Risk-adjusted forecasts of oil prices

Patrizio Pagano and Massimiliano Pisani*
Bank of Italy
Research Department

June 16, 2006

Abstract

This paper documents the existence of a significant forecast error on crude oil futures, particularly evident since the mid-1990s, which is negative on average and displays a non-trivial cyclical component (risk premium). We show that the forecast error on oil futures could have been explained in part by means of real-time US business cycle indicators, such as the degree of utilized capacity in manufacturing. An out-of-the-sample prediction exercise reveals that futures which are adjusted to take into account this time-varying component produce significantly better forecasts than those of the unadjusted futures and random walk, particularly at horizons of more than 6 months.

Keywords: Oil, Forecasting, Futures.
JEL classification: E37, E44, G13, Q4.

---

*We thank Fabio Busetti, Paola Caselli, Stephane Dees, Alberto Locarno, Giuseppe Parigi, Stefano Siviero, and seminar participants at the Bank of Italy and at the Eurosystem Working Group on Forecasting for comments and stimulating discussions. Giovanna Poggi provided valuable research assistance. We are solely responsible for any errors. The opinions expressed in this paper do not necessarily reflect those of the Bank of Italy. Address: via Nazionale 91, 00184 Rome - Italy. E-mail: `patrizio.pagano@bancaditalia.it`; `massimiliano.pisani@bancaditalia.it`

1 Introduction

Although the dependency of global economic activity on crude oil has fallen steadily over the last thirty years, the oil price baseline assumption remains an important variable for all the macroeconomic forecasts of both national and international institutions. For example, forecast of future oil prices are crucial in central bank's monetary policy decisions, because they enter the construction of expected inflation and output-gap (Svensson, 2005). The increase in oil prices in recent months (Figure 1), which has surprised most analysts by its rapidity and intensity, prompts a new call to investigate the validity of the forecasting assumptions.

A commonly used approach to forecast oil prices is based on futures contracts. The notion that the futures price might be the optimal forecaster of the spot price is a by-product of the financial market efficiency hypothesis: the requirement that the average forecasting error is zero is consistent with both efficiency in financial markets (the absence of profitable arbitrage opportunities) and the unbiasedness property of the forecaster (zero forecasting error on average).

However, the possibility of a systematic forecast error of oil futures cannot be excluded if a risk premium exists. In fact, since the oil spot price covaries positively with overall economic activity, this creates an undiversifiable risk for holders of oil who, as a reward, will expect over the holding period an average spot price higher than futures price currently quoted. As a consequence, futures oil prices forecast would yield a significant ex post error.[1]

[1] The benefits of holding oil stocks, usually referred to as "convenience yield", arise from the use of inventories to reduce production and marketing costs and to avoid stock-outs (for more details, see Pindyck, 2001). The size of the convenience yield determines whether the futures price is greater or smaller than the spot price. When the convenience yield is sufficiently high for the spot price to exceed the futures price, the market is described as being in *backwardation.* As emphasized by Gorton and Rouwenhorst (2004) the notion of backwardation should involve a comparison of the futures price to the *expected* spot price in the future, which is unobservable when the futures price is set. In the practice of commodity trading backwardation is commonly used to describe the position of futures prices in relation to *current* spot prices. While backwardation in the former sense is equivalent to the existence of a positive risk premium, backwardation in the latter sense is not. The two definitions of backwardation are often used interchangeably as if they were equivalent. But only backwardation in the first sense refers to the notion of a positive risk premium to investors in commodity futures. "Where the futures contract trades relative to the *current* spot does not directly speak to the presence of a risk premium" (Gorton

Building on a methodology introduced by Piazzesi and Swanson (2004) to explain the excess return on federal funds futures, we document the existence of systematic ex post forecast errors and the fact that these errors have, since the second half of the 1990s, a non-trivial counter-cyclical component that could in part be predicted by using the level of utilized capacity in US manufacturing (a proxy of the conditions of the business cycle).

We assess the forecasting performance of our approach on the basis of an out-of-the-sample prediction exercise. Results show that forecasts adjusted to take into account the time-varying risk premium (so called "risk-adjusted forecasts") display lower mean and root-mean squared errors than the unadjusted futures, the simple constant-adjusted futures and the random walk, particularly at horizons of over 6 months.[2]

The intuition for this result may be that when business conditions are poor, income is low and risk premia must be high to induce substitution from consumption to investment (Cochrane, 2005; Fama and French, 1989). From a different perspective, currently low (high) levels of utilized capacity may signal an increase (decrease) in capacity utilization into the future and therefore – given input complementarity – a rising (decreasing) demand for oil.[3] Hence, when oil demand is bound to increase, this will put pressure on the spot price, and the risk premium required on oil futures is correspondingly high. In turn, when oil demand is expected to remain put or to decrease, the risk premium will be low or even negative.

We are not the first to have found that futures may yield biased forecast of oil prices. For instance, Gorton and Rouwenhorst (2004) show that commodity futures risk premium has been equal in size to the historical risk premium of stocks (the equity premium) and has exceeded the risk premium of bonds. In the framework of the marginal convenience yield, on the basis of estimates of the oil risk-adjusted discount rate, Pindyck (2001) estimates that the 6-month futures contract should under-predict the realized spot price by around 3 to 4.5 per cent.[4] These works concentrate on the average

---

and Rouwenhorst, 2004, p.25).

[2] As in Piazzesi and Swanson (2004) we use in the paper the label "risk-premium" quite loosely, referring to the part of the forecast error that could be predicted.

[3] In a study on the effect of energy price increases on economic activity, Finn (2000) emphasizes the complementarity between capacity utilization and energy consumption.

[4] There are also studies on the efficiency of the oil futures market and on the forecasting properties of the futures that reached opposite conclusions. For example, Chinn et al. (2005) find that over the period January 1999-October 2004 futures prices are unbiased

risk premium, neglecting its time-variability. On the contrary, Moosa and Al-Lougani (1994), focusing on the properties of spot and futures prices in the context of co-integration, find that there is a time-varying risk premium that can be adequately modelled by a GARCH process. Consistently with this result, Considine and Larson's findings (2001) suggest that crude oil assets contain a risk premium that rise sharply with higher price volatility. Similarly to what we do, Coimbra and Esteves (2004) provide evidence of a correlation between oil futures forecast errors and market expectation errors on world economic activity. Yet, none of these papers documents, as we do, the correlation of oil futures forecast errors and real time macroeconomic indicators over the business cycle.

The rest of the paper is organized as follows: in the next Section we document the size of the ex post forecast errors on oil futures, showing that these display a non trivial error component. In Section 2, we document the presence of a structural break in the mid-1990s; we also estimate the relationship linking oil futures forecast errors to business cycle conditions and conduct some robustness analyses. In Section 3 we propose a method to adjust the forecast based on oil futures and evaluate the performance of the risk-adjusted forecasts with respect to the futures and other alternatives. Section 4 contains some concluding remarks.

