Darmstadt Discussion Papers in Economics

Crude Oil Price Determinants

Jochen Möbert

Nr. 186

Arbeitspapiere
des Instituts für Volkswirtschaftslehre
Technische Universität Darmstadt

Applied Research in Economics

Crude Oil Price Determinants*

Jochen Moebert

May 16, 2007

This version August 10, 2007

Abstract

Based on monthly observations, I specify an econometric model capturing the driving forces behind the crude oil price series in recent years. A large set of covariates, such as supply and demand variables as well as futures market variables, is used to test the impact on the crude oil price. Current price movements are a result of scarce refining capacity and speculators betting on higher prices. The results also question OPEC's market power.

JEL Classification: C51, Q41

Keywords: Crude oil market, Econometric modelling

Contact details of the author:

Darmstadt University of Technology, Institute of Economics, Applied Economic Research and Microeconometrics, Marktplatz 15, Residenzschloss, D-64283 Darmstadt, Germany, *moebert@vwl.tu-darmstadt.de*

*I am grateful to David Knorr, Michael Opel, Dominik Roland, Patrick Tydecks, Horst Entorf and Philip Savage for support and comments. Details on EIA data were provided by Joel Lou and Bill Horvath.

1 Introduction

Surprisingly, the number of researchers investigating the impact of crude oil determinants on the series itself is not as large as I expected before this research project started. Most economists are concerned with diminished growth rates, monetary policy, or business cycle costs levied on economies and industries if oil prices soar.[1] Hence, the price series is taken as an explanatory variable but is not the object of interest itself. A noteworthy exception is Kaufmann et al. (2004), which analyzes OPEC's influence on oil prices and thereby develops an important crude oil model. These researchers used quarterly data and investigated the period from 1986 to 2000. I can improve the model with respect to frequency and also analyze a newer time span. Moreover, I have a larger set of covariates which captures nearly all relevant oil price factors I can think of. This data-driven approach fills a gap in the literature and enables me to efficiently analyse both the supply and the demand side.[2]

Reaching an understanding of the oil market requires an investigation of the actions and interaction of market participants such as the Organization of Petroleum Exporting Countries (OPEC), its member nations, industrial nations, and others. Equally important might be reliable estimates of bottlenecks such as lifting capacity, transportation capacity, or refining capacity.[3] The maintenance and extension of these production capacities depend on available technologies. For instance, in the near future modern drilling equipment may increase the amount of crude oil produced and may avoid shortages. In the future, another factor influencing the oil price might be lack of oil reserves.[4]

[1] Cf. Rotemberg and Woodford (1996), Clarida et al. (2000), Hamilton and Herrera (2004), among others

[2] Of course, combining the data analysis with theoretical models is even more appealing. However, the data-driven approach may also improve future theories as well as corresponding econometric models.

[3] Allen (2005) summarizes several arguments on the production capacity but without an econometric analysis. Curlee et al. (1987) analyse several scenarios which might lead to supply disruptions.

[4] How to define and measure oil reserves is described in Haider (2000). There is a heated debate on whether oil reserves are rapidly dwindling, as proposed by some geologists (cf. Campbell and Laherrère (1998), Campbell (2004), Greene et al. (2006)) or, as stated by official institutions such as OPEC as well as its counterpart the International Energy Agency (IEA) whether it will be available several decades (cf. Watkins (2006)). However, from my point of view, relying on official statistics instead of these geological arguments seems appropriate since in the long-run truth-telling is an incentive-compatible choice for OPEC as a whole as well as crude oil demanders. Oil-exporting nations have low incentives to hide a possible lack of crude oil until all reserves are exploited. Otherwise, they miss the chance to sell the scarce oil reserves at even higher prices before oil wells and other oil resources are completely exhausted. Obviously, oil-importing countries also prefer to be informed on oil scarcity to minimize costs arising in the transition process from the oil age to a new energy era. Given this reasoning, an abrupt disclosure of oil shortages is unlikely and would imply that all major market participants are hiding information and are misinformed today. Such coordination failures have arisen in the history of financial markets but the frequency of occurrence of such global misjudgements has been rather small.

The oil price is one of the financial series widely reported on in the media. However, not only one single oil price is fixed every day. Instead, the number of oil grades traded is as large as the number of important oil wells. Oil from different sources differs with respect to the chemical composition, in particular the sulfur content and its density, and oil from less valuable spouters is traded at lower prices. Accordingly, the discounted price captures the lower industrial applicability and, given constant qualities of oil grades, the price discount is stable across time (cf. Hornsell and Mabro (2003), Wang (2003), Bacon and Tordo (2004), among others). Hence, a model explaining a particular oil grade is directly transferrable to other oil grades as well. In standard financial data sheets the oil prices of important wells are listed, e.g. Brent, Nigeria Brass, West Texas Intermediate, among others. In this article the regressand is the price of West Texas Intermediate (WTI), a reference oil for North America. WTI is a high quality oil due to its low sulfur content and its low density.[5] The corresponding futures contracts traded at the NYMEX[6] are the most liquid futures contracts worldwide. For this reason, this data series is especially suited for my purposes since several futures market variables are included in the econometric specifications. The data is available at a monthly frequency where each value is observed on the first day of each month. December 1995 is the first observation month of the sample and the last month is January 2006, such that for each series 122 observations are available.

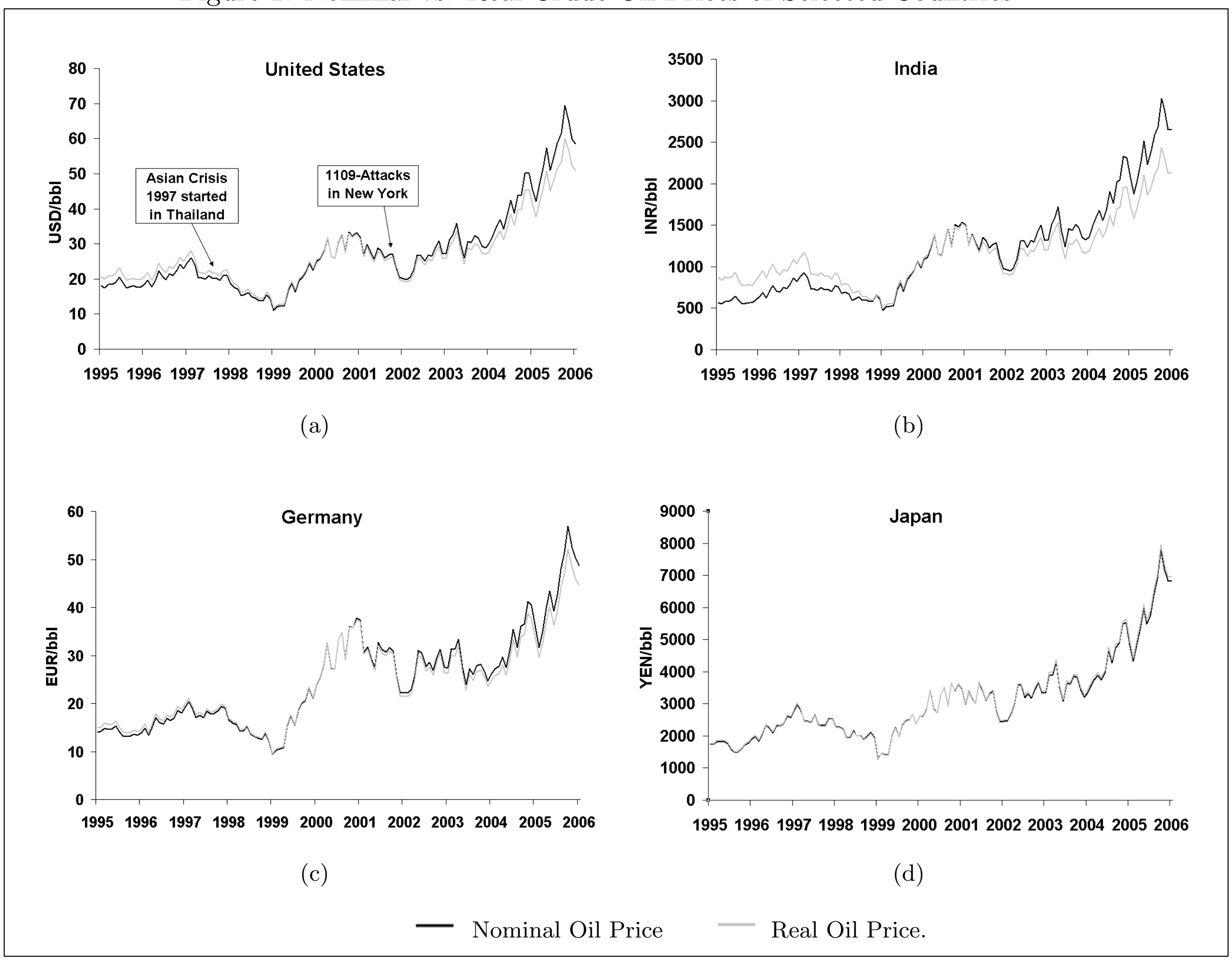
Figure 1(a) points out how volatile the nominal and real USD WTI price were in the last decade. After the Asian crisis, oil prices plummet and oil futures and oil spot markets were not cointegrated as before as shown by Hammoudeh and Li (2004). The price range of the crude oil series indicates the high degree of volatility in this market. After the Asian crisis in 1999 prices had fallen down to around 11 USD before the boom market started and pushed prices up to nearly 70 USD in October 2005. Similar price ranges were observed in Germany and Japan and such large price changes of an important commodity from around six hundred percent readily attract the attention of economists. Also, before the observation period volatile prices were steadily observed in the oil price history. Krichene (2002) investigated the first two phases of the oil market in the last century. The first phase, the noncompetitive nonvolatile crude oil market, ended with the cartelization among OPEC states in 1973.[7] The cartelization

