Energy Economics 32 (2010) 409–417

Contents lists available at ScienceDirect

Energy Economics

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

Macroeconomic factors and oil futures prices: A data-rich model ☆

Paolo Zagaglia

Modelling Division, Sveriges Riksbank, Sweden

ARTICLE INFO

Article history:
Received 26 March 2009
Received in revised form 8 November 2009
Accepted 8 November 2009
Available online 14 November 2009

JEL classification:
C53
E52

Keywords:
Crude oil
Futures markets
Factor models

ABSTRACT

I study the dynamics of oil futures prices in the NYMEX using a large panel dataset that includes global macroeconomic indicators, financial market indices, quantities and prices of energy products. I extract common factors from the panel data series and estimate a Factor-Augmented Vector Autoregression for the maturity structure of oil futures prices. I find that latent factors generate information that, once combined with that of the yields, improves the forecasting performance for oil prices. Furthermore, I show that a factor correlated to purely financial developments contributes to the model performance, in addition to factors related to energy quantities and prices.

1. Introduction

During the past year, oil prices have made the headlines of the financial press almost every day. Since the beginning 2008, the spot price of crude oil traded in the New York Mercantile Exchange (NYMEX) has almost doubled at peak. This has raised serious concerns among market participants and policymakers worldwide. Comments released to the press have often denoted a deep disagreement on the causes of the price spikes and, in general, on the mechanics of oil market.

Bernanke (2008) has represented the central bankers' view in a timely manner, stating that[1]

> "…the price of oil has risen significantly in terms of all major currencies, suggesting that factors other than the dollar, notably shifts in the underlying global demand for and supply of oil, have been the principal drivers of the increase in prices. (…) Another concern that has been raised is that financial speculation has added markedly to upward pressures on oil prices. (…) However, if financial speculation were pushing oil prices above the levels consistent with the fundamentals of supply and demand, we would expect inventories of crude oil and petroleum products to increase as supply rose and demand fell. But in fact, available data on oil inventories show notable declines over the past year."

Since oil commodities are traded through futures and derivatives contracts, market views shape the pricing of oil commodities. In this sense, the financial press has pushed the hypothesis that purely 'financial' considerations, unrelated to 'real' market developments, have been behind the recent spikes (see Chung, 2008 and Mackintosh, 2008).

The distinction between financial and real determinants of oil prices in the long run is also present in the academic literature. A large number of papers suggest that oil prices are mainly driven by demand and supply considerations. For instance, Kilian (2008b) suggests that a proper measurement of the business cycle effects of energy prices requires disentangling the role of demand supply shocks in energy markets. Kilian (2008a) decomposes the real price of crude oil into supply shocks, shocks to the global demand for industrial commodities, and demand shocks that are idiosyncratic to the oil market. The role of energy quantity factors is stressed also in Alquist and Kilian (2008), who show that spread between oil futures prices of different maturities are related to uncertainty about supply shortfalls.

The literature on the financial determinants of oil prices has produced various contributions on the role of market uncertainty and volatility for oil pricing. Askari and Krichene (2008) model the jump intensity of daily crude oil prices between 2002 and 2006. They find that measures of market expectations extracted from call and put option prices have incorporated no change in underlying fundamentals in the

☆ This work started when I was a visiting scholar in the Research Unit of the Bank of Finland, whose hospitality is greatly acknowledged. I am deeply grateful to the editor (Richard Tol) and to two anonymous referees for very thorough and helpful comments. I also acknowledge the assistance of Vincent Brousseau and Tarja Yrjölä in the construction of the dataset. Efrem Castelnuovo, Juha Kilponen, Massimiliano Marzo and Alistair Milne contributed with relevant suggestions and constructive criticism. The views expressed here are my own and should not be interpreted as reflecting the views of the Executive Board of Sveriges Riksbank. This research was carried out by the author for purely academic purposes and without financial support from any private institution.

E-mail address: paolo.zagaglia@gmail.com.
URL: http://pz.pcriot.com.

[1] A similar argument is discussed by Trichet (2008).

0140-9883/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.eneco.2009.11.003

short term. Chong and Miffre (2006) document the presence of a significant pattern of risk premia earned by investors on a number of commodities futures since 1979, including crude oil. Gorton et al. (2007) show that, although commercial positions on oil futures are correlate with inventory signals, they do not determine risk premia.

The presence of two opposing views on price formation in the oil market over the long run implies that a number of key questions are not dealt with in the literature. The issue of causality between spot and futures prices across the maturity structure is largely unsettled. Suppose that oil futures contain information about spot prices. Omitting futures prices would bias the results in favour of a strong role for demand–supply factors to drive the spot price. Moreover, the role of macroeconomic factors for the dynamics of oil prices is typically studied in isolation from the conditions prevailing in financial markets.[2]

In this paper I study whether the interplay between real and financial factors can play a systematic role for explaining oil prices changes over a long time period. I exploit the information from a large panel to investigate the sources of changes in the term structure of futures prices for WTI crude oil. Like Bernanke et al. (2008), I extract common factors form the large panel dataset, and I model the joint dynamics of the factors and the oil prices in a 'Factor-Augmented' Vector Autoregression (FAVAR). The factors mimic the drivers of oil prices that are 'latent,' in the sense that they are not directly observed by the econometrician from the information set. In standard Vector Autoregressive (VAR) models, the econometrician is required to choose what observable variables best represent theoretical concepts, such as supply and demand. The supply of oil can be measured with data on oil production. However, these data series are affected by measurement errors of different types, for instance arising from aggregation. As argued by Bernanke et al. (2008), the use of sparse information in the form of factors extracted from a large dataset mitigates this problem.

This modelling strategy has already been applied by Ludvigson and Ng (forthcoming) and Mönch (2008) for the construction of pricing models for the yield curve of government bonds, and it presents several advantages. The model can capture the interdependence between oil price changes and the factors of different nature. The FAVAR also allows to model jointly the relevant maturities of oil futures prices in a flexible way. It should be stressed that the literature features a long list of contributions on the role of unobservable factors for oil price dynamics.[3] These contributions differ from the present paper in two dimensions. The factors are typically meant to drive the time-varying volatility of observed at a daily frequency. Instead here I use monthly data, and I abstract from the role of high-frequency price movements.

The panel dataset from which I extract common components include over 200 data series with detailed information on energy demand and supply, energy prices, macroeconomic and financial variables. I show that a latent factor correlated with the open interest on oil futures prices contributes significantly to the joint model of the oil price returns. This appears to corroborate the conjecture of Trichet (2008) on the financial determinants of oil prices. The other factors are strongly correlated with data on energy quantity and prices, as typically suggested by the macroeconomics literature. I find that augmenting the information from the term structure of oil futures prices with latent factors improves the forecasting performance of the model.