2 Forecast errors on crude oil futures

In the following analysis we use oil price futures on the WTI grade. They trade on the New York Mercantile Exchange (NYMEX) and are settled each month. The contract provides for the physical delivery of 1,000 barrels of oil at any point during the settlement month. Any one of several different types of crude oil can be delivered, but WTI is usually chosen. The trading began in 1983 initially with a delivery period of up to six months, which was later gradually extended in line with a substantial increase in the volume of contracts traded.

Let $f_t^{(n)}$ denote the oil price implicit in the futures contract expiring in month $t+n$; we will refer to n as the n-month-ahead contract. Let also p_{t+n}

predictors of crude oil, even if futures typically explain only a small proportion of the variability in oil price movements. However, using the same methodology as Chinn et al. (2005), over the April 1989-December 2003 period, Chernenko et al. (2004) find mixed evidence on the existence of risk premia associated with oil futures.

denote the ex post realized spot oil price in month $t+n$, measured in US dollars.

We can define the ex post realized forecast error as:

$$fe_{t+n}^{(n)} = f_t^{(n)} - p_{t+n}. \tag{1}$$

Under the expectations hypothesis $f_t^{(n)} = E_t\left(p_{t+n}\right)$, that is futures are equal to the expected future spot prices. Equation (1) therefore represents the forecast error. If the futures is an unbiased predictor of subsequent oil prices then the average forecast error should be zero. If this is not the case, we interpret the average forecast error as (minus) the risk premium; as a consequence, futures-based forecasts should be opportunely adjusted.

2.1 Bias

We test the assumption of no bias by running the following regression on all the contract horizons up to twelve months:

$$fe_{t+n}^{(n)} = \alpha^{(n)} + \varepsilon_{t+n}^{(n)}, \tag{2}$$

where α is a constant measuring the average ex post realized forecast error and ε is an error term.

We estimate regression (2) over the sample period for which futures data are available at all maturities up to one year, that is from January 1986 to December 2004. To compute forecast errors we take the simple average of futures daily quotations in the third week of each month t. The choice is suggested to avoid possible daily outliers. The week selected is the third because, as it will be clear below, it is the closest to the release of relevant business-cycle indicators. However, all the results also hold true when we sample the data on a particular day (the 15^{th}) of each month.

Given that futures contract overlap induces heteroskedasticity and autocorrelation, we compute standard errors using the Newey-West procedure, allowing for a $n-1$ Bartlett window.

Table 1 presents the results. As it can be seen, the futures is not an unbiased predictor: the value of the constant at each forecast horizon n is significantly negative, ranging from 15 cents to \$3.1, and longer-horizon contracts display larger forecast errors.

The simple average forecast errors displayed in Table 1 imply that a six-month contract under-predicts the realized spot by \$1.8, or around 8 per cent if evaluated at the mean price of the sample.

2.2 Capacity utilization

Up to this point we have documented the presence of a significant forecast error in oil price futures. As suggested by a large literature on financial markets (e.g. Cochrane, 2005, for a survey) this could be reconciled with the presence of a risk premium, which could also be time-varying. For instance, a number of studies (e.g. Cochrane and Piazzesi, 2005) show that excess returns on US treasuries are high in recessions and low in booms. As suggested by other works (Moosa and Al-Lougani, 1994; Coimbra and Esteves, 2004) excess returns on oil may display time-varying risk premia too.

To investigate whether business cycle phases help in explaining realized futures-based forecast errors on oil prices, we run the following regression:

$$fe_{t+n}^{(n)} = \alpha^{(n)} + \beta^{(n)} UCap_{t-1} + \varepsilon_{t+n}^{(n)}, \tag{3}$$

where $UCap$ is the degree of capacity utilization in US manufacturing, which is a proxy of the US business cycle. We focus on the US since it is the largest world oil consumer. Furthermore, we want to use data known to market participants at the time the future contact is subscribed, that is in month t. Since US capacity utilization values are released by the Federal Reserve around the 15^{th} day of each month for the previous month, we date our $UCap$ variable as $t-1$. Furthermore, these figures are subject to several backward revisions; so, to use in our regressions the values known to investors at the time of the contract subscription, we use a real-time series.[5] Table 2 reports the results. Forecast errors and utilized capacity in US manufacturing are positively related. The estimated slope coefficient is positive, increases with the maturity of the contract and is statistically significant from the 4-month horizon on. The positive sign on the slope coefficient (β) suggests that when business conditions are poor – as indicated by a low level of capacity utilization – income is also low and risk premia must be high to induce substitution from consumption to investment. Yet a positive β also implies that the lower is currently utilized capacity the higher is the oil price expected $n-$periods ahead, that is:

$$E_t \left[p_{t+n} \right] = f_t^{(n)} - \hat{\alpha}^{(n)} - \hat{\beta}^{(n)} UCap_{t-1}. \tag{4}$$

The intuition of the result is based on the mean-reverting behavior of the capacity utilization. In fact, the current degree of capacity utilization is a

[5] See the data appendix for further details.

'predictor' of its future variation. When utilized capacity is low (high), it is expected to grow (decrease) in the following $n-$periods.[6] As a consequence, oil demand is expected to increase (decrease) and, therefore, also the expected oil price.

2.3 Stability analysis

Since the purpose of this paper is to provide forecasts of oil prices, we need to investigate the stability of equation 3. In this section we perform formal tests and look at some suggestive evidence.

As a first step, we recursively estimate equation (3) using the first 60 observations for initialization. Figure 2 shows recursive residuals with the upper and lower 2 (recursively generated) standard error bands when the 3, 6, 9 and 12-month ex post futures forecast errors are used as dependent variable. In order to leave aside the Kuwait invasion we add as a regressor a dummy equal to one for that episode. Residuals outside the bands could be used as an indication of a possible structural break. It is evident that up to the mid-1990s residuals lay strictly inside the bands. On the contrary, the first time when residuals evidently cross the bands is in late 1995.[7] Probably this result is associated to the fact that during the 1990s the oil industry moved to what is referred to as "just-in-time inventories", partly reflecting greater pressure to increase profitability: as a consequence, the average inventory level dropped sharply in late 1995 (see Figure 3) to below the level of the mid-1970s and swings in demand, even predictable ones, started to be met by price changes.[8] Figure 4 shows the increase in oil price volatility: up to 1995 – leaving the period of the Kuwait invasion aside (August 1990-February 1991) – the oil price is roughly stable; on the contrary, afterwards there are both periods of sharp price decreases and periods of very rapid increases. The one-month moving average of daily percentage changes increased from an average of 1.2 per cent in the period 1987-1995 to 1.8 per cent afterwards. The null hypothesis of no difference in the two average values is strongly rejected at any conventional level by a standard F-test. As a final remark, note that the high volatility of the spot price that characterizes the oil market from

[6] For instance, a regression of the 12-month change in capacity utilization on its initial level and a constant yields an estimated slope coefficient of -.29 and a standard error of .06. Similar results hold on the other horizons.

