	-4.384**
NOD	$\alpha_0 \neq 0, \alpha_1 \neq 0$	0	-3.467
ΔNOD	$\alpha_0 = 0, \alpha_1 \neq 0$	2	-7.248**
NOP	$\alpha_0 = 0, \alpha_1 \neq 0$	2	4.557
ΔNOP	$\alpha_0 \neq 0, \alpha_1 = 0$	1	-8.983**
OD	$\alpha_0 \neq 0, \alpha_1 = 0$	3	-2.543
ΔOD	$\alpha_0 \neq 0, \alpha_1 = 0$	2	-18.662**
OP	$\alpha_0 \neq 0, \alpha_1 \neq 0$	1	-2.354

Table 3. Continued

Variables	Deterministic components	Number of lags (p)	ADF statistic
ΔOP	$\alpha_0 = 0, \alpha_1 = 0$	0	-5.561**
OQ	$\alpha_0 \neq 0, \alpha_1 = 0$	0	-2.377
ΔOQ	$\alpha_0 = 0, \alpha_1 = 0$	0	-6.676**
OV	$\alpha_0 = 0, \alpha_1 \neq 0$	0	-3.404
ΔOV	$\alpha_0 = 0, \alpha_1 = 0$	1	-6.888**
PC	$\alpha_0 \neq 0, \alpha_1 = 0$	0	-2.121
ΔPC	$\alpha_0 = 0, \alpha_1 = 0$	0	-5.856**
SC	$\alpha_0 = 0, \alpha_1 = 0$	0	-1.034
ΔSC	$\alpha_0 = 0, \alpha_1 = 0$	0	-6.054**

Notes to Table 3. The Augmented Dickey-Fuller test for a unit root is based on the following regression:

$$\Delta y_t = \alpha_0 + \alpha_1 t + \beta_1 y_{t-1} + \beta_2 \Delta y_{t-1} + ... + \beta_p \Delta y_{t-p} + \nu_t$$

Critical values are from MacKinnon (1991, 1996); *p* indicates the augmentation; the selection of *p* is based on the Schwartz Information Criterion; * (**) represents rejection of the null hypothesis of a unit root at the 0.05 (0.01) significance level.

Table 4. Cointegrating vectors and unrestricted cointegration rank tests - Daily data

			Variables			
N° of coint vec	**Trace**	**Max-eig**	**S**	**F1**	**Trend**	**Constant**
None	1284.684 [20.262]	1281.664 [15.892]	1.000	-1.004 (0.001)	$-1.36e^{-6}$ ($5.7e^{-6}$)	0.050 (0.022)
At most 1	3.020 [9.164]	3.020 [9.164]				
			S	**F2**	**Trend**	**Constant**
None	262.712 [15.495]	261.640 [14.265]	1.000	-1.033 (0.005)	-	-
At most 1	1.142 [3.841]	1.142 [3.841]				
			S	**F3**	**Trend**	**Constant**
None	136.235 [20.262]	135.177 [15.892]	1.000	-1.061 (0.012)	-	0.844 (0.263)
At most 1	1.058 [9.164]	1.058 [9.164]				
			S	**F4**	**Trend**	**Constant**
None	99.397 [20.262]	0.845 [9.164]	1.000	-1.089 (0.013)	-	1.313 (0.399)
At most 1	98.552 [15.892]	0.845 [9.164]				

Notes to Table 4. Trace and Max-eig are Johansen's trace and maximum-eigenvalue cointegration tests, respectively; in columns 4-7 we report the estimated (normalized) coefficients of the cointegrating equation $S_t = \beta_1 + \beta_2 t + \beta_3 Fm_t$, where *m* is the maturity (*m*=1,2,3,4); standard errors and 5% critical values for the Johansens's tests are reported in round and square brackets, respectively; $x*e^{-n}$ is equivalent to $x*10^{-n}$.

Table 5. Cointegrating vector coefficients and unrestricted cointegration rank tests - Weekly data

			Variables			
N° of coint vec	**Trace**	**Max-eig**	**S**	**F1**	**Trend**	**Constant**
None	53.379 [12.321]	53.105 [11.225]	1.000	-1.026 (0.005)	-	-
At most 1	0.273 [4.130]	0.273 [4.130]				
			S	**F2**	**Trend**	**Constant**
None	103.506 [12.321]	103.400 [11.225]	1.000	-1.007 (0.002)	-	-
At most 1	0.107 [4.130]	0.107 [4.130]				
			S	**F3**	**Trend**	**Constant**
None	62.082 [12.321]	62.026 [11.225]	1.000	-1.017 (0.004)	-	-
At most 1	0.055 [4.130]	0.055 [4.130]				
			S	**F4**	**Trend**	**Constant**
None	46.180 [12.321]	0.033 [4.130]	1.000	-1.026 (0.007)	-	-
At most 1	98.552 [15.892]	0.033 [4.130]				

Notes to Table 5. See Table 4.