[5] Cf. Energy Information Administration, Annual Energy Outlook 2006.

[6] The New York Mercantile Exchange (NYMEX) is the world's largest commodity market.

[7] In real terms current market prices are still below the oil price in 1973.

Figure 1: Nominal vs. Real Crude Oil Prices of Selected Countries



Own source: Real and nominal crude oil price (monthly series) in local currencies from December 1995 to January 2006. The base year is midyear 2000. Notice the different scales at the ordinates of each subfigure.

triggered the first oil price shock and increased price volatility thereafter in the second phase. In particular, oil prices soar when important political decisions related to the Mideast such as oil shocks and wars[8] or both, appear on the political agenda. However, the development in recent years might differ from previous developments. Many analysts argue that a new period has arisen in which the emergence of the world's largest economies - in particular China and India - have reshaped the landscape. There is the fear that oil resources and capacities will be readily exhausted if heavily populated countries such as India and China strive for western-style living conditions.[9]

From a methodological point of view, an oil market analysis can be fairly simple. The behaviour of market participants is often explainable using standard textbook arguments. The demand side consists of a very large number of countries, companies, and households and can be characterized by large economy models. Of course, the number of suppliers is much smaller. OPEC Countries and large European and American oil companies are the main oil suppliers. Due to the presence of the OPEC cartel and its continuous attempt to reduce the supply OPEC's behavior must be taken into account. However, a section explaining theoretical details (cf. Gray (1914), Hotelling (1931), de La Grandville (1980), Tvedt (2002), among others) on the oil market is omitted. In the following Section the data base and hypotheses are described. First, demand factors, then supply factors and, finally, futures market variables are discussed in the following subsections. Subsequently, Kaufmann's et al. (2004) work is repeated and extended with the data in Section 3. The main model is contained in Section 4, which establishes a market model, and Section 5 concludes.

2 Covariates and Hypotheses

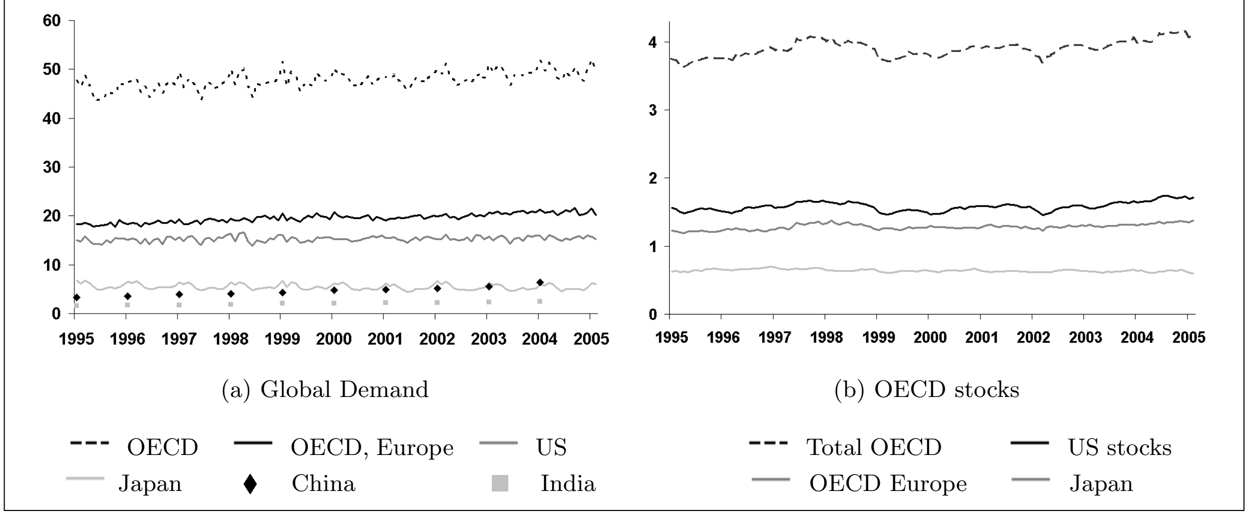
2.1 Demand Factors

An important crude oil variable is the market demand. An increased need for crude oil need shifts the demand curve in a price-quantity diagram to the right and thereby increases the oil

[8]Foster (1996) shows how spot and futures markets were interrelated before the first US-Iraq Gulf war in the nineties.

[9]Cf. Energy Information Administration, Annual Energy Outlook 2006.

Figure 2: Crude Oil Demand (mbd) and Petroleum Reserves (in billion barrels)



(a) Global Demand (b) OECD stocks

Data source: Annual Energy Review 2006 and Monthly Energy Review 2006, Energy Information Administration. Global oil demand is depicted in Figure 3(b).

price if the supply curve is not completely elastic. As shown in Figure 2, the crude oil demand of large industrialised nations and the OECD as a whole rises very slowly. Following the oil price shocks in the seventies, developed countries came to bitterly regret their strong dependence on foreign oil. Thereafter, oil-importing economies tried to reduce oil as an input factor and also began to replace oil with suitable substitutes. These efforts lead to a reduction of consumption per head while the total consumption was still slightly increasing in Western nations. In the sample the minimum demand was 43.7 mbd whereas the maximal value was 51.9 mbd and if a deseasonalized series is used the range is even smaller.[10] A steady demand growth is observable for India and China. India's demand strongly increased from approximately 1.6 mbd to 2.5 mbd and China's demand grew even stronger from around 3.5 mbd in 1995 to considerably more than 7 mbd in 2005. To obtain monthly values India's and China's demand was linearly interpolated. These figures indicate the strong growth of these economies. However, relative to OECD demand their demand is low which points out the large future demand.

[10]In the sample, the mean of all summer months amounts to 46.696 mbd whereas the mean of winter months amounts to 49.143 mbd.

2.2 Petroleum Stocks

After the first oil crisis, OECD member states established an oil stock to prevent unforeseen circumstances cutting down global oil supply. In January 2006, the OECD petroleum stock amounted to 4.1 billion barrels.[11] The development of this variable called OECD_stock, as well as stocks of large nations/regions such as the United States, Japan, and Europe is shown in Figure 2(b). I construct the variable OECD_days introduced by Kaufmann et al. (2004) which is the ratio of OECD_stock in million barrel and OECD_demand in million barrel per day.[12] OECD_days can therefore be interpreted as an indicator of the independence of oil-importing countries from price shocks and OPEC.[13] Kaufmann et al. (2004) expect a negative impact on the price "consistent with results described by Kaufmann (1995) and Balabanoff (1995)... . An increase in stocks reduces real oil price by diminishing reliance on current production and thereby reducing the risk premium associated with a supply disruption." However, from my point of view positive coefficients can also be expected. If petroleum stocks are filled (released) then demand increases (decreases) and crude oil prices might rise (fall). Correspondingly, positive coefficients are measured. Accordingly, no clear sign of the corresponding OECD_stock coefficient is expected.

