Physical Market Determinants of the Price of Crude Oil and the Market Premium

Guillaume Chevillon * Christine Rifflart
ESSEC Business School Observatoire Français des Conjonctures Economiques,
and CREST-INSEE. Fondation Nationale des Sciences Politiques

December 2008

Abstract

We analyze the determinants of the real price of crude oil by means of an equilibrium correction model over the last two decades where we focus on the aspects of the physical market that impact on the clearing price. We find that two cointegrating relations affect the changes in prices: one refers to OPEC’s behavior, attempting to control prices using its market power and quotas; the other to the coverage rate of OECD expected future demand using inventory behaviors. We derive a forecasting equation for the change in oil prices which we use to assess the speculative elements of the price increases of the period 2000-05. We show that worries alien to the physical markets were the causes of the increase in oil prices and we quantify their overall impact.

*Corresponding author. E-mail: chevillon@essec.fr; address: Decision Sciences Department, ESSEC Business School, Avenue Bernard Hirsch, 95021 Cergy-Pontoise cedex, France. This author gratefully acknowledges research support from the Department of Economics, University of Oxford and OFCE. We thank the editor and two anonymous referees for their comments and suggestions.
1 Introduction

From the inception of the 2003 military intervention of allied forces in Iraq, also known as the second Gulf war, until mid-2008, the price of crude oil has witnessed an increase in volatility and an upward trending pattern. Some experts and the media have been led to forecasting a probable end of oil resources in order to explain this situation (Vidal, 2005). This has induced many countries (the U.S. being foremost: see Bezdek, 2007) to develop alternative renewable sources of energy. Yet, it is still unclear how to measure the premium on the price of crude oil that is put by anxious market participants. This paper is a contribution towards this objective.

The standard approach to estimating speculative elements would consist in focusing on pricing from the perspective of volatility in futures markets. Without a baseline model for the future cost of extraction and an estimate of reserves that are yet to be proven, such a methodology might only lead to estimating speculators’s speculations. We propose a different vision: analyze, by means of an integrated-cointegrated (so called equilibrium-, or error-, correction) model, the behavior of the physical determinants of real oil prices (i.e. related to supply and demand for quantities of oil; as Merino and Ortiz (2005), we define them as resulting from the actions of agents involved in production, processing or merchandising of the commodity) and derive their implications for price variations. Any systematic departure from historical patterns is then innovative and can be called a market premium. Using the latter expression, we depart from authors such as Merino and Ortiz (2005) who refer to the seemingly unexplainable deviation between observed and modeled prices as a ‘price premium’. We think that our choice of term allows for a different evaluation of risk and uncertainty by market participants, but their valuation thereof may potentially be correct (we do not take a position on this issue). Hence we do not assume that there is a premium in prices per se, but differences in the perception of market conditions relative to historical standards.

Proponents of the estimation of market premia on oil prices using financial instruments include, inter alia, Barros Luís (2001) who develops a risk aversion measure which is used to assess the risk premium implicit in option future contracts. Also, Sadorsky (2002) considers the opportunity to hold oil options when maximizing a portfolio’s return and uses an ARMAX-ARCH model to forecast that return. As for volatility estimation, Kuper (2002) uses a GARCH framework for estimating markets’
historical uncertainty via price ranges. Directly focusing on futures prices, Fong and See (2002) also estimate a GARCH model but with added Markov regime switching: estimated over 1992-97, it fits the data well and also performs advantageously in out-of-sample forecasting. Working on a sample ranging from 1988 to 2003, Sadorsky (2006) also compares models of conditional volatility in futures markets and shows that univariate GARCH are preferable to other univariate or multivariate specifications, but that when evaluating Value-at-Risk, non-parametric estimation performs well.

Our decision to focus in this paper on linear models turns out quite rare in this line of research. Most of the existent literature that uses linear techniques tends to estimate the interaction between oil price and economic activity in oil importing or exporting countries. For instance, Hamilton (2000) uses the output gap as a basis for assessing the disruptive effects of oil price hikes on agents’ behaviors. Amano and van Norden (1998) prove the existence of a strong relation between real oil prices, exchange rates and terms of trade in major oil importing economies. Also important is our wish to focus on determinants of market prices, not on economic elements presiding over the physical supply side, namely the level of proven reserves, itself function of wells where extraction is economically sound, which in turns depends on the price of oil. The reader is referred e.g. to Farzin (2000) for such an approach.

Surprisingly few authors have focused on crude oil price determination and forecasting. Verleger (1982) was probably the first, but in his time, he could only focus on the link between official and spot (Rotterdam) prices. Among recent papers, Tang and Hammoudeh (2002) use a model whereby OPEC attempts to control prices within a target range: they establish the presence of a non-linear relation between production quotas, inventory levels (or stocks), supply and demand. Unfortunately, their model brings no forecasting power. To our knowledge, Krichene (2002) and Yang et al. (2002) are among the very few who applied an equilibrium correction model (ECM) on oil demand and supply (they focus on the U.S. market). The analysis by Krichene focuses on the whole of the 20th Century: the author derives price elasticities which is unfortunately of no use to our analysis restricted to the last two decades. Krichene (2006b) also considers the impact of monetary policy on demand for oil. As for Yang et al., they first estimate historical volatilities in the U.S. which, in turn, they add to equilibrium correction models for energy commodities such as coal, natural gas and crude oil. This allows them to derive a model where they allow for a switching
price elasticity of demand. More recently Dees et al. (2007) and Kaufmann et al. (2008) also estimate ECM models for either supply and demand or price formation with additional nonlinearities accounting for the post 2002 dynamics that Askari and Krichene (2008) have, inter alia, identified.

In the forecasting literature, Yu et al. (2008) assess neural network models for spot prices. Also, Ye et al. (2002, 2005) derive an easily implementable short-term forecasting model using seasonal effects and inventories. The role of inventories on oil price determination had been stressed by Pindyck (1994, 2001) and assessed in Merino and Ortiz (2005). The latter authors show that the historical connection between stocks and prices appears recently to have altered. In our choice of variables and methodology, we are closest to Merino and Ortiz (2005) and two other related econometric papers: the first, by Kaufmann et al. (2004), assesses the pricing role played by OPEC policies; the second, Dees et al. (2007), mostly focuses on supply and demand equations, although the authors also estimate a pricing model.

Here, we model the historical formation of real oil prices conditional on supply and demand. This allows us to measure a market premium as the deviation of nominal spot prices from our forecasts. The plan of this paper is as follows. We first present the main historical features of crude oil markets in section 2. They, together with economic model for commodities, preside over our choice of variables for our pricing equation; this constitute the purpose of section 3. In section 4 we conduct VAR and cointegration analyses. We assess there the validity of the weak exogeneity assumptions that allows us to focus, in section 5, on a univariate model. We also derive a market premium over the physical clearing price. Finally, section 6 concludes our analysis.