This paper is organized as follows. In Section 2, I outline the structure of the FAVAR model. Section 3 presents the dataset. Section 4 describes the results. Section 5 concludes.

2. The factor-augmented VAR model

The model presented here is based on the assumption that the futures price for one maturity is driven both by the prices of the other maturities, and by macroeconomic shocks. The macroeconomic determinants are proxied by unobservable factors that summarize the common information in a large number of time series. The joint dynamics of the observable an unobservable variables in modelled in the FAVAR model of Bernanke et al. (2008).

The general form of the FAVAR can be written as

$$\begin{bmatrix} F_t \\ Y_t \end{bmatrix} = \mu + \Phi(L)\begin{bmatrix} F_{t-1} \\ Y_{t-1} \end{bmatrix} + \nu_t \qquad (1)$$

where $\Phi(L)$ is a matrix of lag polynomials, and ν_t is a vector of normally-distributed shocks. Y_t is a vector $m\times 1$ of observed variables. The unobservable factors are collected in the $k\times 1$ vector F_t. Eq. (1) states that the dynamics of the factors is affected by its own lags, by the vector of observables, and by the shocks. The model 1 has a variance–covariance matrix Σ.

Eq. (1) cannot be estimated without knowledge of F_t. For that purpose, a large number p of series can be used to extract 'common' factors. The 'information series' are collected in the vector X_t with dimension $p\times 1$. The dynamic factor model of Stock and Watson (2002) can then be used to obtain F_t. This framework assumes that the information time-series X_t are related to the factors F_t and the observed variables Y_t through the observation equation

$$X_t = \Lambda^f F_t + \Lambda^y Y_t + \epsilon_t. \qquad (2)$$

where Λ^F is a $p\times k$ matrix of factor loadings. The measurement equation formalizes the idea that both the oil futures returns and the factors drive the dynamics of the panel dataset. In other words, the factors can be measured with noise from the panel dataset.

Bernanke et al. (2008) propose two methods for estimating the model 1–2. The first one is the 'diffusion index' approach of Stock and Watson (2002), which consists itself of two steps. In the first step, Eq. (2) is used to estimate the unobservable factors F_t through asymptotic principal components. Then, the estimated factor $\hat{F}_t$ is fit to the FAVAR model 1. The second estimation method follows a single-step Bayesian likelihood approach. Bernanke et al. (2008) discuss a Gibbs sampler that approximates the marginal posterior densities of both the factors and the parameters. Since it is not clear a priori which estimation method delivers the results that are most desirable, Bernanke et al. (2008) estimate the model using both approaches, and find that they yield similar outcomes.

In this paper, I apply the two-step estimation procedure. Thus, I extract unobservable factors by using the asymptotic principal component method. Extracting the common factors from the panel dataset consists in recovering the space of X_t spanned by F_t. Denote by V the eigenvectors corresponding to the k largest eigenvalues of the variance–covariance matrix XX'/k. The estimated factors are $\hat{F} = \sqrt{T}V$, where T is the time-series length, and the factor loadings are $\hat{\Lambda}^f = \sqrt{T}X'V$.

3. The dataset

I use monthly data from January 1992 until March 2008 for a total of 193 observations for each series. The vector Y_t consists of returns on the spot price for WTI crude oil traded in the New York Mercantile Exchange (NYMEX), and on futures prices with maturities of 1, 6 and

[2] For instance, Trichet (2008) suggests that the combination of real and financial factors have played a joint role in the recent episodes of oil price movements.

[3] A non-exhausting list consists of Casassus and Collin-Dufresne (2005), Carmona and Ludkovski (2005), Gibson and Schwartz (1990), Pindyck (2004), Postali and Picchetti (2006), Radchenko (2007), Schwartz (1997), and Schwartz and Smith (2000).

12 months.[4] The panel dataset used for the extraction of the factors comprises 239 series that are meant to capture the macroeconomic, financial and geographic forces that move oil prices.[5] The complete list of the series, the sources and the choice of filtering are reported in Appendix A.[6]

Oil prices in the NYMEX respond to global supply and demand factors. Hence, the dataset includes series that are publicly available on petroleum stocks and consumption in the major OECD countries. Since this information is not available for the major emerging economies (Russia, India and China), the industrial production index is used as a proxy for consumption pressures of oil.[7] Instead, crude oil production data account for the entire range of oil producers worldwide. Almost half the series on energy quantities described in Appendix A refer to the U.S. In particular, there is detailed information on the use of all the available energy sources across sectors of the economy, including the energy products derived from petroleum and natural gas. There are indicators on rigging and drilling activities in the U.S., as well as on alternative sources of energy such as ethanol. I use around 50 price indices that are related with U.S. imports and refining.[8] I also control for the role of shipment prices to the Mediterranean Sea and from the Gulf to Northern Europe.

The macroeconomic part of the dataset consists of on measures of monetary aggregates, prices indices, indicators of confidence and bilateral exchange rates for the U.S. economy. Since the stability of the Dollar exchange rate is often pointed to as a key factor for oil prices, I use the global hazard indicator of Brousseau and Scacciavillani (1999). This is a measure of risk in foreign exchange markets calculated from implied volatilities of currency options. Following the lead from the previous quotation of Bernanke (2008), the dataset includes information on the open interest and volumes of traded futures contracts. I also include the prices of different types of crude oil spreads. These spreads are over-the-counter derivatives on the difference between the prices of two oil-related products (i.e., between forward prices of the WTI and the Brent), or between the forward prices at two different maturities (i.e., for the Dubai or the Brent oil). As Marzo et al. (2008) show, these contracts have predictive content for oil futures prices, and can provide information on the speculative motive for trading oil. Finally, the dataset comprises the prices of stocks of major oil companies, and a number of bond and stock indices.

4. Results

4.1. Factor estimates

In the first step of the estimation, I extract common factors from the panel dataset using the static principal components of Stock and Watson (2002) described earlier. Table 1 reports the cumulative fraction of total variance in the panel dataset explained by the first four factors. The cumulative variance of factors 1 to 4 is given by the sum of the largest four eigenvalues of the matrix X_tX_t. The total variance is the sum of the variances of the individual series in the panel X_t, and is measured as the sum of the eigenvalues of X_tX_t. Table 1 shows that the first four factors account for around 50% of the variability contained in the dataset. They exhibit a low degree of persistence. The estimated autocorrelation coefficients are however quite different across factors.

Table 1
Variance explained by the factors.