[7] Recursive residuals on other horizons display a similar behavior.

[8] This point is forcefully emphasized in Lynch (2002).

the mid-1990s is consistent with the evolution of the oil spot price during the same period: a more volatile spot price implies a larger risk for the owners of oil stock, who, as a reward for that risk, expect a higher average spot price than the current futures price; as a result, on average the oil spot price increases over the period 1996-2004.[9]

The low average level of inventories may also have induced a stronger reaction of oil prices to macroeconomic fundamentals. Table 3 reports the pairwise contemporaneous correlation coefficients between changes in oil prices and changes in capacity utilization in manufacturing. As can be seen, in the period 1986-1995 there is no significant correlation between the two variables, while afterwards the correlation coefficients are always positive, and significantly so, at any horizon. At longer horizons the correlation between changes in oil prices and in capacity utilization tops 70 per cent.

We make a more formal analysis on this break by testing for the existence of a structural change in the coefficients of equation (3) across the periods before and after 1995. We run the following regression on all the available contract horizons:

$$\begin{aligned} fe_{t+n}^{(n)} &= \alpha_1^{(n)} D_{86-95} + \alpha_2^{(n)} D_{96-04} + \\ &\quad \beta_1^{(n)} D_{86-95} UCap_{t-1} + \beta_2^{(n)} D_{96-04} UCap_{t-1} + \varepsilon_{t+n}^{(n)}, \end{aligned} \tag{5}$$

where D_{86-95} and D_{96-04} are dummy variables equal to 1 in the sub-periods 1986-95 and 1996-2004, respectively.

As can be seen in Table 4, the F-tests strongly reject the assumptions that the coefficients are stable across the two sub-periods. The hypotheses that the two intercepts and the two slope coefficients in equation (5) are equal in the two sub-periods are rejected at the 5 per cent confidence level at horizons larger than 2, and at the 1 per cent at horizons larger than 3, both singularly and jointly.

Finally, we test for the stability of the variance of regression (3). Results of the Goldfeld-Quandt test on heteroscedasticity of residuals (Goldfeld and Quandt, 1965) and the Hansen test (Hansen, 1992) on stability of the variance of the regression are reported in Table 5. The Goldfeld-Quandt tests whether the variance of residuals in 1986-1995 equals that in 1996-2004: it fails to reject the null hypothesis of homoscedasticity at the 5 per cent level only at horizons 3 and 4. The Hansen test fails to reject the null hypothesis of

[9] For more details on the relationship between first and second moments of oil spot price see Pindyck (2001).

stable variance on the whole sample at horizons 2-4. When performed in each sub-sample, the latter test indicates that while the period 1986-1995 still displays some variance instability, the period 1996-2004 appears much more stable, especially with respect to the whole sample. In fact, in 1996-2004 it fails to reject the null hypothesis of variance stability at all horizons larger than 3. Overall, we take all the above evidence as suggesting that there is a break in the relationship between futures-based ex post forecast errors and business cycle conditions in the sample considered. The larger average error and the significantly higher sensitivity to business cycle conditions from the mid-1990s on lead us to focus the analysis on this sub-period.

Table 6 reports the results of regression (3). As in the whole sample, the slope coefficients become larger with the maturity of the contract and are statistically significant at the 5 per cent level or lower from the 3-month horizon on. The R^2 suggests that the percentage of the variance of the forecast error on oil futures explained by this specification is not trivial, especially at longer horizons. For instance, the model with utilized capacity explains almost 40 per cent of the forecast error on the oil futures contract with 12-month maturity (see Figure 5).

2.4 Robustness analysis

To gauge the robustness of the results we perform some alternative estimations.

First of all, we run the same regression using two other alternative cyclical variables, whose data are available at real time: non-farm payrolls and industrial production. For brevity we report results only for the regression using year-on-year growth in non-farm payrolls; those obtained using industrial production growth are qualitatively similar.[10] As can be seen in Table 7, employment growth is positively associated to the ex post realized forecast error, even if estimated coefficients display larger standard errors than those of the regression estimated using utilized capacity. This is not surprising, given the closer complementarity of utilized capacity in manufacturing and oil consumption.

Second, we check if explicitly forward-looking indicators capture the cyclical variability of the risk premium. To tackle this issue we run equation (3) using as cyclical variable alternatively the Conference Board leading indica-

[10] Results are available on request.

tor or the ISM manufacturing index. Unfortunately, none of these is available with real time data. Results – not reported for brevity but available on request – show that these variables have some explanatory power with respect to oil forecast errors only at short horizons (3 to 6 months).

Finally, another possible driving factor of the forecast error of oil futures could be related to the booming oil demand originating from developing countries. Unfortunately, we lack timely statistics of such countries' business cycles, but in order to capture the boost to world oil demand coming from China – which has become the second largest world oil consumer – we add to equation (3) the growth rate of Chinese industrial production. Since for the third week of each month we have revised data up to the month before last, we date this series $t-2$. Results in Table 8 show no correlation at all between Chinese industrial production and oil futures forecast error.

3 Predictability of oil prices in real time

Having documented the presence of a significant and cyclical forecast error, in this Section we evaluate the forecast of oil prices based on unadjusted futures.

We compare four alternative methodologies to forecast oil prices:

1. a random walk, which implies $E_t\left[p_{t+n}\right] = p_t$;

2. the unadjusted futures: $E_t\left[p_{t+n}\right] = f_t^{(n)}$;

3. the constant-adjusted futures, based on rolling-endpoint OLS estimates on a constant: $E_t\left[p_{t+n}\right] = f_t^{(n)} - \hat{\alpha}^{(n)}$;

4. the risk-adjusted futures, based on rolling-endpoint OLS estimates on a constant and the index of utilized capacity in US manufacturing at time $t-1$: $E_t\left[p_{t+n}\right] = f_t^{(n)} - \hat{\alpha}^{(n)} - \hat{\beta}^{(n)} UCap_{t-1}$.

We perform a set of rolling "out-of-sample" regressions. To obtain rolling-endpoint real-time (or expanding window) forecasts we initialize our estimates using the first 30 observations. We then compute forecasts up to the 12-month horizon, add a new observation and so on.

In sum, the results in Table 6 imply that the adjustment to be made over the futures will be smaller during booms and higher during slacks: that is, to obtain risk-adjusted forecasts we add to the futures a counter-cyclical term.

To gauge a quantitative measure of how different these four forecasting methodologies are, in Table 9 we report some summary statistics on forecast errors. The mean error of the cyclical risk-adjusted forecast is the lowest one, both at short and long horizons. For instance at the 3-month horizon the mean risk-adjusted error is just 45 cents, compared with \$1.3 with the constant adjustment, \$1.4 with the random walk and almost \$2 with the unadjusted futures. At the 9-month horizon the mean forecast error committed by the risk-adjusted futures is still low, 29 cents, compared with more than \$4 for the random walk, \$4.35 for the constant-adjusted futures and almost \$7 for the unadjusted futures.