2 Global oil market features

Over the last two decades, crude oil prices have been determined by market mechanisms, yet this is in an imperfectly competitive market. As with any tradable goods, the price of oil follows the relative evolutions of supply and demand, decreasing when the former abounds, increasing when the market situation tightens. However, oil is no standard commodity: it is marred by political concerns in an oligopolistic market dominated by a strong cartel, the Organization of the Petroleum Exporting Countries

\footnote{Regnier (2007) shows that crude oil prices are more volatile than those of most commodities.}
Worldwide demand still originates chiefly from OECD economies: according to the U.S. Department of Energy, these represented, in 2006, 58% of world consumption (see EIA, 2007). But since the mid-1990s, emerging economies, especially China and India, have seen their consumptions surge. In 2004, China became the second largest consumer worldwide: behind the United States but before Japan. This rapid growth may seem the driving force behind the recent rise in the world’s energy demand, with an increase in demand from fast-developing Asian economies amounting to two thirds of the global increase in energy consumption.

Short term fluctuations in oil consumption—beyond mere seasonal activity—depend inter alia on the intensity of petroleum derivatives in the GDP and on the rate of growth of the latter (see Dees et al., 2007). Price evolutions matter very little here: short-run price elasticity of demand for crude oil is near zero as there are no substitutes to its use that are readily available. Fattouh (2007a, table 1) provides a meta analysis of the short- and long-run price elasticities of oil demand, reporting estimates from seven different papers: the conclusion is that in the short term, demand hardly responds to price variations (elasticities are around -0.05), but this is not true for long-term fluctuations as elasticities range from about -0.6 in the OECD to -0.1 to -0.4 in emerging economies. Industrial economies exhibit a price elasticity with a larger absolute value because, following the two oil shocks in the 1970s, they have developed policies aimed at reducing dependence on energy and diversification away from oil. Indeed, Rogoff (2006, table 4c) reports estimates showing a decrease in the impact of oil prices on GDP in G7 economies; this is in line with his argument that IMF-reported global oil intensity of GDP is in the 2000s 38% lower than in the 1970s. As Rogoff discusses, the 1986 decrease in prices softened the pressure and reduced the effort. Nowadays, according to most estimates the price elasticity of long term world demand is around -0.5, provided that long term has a meaning for a non-renewable source of energy which could deplete rapidly.

On the supply side, the oil sector is still oligopolistic: it is organized around a few global private companies and the OPEC cartel. Although it controls only 40% of production worldwide, OPEC has until recently played a crucial role on price formation owing to its policy for regulating supply. Available reserves and spare productive capacities are chiefly located in the Middle Eastern OPEC countries. Originally created, in 1960, to defend the interests of its members, the cartel has progressively
modified its objectives: moving away from a policy of sustained high prices in the
1970s and 80s to a more consensual policy taking consumer countries into account,
see Fattouh (2007c) for an historical account and contemporary perspectives. Any
OPEC decision to increase or reduce production quotas may hence lower or raise the
price of crude oil. The impact OPEC has had over prices has evolved through time as
shown by Barsky and Kilian (2004), Kaufmann et al. (2004) and Slaibi et al. (2006).
The latter authors discuss in particular how inconsistencies in OPEC’s strategies have
played a role in crude oil price variations. We will return to this below.

Following a sharp reduction in market prices for crude oil in 1986, a new wave of
producers has progressively emerged, in particular in Former Soviet Union countries
(FSU). Until the mid-1990s, OPEC had succeeded in controlling—although some-
what loosely—world prices of crude oil via managing an important fraction of world
production, see Kaufmann et al. (2004) for an econometric analysis of their power.
This has let the price of a barrel of North Sea Brent stable at around US$ 18-19 until
1997, a price then regarded as mostly fair by both OPEC and consumers. Yet, follow-
ing the Asian crisis and the subsequent fall in demand, oil prices collapsed to about
US$ 10 per barrel by the end of 1998 (see Fattouh, 2007a). Hurt by this experience,
OPEC decided to increase transparency and regained power in 1999-2000 as Barsky
and Kilian (2004) discuss. It progressively set up a target band of US$ 22 to 28 a
barrel, based on the OPEC basket of produced crude oil (see Slaibi et al., 2006, table
1). In parallel, quota policy became more reactive so as to follow the variations in
global demand and prevent any excess supply or lack thereof. A group of producer
countries, including Norway, Mexico and Russia, started cooperating with OPEC in
its attempt to control production. Since 2004, however, OPEC seems to have lost
its capacity to control prices: we discuss in section 5.4 how their evolution has been
disconnected from those of relative supply and demand. This has taken place in a
period of geopolitical tensions in the Middle East and global concerns for the long
term production capacities. Yet, oil remains a highly strategic commodity as it is the
main source of energy that sustains economic growth. OPEC owned, at the end of
2007, over 78% of the world’s proven oil reserves.²

What are the main features affecting prices? OPEC’s behavior is crucial since
it represents about 55% of global oil exports according to EIA (2007). Since pro-
duction in this industry is very capital intensive, supply is rigid in the short-term:

²According to OPEC, see http://www.opec.org/library/FAQs/aboutOPEC/q15.htm
non OPEC producers tend to work at full capacity and cannot adjust to take advantage of higher prices. Consumption is affected by seasonal factors, especially in the Northern Hemisphere where winter are cold and summers hot. Hence, consumption of oil derivatives increases in the first and fourth quarters of each year (the United States also consumes a lot of energy in summer for air conditioning purposes). This is anticipated by building refinery inventories of crude oil and petroleum products in the third and fourth quarters (for an analysis of inventories, see Pindyck, 1994 and 2001). Speculations and unusual weather conditions hence generate ample short term volatility.

3 Determinants of the price of crude oil

To derive an econometric model for the evolution of oil prices, we need to understand what factors may intervene. Oil is a commodity which is traded in an imperfectly competitive market, with few suppliers and many consumers. This oligopolistic structure implies that the price does not simply react to demand and supply as with many exhaustible commodities. Indeed, according to Hotelling (1931), the price—net of the constant marginal extraction cost—in a perfectly competitive market should be in line with the interest rate, i.e. the discount rate that stabilizes the value of future receipts so that the sacrifice of future generations who will not consume this resource is taken into account. There is a number of issues that can lead to departure from Hotelling’s rule; most of them (inter alia externalities, rising marginal extraction costs, risk aversion, imperfect information) are present in oil markets which are far from perfect. Analyzing the latter was the object of significant interest in the 1970s: see for instance Ulph and Folie (1980) and references therein. Fattouh (2007a) presents a review of recent structural analyses.

There exist many theoretical models for the determination of world supply, demand and prices of oil, many of which originated from Nordhaus (1973) (see Choukri (1979) for an early review of the pre-1979 shock theories). Following Griffin (1985) and the recent evidence in Ramcharran (2001, 2002), we believe in a model consisting of a mix of cartel and competitive market practices for the level of production.