Table 6. Cointegrating vector coefficients and unrestricted cointegration rank tests - Monthly data

			Variables			
N° of coint vec	**Trace**	**Max-eig**	**S**	**F1**	**Trend**	**Constant**
None	52.773 [12.321]	52.766 [11.225]	1.000	-1.001 (0.0003)	-	-
At most 1	0.007 [4.130]	0.007 [4.130]				
			S	**F2**	**Trend**	**Constant**
None	26.333 [12.321]	26.328 [11.225]	1.000	-1.011 (0.004)	-	-
At most 1	0.006 [4.130]	0.006 [4.130]				
			S	**F3**	**Trend**	**Constant**
None	24.492 [12.321]	24.472 [11.225]	1.000	-1.021 (0.007)	-	-
At most 1	0.020 [4.130]	0.020 [4.130]				
			S	**F4**	**Trend**	**Constant**
None	23.417 [12.321]	23.368 [11.225]	1.000	-1.029 (0.010)	-	-
At most 1	0.049 [4.130]	0.049 [4.130]				

Notes to Table 6. See Table 4.

Table 7. Cointegrating vector coefficients and unrestricted cointegration rank tests - Quarterly data

			Variables			
N° of coint vec	**Trace**	**Max-eig**	**S**	**F1**	**Trend**	**Constant**
None	27.137 [12.321]	26.919 [11.225]	1.000	-1.001 (0.0003)	-	-
At most 1	0.218 [4.130]	0.218 [4.130]				
			S	**F2**	**Trend**	**Constant**
None	26.067 [12.321]	25.933 [11.225]	1.000	-1.009 (0.003)	-	-
At most 1	0.134 [4.130]	0.134 [4.130]				
			S	**F3**	**Trend**	**Constant**
None	27.957 [12.321]	27.796 [11.225]	1.000	-1.016 (0.006)	-	-
At most 1	0.161 [4.130]	0.161 [4.130]				
			S	**F4**	**Trend**	**Constant**
None	28.600 [12.321]	28.409 [11.225]	1.000	-1.023 (0.007)	-	-
At most 1	0.191 [4.130]	0.191 [4.130]				

Notes to Table 7. See Table 4.

Table 8. Estimates, diagnostic tests and forecasting measures for model: $\Delta S_t = \alpha + \beta \Delta F_t + \gamma ECT_{t-1} + \varepsilon_t$ - Daily data

Model	Futures	*F1*	*F2*	*F3*	*F4*
Estimation sample 1986.01 2003.12					
α		0.003 (0.004)	0.006 (0.007)	0.004 (0.005)	0.003 (0.005)
β		0.934** (0.006)	1.038** (0.009)	1.131** (0.011)	1.158** (0.012)
γ		-0.546** (0.012)	-0.136** (0.007)	-0.064** (0.005)	-0.043** (0.004)
Adj-R^2		0.839	0.732	0.707	0.669
BG test		17.010**	166.637**	164.435**	157.757**
W test		496.864**	962.135**	707.529**	1083.623**
Forecasting sample 2004.01 2005.12					
Static Forecasts	MAPE	0.354	0.476	0.541	0.581
	MAE	0.191	0.260	0.297	0.320
	Theil	0.004	0.004	0.005	0.005
	RMSE	0.425	0.503	0.542	0.567
Dynamic Forecasts	MAPE	10.122	25.783	21.525	20.192
	MAE	5.921	14.872	12.770	11.979
	Theil	0.060	0.145	0.127	0.120
	RMSE	7.147	18.561	15.954	14.968
Rolling forecasting window: two years					
Rolling Forecasts (Static)	MAPE	0.547	0.748	0.801	0.841
	MAE	0.144	0.194	0.209	0.221
	Theil	0.005	0.006	0.007	0.007
	RMSE	0.291	0.348	0.371	0.398
Rolling Forecasts (Dynamic)	MAPE	14.535	45.983	37.017	29.486
	MAE	4.009	10.603	8.546	6.913
	Theil	0.090	0.249	0.199	0.177
	RMSE	4.755	12.410	10.001	8.099

Notes to Table 8. Standard errors are reported in parentheses; * (**) represents 0.05 (0.01) significance level; BG is the Breusch-Godfrey test for residual autocorrelation of order 4; W is the White test for heteroskedasticity; MAPE = Mean Absolute Percentage Error; MAE = Mean Absolute Error; Theil = Theil Inequality Coefficient; RMSE = Root Mean Squared Error.