2.3 Supply Factors

Production Variables

The production process in the petroleum industry is often separated into an upstream, midstream, and downstream segment. The upstream segment summarizes exploration and production activities. The production costs vary across different regions and production

[11] "Petroleum stocks include crude oil (including strategic reserves), unfinished oils, natural gas plant liquids, and refined products.", Energy Information Administration/Monthly Energy Review December 2006, p.156.

[12] Throughout the whole paper, all variables except the oil price itself are identifiable through underlines.

[13] Other variables often discussed in the media are US petroleum stocks. They account for approximately 1.7 billion barrels or 40% of total OECD stocks. Due to the exceptional position of the United States, US oil stocks are steadily observed by oil analysts. Total US oil stocks can be divided into several categories. For instance, the strategic petroleum reserve (SPR) merely consists of crude oil and amounts to around 0.7 billion barrels. SPR stocks are under the control of the federal government whereas other stocks exist due to the refining process or are part of the supply chain. The sum of US crude oil stocks amount to more than fifty percent of the total US stocks. The non-crude oil stocks consists of heating oil, gasoline, and other oil products. All three series evolved similarly in the observed period. They exhibit a low degree of variability and show a small upward trend at the end of the sample.

methods. Producers in the Middle East have the lowest costs of approximately 4 US dollar per barrel.[14] International oil majors declare production costs of between 6 and 12 US dollars per barrel in their annual reports (cf. Birol and Davie (2001)). The production of unconventional oil lies in a larger cost range but is often higher than the conventional production. Due to technological progress all exploration costs, in particular costs of offshore explorations, have decreased in recent decades. Accordingly, the exploratory success rate offshore has steadily increased (cf. Forbes and Zampelli (2000)). Unfortunately, the number of public series containing information on the upstream segment is small. Oil majors offer some information on exploration costs in their company reports. However, the construction of variables covering global exploration activities is very demanding since all or at least much company information must be aggregated. The only upstream variable I include is PROD_rig which counts the number of active drilling rigs.[15] I expect that more active drilling rigs reduce the crude oil price.

The midstream[16] contains the transportation sector. Crude oil is often shipped by large oil carriers. One important ship class is the 'very large crude-oil carriers' (VLCC) which have a deadweight tonnage from 200,000 dwt to 315,000 dwt.[17] There also exists the ultra class with even higher tonnage. Some tankers of the ULCC-class (ultra large crude-oil carriers) have more than 500,000 dwt. In total, in the mid-nineties 427 VLCCs and ULCCs carried crude oil across oceans whereas 472 oil carriers were used at the end of the observation period. However, there are many more VLCCs than ULCCs. Therefore, I take into account the charter prices measured in thousand USD per day of this class. In the last years of the sample, a strong price increase is shown in Figure 3(a), which possibly reflects the shortage of global tanker capacity. The extension of the tanker fleet requires large capital investments and time since the production of large carriers lasts several years. Therefore, the transportation process can be a bottleneck in the production process (cf. Brook et al. (2004)). However, the shipping costs per barrel are still fairly low and only around a few cents. I conjecture that higher charter prices have a modest positive impact on the WTI series.[18]

The downstream segment cracks oil into finished products such as gasoline, heating oil, and

[14] In Birol and Davie (2001) the original unit is barrel of oil equivalent (boe), to be exact.

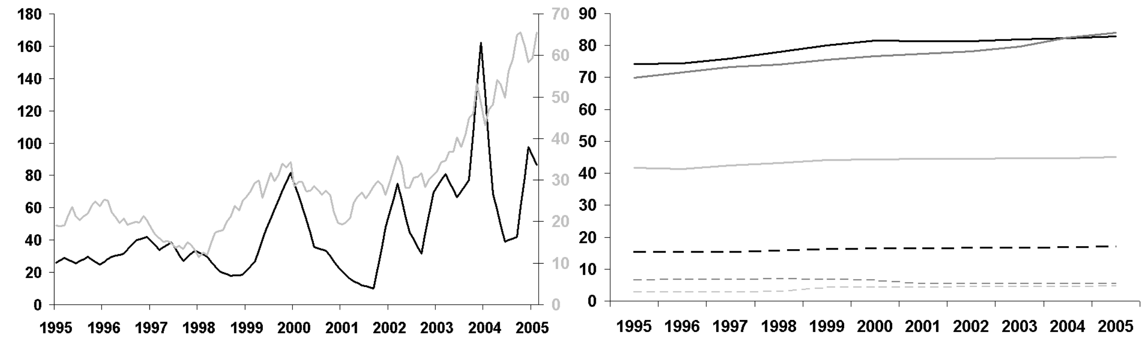
[15] The data was provided by www.bakerhughes.com.

[16] In some articles only upstream and downstream segments are distinguished. Then, the transportation sector is part of the upstream sector.

[17] dwt abbreviates 'deadweight tonnage', which measures the cargo capacity of a carrier.

[18] Of course, simultaneity might be an issue. However, I will skip this issue until Section 4.

Figure 3: Oil Production - Refining Capacities and Very Large Crude-Oil Carriers (VLCC)



(a) VLCC

(b) Refining Capacity and Demand (in mbd)

VLCC rates (in 1000 USD/day)
WTI spot oil price

Global Capacity
Global Demand
OECD-Capacity
US-Capacity
Russia's Capacity
China's Capacity

Data Source: (a) P.F. Bassøe AS, www.pfbassoe.no. (b) EIA, Annual Energy Review 2006, Table 11.9 World Crude Oil Refining Capacity, 1970-2005.

other products. This cracking process takes place in refineries mostly located in industrialized countries. Figure 3(b) shows the global refining capacity in million barrels per day (mbd) and the capacity of countries with large capacities. The global refining capacity is defined as the maximum amount of crude oil which can be processed in a calendar year divided by the number of days in the corresponding year. This data series is therefore an annual average and each reported value is assumed to be the midyear value. Then, monthly data were calculated through linear interpolation. Errors arising by interpolating the data are probably negligible since refining capacity is a very stable time series as shown in Figure 3(b). No refining capacity value was available for the last months until January 2006 since the EIA-data series only lasts until 2005. I completed the series by adding figures from the Oil&Gas Journal (OGJ), where end of year data on global refining capacity is reported until the end of 2005.[19] Besides the low variability of refining capacity, it is noticeable that the global demand exceeded global refining capacity in 2004. Hence, this scarcity of global free refining capacity measured by the variable PROD_fref might be one reason for rising oil prices, and a positive impact on the WTI series is expected.

[19] Cf. Nakamura (2005) in the OGJ. Furthermore, the primary source of EIA refining capacity is the OGJ. Until January 2007, EIA did not update their refining capacity. Hence, this data series is still the most current figures.

OPEC Variables

Another important and intensively debated supply side factor is the behaviour of OPEC-countries which try to shorten their supply to push up the market price and thereby earn profits (cf. Griffin (1985), Jones (1990), Ramcharran (2002), De Santis (2003), Kaufmann (2004), among others). OPEC uses several instruments to impinge on the oil market. Since 1983 OPEC has announced the production quota of all its member countries which measures the overall OPEC supply. In addition, to monitor the price directly, OPEC introduced a target corridor between 22 and 28 US dollars in March 2000. It was planned that prices above the corridor increase OPEC's production. However, after market prices continued to rise above 28 US dollars for more than a year OPEC suspended its corridor in June 2005. However, they announced the reintroduction of a new corridor when the market settles down again. It seems that OPEC countries lost power and were pleasantly surprised by recent market developments.[20]

Another factor diminishing the power of OPEC is the discipline among its members. Each of the current eleven OPEC-countries always has incentives to sell more oil than agreed among OPEC members.[21] Correspondingly, each country gains at the expense of the other OPEC members since selling larger amounts of crude oil reduces the market price and harms the other OPEC members. Several strategies are available to OPEC members to gain from the artificial scarcity the OPEC cartel is maintaining. One strategy is to disregard OPEC's production agreement and produce more and sell more oil than the agreement allows. Another somewhat more subtle way to violate OPEC agreements is the redefinition of oil sands as oil reserves. Given OPEC's regulations - for more than two decades OPEC's production quotas have been linked to reserves - this redefinition automatically allocates higher production quotas to these countries.