---

3See the negative seasonal dummies for the second and third quarters in table 6.b of Dees et al. (2007). The U.S. demand for oil is also large in the third quarter, this reflects hotter summers than in European OECD countries.
Given that models of target revenue or property rights could be rejected on empirical grounds by Griffin and that non-OPEC producers have increased their significance since the mid-1980s, we do not consider these types of models. In contrast, we use a demand and supply model for commodities, in the line of Bacon (1991), Dees et al. (2007) and Krichene (2007), but where, contrary to these authors, we focus not on the determinants of supply and demand, but solely on the physical market price of crude oil. This amounts to conditional modeling of the short-term fluctuations. The effect of prices on demand and supply appears through an equilibrium correction model which hence combines both shorter- and longer-run effects. The rationale for their interaction stems from Hubbard (1986) who shows how intertwined they must be. In line with Pindyck (1994) and Pindyck (2001), we also consider the role of inventories. Distortions to market behavior caused by OPEC and geopolitical considerations, as studied by Fattouh (2007c), Kaufmann et al. (2004) and Slaibi et al. (2006) lead us to model the implicit target price of OPEC and its quota policy and to include dummy variables for external events such as the first and second Iraq wars. As noted by Geroski et al. (1987), we also need to take account of the changes in OPEC pricing conduct. We therefore decided to restrict our analysis to the post-1987 era since the market experienced major changes at the time; in particular, market forces began to gain significance. This implies that our analysis will be restricted to a small sample of 68 quarterly observations ranging from 1989(1) to 2005(4). Consequently, we must adopt a parsimonious representation and we focus on the variables that necessarily impact the price.

In the simplest commodity model, clearing of the physical market would normally lead to a real price \( P_t \) being set so that demand equals supply. Unfortunately, the market for crude oil is imperfect. Global supply originates from non-OPEC countries, which produce at full capacity at all prices, and from OPEC, where production is regulated to maintain the price of a basket of various types of crude oil within a target band. Indeed, almost all spare production capacities are located in the cartel’s countries and they—Saudi Arabia in particular—are often regarded as the supplier of last resort.

We hence model the supply side of the market via the production quotas of OPEC and the short run fluctuations that inventories allow. We denote crude inventory levels in the OECD by \( S_t^{OECD} \) and outside the OECD, by \( S_t^{non\ OECD} \). The benefit in considering these variables rather than production lies in the correlation between
supply and demand: here we are in effect orthogonalizing the specific effect of supply with respect to demand, so that multicollinearity is reduced in our regressions.

Reflecting the differences in energy intensity of GDP between industrial and emerging economies, we separate the demand side into those originating from OECD, $D_t^{OECD}$, and from non-OECD, $D_t^{non\ OECD}$, countries. Demand originating from the OECD has historically been the driving force behind market evolutions but non OECD economies have recently been consuming increasingly. Quarterly production and demand series were obtained from the International Energy Agency. We do not use strategic reserves as these are often very difficult to measure as governments are loath to report them.

We now review the other variables that we include in the model, see also table 2 for definitions and data sources. In the rest of the paper, a lower case denotes the natural logarithm of the variable in upper case.

3.1 What price?

In our attempt to relate the quarterly price of crude oil to the situation of the physical market, we need to choose the relevant variables. First, what price should we analyze? There exist several references in world exchanges: North Sea Brent, West Texas Intermediate (WTI) and Dubai. Each refers to an oil of high quality with a specific production or trading location. The WTI and Dubai prices are chiefly traded in the United States and Asia, whereas the North Sea Brent is often used as the world reference. In the London-based ICE Futures exchange (formerly known as the International Petroleum Exchange, or IPE), the Brent is used to specify the price of two thirds of crude oil exchanged worldwide. In this paper, we therefore focus on the spot price of a barrel of Brent, i.e. the price negotiated for immediate delivery on the physical market. Although the different prices do not fluctuate perfectly in line over time, they are strongly correlated and modeling one or the other should not impact our analysis: Fattouh (2007b) compares their dynamics and concludes that most varieties cointegrate in threshold models. The Brent-WTI price differential is

---

4 The statistics of the International Energy Agency for 2006, see IEA (2006), indicate that energy intensities of GDP (in tons of oil equivalent per million constant 2000 US$) were 265, 292 and 393 in high, middle and low income countries respectively.

5 This was discussed by N. Tanaka, Executive Director of the International Energy Agency at the Ministerial Session of the 9th International Oil Summit in Paris on April 10, 2008.
stationary with no need for thresholds.

In this paper, we model the price for oil conditional on the levels of demand and supply. Hence we do not model how higher prices may render more expensive drilling locations profitable. Hence, in line with a measure of supply and demand (they are expressed in millions of barrels per day, averaged over a quarter), we must transform the price from its nominal to real value. Because crude oil imports are not final goods and constitute an input to an economy’s production processes, it is natural to deflate the nominal price of crude oil with respect to the price of production goods in importing economies: we use the GDP deflator for the G7 group of countries. The resulting real price \( P_t \) is the ratio of the average quarterly nominal spot price of a barrel of Brent over the chosen GDP deflator. It is therefore related to the need for oil in the production process of industrialized economies. It may be questioned whether this choice of deflator is appropriate. It is indeed common to attribute the price increases of the new millennium to demand originating from fast-developing economies. As these nations experience higher rates of inflation than observed in the OECD, it might seem reasonable to use their deflators also. Two issues then arise. One concerns the availability and reliability of the necessary data: it is notoriously difficult to measure quarterly GDP deflators of countries developing so fast as China, see e.g. The Economist (2008) for misreporting in GDP figures, and even more so, to obtain accurate historical data. The second is about their relevance. We choose to use the deflator of industrialized economies since it allows to consider both demand for oil originating from their industries and the demand for the goods they produce coming from oil exporting economies. This choice of deflator, with a lower rate of increase, should over-emphasize the impact that demand for oil in non-OECD economies has on the increase in oil prices.

---

6Ideally, we would have used the deflator for all OECD economies, i.e. the main consumers of oil. Owing to the lack of data for recent periods, we used the deflator for the G7. The latter closely follows the OECD deflator for the available periods.

7Underestimating the rate of growth of the GDP deflator of emerging economies implies that changes in the real price of oil that they face is overestimated since the real price is the ratio of nominal price over deflator. Hence the parameter that links changes in real oil price (\( \Delta p \)) to the corresponding change in demand (\( \Delta d \)), i.e. \( \partial p/\partial d \), is underestimated for non-OECD countries, but not for those belonging to the OECD.
3.2 OPEC behavior

To analyze OPEC behavior, we resort to two distinct variables: one is their official production quota ($Q_t$), the other is their nominal price target. In contrast to Kaufmann et al. (2004) and Dees et al. (2007) we do not explicitly consider the wedge between OPEC production and quotas (cheating from member states): it is included in inventory variations and quotas are used a policy variable.

Until 2000, the nominal price target was implicit, see Tang and Hammoudeh (2002). It was established to achieve domestic social needs, allow economic development and prevent the expansion of other oil producers. Because the marginal cost of production is lower for OPEC than for other producers, they possess a loose capacity to gear prices downwards. At the beginning of 2000, they decided to announce a target price at the end of their meetings and they delineated a rule which would play automatically on production quotas, although the latter has never been put in place. Reference prices are based on a basket of seven types of crude oil that member countries produce. This basket price was at the time expected to fluctuate between US$ 22 and 28 per barrel: see Slaibi et al. (2006). The nominal target price that we use is therefore set to US$ 18 prior to March 2000, shifting then to US$ 25 until the second quarter of 2004. To reflect OPEC’s acknowledgement of their loss of marketing power, we set the nominal target to the actual price from the third quarter of 2004 onwards (we use the OPEC bulletin as a reference). We deflate this nominal target using the G7 GDP deflator. This results in an OPEC target for the real price of crude oil that we denote by $P_t^*$. 