Table 9. Estimates, diagnostic tests and forecasting measures for model:
$\Delta S_t = \alpha + \beta \Delta F_t + \gamma ECT_{t-1} + \varepsilon_t$ - Weekly data

Model / **Futures**		***F1***	***F2***	***F3***	***F4***
Estimation sample 1986.01 2003.12					
α		-0.002 (0.012)	0.004 (0.011)	0.005 (0.013)	0.004 (0.014)
β		0.981** (0.008)	1.037** (0.009)	1.087** (0.010)	1.133** (0.012)
γ		-0.175** (0.018)	-0.302** (0.022)	-0.139** (0.016)	-0.091** (0.013)
Adj-R^2		0.924	0.934	0.912	0.896
BG test		28.736**	121.291**	69.866**	51.282**
W test		161.168**	163.718**	193.177**	194.251**
Forecasting sample 2004.01 2005.12					
Static Forecasts	MAPE	1.203	0.526	0.766	0.890
	MAE	0.591	0.251	0.364	0.428
	Theil	0.008	0.006	0.007	0.007
	RMSE	0.837	0.603	0.693	0.741
Dynamic Forecasts	MAPE	18.076	5.249	6.263	7.142
	MAE	8.813	2.918	3.513	4.016
	Theil	0.103	0.037	0.048	0.054
	RMSE	9.476	3.793	4.965	5.696
Rolling forecasting window: two years					
Rolling Forecasts (Static)	MAPE	1.516	0.869	1.139	1.262
	MAE	0.411	0.209	0.290	0.325
	Theil	0.009	0.007	0.009	0.010
	RMSE	0.575	0.389	0.471	0.512
Rolling Forecasts (Dynamic)	MAPE	26.625	13.395	14.358	14.180
	MAE	6.909	2.575	3.115	3.212
	Theil	0.194	0.066	0.076	0.077
	RMSE	8.204	2.994	3.663	3.807

Notes to Table 9. See Table 8.

Table 10. Estimates, diagnostic tests and forecasting measures for model:
$\Delta S_t = \alpha + \beta \Delta F_t + \gamma ECT_{t-1} + \varepsilon_t$ - Monthly data

Model / **Futures**		***F1***	***F2***	***F3***	***F4***
Estimation sample 1986.01 2004.01					
α		0.002 (0.005)	0.002 (0.022)	0.003 (0.030)	0.001 (0.036)
β		1.003** (0.003)	1.069** (0.013)	1.138** (0.020)	1.220** (0.026)
γ		-0.788** (0.067)	-0.162** (0.070)	-0.141** (0.035)	-0.123** (0.032)
Adj-R^2		0.998	0.968	0.941	0.915
BG test		4.759	2.252	0.525	1.187
W test		10.253	41.498**	57.247**	48.465**
Forecasting sample 2004.02 2006.01					
Static Forecasts	MAPE	0.171	0.730	1.039	1.460
	MAE	0.081	0.363	0.524	0.741
	Theil	0.001	0.005	0.007	0.009
	RMSE	0.140	0.475	0.698	0.943
Dynamic Forecasts	MAPE	0.575	4.450	6.677	9.762
	MAE	0.322	2.396	3.694	5.414
	Theil	0.004	0.027	0.044	0.063
	RMSE	0.444	2.812	4.614	6.750
Rolling forecasting window: two years					
Rolling Forecasts (Static)	MAPE	0.231	1.152	1.559	1.852
	MAE	0.060	0.314	0.424	0.507
	Theil	0.002	0.007	0.010	0.012
	RMSE	0.085	0.400	0.538	0.648
Rolling Forecasts (Dynamic)	MAPE	0.763	2.565	4.218	5.305
	MAE	0.215	0.731	1.168	1.511
	Theil	0.005	0.015	0.025	0.031
	RMSE	0.261	0.874	1.421	1.853

Notes to Table 10. See Table 8.

Table 11. Estimates, diagnostic tests and forecasting measures for model:
$\Delta S_t = \alpha + \beta \Delta F_t + \gamma ECT_{t-1} + \varepsilon_t$ - Quarterly data

Model / Futures		***F1***	***F2***	***F3***	***F4***
Estimation sample 1986.01 2003.12					
α		0.002 (0.006)	0.009 (0.043)	0.009 (0.067)	0.005 (0.085)
β		1.006** (0.002)	1.058** (0.016)	1.120** (0.027)	1.186** (0.036)
γ		-0.656** (0.116)	-0.373** (0.092)	-0.329** (0.085)	-0.314** (0.080)
Adj-R^2		0.999	0.985	0.963	0.941
BG test		9.383**	0.6148*	7.673*	9.034*
W test		9.609	12.120	11.126	11.417
Forecasting sample 2004.01 to 2005.12					
Static Forecasts	MAPE	0.121	1.107	2.001	2.845
	MAE	0.060	0.551	1.001	1.436
	Theil	0.001	0.007	0.011	0.016
	RMSE	0.107	0.681	1.196	1.705
Dynamic Forecasts	MAPE	0.398	3.777	6.863	9.665
	MAE	0.220	2.052	3.733	5.274
	Theil	0.003	0.023	0.040	0.057
	RMSE	0.275	2.401	4.367	6.211
Rolling forecasting window: two years					
Rolling Forecasts (Static)	MAPE	0.163	1.179	1.810	2.259
	MAE	0.042	0.316	0.491	0.620
	Theil	0.001	0.007	0.011	0.014
	RMSE	0.052	0.401	0.623	0.793
Rolling Forecasts (Dynamic)	MAPE	0.288	2.578	4.079	5.119
	MAE	0.082	0.660	1.070	1.370
	Theil	0.002	0.014	0.023	0.029
	RMSE	0.099	0.777	1.268	1.642