A similar conclusion can be drawn on the basis of root mean squared errors, which for the risk-adjusted forecast are typically below those implied by the other three forecasting techniques considered; this is always true from the 4-month horizon on.

To check whether these root mean squared errors are also statistically significantly lower than their counterparts obtained with the other three methodologies we perform a Diebold-Mariano test (Diebold and Mariano, 1995) by running regression (3) on a moving window of 30 observations. P-values of the test statistics based on a pair-wise comparison of root mean squared errors are reported in Table 10.[11] Root mean squared errors of the cyclical adjusted futures are lower than those of the unadjusted futures after $n = 3$, and lower than those of the constant-adjusted futures and of the random walk after $n = 4$: the risk-adjusted forecast statistically out-performs the constant-adjusted at the 10 per cent or lower level from horizon 6 on, the unadjusted futures from horizon 7 on and the random walk from horizon 8 on.

Finally, Table 9 also reports the n-th autocorrelation (ρ_n) for the n-month-ahead forecast, which for an efficient forecast should be as close to zero as possible. Here results are just a little less clear cut: risk-adjusted forecasts are less autocorrelated than the constant-adjusted and the unadjusted futures in 10 out of the 12 horizons, while with respect to the random walk risk-adjusted forecasts display smaller autocorrelation in 7 horizons and

[11] The Diebold-Mariano test-statistic is the "t-stat" in a regression of the differences in the root mean squared errors for two alternative forecasts on a constant.

larger in 5.

One may also assess how different these methodologies are by looking at Figure 6. It shows forecasts of oil prices in two illustrative months, January 1997 and September 2003. In the upper panel data show that in January 1997 the spot oil price was around \$26 and, according to the futures, oil prices were expected to decline to just over \$20 by January 1998. Demand was very high and utilized capacity in manufacturing was running well above the historical average, at almost 83 per cent: margins for further increases in capacity utilization were limited and, therefore, the risk premium required over the futures would have been very low and the risk-adjusted forecast virtually indistinguishable from the futures itself. The constant-adjusted forecast would have signaled roughly constant prices. The realized spot did indeed decline (to \$16.3).

In summer 2003 (Figure 6, lower panel) oil prices were stable at around \$30. In September the futures signaled a weak decline in the price to just below \$26 in the following 12 months. The recovery out of the recession in 2001 was not yet firmly established and the capacity utilization index was still relatively low, at around 73 per cent. Margins for a pick up in demand were correspondingly high and the risk premium was sizeable: the risk-adjusted forecast would have signaled an oil price as high as \$38.4 by September 2004. Note that not taking into account the cyclical factor – as the constant-adjusted forecast does – would have yielded just slightly increasing oil prices. Indeed oil prices did rise and at the end of the horizon were at around \$45.

If the forecast error could have been significantly reduced by investors exploiting available information on the US business cycle, as we have shown, the question that naturally arises is why they did not do so. The US Commodity Futures Trading Commission (CFTC) provides aggregate data on most of the long and short positions held in futures markets, divided into hedging and non-hedging categories. An analysis of net long positions held by non-commercial traders, usually referred to as "speculators", reveals that both in 1999-2000 and in late 2003 these positions were largely positive, signaling expectations of rising oil prices, which effectively were realized in the following months.[12] Therefore, it is possible that this category of market

[12] Formal analysis of energy futures markets (Sanders et al., 2004) reveals that positive futures returns Granger-cause increases in the net long positions held by reporting non-commercial traders. There is no consistent evidence that traders' net long positions contain

participants was aware of this risk premium and provided an insurance to (hedging) commercial market participants. This notwithstanding, a significant part of the premium was not competed away. A possible explanation is that non-commercial traders represent a small percentage (just a little over 10 per cent) of all open interest, since they trade mainly in over-the-counter markets. Alternatively, as suggested by Piazzesi and Swanson (2004) in the context of futures on federal funds, the futures market may not be perfectly competitive or non-commercial traders may themselves be risk adverse.

4 Concluding remarks

This paper documents that crude oil futures display a significant ex post forecast error, which is negative on average. We also show that this forecast error has a non trivial cyclical component which can be, in part, explained by means of real-time US business cycle indicators, such as the degree of utilized capacity in manufacturing. Results appear robust to various checks such as the use of alternative US business cycle indicators and the inclusion of variables which proxy the effect of increased oil demand in China.

Adjusting the oil price forecast embedded into futures to take account of this time-varying risk premium yields "risk-adjusted" forecasts which perform extremely well in periods both of "bear" and of "bull" oil markets. More formally, with an out-of-the-sample prediction exercise we show that the forecast adjusted for a time-varying risk premium - linked to the US business cycle - performs significantly better than the unadjusted futures, the simple constant-adjusted futures and the random walk, particularly at horizons longer than 6 months.

Our results have crucial implications for policy analysis and economic modelling. First, they point out that futures should be used with caution in predicting oil prices which, in turn, affect inflation and output gap forecasts, the two variables that, according to modern economic theory, are crucial for monetary policy decisions. Second, our results show that futures-based forecasts of oil prices, adjusted for time-varying risk premia, may be exploited to identify unexpected oil price changes ("shocks"), which are often used in the context of dynamic macro analyses. We leave this topic for future research.

---

any general predictive information about market returns, that is net long positions do not generally lead market returns.

Appendix: Data Sources

WTI oil spot and futures prices: *Thomson Financial Datastream.*

Real-time indicators (capacity utilization, non-farm payrolls, industrial production): *Federal Reserve Bank of Philadelphia (www.phil.frb.org/econ/forecast/reaindex.html)*

Chinese industrial production: *Thomson Financial Datastream.*

Oil stocks: *U.S. Department of Energy, Energy Information Administration (www.eia.doe.gov/dnav/pet/pet_stoc_wstk_dcu_nus_w.htm)*

References

[1] Chernenko, Sergey V., Krista B. Schwarz and Jonathan H. Wright (2004) "The Information Content of Forward and Futures Prices: Market Expectations and the Price of Risk", *Board of Governors of the Federal Reserve System, International Finance Discussion Papers, No. 808.*

[2] Chinn, Menzie, Michael LeBlanc, and Olivier Coibion (2005) "The Predictive Content of Energy Futures: An update on Petroleum, Natural Gas, Heating Oil and Gasoline", *NBER working paper No. 11033.*

[3] Cochrane, John H. (2005), "Financial Markets and the Real Economy", *NBER working paper No. 11193.*

[4] Cochrane, John H and Monika Piazzesi. (2005), "Bond Risk Premia", *American Economic Review,* vol. 95, pp. 138-160.

[5] Coimbra, Carlos and Paulo Soares Esteves (2004), "Oil Price Assumptions in Macroeconomic Forecasts: Should We Follow Futures Market Expectations?", *OPEC Review,* vol. 28, pp. 87-106.