3.3 Data analysis

We present, in figure 1, quarterly observations of the natural logarithm of the variables described above. Four variables are clearly trending: stocks and demands, with seasonal patterns exhibited by demands—although the seasonality in ‘non OECD’ demand is less regular—and less obviously by OECD stocks. Non OECD stocks are much smoother. This reflects the variable’s larger growth: exp(.88) corresponds to an increase of 142% compared to 20 to 30% for the other three variables. Quotas and prices also exhibit increasing patterns over the sample, especially since 2002 for the latter. Augmented Dickey-Fuller tests confirm this aspect: allowing for a constant, a trend and seasonals, the null of a unit root cannot be rejected (with the exception
of \( p_t \) and \( p \)-values are all larger than 20\%. The case of \( p_t \) is different in that the \( p \)-values vary between 2\% (with three lags) to 98\% (with four lags); yet for five lags—the number that minimizes the Akaike Information Criterion—the \( p \)-value is 8\%. Given the clear upward movement in prices toward the end of sample, we feel confident that \( p_t \) is non-stationary. Subsequent testing for the first differences show that none of the variables can be assumed \( I(2) \), individual nulls being rejected with \( p \)-values lower than \( 10^{-4} \).
4 Cointegration analysis

4.1 A VAR representation

Since we chose to restrict our attention to a sample consisting of 68 observations, we are constrained in our analysis and cannot resort to general unrestricted VAR modeling, preferring instead parsimonious representations. Yet this does not preclude analyses on sub-versions of the model. First, we analyze the variables presented previously by means of a VAR(1) where we include an unrestricted constant. A graphic analysis of the fit is presented in figures 2 and 3 where we discarded the use of non OECD inventories as this variable does not seem to improve the estimation but lowers the precision (via increased variance of the estimators). We hence keep $s_{t}^{non\text{ OECD}}$ for single equation modeling.

Although it enters the VAR, we do not present here the results for $p_{t}^{*}$ to make the figures more readable. Despite a few large outliers, the fit is reasonable—although the lack of seasonal patterns for the fitted OECD demand would require, if we were to specifically model the latter, additional fourth-order lags as regressors. In the analyses below, we follow Barsky and Kilian (2004) who stressed the impact of wars in the Middle-East on the oil market: we include indicator variables to account for the beginning of the first Iraq war (I90q3 and I90q4), the terrorist attacks of September 11, 2001 and ensuing Afghan war (I01q4 as the terrorist attack occurred towards the very end of the third quarter and were soon followed by the inception of the Afghan war) and the beginning of the second Iraq war (I03q2). When performing a battery of specification tests for the VAR(1) representation, we notice that seasonality induces slight fourth-order residual autocorrelation. Adding fourth-order lags of the endogenous variables would correct for this but, due to the available sample size, we cannot perform our cointegration analysis on a VAR(4). Instead, since seasonality is a stationary feature and does not change the cointegrating vectors, we leave this issue aside until we focus on the congruent stationary mapping. We will show that annual patterns can easily be accounted for at that stage (see section 5.1). We conclude that the variables presented here constitute a reasonably congruent (see Hendry, 1997) representation of the data generating process.

We now perform cointegration tests and we use Johansen’s trace test (see Joh-
Figure 2: Fit of a VAR(1) model on the six variables $p, p^*, q, s^{OECD}, d^{OECD}, d^{non \, OECD}$. The figure reports results for $(p, q)$. In the third and fourth columns, $r_{x_t}$ refers to the residuals in the equation for the variable $x_t$.

Figure 3: Fit of a VAR(1) model on the six variables $p, p^*, q, s^{OECD}, d^{OECD}, d^{non \, OECD}$. The figure reports results for $(s^{OECD}, d^{OECD}, d^{non \, OECD})$. In the third and fourth columns, $r_{x_t}$ refers to the residuals in the equation for the variable $x_t$. 

14
<table>
<thead>
<tr>
<th>Maximum Rank</th>
<th>Log-Likelihood</th>
<th>Eigenvalues</th>
<th>Trace Statistic</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>828.0655</td>
<td>–</td>
<td>180.84</td>
<td>[0.000] **</td>
</tr>
<tr>
<td>1</td>
<td>859.9639</td>
<td>0.61411</td>
<td>117.04</td>
<td>[0.001] **</td>
</tr>
<tr>
<td>2</td>
<td>886.8835</td>
<td>0.55227</td>
<td>63.204</td>
<td>[0.150]</td>
</tr>
<tr>
<td>3</td>
<td>900.8864</td>
<td>0.34164</td>
<td>35.198</td>
<td>[0.442]</td>
</tr>
<tr>
<td>4</td>
<td>910.1124</td>
<td>0.24073</td>
<td>16.746</td>
<td>[0.666]</td>
</tr>
<tr>
<td>5</td>
<td>916.2178</td>
<td>0.16661</td>
<td>4.5354</td>
<td>[0.852]</td>
</tr>
<tr>
<td>6</td>
<td>918.2215</td>
<td>0.058059</td>
<td>0.52792</td>
<td>[0.467]</td>
</tr>
</tbody>
</table>

Table 1: Cointegration test statistics, over the sample 1989(2)-2005(4), for the VAR(1) described in section 4.

hansen, 1991 and 1996) which leads to rejecting the presence of fewer than one stationary relation (at a p-value lower than 1%) and to accepting the existence of two cointegrating vectors as in table 1.

We also test the same model when allowing also for a linear trend to enter the cointegrating space. This leads to the same conclusion of a cointegration rank of 2. Under the assumption that one cointegrating vector relates \( p_t \) to \( p_t^* \), and the other, \( s_t^{OEC} \) to \( d_t^{OEC} \), we impose identifying hypotheses of opposite unit coefficients for \( p_t \) and \( p_t^* \) in the first cointegrating vector, and for \( s_t^{OEC} \) and \( d_t^{OEC} \) in the second. This feature is not rejected by the data and we proceed to further reduction, progressively eliminating the remaining variables from the cointegration space. We eventually identify the following stationary relations:

\[
\begin{align*}
    c_{1,t} &= p_t - p_t^* + 0.63q_t \\
    c_{2,t} &= s_t^{OEC} - d_t^{OEC} + 0.00172t
\end{align*}
\]

These vectors are presented—demeaned with superscript ‘\( \mu \)’, i.e. \( c_{i,t}^{\mu} = c_{i,t} - T^{-1} \sum_{j=1}^{T} c_{i,j} \) for \( i = 1, 2 \) over a sample of size \( T \)—in figure 4. Their interpretation is: \( c_1 \) OPEC attempts to control the difference between the observed price and its target by modifying its quota; and \( c_2 \) there is a link between OECD stocks and demand, which corresponds loosely to a coverage rate—we will define it properly below—and their difference (in logs) has been decreasing at a constant rate over the last fifteen years, hence the necessity to add a linear trend for stationarity.
Figure 4: Cointegrating vectors corrected for the mean. $c_{1,t}^\mu$ links OPEC quotas to the difference between price and target. $c_{1,t}^\mu$ relates OECD stocks and demand, to the continuous decrease in their ratio.
4.2 Robustness checks