Factor	R^2	First autocorr. coeff.
1	0.1823	0.3241
2	0.3065	−0.4735
3	0.3961	0.4727
4	0.4758	0.1521

I include the first four factors in the FAVAR model for two reasons. Testing for the optimal number of factors using the statistical framework of Bai and Ng (2008) points in favour of this number of factors. On more general grounds, the VAR model presents a tradeoff between parsimony and fit. Various experiments suggest that the results are qualitatively unaltered by the inclusion of additional factors.

The factors extracted from the panel have no structural interpretation unless identifying assumptions are imposed. In order to provide some understanding on the information they convey, I regress the factors on the variables of the panel. Table 2 reports the variance explained by the five series that are most correlated with the factors. The first factor is strongly correlated with a price index of crude oil imports for the U.S. This can be interpreted as a cost indicator of the price pressure on oil futures. The second and third factors are related to stock volumes of oil-related products. This has to do with the intermediate demand for crude oil. Finally, the fourth factor is correlated with a purely financial variable that is disconnected from real developments in oil markets. This provides support to the claim that financial factors contribute to the determination of oil prices.

Table 2
Share of explained variance of highly-correlated series.

	R^2
Factor 1 (18% to total variance)	
Landed cost of crude oil imports from all non-OPEC countries	0.85
Average F.O.B. cost of crude oil imports from all non-OPEC countries	0.83
Refiner price of no. 2 diesel fuel for resale	0.80
Landed cost of crude oil imports from Mexico	0.80
Refiner price of no. 2 diesel fuel to end users	0.79
Factor 2 (12% to total variance)	
Motor gasoline stocks (including blending components and gasohol)	0.84
Volume of crude oil futures 3-month contracts	0.46
Other petroleum products stocks	0.38
Finished motor gasoline imports	0.31
Petroleum stocks, other OECD	0.27
Factor 3 (9% to total variance)	
Propane/propylene product supplied	0.31
Refiner price of kerosene to end users	0.18
Liquefied petroleum gases product supplied	0.17
Refiner price of propane (consumer grade) for resale	0.12
Refiner price of kerosene for resale	0.11
Factor 4 (8% to total variance)	
Open interest on 12-month crude oil futures	0.19
Open interest on 6-month crude oil futures	0.18
Jet fuel refinery net production	0.17
Total petroleum refinery and blender net inputs	0.17
Volume 6-month crude oil futures	0.15

Legend: This table reports R^2 of univariate regressions of factors on macro variables. I report the five variables with the highest correlation with the factors.

[4] Returns are computed as the first difference of the logarithm of the oil prices.

[5] One referee notes that the number of series included in the panel is very large. Boivin and Ng (2005) show that adding series of the same type of data can have negative effects on the predictive performance of the model. They find that using fewer pre-screened series can generate predictive gains. However, as noted by Bernanke et al. (2008), this step of pre-screening is largely *ad hoc*. Furthermore, the cost of using a large panel is typically marginal.

[6] As pointed out by one referee, the dataset does not include the recent developments in oil and gold prices since April 2009. However the paper focuses on the systematic relation between financial and macroeconomic factors over a long time span, rather than on the interpretation of single episodes.

[7] This is based on the assumption that industrial production drives the demand for oil. In this sense, the cyclical properties of oil demand can be approximated by the information on industrial production.

[8] It should be stressed that the correlation between the price indices and the oil price returns does not matter in the construction of the FAVAR model. Rather, these price indices can matter because they can explain part of the overall variability within the panel. As indicated earlier, in this paper I maintain an agnostic stance on the information provided by the panel. I include an information set as large as possible and I let the principal components pick up the sources of variability that matter the most within the panel.

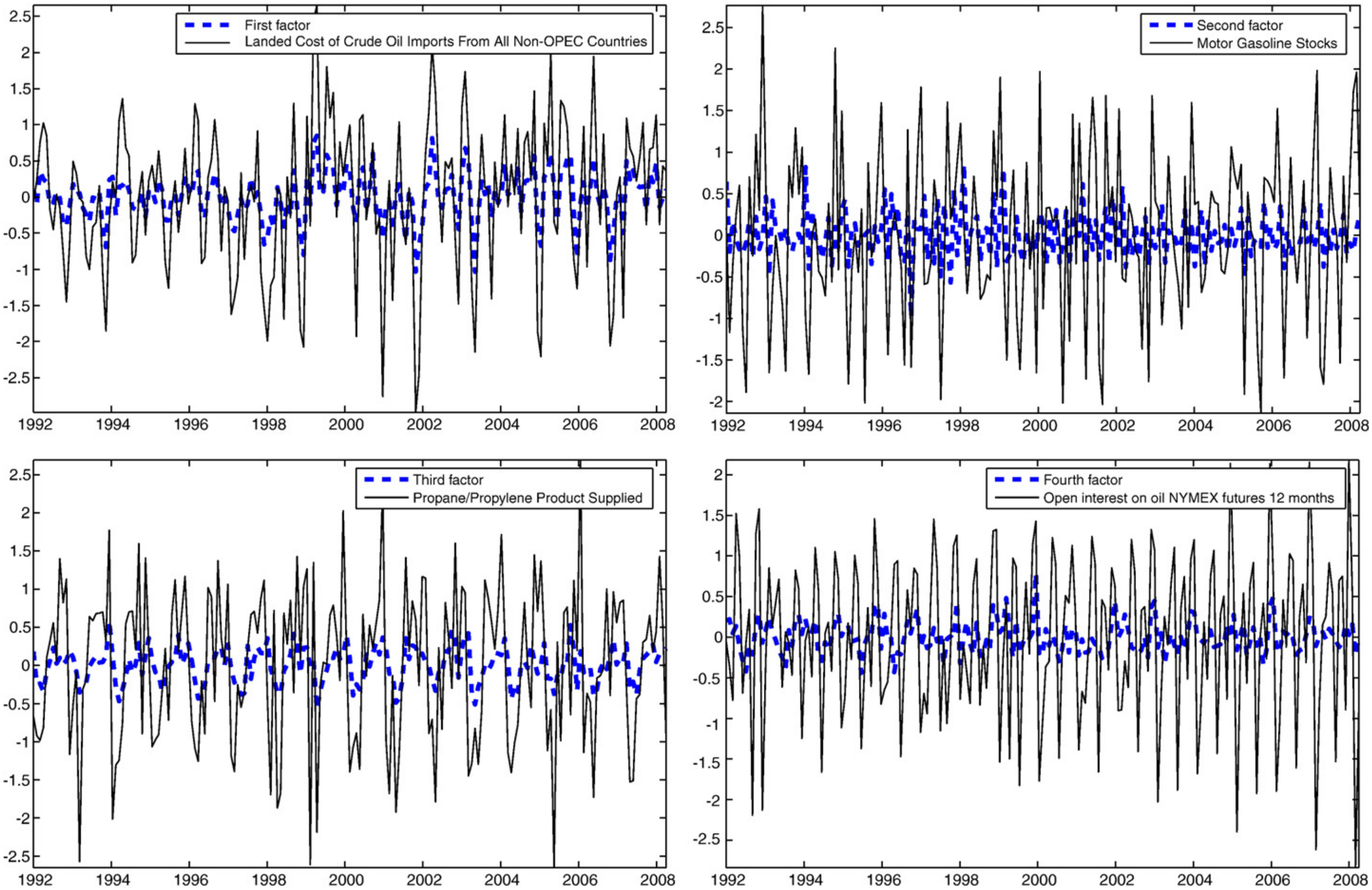

Fig. 1. Estimated factors and correlated series.