Both strategies decrease the oil price and threaten the cartel stability (cf. Jabir (2001)). A factor which can be used as a proxy for cartel stability is the free production capacity of oil facilities in each OPEC country. Non-producing capacities reduce the earnings because staff and

[20] Matthies (2005) speculated on a new unofficially announced corridor above 50 dollars. However, his argumentation relies on the declaration of the OPEC president, who announced that "the basket price would have to remain above 50 dollars for at least seven days before OPEC ministers would start discussing any increase in production quotas." I think this statement actually indicates the uncertainty among OPEC leaders on new market structures.

[21] cf. Jones (1990), Adelman (1993), among others. Adelman uses the following phrase to describe the existing threat to OPEC's cartel stability. "The cycle will continue: meetings–quotas–firm prices–cheating–price declines–threats and promises–meetings ... with here and there some drastic political-military moves."

capital must still be paid whereas no revenues are generated. Hence, increasing free production capacities might ultimately lead to the violation of OPEC-rules. However, the maintenance of free capacity is necessary to increase oil supply in order to cut prices and to maintain OPEC's official price ranges. Both arguments together imply that there is an optimal free capacity rate for OPEC as a whole as well as each OPEC member state.

Finally, in recent years OPEC's power has been diminished by the increasing availability of energy substitutes. As already mentioned, in many countries the per head consumption decreased due to substitution in industrialised countries after the oil price shocks in the seventies.[22] There is some indication that a large movement is getting under way which might reduce the oil dependence even further. In US states - especially in California - green technologies are gaining ground due to recent policy decisions and growing global evidence for large costs levied upon each of us through climate change. This recent development might be a threat to oil-exporting countries since many of them strongly depend in particular on US demand.

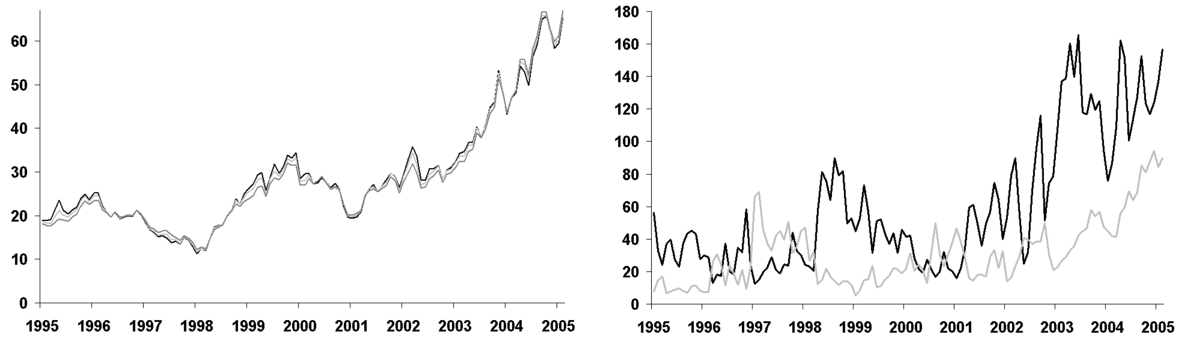
To grasp OPEC's influence on the oil price I specify the same set of OPEC variables as Kaufmann et al. (2004). OPEC_capu measures the utilization of production capacity of all OPEC members. OPEC_quota measures the ex-ante concerted production quotas in mbd and OPEC_cheat is the ex-post OPEC supply minus OPEC_quota in mbd. I expect that a higher utilisation of OPEC production capacity increases the price since possible shocks, either demand shocks or temporary non-availability of non-OPEC production capacities, increase the risk of supply shortages. In contrast, higher production quotas imply a higher supply and might reduce prices. Finally, if OPEC members strongly violate the production quotas, cartel agreements are of no relevance and prices might fall.[23]

2.4 Futures Market Factors

In the recent decades, the turnover of oil derivatives continuously increased and impact on the spot market as well. Someone who needs oil in a future period has the option to purchase today and store until the actual demand arises or to purchase a futures contract. Hence,

[22] Notice, only the per head consumption is nonincreasing whereas the total demand rose in last decades.
[23] Here, my expectations are in accordance with Kaufmann et al. (2004).

Figure 4: Futures Prices and Traded Contracts at the NYMEX



(a) Spot and Futures Prices

(b) Number of Traded Contracts

Spot Price
FUT2
FUT4

Long Positions (in thousands)
Short Positions (in thousands)

Data Source: EIA, Monthly Energy Review 2006 and Bloomberg. The variable FUT1 is not shown in figure (a) since this series is nearly indistinguishable from the spot price itself.

lagged futures prices might increase prediction accuracy since they may indicate expectations of market participants. I will test whether the heavily traded lagged futures series of WTI contain information on the spot price. Three different series are used and each futures variable is denoted FUTq_p where q is the term to maturity and p is the lag length. The maturities of 1-month, 2-months, and 4-months are available. I will test several lag structures. FUT2 and FUT4 as well as the crude oil spot price are shown in Figure 4. Notice that the variable FUT1 is nearly identical to the WTI spot series and therefore omitted. Typically, futures prices are quoted below spot prices due to the convenience yield and market volatility (cf. Litzenberger and Rabinowitz (1995), Pindyck (2001), among others). Such a situation is termed as backwardation and the reverse, but seldom observed situation is called contango.[24] Within the observation period it is only during the Asian crisis, after 9/11, and at the end of the observation period that a contango emerges where all three futures series were more expensive than spot prices.

In addition to the futures prices themselves, the number of long and short positions of the WTI crude oil futures contracts at the NYMEX are used as an explanatory variable. The variable FUT_trade shown in Figure 4(b) is the difference between both long and short positions and may capture expectations of market participants. For all futures market variables positive

[24](We omit to distinguish strong backwardation, weak backwardation, and contango as done in many applied articles, e.g. Considine and Larson (2001), explaining the spread between spot and futures prices.

coefficients are expected since higher futures prices or a higher demand of long positions relative to short positions should raise the price.

2.5 Summary of Hypotheses

Finally, Table 1 summarizes all variables, their meanings, and expected signs of corresponding coefficients discussed in this section. In some cases, this variable set shown in Table 1 is extended. I derive indicator variables or combine these variables in the econometric analysis in Sections 3 and 4. In addition to the variables described above, four indicator variables measuring calendar effects or specific events have been added.

Table 1: Summary of Hypotheses

Category	*Variable*	*Description*	*Sign*
Demand Variables			
Q	Q_glob	global crude oil demand in mbd	+
	Q_oecd	crude oil demand in mbd of OECD	+
	Q_indi	crude oil demand in mbd of India	+
	Q_chin	crude oil demand in mbd of China	+
OECD	OECD_days	ratio of OECD's oil stocks divided by Q_oecd	+/−
Supply Variables			
PROD	PROD_rig	Number of rigs	+
	PROD_vlcc	Charter prices measured in thousand USD per day of very large crude-oil carriers (VLCC)	+
	PROD_fref	Measures the available free global refining capacity	−
OPEC	OPEC_capu	OPEC supply relative to OPEC's production capacity	+
	OPEC_quota	OPEC's concerted production in mbd	−
	OPEC_cheat	OPEC production in mbd minus OPEC_quota	−
Futures Market Variables			
FUT	FUT_trade	Long minus short positions at NYMEX	+
	FUT1_q	1 month futures q month lagged	+
	FUT2_q	2 month futures q month lagged	+
	FUT4_q	4 month futures q month lagged	+
Indicator Variables			
I	I_spring	Indicates spring months - March, April, May	+/−
	I_summer	Indicates summer month - June, July, August	+/−
	I_event(+)	Indicates positive events on the supply side such as US SPR realeses, Iraqi oil fields were not destroyed during military action in 2003, etc.	−
	I_event(−)	Indicates negative events at the supply side such as hurricane Katrina and Rita, etc.	+

Abbreviations: Q=demanded quantity, PROD=production, FUT=futures variables, and I=indicator variables.