Given the low number of available observations, the multivariate Johansen test may not be the most efficient, although Coombs and Algina (1996) show that a ratio of more than 10 observations per dependent variables imply that the test controls size reasonably when the data are normally distributed. The latter assumption is subject to the outliers that wars caused, yet Gonzalo (1994) also showed that the test is robust to failure of normality. To ensure the robustness of our results, we also perform the single equation equilibrium-correction test of Banerjee et al. (1998). This has been recommended by Ericsson and MacKinnon (2002) in small samples (the technique is subject to weak exogeneity assumptions which we show to hold in the next subsection). The estimated equations are, first, corresponding to (1):

\[ \Delta p_t = \alpha_p + \lambda_p p_{t-1} + \lambda_p^* p_{t-1}^* + \lambda_q q_{t-1} + \mu_p^* \Delta p_{t-1}^* + \mu_q \Delta q_t \]

where we cannot reject \( \mu_q = \mu_p^* = \mu_1 = 0 \) and impose it for parsimony. The resulting \( t \)-statistic for the null hypothesis that \( \lambda_p = 0 \), i.e. of no cointegration involving \((p_t, p_{t}^*, q_t)\) is -3.57 which Ericsson and MacKinnon (2002), using their accompanying files, associate with a \( p \)-value of 4.9% so that absence of cointegration is rejected at the 5% level. Rejection is here marginal, yet present in this small sample. Now, for the second cointegrating vector, we estimate

\[ \Delta s_{t}^{OEC} = \alpha_s + \alpha_r (t - 1) + \lambda_s s_{t-1}^{OEC} + \lambda_d d_{t-1}^{OEC} + \mu_d \Delta d_{t-1}^{OEC} \]

and obtain a \( t \)-statistic of -4.13 for \( \lambda_s \), i.e. a \( p \)-value of 2.1% for the null of no cointegration. We are therefore confident that equations (1) and (2) identify the two cointegrating vectors in the system.

4.3 Weak exogeneity

Given the low number of observations at our disposal, we aim to perform a univariate analysis of the determinants of the real price of crude oil. For this to be possible, we need to ascertain that we are entitled to model the distribution of the variable of interest without having to consider the marginal distributions of the conditioning variables. This can be done provided that the regressors are \textit{weakly exogenous} for the parameters of the conditional regression, see Engle et al. (1983). In addition to parameters being variation free (which we assume here), we need to assess that
the factor loadings—or impact coefficients, see Johansen (1996)—of the cointegrating vector $c_1$ are zero for all the equations of the VAR, except that which governs $p_t$ (and necessarily $p^*_t$ given its definition from 2004 on). We proceed to a Likelihood Ratio test for these restrictions. The log-likelihood of the unrestricted model is $l_{UR} = 797.048$; imposing 3 impact coefficients (those of inventories and both demands) to be zero leads to $l_R = 794.278$. The statistic is then $2(l_{UR} - l_R) = 5.54$ which we need comparing to the critical values of a $\chi^2(3)$: this is associated with a $p$-value of 14%. Restricting also the factor loading of $q_t$ leads to a $p$-value of 8%. It seems difficult to consider that OPEC quotas are weakly exogenous for the setting of prices, but we yet consider that we can marginally disregard the process governing the behavior of OPEC since quotas are not modified often can be considered as given from a quarter to the next. This therefore allows for a univariate analysis of price determination, which we do in the next section.

Weak exogeneity of prices and quotas for the parameters of the supply and demand equations are another key assumption that is necessary for the single-equation test of equation (4). It is an assumption that allows for maximum likelihood, or least squares, estimation of an equilibrium correction mechanism with no need for instrumental variables. A likelihood ratio test of the hypothesis of zero impact coefficients of (2) onto the equations for $(p_t, p^*_t, q_t)$ in the VAR yields $l_R = 796.000$ and $p$-value of 55%. Weak exogeneity here does not imply that prices do not have a long-term impact on production levels, such as in the target revenue theory tested by Ramcharran (2001, 2002), or consumption (see Yang et al., 2002, and Dees et al., 2007), simply that when considering quarterly variations, we can condition price evolutions on the other variables without modeling the mechanism that controls the formation thereof.

5 Univariate analysis

5.1 An equilibrium correction model

Under weak exogeneity of the regressors for the parameters of the conditional regression, we proceed to univariate modeling of the changes in oil price. We also create an “expected demand coverage rate” variable in order to assess whether the current level of stocks is sufficient to satisfy expected demand for the next period given the highly seasonal pattern of the latter. This variable is constructed as the ratio of
current crude oil stocks in the OECD countries over expected OECD demand for
the next quarter, computed as the current value to which we add the previous year’s

**corresponding quarter-to-quarter variation.** This is thus defined as \( d_{t+1|t} = E_t[d_{t+1}] \)

following the linear rule:

\[
d_{t+1|t} = d_t + L^4 \Delta d_{t+1} = d_t + \Delta d_{t-3},
\]

where \( d_t \) stands for \( d^{OECD}_t \). The expected demand coverage rate hence becomes

\[
(CR_t) = \frac{S^{OECD}_t}{D^{OECD}_t},
\]

and \( cr_t = \log CR_t \). This new variable closely mimics \( c^\mu_{2,t} \) but with a forward looking

element. We proceed to a General-to-Specific estimation of a regression of the quar-

terly change \( \Delta p_t \) over the variables that we defined in the previous section, including

demeaned cointegrating vectors, together with \( cr_t \). All the stochastic regressors

are entered with lags to obtain a forecasting model. For thoroughness of the reduction

procedure, we use the PcgGets software (see Hendry and Krolzig, 2001) and check for

subsample constancy of the selected final equation. We also identify the potential

outliers that correspond to wars in the Middle East. The resulting model for the

physical determination of crude oil prices is as follows (standard errors in brackets):

\[
\Delta \hat{p}_t = -0.309 \text{cr}_{t-1} - 2.11 \Delta s^{non}_{t-1}^{OECD} + 4.53 \Delta d^{non}_{t-1}^{OECD} - 3.29 \Delta^2 d^{non}_{t-1}^{OECD}
\]

\[
- 0.203 c^\mu_{1,t-1} - 2.28 \Delta t c^\mu_{2,t-1} + 1.31 c^\mu_{2,t-1}
\]

\[
+ 0.34 I90q3_t + 0.528 I90q4_t - 0.259 I01q4_t - 0.36 I03q2_t
\]

with estimation statistics:

\[
\text{RSS} = 0.3462, \quad \hat{\sigma} = 0.0816, \quad R^2 = 0.742, \quad \bar{R}^2 = 0.693,
\]

\[
\ln L = 163.920, \quad AIC = -4.855, \quad HQ = -4.707, \quad SC = -4.480,
\]

\[
T = 63, \quad p = 11, \quad F_{pNull} = 0.0000, \quad F_{pConst} = 0.0000;
\]

where \( \bar{R}^2 \) is adjusted \( R^2 \), \( \ln L \) is the log-likelihood, \( F_{pNull} \) and \( F_{pConst} \) are tests for

the joint nulls that, respectively, no regressor parameters and only a constant are
significant. It may be questioned whether equation (7) does not suffer from small sample bias since we use 11 regressors for 63 observations. In reality, since four of the explanatory variables are dummies, their impact is solely to offset the effect of the corresponding observation on estimation. The model hence consists of 7 regressors for 59 observations, a ratio of 8.4 observation per regressor. Although low, this ratio is quite standard in dynamic macroeconomics: it corresponds for instance to estimating a bivariate VAR(4) (with a constant) or a trivariate cointegrated VAR(1) using quarterly data over 1970(1)-2005(1).