Notes to Table 11. See Table 8.

Table 12. Estimates, diagnostic tests and forecasting measures for model:

$S_t = \alpha + \beta F_{t-i} + \varepsilon_t$ - Monthly data

Model \ Futures		F1	F2	F3	F4
Estimation sample 1986.01 2003.12					
α		1.042 (0.503)*	2.124 (0.795)**	3.031 (0.994)**	3.872 (1.166)**
β		0.955 (0.023)**	0.915 (0.037)**	0.883 (0.047)**	0.852 (0.055)**
Adj-R^2		0.888	0.741	0.624	0.524
BG test		8.753**	108.002**	140.71**	151.76**
W test		27.576**	8.593*	3.281	4.430
Forecasting sample 2004.01 to 2005.12					
Static Forecasts	MAPE	6.313	10.539	13.018	16.107
	MAE	3.185	5.209	6.373	7.809
	Theil	0.041	0.061	0.078	0.095
	RMSE	3.995	5.915	7.397	8.841
Rolling forecasting window: two years					
Rolling Forecasts (Static)	MAPE	4.399	6.643	8.078	9.541
	MAE	1.350	2.020	2.371	2.740
	Theil	0.026	0.039	0.047	0.054
	RMSE	1.653	2.409	2.876	3.201

Notes to Table 12. See Table 8.

Table 13. Estimates, diagnostic tests and forecasting measures for the structural and mixed models - Monthly data

Variables		**Model (11)**	**Model (12)**
Estimation sample 1993m01 2003m12			
Constant		-0.392** (0.122)	35.885** (6.224)
F_t		1.005** (0.002)	-
OC_t		0.005* (0.002)	-
RIS_{t-1}		$-4.02e^{-6}$* ($1.89e^{-6}$)	$-1.25e^{-5}$** ($3.39e^{-6}$)
RIS_{t-12}		-	$-2.42e^{-5}$** ($3.32e^{-6}$)
WP_{t-1}		-	-0.752** (0.150)
PPI_t		-	0.254** (0.072)
$D01_t$		-	-4.721** (1.299)
$S99_t$		-0.040* (0.020)	9.093** (0.809)
Adj-R^2		0.999	0.722
BG test		7.731*	95.078**
W test		1.312	43.023**
Forecasting sample 2004m01 2005m08			
Static Forecasts	MAPE	0.379	37.208
	MAE	0.186	19.304
	Theil	0.002	0.271
	RMSE	0.240	21.487
Rolling forecasting window: two years			
Rolling Forecasts (Static)	MAPE	0.289	18.753
	MAE	0.068	4.384
	Theil	0.002	0.108
	RMSE	0.087	5.065

Notes to Table 13. See Table 8; models (11) and (12) are described in Section 4.2.

Table 14. Estimates, diagnostic tests and forecasting measures for structural and mixed models - Quarterly data