[6] Considine, Timothy J. and Donald F. Larson (2001) "Risk Premiums on Inventory Assets: The Case of Crude Oil and Natural Gas", *The Journal of Futures Markets*, vol. 21, pp. 109-126.

[7] Diebold, Francis X. and Roberto S. Mariano (1995), "Comparing predictive accuracy", *Journal of Business and Economic Statistics,* vol. 13, pp. 253-263.

[8] Fama, Eugene F. and Kenneth R. French (1989), "Business conditions and expected returns on stocks and bonds", *Journal of Financial Economics,* vol. 25, pp. 23-49.

[9] Finn, Mary G. (2000), "Perfect Competition and the Effects of Energy Price Increases on Economic Activity", *Journal of Money Credit and Banking,* vol. 32, pp. 400-416.

[10] Goldfeld, Stephen M. and Richard E. Quandt (1965) "Some Tests for the Homoscedasticity", *Journal of the American Statistical Association*, vol. 60, pp. 539-547.

[11] Gorton, Gary and K. Geert Rouwenhorst (2004), "Facts and Fantasies about Commodity Futures", *NBER working paper No. 10595.*

[12] Hansen, Bruce (1992), "Testing for parameter instability in linear models" *Journal of Policy Modeling,* vol. 14, pp. 517-533.

[13] Lynch, Michael C. (2002), "Causes of oil price volatility", paper presented at the Eight International Energy Forum, Osaka, Japan.

[14] Moosa, Imad A., and Nabeel E. Al.Loughani (1994) "Unbiasedness and Time Varying Risk Premia in the Crude Oil Futures Market". *Energy Economics*, vol. 16, pp. 99-105.

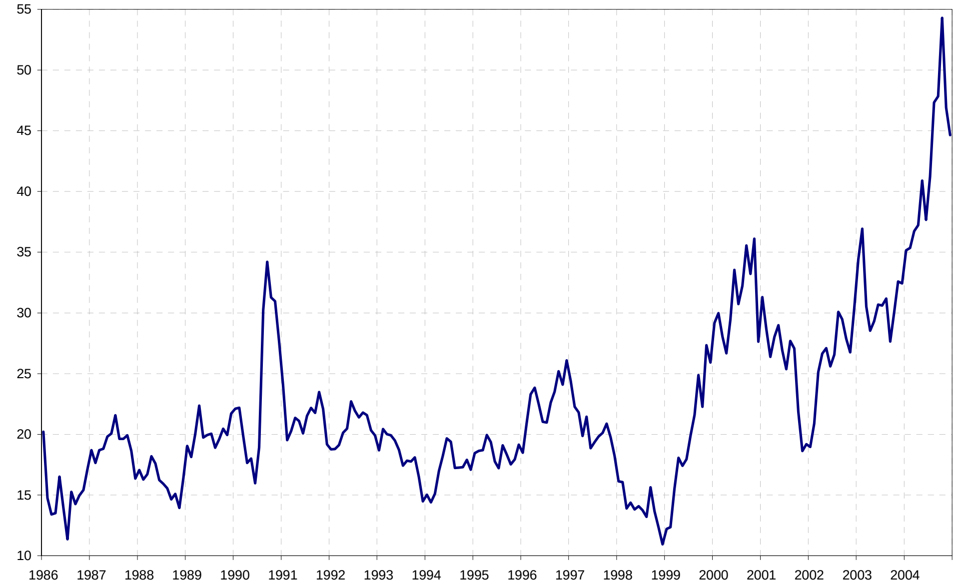
[15] Piazzesi, Monika, and Eric Swanson (2004) "Future Prices as Risk-Adjusted Forecasts of Monetary Policy", *NBER working paper No. 10547.*

[16] Pindyck, Robert (2001) "The dynamics of commodity spot and futures markets: a primer", *The Energy Journal*, vol. 22, pp. 1-29.

[17] Sanders, Dwight R., Keith Boris and Mark Manfredo (2004) "Hedgers, funds, and small speculators in the energy futures markets: an analysis of the CFTCs Commitments of Traders reports", *Energy Economics,* vol. 26, pp. 425-445

[18] Svensson, Lars E. O. (2005) "Oil prices and ECB Monetary Policy", *manuscript*, www.princeton.edu/~svensson/papers/ep501.pdf

Figure 1: WTI oil price



Notes: US dollars per barrel. Monthly observations. Each observation is the simple average daily spot prices during the third week of the month.

Figure 2: recursive residuals from regression (3) and 2-standard error bands at different horizons