Fig. 1 plots the estimated factors together with the most correlated series of the panel dataset. The factor loadings are plotted in Fig. 2. I break down the contribution to the loadings of each factor by three groups of series, divided into energy prices, energy quantities and macro and financial data. The contributions to the factors differ largely across series. Energy prices provide the largest contribution to the first factor. Energy quantities instead account for the largest weights in the second and third factors. Finally, macro and financial series determine the largest fraction of the fourth factor. These considerations support the economic interpretation of the factors discussed earlier.

4.2. Preliminary evidence on the role of factors

In order to understand the relation between the factors and the return on oil futures prices at different maturities, I report the correlation between the oil price returns and the lagged factors in Table 3. The correlations differ in terms of size and sign across factors. The first factor has a large and positive contemporaneous correlation with all the returns. This is consistent with the interpretation of measure of cost pressure on futures prices. The other three facts are less correlated with the returns. The sign of the contemporaneous correlation on the second and third factors is negative, in agreement with the idea that available stocks provide a buffer to prices. The third factor is however weakly correlated, with the magnitude of correlation increasing at longer lags. Overall, this preliminary evidence suggests that not all the factors have predictive power for the returns at various lags.

To explore further this issue, I estimate unrestricted regressions of the returns on the factors, which takes the form

$$Y_t = \mu + \Lambda F_t + \epsilon_t. \quad (3)$$

Table 4 reports the parameters estimates and the fraction of explained variation. Two observations arise. The first one concerns the fact that only the first and the fourth factors have statistically significant coefficients for regressions of all the returns. The estimated coefficients are the signs one would expect from the correlation analysis. The second consideration is that the regressions explain large fractions of the variation in the returns of up to 6 months of maturity. Moreover, the longer the maturity, the more limited the scope of the factors for explaining the dynamics of the returns. For 1-year oil futures, the joint predictive power of the factors becomes low, as the R^2 declines to approximately 12%.

4.3. Parameter estimates

Before estimating the factor-augmented VAR, I evaluate the persistence of all the variables. In order to investigate the null of a unit root, I run the tests proposed by Dickey and Fuller (1979) and Phillips and Perron (1988). Instead of relying on the standard formulation of these tests, I apply the state-of-the-art modifications proposed by Perron and Ng (1996), Perron and Ng (2001).[9] Table 5 reports the test statistics. The results indicate that the null of a unit root is rejected for all the variables. Hence, all the series to be included in the FAVAR can be modelled as stationary variables.

The specification of the FAVAR model 1 is evaluated through four types of diagnostic tests. For the choice of the lag length, I compute

[9] These are based on the use of Generalized Least Squares detrended data for the estimation of the spectral density matrix at zero frequency, and on the computation of a class of improved selection criteria for the choice of the order of the underlying autoregression. Perron and Ng (1996) shows that both aspects improve the small-sample properties of the tests.

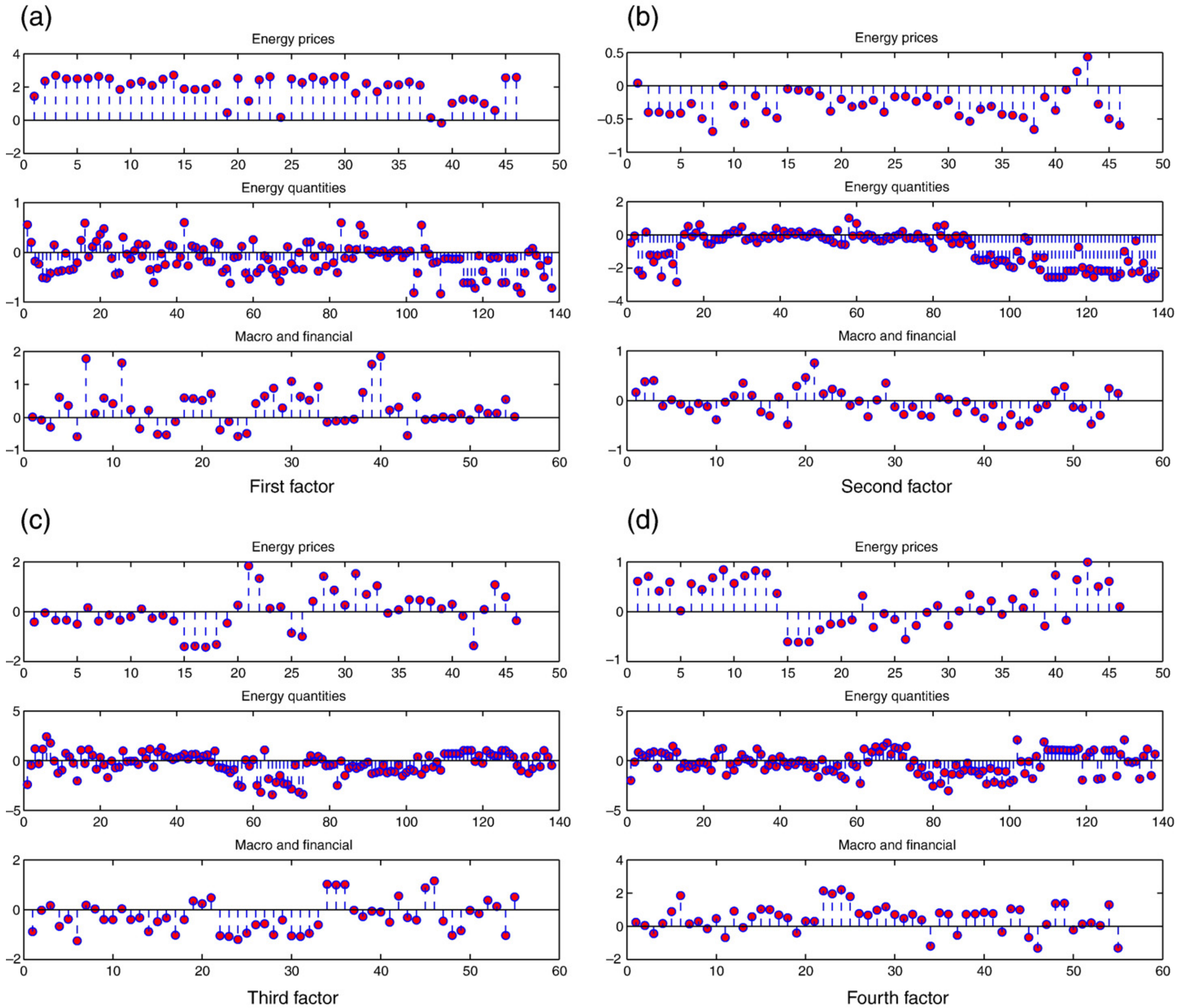

Fig. 2. Factor loadings for types of data series in the panel.