Table 2: Tests Statistics for Annual and Seasonal Unit Roots

Test	ADF_00	HEGY_00	HEGY_00	ADF_06	HEGY_06	HEGY_06
Coeff	γ	π_1	π_2	γ	π_1	π_2
WTI	-0.416	1.085	0.076	-0.005	2.204	-1.173
OECD_days	-2.876$^+$	-1.233	0.233	-4.353**	-2.229	-0.314
OPEC_quota	-2.082	-0.137	-0.463	-1.721	1.071	-2.157*
OPEC_cheat	-1.691	1.170	-0.298	-2.252	0.553	-1.360
OPEC_capu	-1.700	-0.503	-1.100	-2.783$^+$	0.534	-1.130

Own Source: ADF test include four lags. The ADF column tests the full sample while ADF_00 tests for unit roots in the subsample until September 2000. The same holds for HEGY and HEGY_00. For HEGY tests a maximal lag length of 12 months was used. Critical values of t-statistics for π_1 and π_2 are reported in Franses and Hobijn (1997). 1%, 5%, 10% significance levels are labelled by **, *, $^+$.

3 Cointegration Analysis

In this Section the analysis done by Kaufmann et al. (2004) is replicated.[25] First, the degree of integration of the variables is tested. To test for unit roots I performed the Augmented Dickey-Fuller (ADF) test as well as the HEGY unit root test.[26] π_1 in Table 2 tests for unit roots, whereas π_2 also tests for seasonal unit roots. The null hypothesis in the ADF and HEGY is that each series is nonstationary. Since most test statistics do not reject the null, most series are nonstationary. In addition, to compare the results with the one found by Kaufmann et al. (2004) unit roots were also tested for a subsample lasting until September 2000.[27] I used the suffix _00 in this Section to indicate the restricted sample and _06 for the full sample.

Results shown in the columns OLS_00, DOLS_00, and FIML_00 in Table 3 are quite similar to the results presented in Kaufmann et al. (2004). Except for OPEC_cheat, the signs of the coefficients are the same and I also find significant coefficients. Additionally, the adjustment rate α is negative and significant only in the FIML_00 specification, whereas in Kaufmann et al. (2004) all α are significant. However, the full-sample regressions shown in the _06 columns report different results. In particular, the coefficients of OPEC variables change their sign. Furthermore, the adjustment rate is positive in two regressions.[28]

[25] Therefore, in Section 3 econometric papers explaining estimation methods are not mentioned here but can be found in Kaufmann et al. (2004).

[26] Hylleberg et al. (1990).

[27] The quarterly data set of Kaufmann et al. (2004) lasts until the third quartal 2000.

[28] All lag lengths are chosen using the Akaike information criterion.

Furthermore, the results in both samples strongly depend on the chosen specification. For example, I found no clear evidence for stationary OLS and DOLS residuals, which indicates that no cointegration relation might exist. The rejection of these unit roots depends on the chosen unit root test as well as on the number of lags included. Hence, these tests are not robust to different specifications.[29] Also, the results of the vector error correction model (VECM) are not robust to other specifications. For the 00-sample, the null of no cointegration is clearly rejected[30] by the trace statistic $\lambda_J = 78.88$ and $\lambda_{SL} = 63.61$. However, the second but not the third null hypothesis can also be rejected, indicating a cointegration rank of two.[31] Moreover and again, the cointegration rank estimated by the trace statistics depends on the number of chosen lags. The estimation results reported in the FIML_06-column in Table 3 are performed under the assumption of a single cointegration rank, though again both rank tests do not clearly confirm this number of cointegration relations for the full sample.[32] The results are different relative to the findings of Kaufmann et al. (2004) and relative to the FIML_00 column. In particular, the signs of the OPEC_capu and OPEC_quota coefficient are reversed.

The different results found for both samples are not surprising because, as shown in Figure 1, the WTI time series is relatively stable until 2000. Subsequently, the oil price was strongly boosted and has at least doubled. To test for structural differences before and after 2000 I perform Chow tests after the cointegration regression of the full sample. Both the Chow sample split and the Chow break point test indicate structural breaks as shown in Figure 5. The Chow break point test indicates a high instability of the crude oil price model, whereas the Chow sample split in particular suggests the end of 2001 as possible structural break point. Hence, the most sensible timeframe is the period after September 2001. The 9/11 event definitely changed the world and rattled financial markets. Moreover, the Chow break point test indicates a greater degree of stability in 2002. For the development of an econometric oil price model

[29] For instance, Phillips-Perron tests usually reject the null for most lag length and a cointegration relation is implied. Yet, other unit roots fail to reject the null.

[30] Cf. Johansen (1995), Saikkonen and Lütkepohl (2000) abbreviated by J and SL.

[31] The corresponding p-values for the Johansen test with one lag are p(r=0)=0.0341, p(r=1)=0.0812, p(r=2)=0.2365 and for the Saikkonen and Lütkepohl test p(r=0)=0.0231, p(r=1)=0.0563, p(r=2)=0.2851. In both tests the chosen lag length is one due to Hannan-Quinn and Schwarz criterion whereas the Akaike criterion proposes six lags. Maximal lag length was six. I included a constant but no other deterministic terms.

[32] I found either two or three cointegration ranks independent of the chosen test (Johansen vs. Saikkonen and Lütkepohl and independent of the specification constant or 'constant and trend').

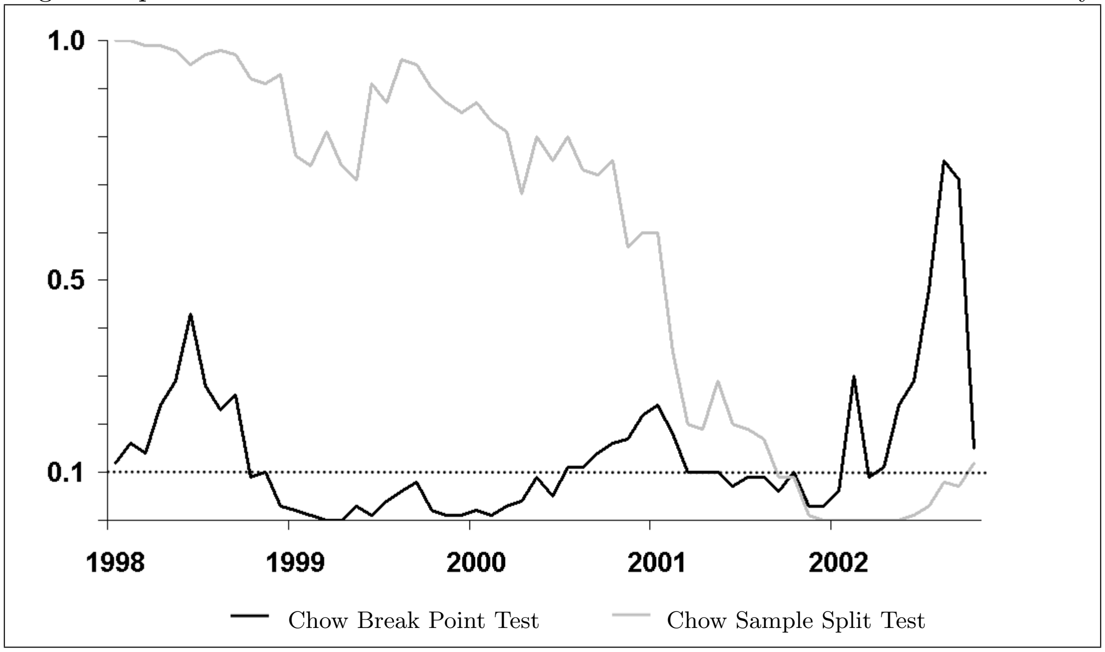
Table 3: Estimates for Price Equation

Dependent Variable: P (Crude Oil Price)

Test	OLS_00	DOLS_00	FIML_00	OLS_06	DOLS_06	FIML_06
Cointegrating Relation						
Constant	90.61**	122.48**	-83.92**	40.88	141.25^{+}	-309.63**
	(3.66)	(7.42)	$\chi^2(1)$ 17.96	(0.77)	(1.90)	$\chi^2(1)$ 17.54
OECD_days	-0.96**	-1.37**	1.03**	-1.20**	-2.16**	2.67**
	(4.42)	(6.57)	$\chi^2(1)$ 279.46	(2.87)	(2.84)	$\chi^2(1)$ 16.08
OPEC_capu	58.32*	80.71**	-214.17**	-92.61*	-120.65*	185.48**
	(2.16)	(3.85)	$\chi^2(1)$ 41.52	(2.27)	(2.61)	$\chi^2(1)$ 11.28
OPEC_cheat	1.54^{+}	1.37*	13.19**	0.03	-2.51*	10.21**
	(1.74)	(2.59)	$\chi^2(1)$ 87.07	(0.02)	(1.27)	$\chi^2(1)$ 26.96
OPEC_quota	-2.15^{+}	-2.84*	6.21**	6.69**	7.12**	-4.68**
	(1.79)	(2.63)	$\chi^2(1)$ 49.59	(4.27)	(4.57)	$\chi^2(1)$ 9.04
R^2	0.491	0.878		0.523	0.683	
$\overline{R}^2$	0.420	0.788		0.500	0.605	
DW	0.996	1.245		0.385	0.497	
Short Run Dynamics						
Adjustment rate (α)	-0.035	-0.288	-0.57**	0.062*	0.033	-0.03**
	(0.58)	(1.26)	(11.33)	(2.02)	(0.84)	(1.42)
Number of Lags(s)	0	6	6	3	3	3