Variables		Model (13)	Model (14)	Model (15)	Model (16)
Estimation sample 1993Q1 2003Q4					
Constant		8.278** (1.040)	13.397* (5.052)	8.689 (8.333)	9.942 (7.789)
$\ln(F_t)$		0.991** (0.016)	-	-	-
$\Delta(\ln(F_t))$		0.088** (0.029)	-	-	-
S(t-1)		-	-	-	0.243* (0.099)
$\ln(S_{t-1})$		-	0.664** (0.097)	-	-
$\ln(NOD_t+OD_t)$		-	0.964* (0.419)	-	-
$\Delta(\ln(NOD_t+OD_t))$		0.3367* (0.167)	-	-	-
NOD_t+OD_t		-	-	0.384** (0.117)	0.294* (0.115)
$\ln(IS_t)$		-1.049** (0.128)	-2.101** (0.645)	-	-
$\Delta(\ln(IS_t))$		0.686** (0.160)	1.669* (0.694)	-	-
RIS_{t-1}		-	-	-0.032** (0.003)	-0.025** (0.004)
RIS_{t-4}		-	-	-0.017** (0.003)	-0.010* (0.004)
OP_{t-1}		-	-	-0.675** (0.235)	-0.646** (0.219)
$D01_t$		-	-	-5.378** (1.169)	-4.991** (1.100)
$S99_t$		-	-	8.303** (0.808)	6.691** (0.998)
Adj-R^2		0.995	0.836	0.933	0.942
BG test		0.641	2.940	0.919	1.532
W test		13.689	7.315	7.463	8.330
Forecasting sample 2004Q1 2005Q4					
Static Forecasts	MAPE	2.708	16.187	27.646	30.649
	MAE	1.304	8.247	14.172	15.937
	Theil	0.015	0.115	0.209	0.218
	RMSE	1.471	9.967	16.876	17.926
Dynamic Forecasts	MAPE	-	27.646	-	37.653
	MAE	-	14.172	-	19.715
	Theil	-	0.209	-	0.286
	RMSE	-	16.876	-	22.446
Rolling forecasting window: two years					
Rolling Forecasts (Static)	MAPE	1.718	14.4273	24.311	16.193
	MAE	0.486	3.598	6.426	4.081
	Theil	0.011	0.080	0.146	0.092
	RMSE	0.596	4.180	7.481	4.692
Rolling Forecasts (Dynamic)	MAPE	-	23.348	-	23.575
	MAE	-	5.925	-	6.242
	Theil	-	0.135	-	0.142
	RMSE	-	6.872	-	7.275

Notes to Table 14. See Table 8; models (13)-(16) are described in Section 4.2.

Table 15. Short-run and long-run marginal effects

Model (11): $S_t = \alpha + \beta F_t + \gamma OC_t + \delta RIS_{t-1} + S99 + \varepsilon_t$		
Explanatory Variables	**Short-run Marginal Effects**	**Long-run Marginal Effects**
OC_t	0.005 * (0.002)	-
RIS_{t-1}	-4.02e^{-6} * (189e^{-6})	-
Model (12): $S_t = \alpha + \beta RIS_{t-1} + \gamma WP_{t-1} + \delta RIS_{t-12} + \phi PPI_t + S99 + D01 + \varepsilon_t$		
WP_{t-1}	-0.752 ** (0.150)	-
PPI_t	0.254 ** (0.072)	-
RIS_{t-1}	-1.25e^{-5} ** (3.39e^{-6})	-3.67e^{-5} ** (4.43e^{-6})
RIS_{t-12}	-2.42e^{-5} ** (3.32e^{-6})	

Notes to Table 15. Standard errors are reported in parentheses; * (**) represents 0.05 (0.01) significance level; long-run marginal effects are calculated as $LRME = \frac{\beta + \delta}{1-\rho}$, where ρ is the coefficient of the lagged dependent variable, if present.

Table 16. Short-run and long-run elasticities

Model (13): $\ln(S_t) = \alpha + \beta \ln F_t + \gamma \Delta \ln F_t + \delta \Delta \ln(NOD_t + OD_t) + \phi \ln IS_t + \vartheta \Delta \ln IS_t + \varepsilon_t$		
Explanatory Variables	**Short-run Elasticity**	**Long-run Elasticity**
$\ln(IS_t)$	-1.049 ** (0.128)	-
Model (14): $\ln(S_t) = \alpha + \delta \ln(NOD_t + OD_t) + \phi \ln IS_t + \vartheta \Delta \ln IS_t + \varphi \ln S_{t-1} + \varepsilon_t$		
$\ln(NOD_t + OD_t)$	0.964 * (0.419)	2.871** (0.635)
$\ln(IS_t)$	-2.101 ** (0.645)	-6.256 ** (0.391)

Notes to Table 16. Standard errors are reported in parentheses; long-run elasticities are calculated as $LRE = \frac{\beta}{(1-\rho)}$, where ρ is the coefficient of the lagged dependent variable, if present.

Table 17. Short-run and long-run marginal effects

Model (15): $S_t = \alpha + \beta RIS_{t-1} + \delta RIS_{t-4} + \phi OP_{t-1} + \phi(NOD_t + OD_t) + S99 + D01 + \varepsilon_t$		
Explanatory Variables	**Short-run Marginal Effects**	**Long-run Marginal Effects**
$(NOD_t + OD_t)$	0.384 ** (0.117)	-
OP_{t-1}	-0.675 ** (0.235)	-
RIS_{t-1}	-0.032 ** (0.003)	-0.049 ** (0.0034)
RIS_{t-4}	-0.017 ** (0.003)	
Model (16): $S_t = \alpha + \beta RIS_{t-1} + \delta RIS_{t-4} + \phi OP_{t-1} + \phi(NOD_t + OD_t) + \varphi S_{t-1} + S99 + D01 + \varepsilon_t$		
$(NOD_t + OD_t)$	0.294 * (0.115)	-
OP_{t-1}	-0.646 ** (0.219)	-
RIS_{t-1}	-0.025 ** (0.004)	-0.046 ** (0.0028)
RIS_{t-4}	-0.010 ** (0.004)	

Notes to Table17. See Table 15.