Table 3
Correlations between factors and returns.

	Factor 1	Factor 2	Factor 3	Factor 4
(a) Contemporaneous correlation				
1 month	0.896	−0.129	−0.019	0.109
3 months	0.887	−0.143	−0.023	0.119
6 months	0.859	−0.085	−0.126	0.111
12 months	0.299	−0.079	−0.051	0.122
(b) Correlation with 1-month lagged factors				
1 month	0.146	0.067	0.015	0.187
3 months	0.153	0.071	0.018	0.182
6 months	0.202	0.017	−0.056	0.096
12 months	0.144	−0.093	−0.131	0.0006
(c) Correlation with 6-month lagged factors				
1 month	−0.005	−0.009	0.154	0.015
3 months	0.005	−0.001	0.154	0.014
6 months	0.049	0.041	0.143	0.049
12 months	−0.035	0.084	0.033	0.180

Table 4
Unrestricted regressions of returns on factors.

	1 month	3 months	6 months	12 months
Factor 1	0.887	0.900	0.859	0.298
	[0.030]	[0.029]	[0.036]	[0.067]
Factor 2	−0.143	−0.133	−0.084	−0.079
	[0.035]	[0.028]	[0.035]	[0.068]
Factor 3	−0.023	−0.068	−0.126	−0.050
	[0.029]	[0.031]	[0.033]	[0.069]
Factor 4	0.119	0.098	0.110	0.122
	[0.031]	[0.029]	[0.035]	[0.069]
R^2	0.823	0.843	0.774	0.113

Legend: Brackets report standard errors. Constants are omitted.

standard information criteria, such as the Akaike and Bayesian criteria. I also use the standard tests for serial correlation described in Davidson and MacKinnon (1993, p. 359), the test for Autoregressive Conditional Heteroskedasticity (ARCH) effects outlined in Davidson and MacKinnon (1993, p. 355), and the heteroskedasticity-robust Chow test for parameter stability discussed in Davidson and MacKinnon (1993, p. 378).

Table 5
Unit-root tests.

	Returns				Factors			
	1 month	3 months	6 months	12 months	1	2	3	4
Phillips–Perron Mz_{α}	−94.742	−94.908	−93.516	−42.113	−122.625	−16.542	−20.005	−15.260
Phillips–Perron Mz_{t}	−6.881	−6.888	−6.836	−4.586	−7.830	−2.679	−2.888	−2.871
Sargan–Bhargava	0.072	0.097	0.072	0.108	0.064	0.056	0.144	0.126
Mod. point-optimal	0.260	0.258	0.266	0.588	0.199	0.329	1.033	0.837

Legend: The auxiliary models include a constant. The lag lengths are chosen using the modified BIC discussed in Perron and Ng (2001). The modified Phillips–Perron are all outlined in Perron and Ng (1996), the point-optimal test is from Elliott et al. (1996) and is amended in Perron and Ng (2001) together with Sargan and Bhargava (1993) test. All the test statistics are significant at the 5% level.

The Chow test is evaluated from January 2006. Table 6 reports the results. Panel (a) support a FAVAR model with two lags. Panel (b) indicates that the FAVAR(2) is well specified. The model has stable parameters, and there are neither signs of serial correlation nor ARCH in the residuals.

The estimates of the FAVAR model are detailed in Table 7. The upper part of the diagonal of the coefficient matrix Γ_1 suggests that only the first fact displays a certain degree of persistence. Additional evidence on the relation between the factors and the returns can be obtained through pairwise Granger-causality tests in the VAR. These are F tests for zero restrictions on the lagged coefficients of a variable onto another. Table 8 reports the test statistics and the p-values for the null of Granger causality of the factors for the returns, and vice versa. The first panel shows that not all the factors have predictive power for the returns. In this sense, the most important factor is the first one. The second factor, instead, does not Granger-cause any of the returns. Since these are bivariate tests, they provide no information on the indirect relation between variables. For instance, the second factor might Granger-cause another factor, which can in turn have predictive power on the returns. Interestingly, the second panel shows that the returns Granger cause of the factors. This highlights one of the advantages of the modelling strategy pursued in this paper, namely capturing the interaction between the observable and non-observable variables.

Fig. 3 plots the fitted series in-sample. The fitted series do not succeed in capturing the large variation that characterizes the historical data. However, they fit the peaks relatively well. In the case of the returns on 1-year futures, the model replicates the large swing of the sample that takes place in 1996–1998.

4.4. Out-of-sample forecasts

In this section, I compare the performance for out-of-sample forecasts from the FAVAR with that of alternative models. In particular, the competitor models are

- a VAR of returns only

$$\hat{Y}_{t+h|t} = \hat{\mu} + \hat{\Lambda} Y_t, \tag{4}$$

- a factor-only VAR

$$\hat{Y}_{t+h|t} = \hat{\mu} + \hat{\Lambda} F_t, \tag{5}$$

- a random walk

$$\hat{Y}_{t+h|t} = Y_t. \tag{6}$$

The forecasting exercise is run as follows. I initialize the parameter estimates on data until December 2002. The forecasts are then

Table 6
Diagnostic tests of the FAVAR model.