Own Source: OLS and DOLS estimation routines use Newey-West standard errors. 1%, 5%, 10% significance levels are labelled by **, *, $^{+}$. Absolute t-values are in parentheses below the coefficients. $\overline{R}^2$ is the adjusted R^2 and DW is the Durbin-Watson statistic. Cointegration coefficients are normalized such that the coefficient of the lagged price is one. The coefficients and statistics of the indicator variables (quartals and specific events) are not reported. The number of lags(s) is chosen using the Akaike information criterion. The maximal lag length was restricted to six. For DOLS two leads and two lags are included. VECM are estimated under the assumption of a single cointegration relation. The coefficient of the lagged price variable is 1.00. For both VECMs an intercept and a trend are included as a deterministic term whereas the error correction term includes only a constant.

Figure 5: p-Values of Structural Break Tests under the Null of Parameter Constancy



Own Source: The ordinates show bootstrapped p-values. Points below the 0.1 threshold imply structural breaks. Observations before and after the shown time span are excluded in the structural break analysis due to the lack of degrees of freedom.

explaining the recent market situation I neglect the data before January 2002 and focus on the 49 observations[33] until January 2006.[34]

The Chow tests indicates that the assumption of parameter constancy is not maintainable. After the reduction of the sample, I continued the level analysis. However, my findings were not convincing due to the lack of robustness. Several cointegration relations, Granger causality tests, and other standard time series methods were tested but failed to supply an adequate specification. Similar to the results reported in Table 2 and Table 3 all estimations strongly depend on the chosen lag length, and small model specifications considerably changed the results. Therefore, I turn to the analysis of a market model in the next section.

[33]Note, the complete sample of Kaufmann et al. (2004) had also around 50 observations.

[34]Under the assumption that September 2001 was an important date which changed the market structure Kaufmann et al. (2004) cannot be used for today's analysis. Also, the extension Dees et al. (2005) using quarterly data until the end of 2002 is then not applicable to current market conditions.

4 Market Model

Here, a market model for the demand and supply of oil is specified. To circumvent spurious regression results first differences of variables are used as covariates. All variables were found to be integrated of order one. The analysis of differenced series can be seen as a drawback relative to a level analysis. However, it is always possible to reconstruct the level series from the differences. The main reason for choosing a model in first differences is its simplicity which allows us to understand and extend the specifications. Last but not least, the explanatory power of the model makes me very confident that a well-founded research model has been established.

The demand and supply functions are modelled as price and quantity equations. First, I argue that the current demand for oil is independent of the current oil price. Hence, the simultaneity issue must not be taken into account here. Subsequently, at the end of this section this assumption is justified by applying appropriate techniques. More importantly, sound economic reasoning supports the assumption. If production plans of companies budget oil as an input factor, then the optimal moment of purchase might be considered but not the trade per se. Similarly, many commuters have no choice and pay almost any fuel price to get to work. In the long-run firms and households will reduce the demand for oil by buying more efficient machineries and cars or by simply replacing petrol engines. Therefore, the oil demand might depend on past demand variables to control for certain demand patterns. In addition and as shown in Equation (1), the indicator variable I_spring was found to be significant. In the months March, April, and May the demand for oil was significantly smaller than in the rest of the year. The indicator function I_events(+) is also strongly significant. The quantity function is estimated by OLS, where I used lags of one and twelve months. OLS is applicable here because

$$\Delta Q\_glob_t = \underset{(5.253)}{0.788} - \underset{(3.145)}{0.320}\, \Delta Q\_glob_{t-1} + \underset{(2.735)}{0.417}\, \Delta Q\_glob_{t-12} - \underset{(2.782)}{2.408}\, I\_event(+) - \underset{(4.621)}{2.040}\, I\_spring \quad (1)$$

$$N = 49, \qquad R^2 = 0.604, \qquad \overline{R}^2 = 0.568, \qquad DW = 2.57 \qquad |t - values|\ in\ ()$$

lagged variables and indicator variables are predetermined and endogenous variables are not present.[35]

[35] Notice, the indicator variables increase the fit but change the results only slightly. Without the indicator variables, the coefficients are -0.254 (2.492) for ΔQ_glob_{t-1} and 0.755 (5.868) for ΔQ_glob_{t-12}.

The inclusion of the lagged demand variables is preferable from an economic point of view but due to the presence of autocorrelation might produce poor econometric results. The Breusch-Godfrey test clearly rejects the null of no-autocorrelation for any number of lagged errors. In general, this situation can lead to inconsistent coefficients. Therefore, I estimate the model without the lagged variables to test the impact on the indicator variables. The coefficients

$$\Delta \text{Q\_glob} = \underset{(5.873)}{0.942} - \underset{(14.372)}{2.686}\, I\_events(+) - \underset{(10.495)}{2.730}\, I\_spring$$

$$N = 49, \qquad R^2 = 0.425, \qquad \overline{R}^2 = 0.400, \qquad DW = 2.76 \qquad |t - values|\ in\ ()$$

of the second model are similar to the first specification since the signs of the indicator variables are unaffected and magnitudes of coefficients and p-values change only slightly. Therefore, at least the regression coefficients of the indicator variables seem quite reliable. Due to the presence of negative autocorrelation - as opposed to positive autocorrelated residuals - I can hope that the error is negligible.

In addition to the reported specifications, I tested several different lag structures for ΔQ_glob, however, the best fit was found for lags of one and twelve months. The significant and negative coefficient of ΔQ_glob_{t-1} and positive coefficient of ΔQ_glob_{t-12} show that changes in demand are negatively influenced by demand changes in the preceding month and positively influenced by demand changes a year previously. A reasonable interpretation of this finding is that potential demanders have some choice whether they fill their oil stocks in this or the preceding month, which might explain the negative coefficient of ΔQ_glob_{t-1}. The one-year effect might indicate that firms as well as some household current plans often rest upon previous year's plans. If most demanders in an economy act in accordance with this assumption and plans are typically specified at a definite month of a year then a positive coefficient of ΔQ_glob_{t-12} is likely to emerge.

Price Equation

After the specification of the quantity equation I turn to the development of a price model. Before I present the estimation results, there is a description of main findings. One of the main driving forces is the increased number of speculators who are not interested in the commodity itself, but in buying low and selling high. As shown in Figure 4(b), futures traders heavily increased their long positions, thereby boosting the futures price, and contemporaneously the spot price rises. However, I believe that speculators are attracted by economic factors and not by unfounded gambling. In particular, there are several possible paucities in the production process of oil. This risk is latent due to political instability in the Arab and Persian World as well as the impending lack of free capacities. The increasing oil demand of large economies such as China and India are per se of minor importance for the observed upward trend in recent years. However, the rather small increase in the global oil demand due to the economic growth in these emerging countries may have utilized large parts of the otherwise free refining capacities, as shown in Figure 3(b). Hence, although the produced crude oil satisfied the crude oil demand, there was the risk of a supply scarcity of refined products. At the end of the observation period global oil demand exceeded the available refining capacities. Hence, this paucity of refined products increased their market value and, due to the high cointegration of refined products and crude oil (cf. Gjolberg and Johnsen (1999), Asche et al. (2003), Hammoudeh and Li (2004), among others) the WTI price series also rose.[36]

Figure 6 underpins this argumentation where the utilitsation of refining capacity is restricted to 100%.[37] These utilisation diagrams indicate how often refining facilities were working to capacity. In particular, in recent months the refining process seems to have been a bottleneck. It is obvious that the crude oil price increased when the refining capacity utilisation was around 100%. Subfigure (b) shows also that the OPEC production capacity was utilised by around 95% and that prices heavily increased before the capacity limit have been reached. This might indicate that the influence of this OPEC variable has diminished. Furthermore, it seems reasonable that this scarcity was easily predictable due to the high investment costs necessary for installing new refineries. Hence, institutional investors had incentives to invest in the oil market and bought long futures contracts.