Testing for misspecification of (7), we cannot reject the absence of residual autocorrelation ($F_{AR(1-4)} = 1.074$, $p$-value of 38%), of ARCH effect in residuals ($F_{ARCH(1-4)} = .686$, $p$-value of 61%), of heteroskedasticity ($\chi^2_{hetero} = 22.96$, $p$-value of 19%), or the normality of the residuals ($\chi^2_{normality} = .125$, $p$-value of 91%). Thus, this model satisfies all the specification tests, has a $\bar{R}^2$ of about 70%, and we can deem it congruent.

What is the model of price formation that equation (7) presents? Prices react to market forces observed via discrepancies between supply and demand: when inventories (i.e. cumulated past supply in excess of demand) have been increasing outside the OECD ($\Delta s_{t-1}^{non\ OECD} > 0$), and when they are high in the OECD compared to expected demand ($cr_{t-1} > 0$), prices fall. Wedges between supply and demand in the OECD also appear via $c_2^t$, but with a nonlinear effect: this variable represents the (highly seasonal) deviation of OECD inventories from their long term decreasing trend. An excess in $c_2^t_{t-1}$ directly implies a lower price next period (coefficient of $1.31 - 2.28 = -0.97$) with a reinforced impact if inventories are high compared to the previous year (the large negative coefficient of $\Delta_4 c_2^t_{t-1}$). Inventory variations play a more significant role in OECD economies as these industrialized economies have a more stable average consumption but with strong (mostly predictable) seasonality. By contrast, inventories in non OECD economies, historically lower, have caught up in the 1990s with the levels observed in the OECD but have remained almost constant since. Outside the OECD, demand has kept increasing and its effect appears through a large positive coefficient. The link we estimate is positive but shows a staggered transmission of increases in OECD demand, with a more pronounced effect two periods ahead (it rewrites as $1.24 \Delta d_{t-1}^{non\ OECD} + 3.29 \Delta d_{t-2}^{non\ OECD}$). Equation (7) also shows via an equilibrium correction feature involving $c_1^t_{t-1}$ that OPEC achieves its aims at controlling prices.

In addition to market effects, we identify indicator variables for wars. The first two
(I90q3t and I90q4t) correspond to the invasion of Kuwait and following disruptions that entailed spikes in prices. By contrast, the second invasion of Iraq was largely anticipated. The war against the Iraqi regime started on March 20, 2003 and lasted until President G. W. Bush declared an “end of major combat operations” on May 1, 2003. The relative swiftness of the operations and the quick fall of Iraq’s regime turned out to calm markets (see Barsky and Kilian, 2004, for an appraisal of the effect of this war) although the actual tensions on production capacities remained. This explains why we find a negative impact of I03q2t. The other indicator that we identify concerns the events surrounding the terrorist attacks of September 11, 2001 and the ensuing Afghan war which also coincides with a business cycle trough in the U.S. (in November 2001, according to the NBER). Barsky and Kilian (2004) discuss this at length: despite conventional wisdom that Middle East turmoils and low reserves (in addition to a fierce winter in North America) should push the prices up, the decrease in prices accelerated in 2001. The specifics of the last quarter of 2001, where prices fell more than demand, despite the low inventories were not captured by our model: this explains why we resorted to a dummy variable and why the latter has a negative impact.

We also considered whether the model is constant over the sample. Figures 5 and 6 record parameter estimates and corresponding t-statistics obtained from recursive estimation with an initializing sample of 24 observations. We observe that all parameters are reasonably constant and significant in-sample, with the possible exception of the coefficients of $\Delta d_{t-1}^{non-OECD}$ and $c_{2,t-1}^{\mu}$ which are stable yet barely significant until the latest observations. We will return to this when we consider alternative parsimonious models in the next subsection. This is not of too much concern since both regressors also enter via their second difference and four-quarter change, respectively, with strongly significant parameters: the staggered estimated effect may be suffering from a lack of precision, but the overall effect is robust.

The model presented above in equation (7) constitutes a valid representation for the determination of the clearing price of crude oil according to the physical market. This accounts for about 70% of price variations, leaving the remaining 30% to financial markets and unexpected innovations. Figure 7 presents the corresponding forecasts for the price level.
Figure 5: Recursive parameter estimates in equation (7) and corresponding $t$-statistics, with an initialization sample of 24 observations.
Figure 6: Recursive parameter estimates in equation (7) and corresponding $t$-statistics, with an initialization sample of 24 observations.
Figure 7: Fit of the equilibrium correction mechanism for the level of oil prices (upper panels), and corresponding forecast errors (lower panels).
5.2 Alternative specifications

To prevent overfitting our model, we attempt to determine, via a smaller version thereof, which variables account for most of the fit. This is presented in (8) with a fit of $\bar{R}^2 = 63\%$ but specification tests lead to rejecting the hypothesis of non-autocorrelated errors with a $p$-value of 0.7\%. This does not need to be a problem if we follow Hendry and Krolzig (2005) who hold that when performing multiple, say $n$, specification tests at a nominal size $\eta$, the probability of rejecting at least one is $1-(1-\eta)^n \approx n\eta$ so that congruency tests should be carried at the 1\% level to avoid almost one-hundred percent chance of model rejection and yet some individual test rejection is to be expected. This smaller model (8) which excludes non-linear patterns therefore constitutes a satisfactory yet non-congruent representation:

$$
\Delta \hat{p}_t = -0.488 \text{cr}_{t-1} - 1.91 \Delta s_{t-1}^{non \text{ OECD}} - 1.44 \Delta 4\mu_{t-1} - 0.23 \mu_{1,t-1} \\
+ 0.411 (I90q3_t + I90q4_t) - 0.292 (I01q4_t + I03q2_t)
$$

Among the explanatory variables that we also considered, we tried to establish whether the price was reacting to movements in its currency of denomination. For this purpose, we constructed an exchange rate for the United States dollar with respect to a basket of currencies, namely the Euro, the Pound Sterling and the Japanese Yen, where the weights are chosen to follow those of the respective currencies in the valuation of the IMF Special Drawing Rights (SDRs). But this variable seemed not to matter, confirming the results in Yousefia and Wirjanto (2004).