	(1) Lag length choice							
	1 lag	2 lags	3 lags	4 lags	5 lags	6 lags	7 lags	8 lags
AIC	−38.015	−43.116	−42.908	−42.880	−42.715	−42.709	−42.613	−41.904
BIC	−29.080	−30.209	−29.910	−29.871	−29.705	−29.664	−28.900	−28.870
	(2) Diagnostics for FAVAR(2)							
	Factor 1	Factor 2	Factor 3	Factor 4	1 month	3 months	6 months	12 months
Serial correlation:								
AR(1)	0.657	0.120	0.073	0.208	0.279	0.496	0.101	0.086
	[0.274]	[0.686]	[0.879]	[0.591]	[0.570]	[0.435]	[0.692]	[0.783]
AR(6)	1.304	4.583	4.951	4.988	5.290	2.750	5.046	4.929
	[0.958]	[0.271]	[0.160]	[0.158]	[0.117]	[0.521]	[0.156]	[0.168]
Heteroskedasticity:								
ARCH(1)	0.657	0.120	0.173	0.208	0.276	0.496	0.101	0.086
	[0.290]	[0.910]	[0.886]	[0.801]	[0.710]	[0.480]	[0.959]	[0.983]
ARCH(6)	3.054	2.983	2.830	1.984	2.064	2.750	1.987	5.120
	[0.250]	[0.271]	[0.305]	[0.408]	[0.397]	[0.319]	[0.401]	[0.168]
Stability:								
Chow test	8.026	8.730	8.529	8.904	4.332	6.140	6.373	6.809
	[0.255]	[0.210]	[0.238]	[0.196]	[0.471]	[0.390]	[0.372]	[0.355]

Legend: AIC and BIC denote, respectively, the Akaike and Bayesian information criteria. The test for serial correlation is from Davidson and MacKinnon (1993, p. 359). The test for ARCH effects is described in Davidson and MacKinnon (1993, p. 557). The stability test is the heteroskedasticity-robust Chow test outlined in Davidson and MacKinnon (1993). Brackets report significance levels.

Table 7
Parameter estimates of the FAVAR model.

	$\Phi(1)$							
Factor 1	0.657	−0.120	0.073	0.208	0.279	−0.496	−0.101	0.086
	[0.193]	[0.086]	[0.079]	[0.071]	[0.629]	[0.635]	[0.192]	[0.083]
Factor 2	−.0004	−0.583	−0.036	−0.224	−0.194	0.459	−0.446	0.123
	[0.158]	[0.071]	[0.065]	[0.058]	[0.517]	[0.521]	[0.156]	[0.068]
Factor 3	−0.387	−0.096	0.571	0.161	0.874	−0.680	−0.0008	0.009
	[0.174]	[0.078]	[0.072]	[0.064]	[0.569]	[0.575]	[0.175]	[0.075]
Factor 4	−0.243	−0.018	−0.031	0.126	0.302	−0.203	−0.016	−0.010
	[0.195]	[0.087]	[0.081]	[0.072]	[0.636]	[0.642]	[0.195]	[0.083]
1 month	0.588	−0.016	0.065	0.225	0.632	−0.962	−0.141	0.085
	[0.207]	[0.092]	[0.086]	[0.076]	[0.675]	[0.681]	[0.206]	[0.089]
3 months	0.566	−0.002	0.062	0.220	1.315	−1.639	−0.118	0.084
	[0.204]	[0.091]	[0.084]	[0.075]	[0.666]	[0.672]	[0.204]	[0.087]
6 months	0.534	−0.032	−0.042	0.126	0.798	−0.889	−0.194	−0.083
	[0.211]	[0.094]	[0.087]	[0.078]	[0.687]	[0.693]	[0.211]	[0.091]
12 months	0.311	−0.061	−0.088	0.043	−0.476	0.169	0.382	−0.523
	[0.186]	[0.083]	[0.077]	[0.068]	[0.607]	[0.612]	[0.186]	[0.081]
	$\Phi(2)$							
Factor 1	−0.319	0.031	−0.223	0.049	0.976	−0.953	0.097	−0.081
	[−0.014]	[0.079]	[0.080]	[0.072]	[0.636]	[0.631]	[0.196]	[0.085]
Factor 2	−0.015	−0.145	0.156	0.322	0.168	−0.031	−0.414	0.027
	[0.138]	[0.064]	[0.066]	[0.052]	[0.523]	[0.518]	[0.161]	[0.070]
Factor 3	−0.167	−0.327	−0.144	−0.081	0.511	−0.160	−0.056	−0.034
	[0.152]	[0.071]	[0.072]	[0.065]	[0.576]	[0.571]	[0.178]	[0.078]
Factor 4	0.228	−0.221	0.405	−0.009	−0.966	0.689	0.064	−0.061
	[0.170]	[0.079]	[0.081]	[0.073]	[0.643]	[0.637]	[0.199]	[0.087]
1 month	−0.131	0.064	−0.234	0.048	1.179	−1.261	0.072	−0.067
	[0.181]	[0.084]	[0.085]	[0.077]	[0.682]	[0.676]	[0.211]	[0.091]
3 months	−0.141	0.076	−0.222	0.044	1.493	−1.607	0.119	−0.070
	[0.178]	[0.083]	[0.087]	[0.076]	[0.673]	[0.667]	[0.208]	[0.091]
6 months	−0.328	0.054	−0.158	0.017	1.212	−1.118	0.170	−0.182
	[0.184]	[0.086]	[0.087]	[0.079]	[0.695]	[0.689]	[0.215]	[0.094]
12 months	0.038	−0.102	−0.059	0.052	−0.079	0.081	0.021	−0.460
	[0.162]	[0.076]	[0.077]	[0.069]	[0.614]	[0.609]	[0.190]	[0.083]
	Σ							
Factor 1	0.797							
Factor 2	−0.029	0.538						
Factor 3	0.002	−0.072	0.653					
Factor 4	0.138	−0.026	−0.116	0.815				
1 month	0.774	−0.117	−0.029	0.225	0.916			
3 months	0.759	−0.122	−0.036	0.234	0.899	0.893		
6 months	0.764	−0.074	−0.105	0.226	0.855	0.850	0.951	
12 months	0.288	−0.045	0.014	0.144	0.322	0.319	0.390	0.742

Legend: Brackets report standard errors. Constants are omitted.

Table 8
Pairwise Granger-causality *F* tests.

	Factor 1	Factor 2	Factor 3	Factor 4
Does not Granger-cause				
1 month	2.685	0.405	3.913	3.405
	(0.071)	(0.667)	(0.022)	(0.035)
3 months	3.661	0.450	3.794	3.139
	(0.028)	(0.638)	(0.024)	(0.046)
6 months	3.717	0.039	1.879	1.245
	(0.026)	(0.962)	(0.155)	(0.204)
12 months	15.351	1.789	2.727	0.831
	(7e−7)	(0.169)	(0.068)	(0.437)
	1 month	3 months	6 months	12 months
Does not Granger-cause				
Factor 1	0.281	0.100	1.065	2.381
	(0.756)	(0.905)	(0.347)	(0.095)
Factor 2	6.512	6.550	12.584	2.521
	(0.002)	(0.002)	(7e−6)	(0.083)
Factor 3	2.214	2.475	1.729	0.524
	(0.112)	(0.009)	(0.180)	(0.592)
Factor 4	1.833	1.901	2.937	2.040
	(0.162)	(0.152)	(0.055)	(0.132)

Legend: This table reports pairwise *F* statistics and their *p*-values (in brackets).

computed for various horizons, and the model estimates are updated recursively by estimating with one additional data-point at the time.