The following regression analyses may be even more convincing than the verbal arguments and the data description. All results use the WTI price series as dependent variable. The first regressions results of specification (A1) is shown in Table 4. The first differences of global oil demand Q_glob and the first difference of the variable FUT_trade is included. Furthermore, three

[36] It may be argued that the crude oil price rise due to a higher world crude oil demand. However, the world crude oil demand is a result of the increased demand for refined products.

[37] At the start of 2006 the lack of refining capacity was quickly remedied and production was extended to 85.127 mbd. Hence, given the linear interpolation of the yearly refining capacity some of the values in 2005 and 2006 might be error-prone. Therefore, I restrict the capacity utilisation to 100%.

Table 4: Modelling the Price Equation

Dependent Variable: ΔWTI (Ordinary Least Squares)

Variable	(A1)	(A2)	(A3)	(B)	(C1)	(C2)
const	-0.908^{+}	-0.373	-0.184	-1.001*	-0.944^{+}	-0.931
	(1.83)	(0.98)	(0.52)	(2.10)	(1.74)	(2.01)
ΔQ_glob	-0.466*	-0.420^{+}	-0.487*	-0.434^{+}	-0.491*	-0.492*
	(2.19)	(1.95)	(2.18)	(1.91)	(2.27)	(2.33)
ΔFUT_trade	0.793**	0.687**	0.661**	0.787**	0.822**	0.820**
	(6.03)	(4.68)	(4.68)	(5.99)	(5.33)	(5.75)
PROD_I97ref	2.132**			1.879**	2.13**	2.114**
	(3.72)			(3.32)	(3.40)	(4.00)
PROD_I98ref		1.686**				
		(3.22)				
PROD_I98ref			1.714*			
			(2.62)			
I_event(+)	-2.249**	-1.868	-2.335^{+}	-1.948*	-2.458**	-2.444**
	(2.84)	(1.50)	(1.94)	(2.14)	(2.99)	(3.07)
I_event(−)				0.956^{+}		
				(1.90)		
I_summer	1.906**	1.609*	1.912*	1.858**	2.000**	1.965**
	(2.95)	(2.34)	(2.25)	(3.17)	(2.86)	(3.23)
ΔOPEC_capu					0.248	0.201*
					(1.04)	(2.146)
ΔOPEC_cheat					-0.164	
					(0.20)	
ΔOPEC_quota					-0.18	
					(0.19)	
N	49	49	49	49	49	49
R^2	0.601	0.571	0.562	0.619	0.615	0.615
$\overline{R}^2$	0.555	0.522	0.511	0.564	0.538	0.560
DW	2.182	2.145	1.905	2.301	2.270	2.270

Own Source: *Q_glob* is the traded quantity, FUT_trade is the difference between long and short contracts on the futures market in 10.000 contracts, PROD_I97ref (_I98ref, _I99ref) is an indicator variable which is one if the utilized refining capacity is larger than 97% (98%, 99%), I_events(+) is an indicator variable for positive and I_events(−) an indicator variable for negative events. OPEC_capu is the utilized OPEC capacity in percent, OPEC_cheat is OPEC's production above arranged levels, OPEC_quota the arranged production level among OPEC members. N is the number of observations, $\overline{R}^2$ is the adjusted R^2 and DW is an abbreviation for the Durbin-Watson Statistic, absolute t-values in (). 1%, 5%, 10% significance levels are labelled by **, *, $^{+}$.

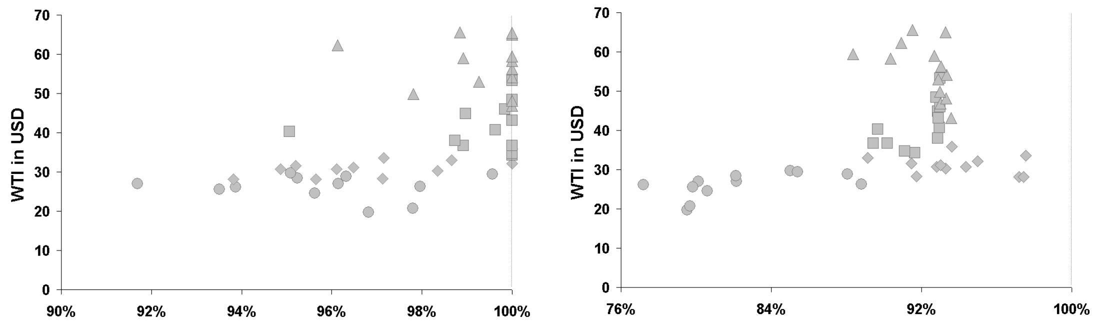
Table 5: Modelling the Price Equation (cont.)

Dependent Variable: ΔWTI (Ordinary Least Squares)

Variable	(D1)	(D2)	(E1)	(E2)
const	-0.776^{+}	-0.390	-0.492	0.318
	(1.77)	(0.83)	(1.36)	(0.57)
ΔQ_glob	-0.431*	-0.409*	-0.412*	-0.379*
	(2.24)	(2.08)	(2.34)	(2.36)
ΔFUT_trade	0.743**	0.742**	0.774**	0.654**
	(4.84)	(5.12)	(4.58)	(3.86)
PROD_I97ref	2.05**	2.389**	1.754**	1.545**
	(4.32)	(3.70)	(4.33)	(3.66)
I_event(+)	-2.244**	-2.018*	-1.881*	
	(3.23)	(2.27)	(2.25)	
I_summer	1.977**	1.787**	1.781**	1.779**
	(3.48)	(3.42)	(3.99)	(3.52)
ΔOPEC_capu	0.196*	0.172*	0.327**	0.279*
	(2.28)	(2.09)	(2.79)	(2.40)
ΔPROD_rig	-0.043*	-0.043*	-0.049*	-0.063**
	(2.11)	(2.15)	(2.54)	(3.08)
ΔPROD_vlcc		-0.010		
		(0.75)		
ΔSPREAD1_2			-6.355**	-7.324**
			(3.01)	(3.19)
FUT1_Icont_6				-1.216^{+}
				(1.78)
N	49	49	49	49
R^2	0.658	0.664	0.707	0.724
$\overline{R}^2$	0.599	0.597	0.648	0.668
DW	2.300	2.343	2.364	2.297

Own Source: PROD_rig is the number of rigs, PROD_vlcc charter prices of very large crude oil carriers, SPREAD1_2 is the two-month lagged difference between spot price and one month futures, and FUT1_Icont_6 is a six month lagged indicator variable which indicates a contango situation at the futures market. 1%, 5%, 10% significance levels are labelled by **, *, $^{+}$.

Figure 6: Price-Diagrams of Capacity Utilisation Variables



(a) Refining Capacity Utilisation (b) OPEC Capacity Utilisation

○ 2002 ◇ 2003 □ 2004 △ 2005+

Own source: (a) Refining capacity utilization is the global demand relative to refining capacity in mbd. (b) OPEC capacity utilization is calculated as OPEC production relative to OPEC's capacity in mbd.

indicator variables PROD_I97ref, I_event(+), and I_summer are used as covariates. The variable PROD_I97ref is an indicator variables which is one if the utilized refining capacity is above 97%. The indicator variable I_event(+) captures specific occurrences such as the information that Iraqi oil fields had not been destroyed as had been feared during military actions in March 2003 or the SPR release authorized by President Bush. All covariates are significant in specification (A1). The negative sign of Q_glob indicates that rising crude oil markets were not directly driven by the demand side. FUT_trade shows that if the difference of long minus short positions increase in the futures market then the oil price is positively affected. This sales variable is preferable to futures price variables since spot and futures price series are closely coupled due to directly observable convenience yield and market volatility (cf. Considine and Larson (2001), Bahram et al. (2006), among others). The positive coefficient of PROD_I97ref indicates rising prices if the utilisation of refining capacity was above 97%. Positive events reduce the price and rising prices were especially observed in the summer months.[38]

In specification (A2) and (A3) I replace PROD_I97ref by PROD_I98ref or PROD_I99ref. Again, the coefficient of the refining capacity variable is positive and strongly significant. The magnitude of these coefficients decreases relative to specification (A1). A reasonable interpretation might be that it is very easy to predict the impending lack of refining capacity

[38]Notice, the estimation without the indicator variables change the results only slightly. The results are -0.291 (1.468) for ΔQ_glob, 0.830 (6.248) for ΔFUT_trade, and 1.731 (2.848) for PROD_I97ref. Thus, the refining capacity and the futures variable are still significant.

and that the actual further reduction of free refining capacity from three percent to even less was expected and caused no larger price change.