Table 18. Estimates, diagnostic tests and forecasting measures for model: $S_t = \alpha + S_{t-1} + \varepsilon_t$

Model / Frequencies		Daily	Weekly	Monthly	Quarterly
Estimation sample 1986.01 2003.12					
α		0.005 (0.009)	0.011 (0.044)	0.067 (0.124)	0.219 (0.347)
Adj-R^2		0.991	0.930	0.886	0.686
BG test		13.650**	72.576**	15.569**	2.980
W test		114.571**	18.310**	35.482**	12.329**
Forecasting sample 2004.01 2005.12					
Static Forecasts	MAPE	1.710	2.897	5.546	8.345
	MAE	0.935	1.438	2.815	4.048
	Theil	0.011	0.018	0.035	0.051
	RMSE	1.238	1.853	3.495	4.863
Dynamic Forecasts	MAPE	11.697	31.280	31.298	32.378
	MAE	6.831	16.702	16.682	16.863
	Theil	0.078	0.233	0.231	0.234
	RMSE	8.311	19.277	19.119	18.998
Rolling forecasting window: two years					
Rolling Forecasts (Static)	MAPE	1.918	3.689	6.488	9.976
	MAE	0.509	0.940	1.686	2.559
	Theil	0.013	0.033	0.039	0.061
	RMSE	0.700	1.583	2.107	3.197
Rolling Forecasts (Dynamic)	MAPE	26.459	25.212	25.239	25.890
	MAE	6.937	6.555	6.601	6.787
	Theil	0.166	0.157	0.157	0.159
	RMSE	8.268	7.813	7.814	7.913

Notes to Table 18. See Table 8.

Table 19. Estimates, diagnostic tests and forecasting measures for model: $S_t = \alpha + \rho S_{t-1} + \varepsilon_t$

Model / Frequencies		Daily	Weekly	Monthly	Quarterly
Estimation sample 1986.01 2003.12					
α		0.061 (0.031)	0.113 (0.108)	1.115* (0.502)	3.264* (1.428)
ρ		0.997** (0.001)	0.997** (0.004)	0.950** (0.023)	0.855** (0.066)
Adj-R^2		0.991	0.980	0.888	0.702
BG test		12.534**	54.286**	16.837**	1.635
W test		155.566**	5.800	27.785**	8.258*
Forecasting sample 2004.01 2005.12					
Static Forecasts	MAPE	1.726	2.911	6.424	14.343
	MAE	0.944	1.446	3.241	6.994
	Theil	0.011	0.019	0.042	0.084
	RMSE	1.243	1.858	4.094	7.709
Dynamic Forecasts	MAPE	24.830	30.872	40.983	42.092
	MAE	14.727	16.492	21.722	21.814
	Theil	0.182	0.230	0.316	0.318
	RMSE	17.758	19.055	24.608	24.321
Rolling forecasting window: two years					
Rolling Forecasts (Static)	MAPE	1.936	4.065	6.785	12.697
	MAE	0.514	1.056	1.783	3.232
	Theil	0.013	0.035	0.041	0.072
	RMSE	0.702	1.668	2.197	3.724
Rolling Forecasts (Dynamic)	MAPE	25.169	26.094	22.733	22.953
	MAE	7.307	7.585	6.419	6.474
	Theil	0.165	0.170	0.144	0.413
	RMSE	8.586	8.770	7.527	7.475

Notes to Table 19. See Table 8.

Figure 1. WTI spot price for the period January 1986 - December 2005 (monthly data)