Table 9 reports the root mean squared errors (RMSEs). Table 10 lists the squared errors relative to those of a random walk. The FAVAR generates the best forecasts for 1- and 3-month and 1-year returns at short predictive horizons. The VAR with returns only is instead the best predictor for returns 6 months ahead. For forecast horizons longer than 3 months, the FAVAR generates the same squared errors of either the VAR with returns only or the factor-only model. However, the squared errors generated by the models are rather close. This means that no major reduction in RMSE are obtained from choosing the best-performing model. To summarize, the joint information from factors and returns improves to a limited extent the predictive power for the returns at short horizons.

5. Conclusion

This paper models the dynamics of the term structure of oil futures prices by using information from a panel dataset including over 230 series with global macroeconomic indicators, financial market indices, quantities and prices of energy products. I estimate a Factor-Augmented

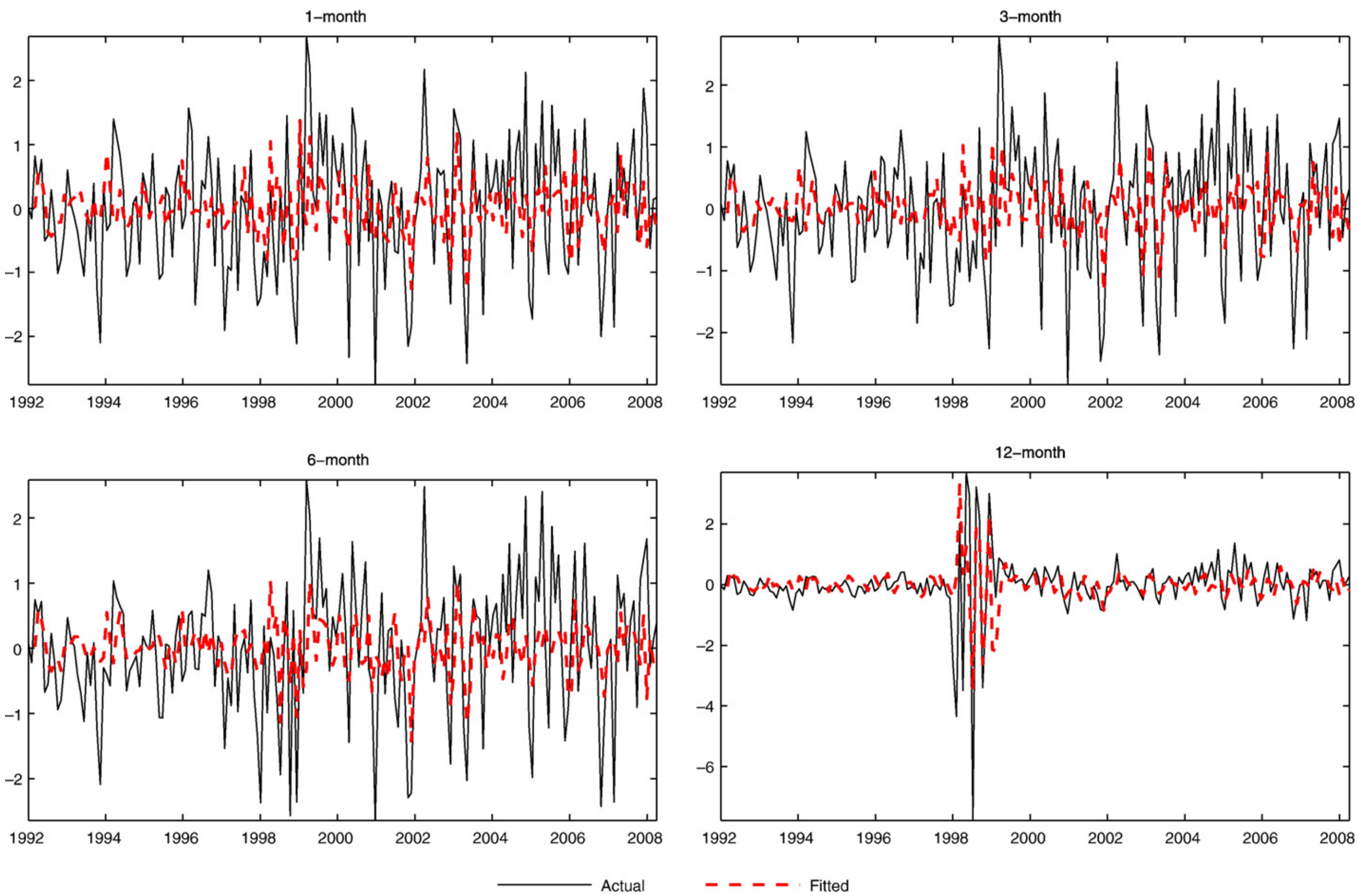

Fig. 3. Observed and model-implied returns.

Table 9
Out-of-sample forecasts: RMSEs.

Horizon	1 month	3 months	6 months	12 months	1 month	3 months	6 months	12 months
	(a) FAVAR				(b) VAR with returns only			
1	1.0353	1.0395	1.1015	1.1642	1.0514	1.0651	1.0787	1.1898
2	1.0161	1.0215	1.0628	1.1268	1.0209	1.0272	1.0709	1.1267
3	0.9732	0.9844	1.0321	1.0997	0.9742	0.9865	1.0191	1.1108
4	0.9809	0.9931	1.0379	1.1159	0.9807	0.9953	1.0301	1.1184
5	0.9760	0.9890	1.0376	1.1169	0.9761	0.9907	1.0309	1.1226
6	0.9849	0.9978	1.0455	1.1242	0.9848	0.9996	1.0383	1.1302
	(c) Factor-only model				(d) Random walk			
1	0.7305	0.7162	0.7292	1.6776	1.4550	1.4833	1.5277	0.6941
2	0.6564	0.6437	0.6318	1.4703	1.5684	1.6054	1.6961	0.7690
3	0.6898	0.6543	0.6454	1.5050	1.4142	1.5073	1.6014	0.7315
4	0.6949	0.6848	0.6784	1.5587	1.4124	1.4507	1.5303	0.7157
5	0.7905	0.7842	0.7736	1.7813	1.2353	1.2614	1.3414	0.6270
6	0.6557	0.6391	0.6331	1.4641	1.5023	1.5614	1.6513	0.7678

Legend: This table reports the root mean squared errors of out-of-sample forecasts at various horizons.