5.3 Contributions

In this section, we compute the contributions of the different variables to the change in oil prices, as a result of equation (7). Because we deal with an equilibrium correction model, we need to take full account of the dynamics: here we use 2000(1) as a reference and compute the dynamic impact of the regressors from then. We use the fact that the equilibrium correction model takes the form

$$
\Delta p_t = -\alpha (p_{t-1} - w_{t-1}) + \sum z \gamma_z z_t \tag{9}
$$
for some linear combination $w_t$ of the $z$ regressors which we disregard at this stage. The $z_t$ variables might include lags of a unique or several variables. Re-write (9) as $p_t = (1 - \alpha)Lp_{t-1} + \sum z \gamma z_t$ so that the total dynamic impact of one of the $z$ regressors onto $p_t$ is

$$
\sum_{i=0}^{+\infty} ((1 - \alpha)L)^i \gamma z_t = \gamma z \sum_{i=0}^{+\infty} (1 - \alpha)^i L^i z_t
$$

and we define the impact on $\Delta p_t$ as

$$
C_{\Delta p}(z, t) = \gamma z \sum_{i=0}^{\tau} (1 - \alpha)^i L^i \Delta z_t
$$

for some, large enough, $\tau > 0$. Cumulating the $C_z(i)$, starting from some observation of reference $T_0$, provides the cumulated dynamic impact of $z$ on $p$ over $T_0, ..., t$ as

$$
C_p(z, t) = \sum_{i=1}^{t-T_0} C_{\Delta p}(z, T_0 + i)
$$

which we compare to $p_t - p_{T_0}$. For the impact of the cointegrating relation, we simply report the cumulative sum of $-\alpha (p_i - w_i)$ for $i = T_0, ..., t$. Figure 8 records the resulting contributions to changes in $p_t$ since 2000(1). This choice of reference is related to events occurring in the 1997-2000 period. Following a collapse in the price of oil in 1997-98, a very sharp cut in production quotas had been introduced by OPEC starting in July 1998 and lasting until April 1999, in the hope of drying the market out and restoring prices back to the cartel’s targeted fluctuation band. The nominal price of Brent duly responded and increased to US$ 26.8 by the beginning of 2000 (our reference point which most analysts then considered fair), having bottomed at US$ 11 at the end of 1998. From then, three distinct periods appear clearly:

- 2000-2001: after a rise to US$ 30.4 in the third quarter of 2000, prices were driven down by market forces in 2001. In this context, OPEC began to soften its policy: oil production was allowed to increase quickly. OECD inventories nevertheless remained at low levels. Not until 2001 did supply tensions (with respect to demand) progressively disappear. But a slowdown in demand in 2001 alleviated the pressure on prices, leading the cartel to react and cut dramatically its production quotas. The coverage rate of stocks recovered rapidly in 2001, thus explaining the negative contribution to the change in oil price attributed
to OECD supply and demand in 2000 and, particularly, in 2001. Prices declined sharply to US$ 19.3 at the end of 2001 as discussed above. This is captured in our equation by an indicator variable.

- 2002-2003: increasing non OECD demand and low non OECD inventories, combined with a decreasing OECD \( cr \), led to an increase in oil prices. At that stage, OPEC stayed firm and maintained quotas at very low levels to prevent excess supply in 2002. Then, in the first half of 2003, OPEC became much more accommodative. During the period preceding intervention of the U.S. troops in Iraq and throughout the active war, price volatility was high. This explains the presence of the indicator variable in the second quarter of 2003. Indeed, with the combination of a break in Iraqi oil production and of a stronger world demand, markets feared a shortage. To reassure consumers and soften prices, OPEC largely loosened its quotas, and Saudi Arabia committed to satisfy demand. Yet, prices ended up stabilizing at a relatively high level.

- 2004-2005: a strong disconnection emerged, prices exploding while all contributions would have driven them in the opposite direction. The situation in the physical market was improving. Despite a demand shock in 2004, supply stayed above demand. Producers hence managed to accumulate inventories and increase their coverage of expected future consumption. In 2005, a demand slowdown met rising production quotas, the latter even reached an all time (until then) high. Yet, despite vanishing physical tensions and an accommodative OPEC policy, markets did not calm and the price did not decrease.

The discussion above shows that, prior to 2004, the determinants identified by equation (7) explain market prices. Yet, this is no longer so from 2004 on: all the factors that we consider pushed the price downwards but the latter was exploding. In the next section, we analyze this aspect further by quantifying the so-called ‘market premium’ and discussing its origin.

5.4 Market premium

The analysis above stresses, from 2004, a disconnection between the price implied by equation (7) and that which was observed. Such discrepancies are not unexpected since our model only takes into account a few of the potential factors that might drive
Figure 8: Cumulated dynamic contributions of the regressors to the changes in the log of real oil prices. All series are presented as a change since 2000(1). S&D represents the combined effect of supply and demand. The vertical axis records contributions to percent changes of real oil price (unit of logs: 0.1=10%); for $p_t$, the axis directly records the change in logs since 2000(1).
prices. This is exactly the reason why we focused on supply and demand determinants: to identify systematic departures from the implied price. We hence define the *market premium* as the difference between the observed spot—nominal—price and that which can be explained by a historical account of the relation between the real price and the physical market. By *physical* market, we mean the demand and supply of *quantities* of crude oil as defined in the introduction to the paper.

Because the equilibrium correction model considered concerns quarterly changes, the resulting fitted values for the levels are always based on past observations ($\hat{y}_t = y_{t-1} + \Delta y_t$) and hide systematic departures. For this reason, we compute the premium as the difference between the current price and that obtained by cumulating the residuals from estimation as in the previous subsection, starting from the reference of 2000(1). This leads to the series presented in figure 9 where we notice that the premium has indeed been exploding from 2004. As expected, the latter differs strongly from the residuals in figure 7. Here, the idea is to construct a dynamic forecast analysis reflecting the physical elements of the market.

Whereas the level of the premium is stable over 1995–2004, it then increases by US$ 15 until 2005. The average value of US$ 3.5 over 2002-2003, with a peak at 6.5 in 2002(3) is in line with estimates often reported: Barsky and Kilian (2004) discuss an ‘uncertainty premium’ of about US$ 5 to 6 (in real U.S. price of a barrel of crude oil) at that time. The uncertainty amounted to whether a war with Iraq would occur, after being first envisaged in 2002(3), i.e. the first peak. But, contrasting with the fall in oil prices when hostilities were over (captured by a dummy variable in the model), the premium did not decrease. In fact it started drifting upwards. Several factors can explain this, as reported in the monthly Oil Market Reports of the International Energy Agency of the time.

First, the demand shock in 2004 highlights the inadequacy of the supply structure both on the upstream and downstream markets. Over several years, investment had been weak. Shortages of some specific products appeared in 2004. Spare capacities had never been so tight (see the analysis in Kaufmann et al., 2008). Moreover, despite the slowdown in non OECD demand, combined worldwide demand growth was above its recent trend. At the same time, several disruptions were affecting supply from non OPEC countries. Although markets were largely supplied (OPEC had agreed to ease its production quotas), concerns were rising about the difficulties faced by suppliers in coping with such a dynamic demand. The geopolitical context fed these
Figure 9: Dynamic forecast of the nominal crude oil price and associated market premium.
fears. Besides a worsening Iraqi situation, there were spreading tensions in the Middle East, conflicts in Nigeria, political instability in Ecuador, Venezuela and Russia. In this context, the debate on the depletion of world oil reserves emerged. According to some, the world is approaching the all time maximum of oil production, from which prices increase very quickly. Expectations concerning the risk of a global shortage rose, leading to a rise in prices both in spot and futures markets. Volatility was very high in 2004 and 2005 and fostered by a high level of financial liquidity and speculative activity, see Askari and Krichene (2008) for an estimation thereof.