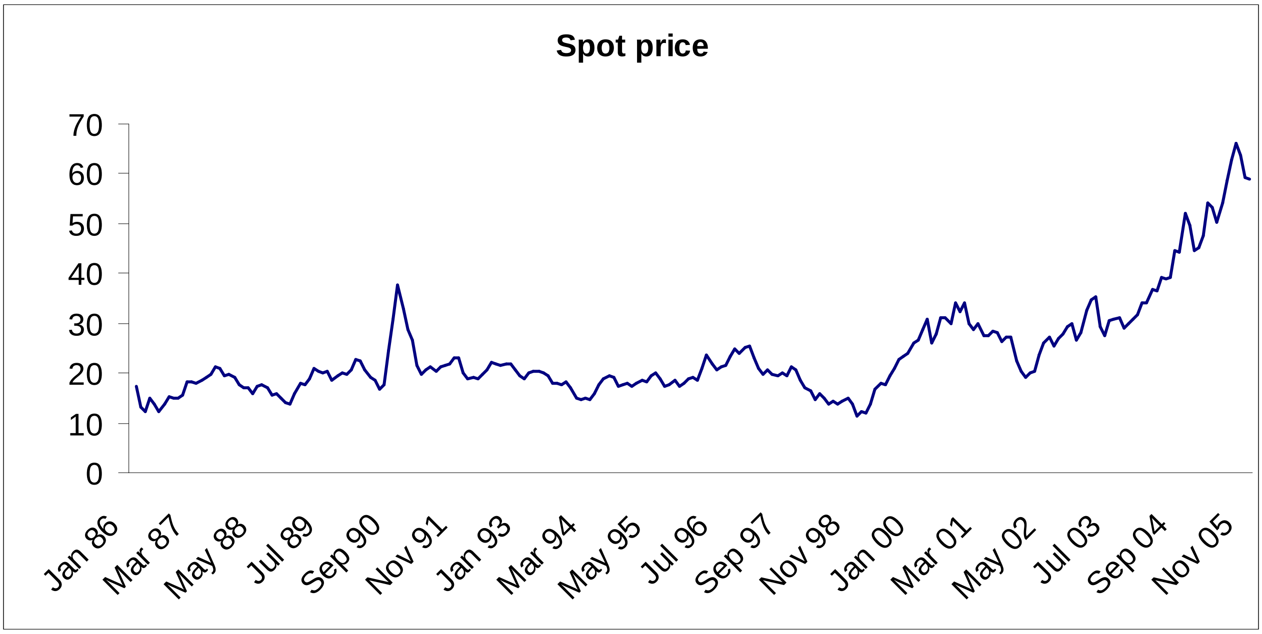


Figure 2. Graphical comparison of models (1) and (10) - Monthly data

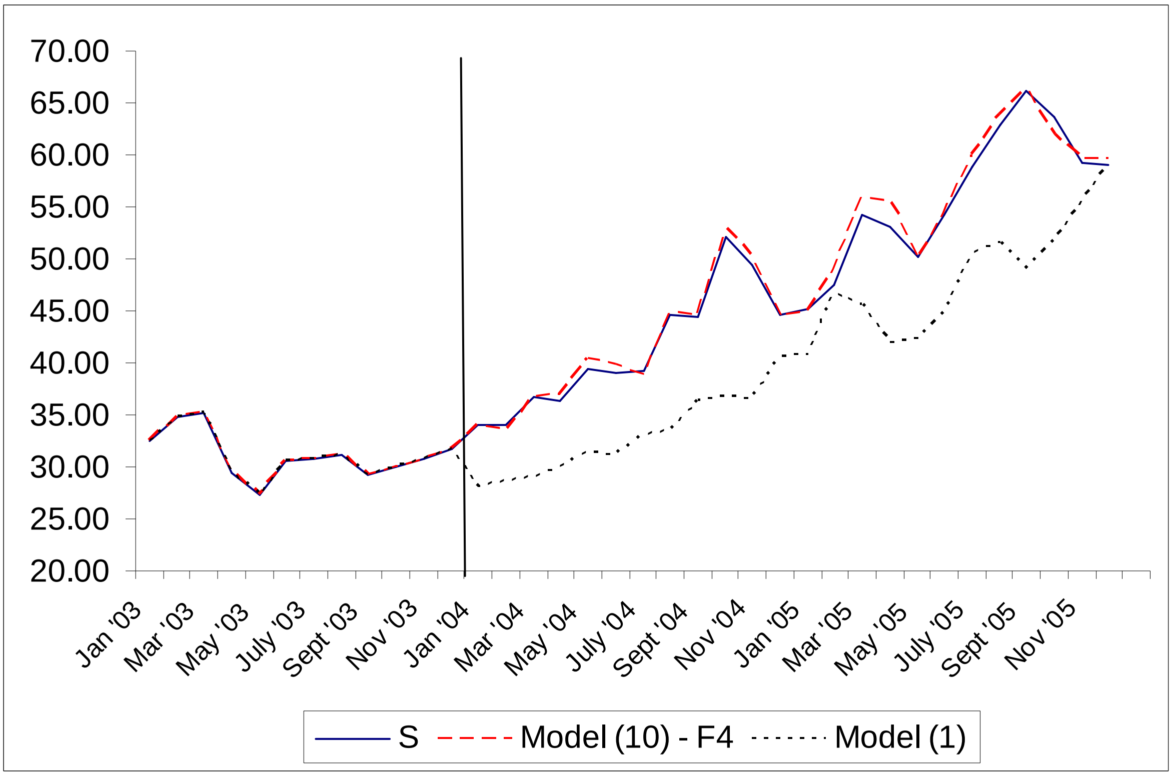
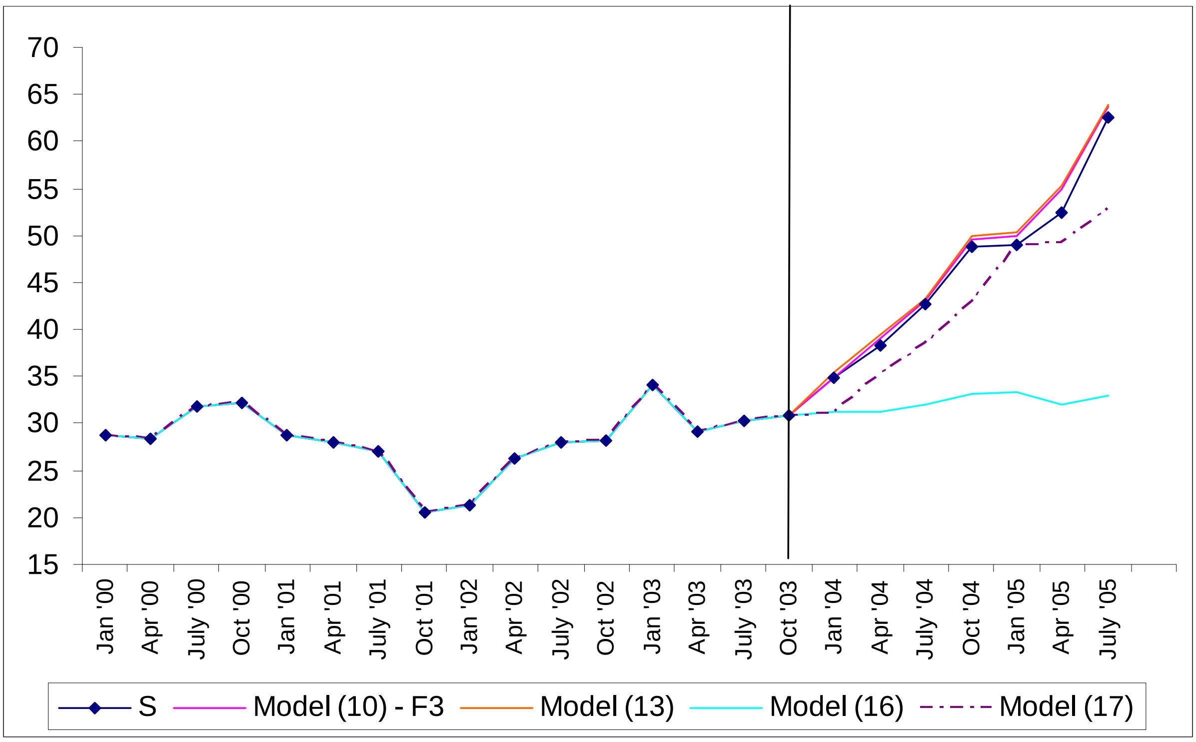