Vector Autoregression with latent factors extracted from the panel. I show that latent factors generate information which, once combined with that of the returns, improves the forecasting performance for oil prices. Furthermore, I find that a factor correlated to purely financial developments contributes to the model performance, in addition to factors related to energy quantities and prices.

The results presented here can be extended in a number of directions. I am planning to use Bayesian model averaging to study the performance of the best-performing subset of factors for forecasting the term structure of oil prices. Moreover, the factors could be used to identify the impact of oil demand and supply shocks. In this sense, it would be important to understand what role purely financial market variables can play for the persistence and magnitude of the estimated shocks.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.eneco.2009.11.003.

Table 10
Out-of-sample forecasts: RMSEs relative to a random walk.

Horizon	1 month	3 months	6 months	12 months	1 month	3 months	6 months	12 months
	(a) FAVAR				(b) VAR with returns only			
1	0.7115	0.7008	0.7210	1.6772	0.7226	0.7181	0.7061	1.7141
2	0.6479	0.6363	0.6266	1.4653	0.6509	0.6398	0.6314	1.4652
3	0.6882	0.6531	0.6445	1.5034	0.6889	0.6545	0.6364	1.5185
4	0.6945	0.6845	0.6783	1.5592	0.6944	0.6861	0.6731	1.5627
5	0.7901	0.7840	0.7735	1.7812	0.7901	0.7854	0.7685	1.7903
6	0.6556	0.6390	0.6331	1.4641	0.6556	0.6402	0.6288	1.4720
	(c) Factor-only model							
1	0.7305	0.7162	0.7292	1.6776				
2	0.6564	0.6437	0.6318	1.4703				
3	0.6898	0.6543	0.6454	1.5050				
4	0.6949	0.6848	0.6784	1.5587				
5	0.7905	0.7842	0.7736	1.7813				
6	0.6557	0.6391	0.6331	1.4641				

Legend: This table reports the root mean squared errors of out-of-sample forecasts relative to the random-walk forecast.

References

Alquist, Ron, and Lutz Kilian (2008), What Do We Learn from the Price of Crude Oil Futures?, unpublished manuscript, University of Michigan.

Askari, Hossein, Krichene, Noureddine, 2008. Oil price dynamics (2002–2006). Energy Economics 30 (5), 2134–2153.

Bai, Jushan, Ng, Serena, 2008. Determining the number of factors in approximate factor models. Econometrica 70 (1), 191–221.

Bernanke, Ben S., 2008. Semiannual Monetary Policy Report to the Congress before the Committee on Banking, Housing, and Urban Affairs. U.S Senate, Washington, July 15.

Bernanke, Ben S., Boivin, Jean, Eliasz, Pitr, 2008. Measuring the effects of monetary policy: a Factor-Augmented Vector Autoregressive (FAVAR) approach. Quarterly Journal of Economics 120 (1), 387–422.

Boivin, Jean, Ng, Serena, 2005. Understanding and comparing factor-based forecasts. International Journal of Central Banking 1 (3).

Brousseau, Vincent, Scacciavillani, Fabio, 1999. A Global Hazard Index for the World Foreign Exchange Markets. ECB Working Paper, No. 1.

Carmona, R., and M. Ludkovski (2005), Spot Convenience Yield Models for Energy Markets, unpublished manuscript, Princeton University.

Casassus, J., Collin-Dufresne, 2005. Stochastic convenience yield implied from commodity futures and interest rates. The Journal of Finance LX (5).

Chong, James, and Joelle Miffre (2006), Conditional Risk Premia and Correlations in Commodity Futures Markets, unpublished manuscript, EDHEC Business School.

Chung, Joanna, 2008. Fund accused of manipulating oil price. Financial Times, July 24.

Davidson, R., MacKinnon, J., 1993. Estimation and Inference in Econometrics. Oxford University Press, Oxford.

Dickey, D.A., Fuller, W.A., 1979. Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association 74.

Elliott, G., Rothenberg, T.J., Stock, J.H., 1996. Efficient tests for an autoregressive unit root. Econometrica 64.

Gibson, R., Schwartz, E.S., 1990. Stochastic convenience yield and the pricing of oil contingent claims. Journal of Finance XLV (3).

Gorton, Gary B., Hayashi, Fumio, Rouwenhorst, K.Geert, 2007. The Fundamentals of Commodity Futures Returns. NBER Working Paper, 13249.

Kilian, Lutz (2008a), Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market, unpublished manuscript, University of Michigan.

Kilian, Lutz (2008b), The Economic Effects of Energy Price Shocks, unpublished manuscript, University of Michigan.

Ludvigson, Sydney C., and Serena Ng (forthcoming), Macro factors in bond risk premia, Review of Financial Studies, forthcoming.

Mackintosh, James, 2008. Hedge funds have worst month since 2000. Financial Times, July 26.

Marzo, Massimiliano, Fabrizio Spargoli, and Paolo Zagaglia (2008), The Information Content of Oil Spreads: How Far Do They Go?, unpublished manuscript, Sveriges Riksbank, January.

Mönch, Emanuel, 2008. Forecasting the yield curve in a data-rich environment: a no-arbitrage Factor-Augmented VAR approach. Journal of Econometrics 146 (1).

Perron, Pierre, Ng, Serena, 1996. Useful modifications to unit root tests with dependent errors and their local asymptotic properties. Review of Economic Studies 63.

Perron, Pierre, Ng, Serena, 2001. Lag length selection and the construction of unit root tests with good size and power. Econometrica 69.

Phillips, P.C.B., Perron, P., 1988. Testing for a unit root in time series regression. Biometrika 75.

Pindyck, R.S., 2004. Volatility and commodity price dynamics. Journal of Fugures Markets 24 (11).

Postali, F.A.S., Picchetti, P., 2006. Geometric Brownian motion and structural breaks in oil prices: a quantitative analysis. Energy Economics 28.

Radchenko, S. (2007), Do Commodity Prices and Volatility Jump?, unpublished manuscript, University of North Carolina at Charlotte.

Sargan, J.D., Bhargava, A., 1993. Testing for residuals from least squares regression being generated by Gaussian random walk. Econometrica 51.

Schwartz, E., 1997. The stochastic behavior of commodity prices: implications for valuation and hedging. Journal of Finance LII(3), 922–973.

Schwartz, E., Smith, J.E., 2000. Short-term variation and long-term dynamics in commodity prices. Management Science 46 (7).

Stock, James H., Watson, Mark W., 2002. Macroeconomic forecasting using diffusion indexes. Journal of Business and Economic Statistics 20, 147–162.

Trichet, Jean-Claude, 2008. Hearing at the Economic and Monetary Affairs Committee of the European Parliament. Brussels, 25 June.