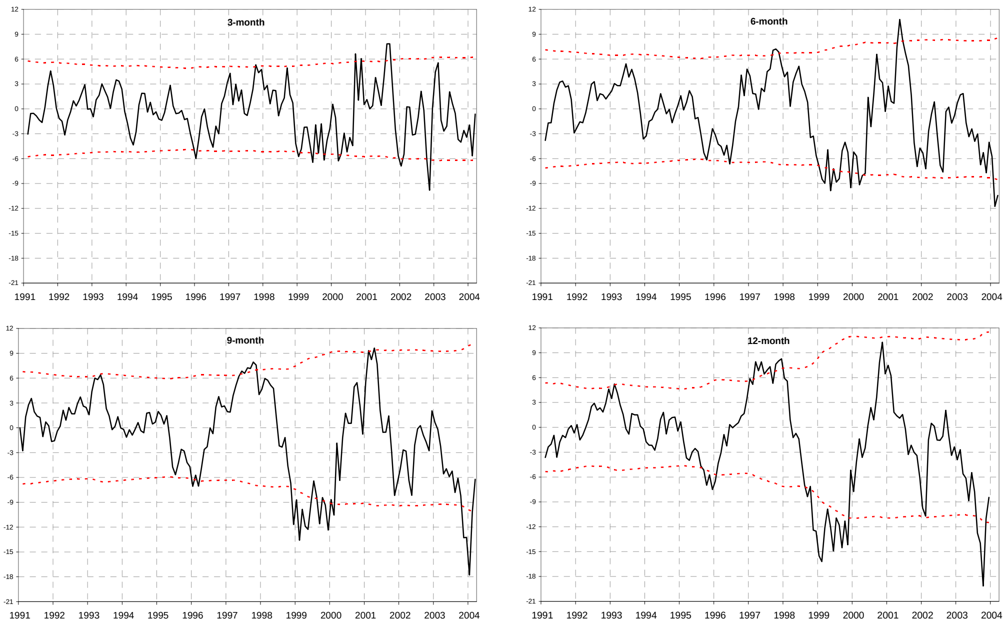
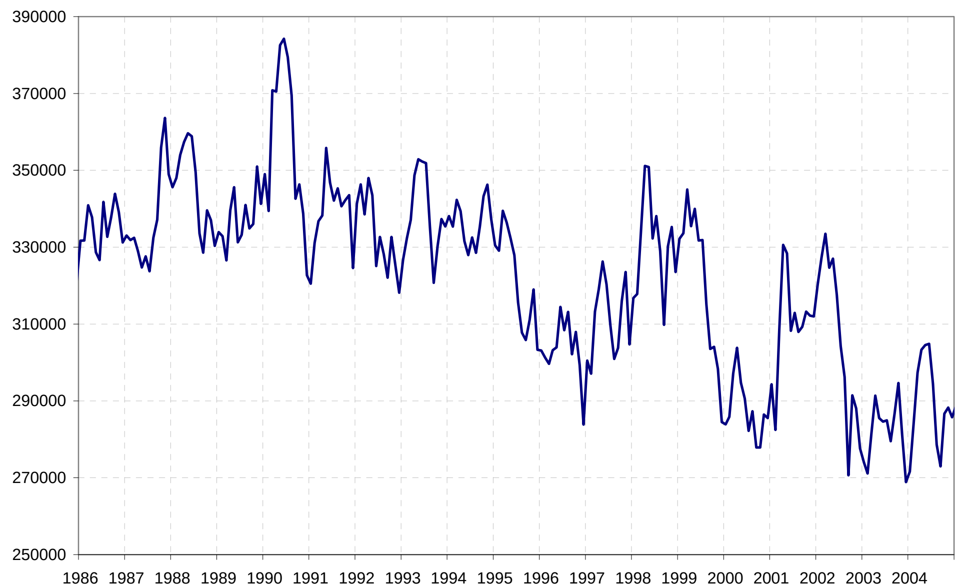
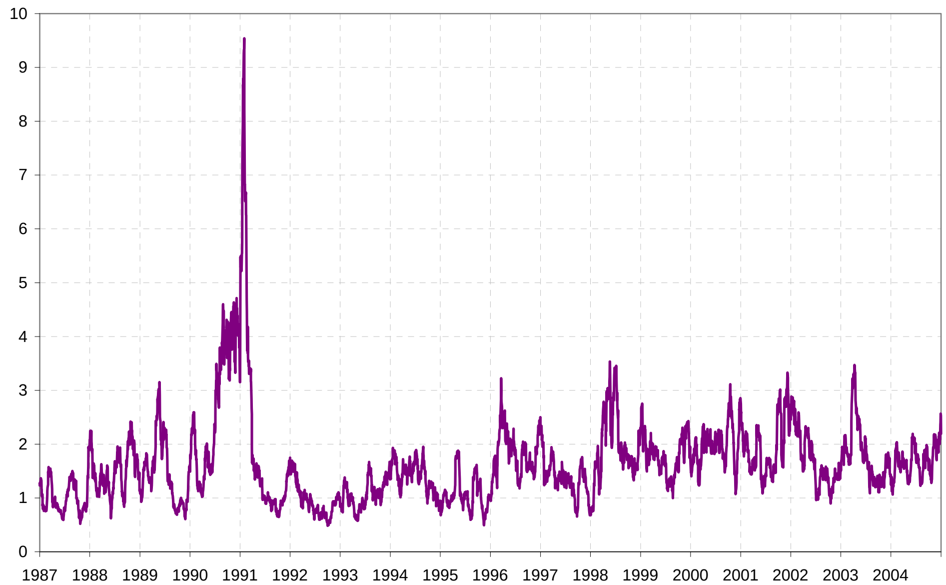


Figure 3: U.S. crude oil ending stocks



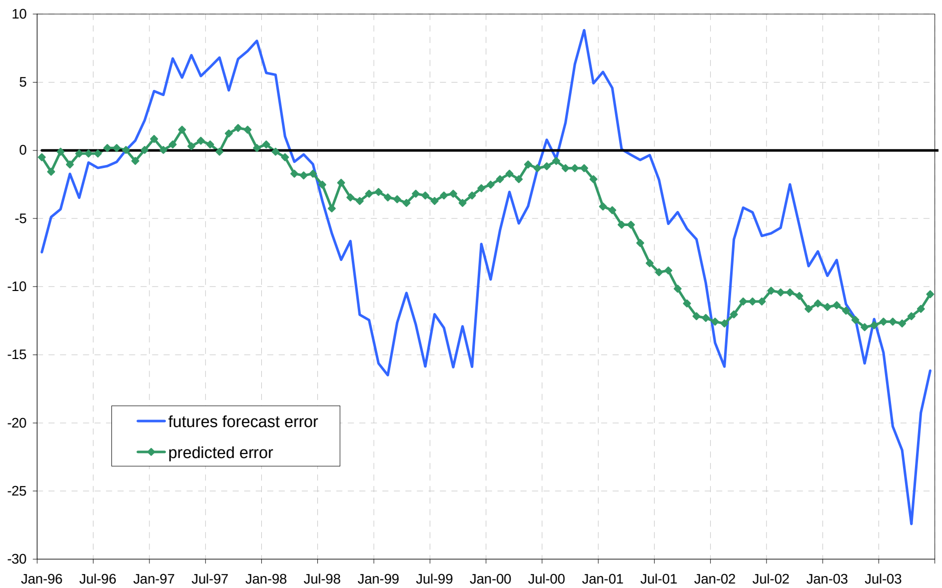
Notes: thousand barrels, excluding SPR.