In addition to the developments affecting oil markets directly, world growth was gathering momentum and the dollar was depreciating. Both effects were fueling expectations of future price rises, see Krichene (2006a, 2008) and references therein. It is still debated whether the loose worldwide monetary conditions were the cause of the spiralling prices of commodities—see Krichene (2006b) for the case of oil—but no matter their causes, oil preceded other commodities in their rises. Returns on commodities, as measured by Standard and Poor’s GSCI index overtook those of the S&P 500 in the third quarter of 2004 (see Standard and Poor’s, 2008) which coincides with the premium, in figure 9, reaching US$10. Merino and Ortiz (2005) show that non-commercial long positions on petroleum related products in the NYMEX futures market—which they deem speculative—increased from 2002 concomitantly with oil prices. Short positions would also need to be considered for a net assessment of speculation. A separate model, extending for instance that of Cortazar and Schwartz (2003) would be necessary for modeling futures prices. Kaufmann et al. (2008) notably derive a pricing equation where they incorporate futures contracts and nonlinear effects of OPEC’s capacity utilization.

All these combined effects cannot be included in our model. Most are transitional, and some could be permanent. In particular, the equation may be underestimating the effect of the demand originating from non OECD economies. As discussed in section 3.1, our choice of GDP deflator might a priori overestimate the impact of $\Delta d_{t-1}^{non \; OECD}$ on $\Delta p_t$, although recursive parameter estimates are stable in figure 5. They do not present any increasing tendency towards the end of the sample, nor do the other estimates. Yet, it is possible that markets may have overreacted to the buoyancy the Chinese economy was exhibiting then. This could explain the ‘speculative’ effects that Merino and Ortiz (2005) observed.

Few of the factors that can justify the rise in recent market premium can be
statistically tested since they appeared almost simultaneously and the available data would not allow to disentangle them yet. This was at a time when risks facing the delivery of crude oil and the uncertainty about future demand were high. Askari and Krichene (2008) show that the dynamics of oil prices have undergone jumps. The increasing premium confirms this aspect and it can be argued that equation (7) then suffers from misspecification. Since the aim of this paper is to measure the deviation from historical drivers and owing to the good fit prior to 2003, we believe instead that a model with regime shift could be more appropriate for accurate tracking of prices. But this would imply using new regressors, such as futures positions taken by non-commercial participants, which do not fall within the scope of the definition that we used for the determination of the physical price. It would also prove a problem since most models based on financial market behaviors focus on higher frequencies, and it is unsure whether the corresponding influences still matter from one quarter to the other.

Finally, we focused, in this paper, on quarterly changes in oil prices conditional on demand and supply. On a quarter to quarter basis, the feedback effect of prices on demand and supply must be low and we explicitly tested for weak exogeneity of the former. Yet, when deriving the implied price dynamics, we run into potential problems since exogeneity may not hold over longer horizons. This seems to us a potential source of misspecification, although the choice of the reference year should mitigate the impact thereof since the premium is computed at most five years away, and we saw that price elasticity of demand is low. If we did not run into the issue of small sample estimation, we may wish to consider instrumental variable estimation to correct for the possible endogeneity of the regressors, or develop a multivariate model for supply, demand and price. With fewer than twenty years of data, we believe that our limited information single-equation approach is more robust to misestimation.

6 Conclusions

In this article, we derived an econometric model for the determination of the real price of a barrel of crude oil. This model is based entirely on physical aspects of the market situation—i.e. quantities produced, stocked, distributed and consumed—and estimated post-1988, when oil prices started being driven by market forces. Two coin-

\footnote{The role of futures is an exception, as shown in Bekiros and Diks (2008).}
Integrating relations matter here, relating to OPEC’s attempts at controlling prices and to a tendency—also observed in other industries—towards a tightening in inventories relative to demand. Using a dynamic forecast derived from the equilibrium correction mechanism, we were able to compute a resulting premium on the price of oil that embodies elements which are not part of historical physical market behavior and may reflect market participants’ perception of uncertainty and risk. This premium accounts for about half the increase in oil price between 2000 and the end of 2005.

Our model also allows for a decomposition of the various contributions of the regressors to the change in prices. We thus observe what factors drive the evolution at any point in time. Between 2004 and 2005, the increase in market price was larger than implied by the physical situation, whereas in the two years before, it was caused by emerging market (i.e. non OECD) demand and a tightening of OPEC quotas. Our model shows that demand from fast-developing economies is not enough to explain the surge in oil prices. In particular, our choice for the transformation from nominal to real price should lead to over-estimating the historical impact of non-OECD demand.

In order to obtain a full understanding of the factors that presided over the recent increase in the market premium, the potential role played by trading strategies such as uncovered futures and options positions that can be deemed speculative would need to be assessed, in relation with other commodities in particular. The impact of sudden fears like the emergence of the ‘end of oil’ debate that followed the increased demand to sustain China’s growth at the turn of the millennium also needs evaluating. Our inability to explain recent hikes from historical patterns is consistent with studies such as Askari and Krichene (2008) who showed that oil price dynamics experienced breaks, or jumps.

As the end of 2008 witnesses a cooling down of the buoyancy of commodity prices, it remains to be seen whether this implies a reversal to historical trends and whether the recent period was just a parenthesis.

References


Fattouh, B. (2007a). The drivers of oil price: the usefulness and limitations of
non-structural model, the demand-supply framework and informal approaches. WPM 32, Oxford Institute for Energy Studies.


<table>
<thead>
<tr>
<th>Variable</th>
<th>Name</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal oil price</td>
<td>Spot price of a barrel of Brent</td>
<td>Thomson Datastream(^1)</td>
<td></td>
</tr>
<tr>
<td>G7 GDP deflator</td>
<td>Constant price, constant PPPs</td>
<td>OECD: Quarterly National Accounts(^2)</td>
<td></td>
</tr>
<tr>
<td>(P_t)</td>
<td>Real oil price</td>
<td>Nominal oil price over G7 GDP deflator</td>
<td></td>
</tr>
<tr>
<td>(D_t^{OECD})</td>
<td>OECD crude oil demand</td>
<td>Millions of barrels per day</td>
<td>IEA: Monthly Oil Market Report</td>
</tr>
<tr>
<td>(D_t^{nom\ OECD})</td>
<td>OECD crude oil demand</td>
<td>Millions of barrels per day</td>
<td>IEA: Monthly Oil Market Report</td>
</tr>
<tr>
<td>(Q_t)</td>
<td>OPEC production quotas</td>
<td>Millions of barrels per day</td>
<td>IEA: Monthly Oil Market Report</td>
</tr>
<tr>
<td>(S_t^{OECD})</td>
<td>OECD crude stocks</td>
<td>Millions of barrels</td>
<td>IEA: Monthly Oil Market Report</td>
</tr>
<tr>
<td>(S_t^{nom\ OECD})</td>
<td>Non OECD crude stocks</td>
<td>Millions of barrels</td>
<td>IEA: Monthly Oil Market Report</td>
</tr>
<tr>
<td>(P_t^*)</td>
<td>OPEC implicit target for real price</td>
<td>Constructed from nominal target: see text</td>
<td></td>
</tr>
</tbody>
</table>

\(^1\): mnemonic OILBRNPQ.Q  
\(^2\): mnemonic (from DRI-WEFA) Q907CMPGDPLA.Q  
\(^\dagger\): International Energy Agency  

Data series are computed as quarterly averages.