Figure 3. Graphical comparison of models (10), (13), (16) and (17) - Quarterly data



NOTE DI LAVORO DELLA FONDAZIONE ENI ENRICO MATTEI

Fondazione Eni Enrico Mattei Working Paper Series

Our Note di Lavoro are available on the Internet at the following addresses:
http://www.feem.it/Feem/Pub/Publications/WPapers/default.htm
http://www.ssrn.com/link/feem.html
http://www.repec.org
http://agecon.lib.umn.edu
http://www.bepress.com/feem/

NOTE DI LAVORO PUBLISHED IN 2007

NRM	1.2007	*Rinaldo Brau, Alessandro Lanza, and Francesco Pigliaru*: How Fast are Small Tourist Countries Growing? The 1980-2003 Evidence
PRCG	2.2007	*C.V. Fiorio, M. Florio, S. Salini and P. Ferrari:* Consumers' Attitudes on Services of General Interest in the EU: Accessibility, Price and Quality 2000-2004
PRCG	3.2007	*Cesare Dosi and Michele Moretto*: Concession Bidding Rules and Investment Time Flexibility
IEM	4.2007	*Chiara Longo, Matteo Manera, Anil Markandya and Elisa Scarpa*: Evaluating the Empirical Performance of Alternative Econometric Models for Oil Price Forecasting

2007 SERIES

CCMP	*Climate Change Modelling and Policy* (Editor: Marzio Galeotti)
SIEV	*Sustainability Indicators and Environmental Valuation* (Editor: Anil Markandya)
NRM	*Natural Resources Management* (Editor: Carlo Giupponi)
KTHC	*Knowledge, Technology, Human Capital* (Editor: Gianmarco Ottaviano)
IEM	*International Energy Markets* (Editor: Matteo Manera)
CSRM	*Corporate Social Responsibility and Sustainable Management* (Editor: Giulio Sapelli)
PRCG	*Privatisation Regulation Corporate Governance* (Editor: Bernardo Bortolotti)
ETA	*Economic Theory and Applications* (Editor: Carlo Carraro)
CTN	*Coalition Theory Network*