Figure 4: WTI oil spot price volatility



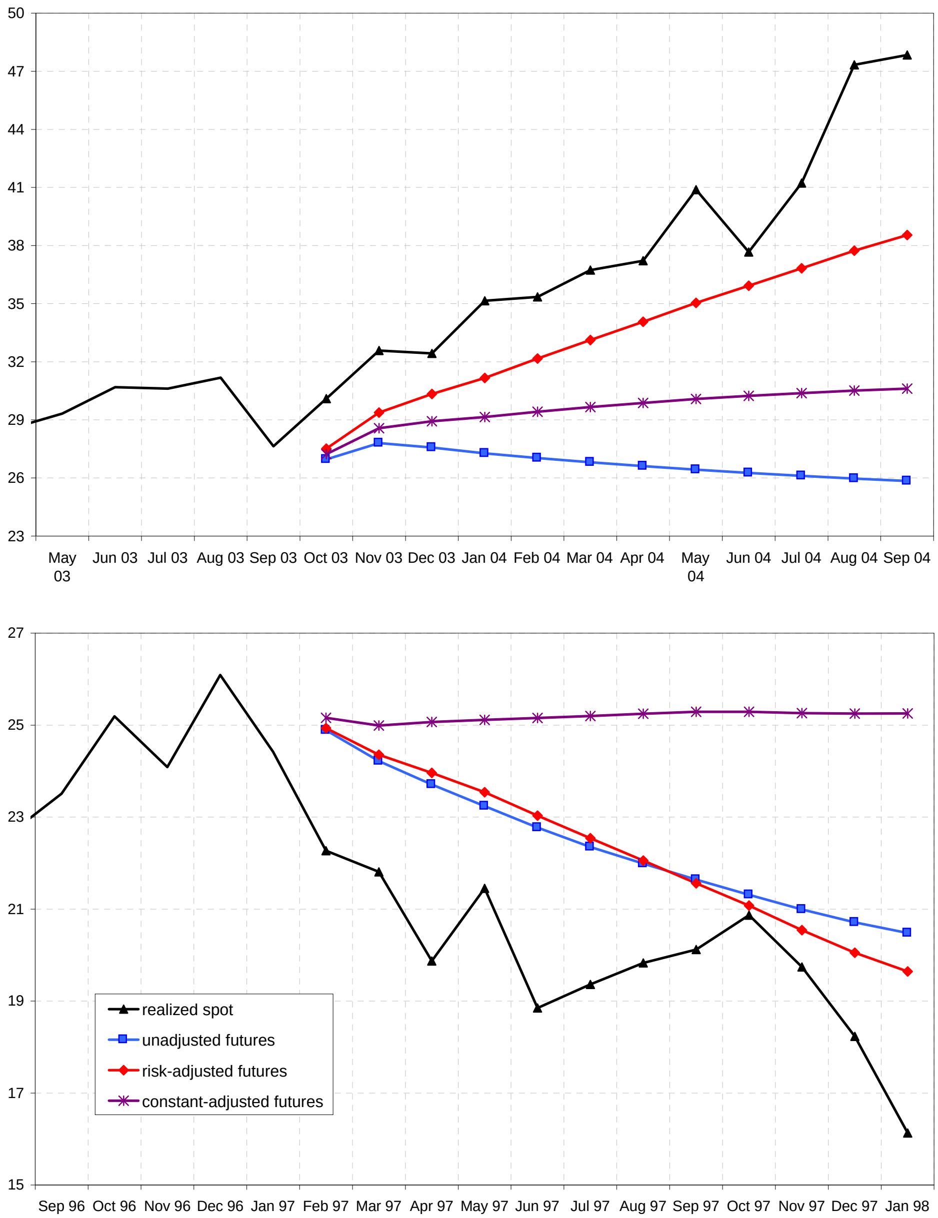
Notes: one-month moving average of absolute daily percentage changes.

Figure 5: oil price ex-post forecast errors and predicted errors (12-month-ahead)



Notes: US dollars per barrel. Monthly observations. The forecast error is defined as in equation (1) using the average of daily futures and spot prices during the third week of each month. Predicted errors are fitted values of equation (3).

Figure 6: oil price forecasts and realized spot prices on two dates



Notes: US dollars per barrel. Risk-adjusted forecasts are computed using estimated coefficients as in Table 6. Constant-adjusted forecasts are computed from the estimated coefficient of equation (2) over the period starting in January 1996.

Table 1: Constant risk premia (whole sample 1986:1-2004:12)

n	constant	
1	-.15	(.15)
2	-.45*	(.25)
3	-.84**	(.35)
4	-1.18***	(.45)
5	-1.48***	(.55)
6	-1.76***	(.65)
7	-2.01***	(.74)
8	-2.24***	(.85)
9	-2.45***	(.94)
10	-2.69**	(1.11)
11	-2.98**	(1.29)
12	-3.14**	(1.38)

Notes: estimation by OLS. Newey-West HAC standard error in parentheses.* denotes significance at 10 per cent; ** denotes significance at 5 per cent; *** denotes significance at 1 per cent.

Table 2: Time-varying risk premia and capacity utilization (whole sample 1986:1-2004:12)

n	constant		capacity$_{t-1}$	
1	-1.54	(3.91)	0.02	(0.05)
2	-6.77	(6.51)	0.08	(0.08)
3	-12.46	(8.35)	0.14	(0.10)
4	-18.25*	(10.09)	0.21*	(0.12)
5	-25.14**	(11.95)	0.29**	(0.15)
6	-32.09**	(13.71)	0.38**	(0.17)
7	-38.41**	(15.49)	0.45**	(0.19)
8	-44.50**	(17.50)	0.53**	(0.22)
9	-49.65**	(19.35)	0.59**	(0.24)
10	-53.63**	(21.51)	0.64**	(0.26)
11	-60.40**	(23.79)	0.72**	(0.29)
12	-64.23**	(25.88)	0.76**	(0.32)

Notes: estimation by OLS. Newey-West HAC standard error in parentheses.* denotes significance at 10 per cent; ** denotes significance at 5 per cent; *** denotes significance at 1 per cent.

Table 3: Correlation between changes in oil price and changes in capacity utilization: break in December 1995

n	1986-1995		1996-2004	
1	0.11	(0.22)	0.27	(0.00)
2	0.13	(0.15)	0.39	(0.00)
3	0.14	(0.13)	0.50	(0.00)
4	0.13	(0.18)	0.56	(0.00)
5	0.11	(0.26)	0.59	(0.00)
6	0.07	(0.44)	0.63	(0.00)
7	0.03	(0.78)	0.66	(0.00)
8	-0.02	(0.80)	0.68	(0.00)
9	-0.07	(0.46)	0.70	(0.00)
10	-0.11	(0.27)	0.70	(0.00)
11	-0.14	(0.13)	0.71	(0.00)
12	-0.16	(0.10)	0.71	(0.00)

Notes: Pair-wise correlation coefficients between the n-period change in oil price and the n-period change in capacity utilization in US manufacturing; p-values of the t-test that the correlation coefficient is equal to zero in parentheses.

Table 4: Tests of parameter stability: 1986-1995 vs. 1996-2004

n	constant	capacity$_{t-1}$	joint test
1	0.16	0.15	0.30
2	0.06	0.06	0.10
3	0.02	0.02	0.02
4	0.01	0.01	0.01
5	0.00	0.00	0.01
6	0.00	0.00	0.00
7	0.00	0.00	0.00
8	0.00	0.00	0.00
9	0.00	0.00	0.00
10	0.00	0.00	0.00
11	0.00	0.00	0.00
12	0.00	0.00	0.00

Notes: p-values of the F-test that the parameters are the same across the two sub-periods (1986-1995 and 1996-2004); based on Newey-West HAC standard errors.

Table 5: Variance stability tests

	Goldfeld-Quandt	Hansen		
n	1986-1995 vs 1996-2004	1986-2004	1986-1995	1996-2004
1	0.00	0.00	0.19	0.00
2	0.03	0.06	0.16	0.00
3	0.06	0.08	0.09	0.04
4	0.08	0.07	0.04	0.07
5	0.04	0.03	0.03	0.09
6	0.00	0.00	0.03	0.13
7	0.00	0.00	0.03	0.12
8	0.00	0.00	0.04	0.11
9	0.00	0.00	0.05	0.10
10	0.00	0.00	0.06	0.08
11	0.00	0.00	0.02	0.07
12	0.00	0.01	0.00	0.05

Notes: p-values. Null hypothesis of the Goldfeld-Quandt test is that the variance is stable across sub-periods. Null hypothesis of the Hansen test is that the variance is stable in the period considered.

Table 6: Time-varying risk premia and capacity utilization (sample 1996:1-2004:12)

n	constant		capacity$_{t-1}$		R^2
1	-4.19	(5.28)	.05	(.07)	.00
2	-11.95	(8.39)	.14	(.10)	.02
3	-20.85*	(10.85)	.25*	(.14)	.05
4	-29.73**	(13.18)	.35**	(.17)	.08
5	-40.18**	(15.76)	.48**	(.20)	.12
6	-50.39***	(18.25)	.61**	(.23)	.15
7	-60.66***	(20.29)	.73***	(.26)	.19
8	-71.13***	(22.47)	.86***	(.29)	.22
9	-80.90***	(24.28)	.98***	(.31)	.26
10	-91.06***	(25.76)	1.10***	(.33)	.31
11	-101.17***	(27.26)	1.23***	(.35)	.35
12	-110.18***	(28.50)	1.34***	(.36)	.38

Notes: estimation by OLS. Newey-West HAC standard error in parentheses.* denotes significance at 10 per cent; ** denotes significance at 5 per cent; *** denotes significance at 1 per cent.

Table 7: Time-varying risk premia and employment growth (sample 1996:1-2004:12)

n	constant		emp. growth$_{t-1}$		R^2
1	-.42	(.35)	12.15	(17.73)	.00
2	-1.24**	(.58)	37.65	(28.68)	.02
3	-2.09***	(.71)	58.91	(37.20)	.03
4	-2.92***	(.87)	84.58*	(46.75)	.05
5	-3.79***	(1.09)	112.15*	(58.78)	.07
6	-4.58***	(1.30)	139.71*	(70.95)	.09
7	-5.36***	(1.50)	168.95**	(84.21)	.11
8	-6.12***	(1.69)	197.51**	(96.68)	.14
9	-6.71***	(1.84)	217.48**	(109.04)	.15
10	-7.29***	(2.00)	238.20*	(121.28)	.17
11	-7.77***	(2.12)	252.20*	(131.68)	.17
12	-8.20***	(2.22)	264.84*	(141.15)	.18

Notes: estimation by OLS. Newey-West HAC standard error in parentheses.* denotes significance at 10 per cent; ** denotes significance at 5 per cent; *** denotes significance at 1 per cent.

Table 8: Time-varying risk premia and Chinese industrial production (sample 1996:1-2004:12)

n	constant		capacity$_{t-1}$		Chinese ind. prod.$_{t-2}$	
1	-3.77	(5.37)	.05	(.07)	-1.73	(6.05)
2	-10.28	(8.97)	.13	(.11)	-6.91	(9.45)
3	-17.76	(11.55)	.23	(.14)	-13.00	(12.70)
4	-26.06*	(13.60)	.33**	(.17)	-15.64	(16.13)
5	-37.03**	(15.51)	.46**	(.20)	-13.70	(16.51)
6	-48.01***	(17.33)	.59**	(.22)	-10.63	(19.38)
7	-60.37***	(20.07)	.73***	(.26)	-1.37	(19.72)
8	-69.81***	(21.31)	.85***	(.28)	-6.51	(23.91)
9	-79.65***	(22.92)	.97***	(.30)	-6.92	(28.43)
10	-89.57***	(24.50)	1.10***	(.32)	-7.74	(30.28)
11	-99.49***	(26.51)	1.22***	(.34)	-9.32	(31.54)
12	-110.79***	(28.66)	1.34***	(.37)	3.68	(31.24)

Notes: estimation by OLS. Newey-West HAC standard error in parentheses.* denotes significance at 10 per cent; ** denotes significance at 5 per cent; *** denotes significance at 1 per cent.

Table 9: Forecasts of oil price (sample 1996:1-2004:12): expanding window

	benchmark			futures based								
	random walk			unadjusted			constant-adjusted			risk-adjusted		
n	ME	SE	ρ_n	ME	SE	ρ_n	ME	SE	ρ_n	ME	SE	ρ_n
1	-.40	2.90	-.14	-.43	2.96	-.13	-.27	2.97	-.13	.13	3.04	-.10
2	-.85	3.87	-.16	-1.14	3.85	-.15	-.73	3.76	-.16	-.11	3.85	-.10
3	-1.38	4.56	-.10	-1.99	4.60	-.04	-1.32	4.35	-.04	-.45	4.41	.00
4	-1.86	5.14	.00	-2.79	5.38	.09	-1.89	4.94	.09	-.65	4.93	.11
5	-2.36	5.64	.06	-3.61	6.15	.14	-2.47	5.51	.16	-.69	5.37	.13
6	-2.81	6.35	.09	-4.37	7.00	.15	-2.98	6.20	.18	-.63	5.89	.11
7	-3.21	6.95	.06	-5.06	7.77	.13	-3.45	6.82	.17	-.41	6.29	.01
8	-3.67	7.57	.01	-5.76	8.46	.09	-3.94	7.36	.13	-.09	6.61	.07
9	-4.03	8.04	-.04	-6.34	9.04	.05	-4.35	7.81	.09	.29	6.85	.05
10	-4.37	8.56	-.13	-6.86	9.58	-.01	-4.74	8.23	.02	.77	7.04	.03
11	-4.70	9.14	-.16	-7.36	10.09	-.08	-5.08	8.65	-.04	1.34	7.22	.02
12	-5.03	9.74	-.21	-7.84	10.61	-.14	-5.46	9.09	-.11	1.70	7.42	.01

Notes: n is the forecasting horizon. ME is the mean error (in US dollars), SE is the root-mean-squared error (in US dollars) and ρ_n is the nth autocorrelation of the forecast error.

Table 10: Forecasts of oil price (sample 1996:1-2004:12): moving window

	benchmark		futures based				
	random walk		unadjusted		constant adj.		risk adj.
n	SE	D-M	SE	D-M	SE	D-M	SE
1	2.90	0.97	2.96	0.97	3.02	0.85	3.18
2	3.88	0.79	3.85	0.88	3.88	0.56	4.07
3	4.56	0.72	4.60	0.76	4.50	0.47	4.63
4	5.14	0.76	5.38	0.63	5.12	0.40	5.10
5	5.64	0.69	6.15	0.38	5.71	0.25	5.46
6	6.36	0.41	7.00	0.14	6.43	0.08	6.03
7	6.95	0.27	7.77	0.07	7.08	0.03	6.48
8	7.57	0.10	8.46	0.02	7.61	0.01	6.71
9	8.04	0.04	9.04	0.01	8.06	0.00	6.87
10	8.56	0.01	9.58	0.00	8.45	0.00	6.98
11	9.14	0.00	10.09	0.00	8.83	0.00	6.94
12	9.74	0.00	10.61	0.00	9.21	0.00	6.96

Notes: n is the forecasting horizon. SE is the root mean squared error (in US dollars), D-M is the P-value of the Diebold-Mariano test: the null hypothesis is no difference in the forecasting precision between the model considered and the cyclically adjusted future; the alternative hypothesis is that the cyclically adjusted future produces better forecasts. The risk adjusted root mean square error is obtained by running regression (3) on a moving window of 30 monthly observations.