To test the effect of hurricanes and local supply disruptions the variable I_event(–) was also included in specification (B). The positive coefficient shows that these negative events had an effect but due to a low significance level and a small magnitude I omit this variable in other specifications.

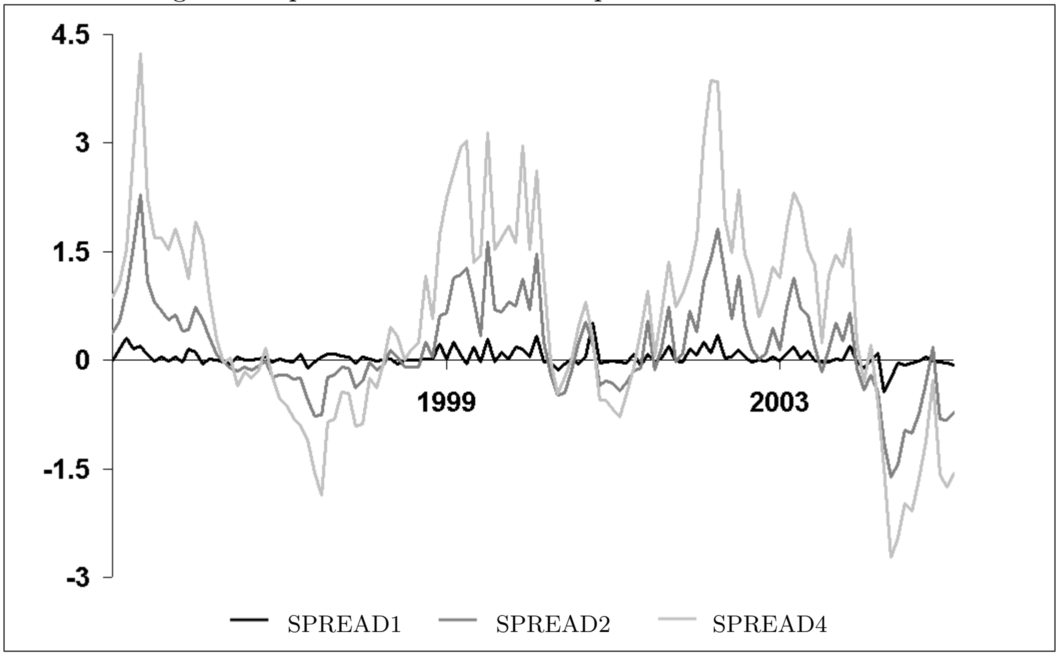
In specification (C1) and (C2), I continue the analysis of OPEC variables of Section 3. All coefficients of the OPEC variables show the expected signs. Yet, if all OPEC variables are included no significant effect is detectable in (C1). However, due to relatively high correlations among these variables I can find a positive effect of the variable OPEC_capu in (C2), which indicates that an increasing capacity utilisation significantly increases the spot crude oil price. However, the overall effect of this variable is small, as shown below, and Figure 6(b) signalizes that OPEC's production was not working to capacity. It is possible that the significant variable hints at the high sensibility of market participants who have a close eye on all aspects of production capacities. Therefore, I keep this variable in further investigations reported in Table 5.

Additional production variables such as PROD_rig and PROD_vlcc are tested in specifications (D1) and (D2). No influence was found for the variable PROD_vlcc whereas PROD_rig was significant. The negative coefficient implies a lower WTI spot price if more drilling rigs have been active. The interpretation of this finding might be similar to other findings at the production side if a larger number of rigs is active the risk of shortages due to wars or natural disasters might be smaller.

In the last two specifications[39], the influence of futures prices is more closely investigated. I found lagged variables which significantly impinge on the crude oil price. In columns (E1) and (E2) I document the most important findings. First, the highly significant variable

[39] I tested several other specifications and variables. For instance, macroeconomic variables such as GDP and FX variables of large and important nations and currencies were included. I also tested the impact of OECD_stock, OECD_days, as well as several variables which capture information on US petroleum stocks (see Footnote 13). However, the quality of the models improved in neither case. For the sake of brevity, the results are not reported here.

Figure 7: Spread in USD of WTI Spot minus Futures Prices



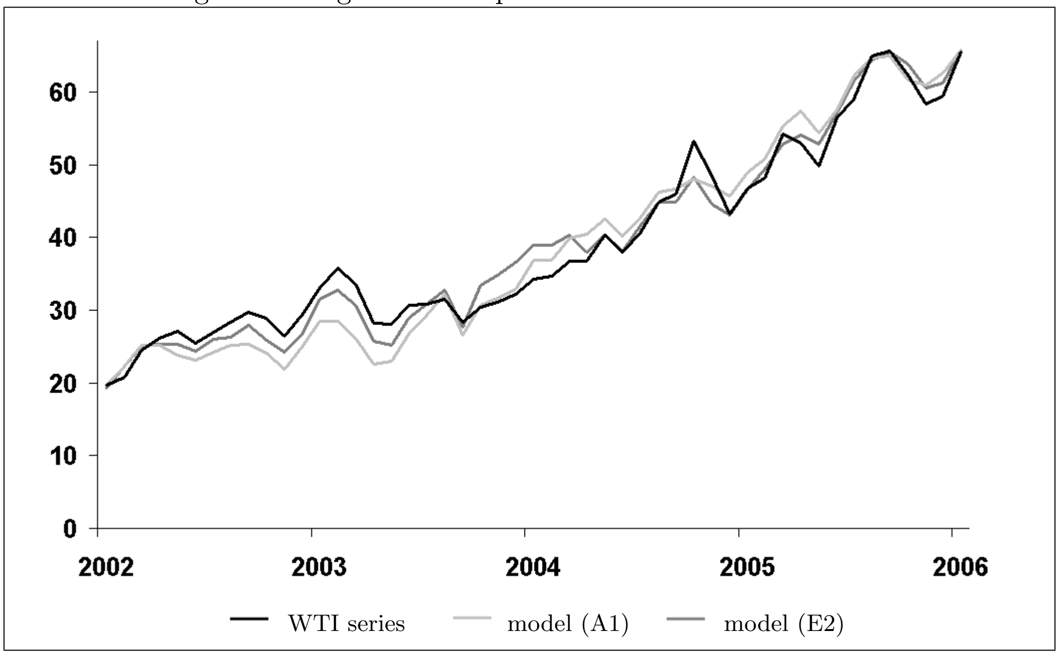
Own Source: SPREADp is the difference of the WTI spot price and the variable FUTp.

ΔSPREAD1_2 is the two-month lagged spread of the spot price minus the one-month futures price. Given the negative coefficient, if two months ago the futures price was above the spot price the variable was negative and the current spot price rises. The large magnitude of the coefficient is a result of the small spreads. If spreads between the spot price and other futures series were used as covariates, the magnitude of the coefficients was still significant but much smaller since the spreads were larger as shown in Figure 7. Finally, I also found a strong negative impact of the indicator variable FUT1_Icont_6 which indicates contangos, i.e. situations where the futures price is above the spot price. This six-month lagged indicator variable significantly reduces the crude oil price. Both futures variables indicate the interplay of the convenience yield and market volatility on the spot price. Obviously and given the theoretical considerations in the literature, a lagged large convenience yield and a high market volatility both seem to have predicting power for future spot prices. Notice that the refining capacity variables are less important if the contango variable is included and the coefficient of the constant also becomes positive. I can therefore conclude that it is not only the number of long- vs. short futures contracts which are of importance in determining the crude oil price development, but also the futures price variables.

To get a better understanding of the models I depicted the estimated price movements of model (A1) and (E2) in Figure 8. Both estimated series closely replicate the original series fairly well. Furthermore, so far I have identified the statistical significance of several variables, however, the magnitude or economic significance of each variable on the crude oil price was not evaluated. To gain intuition on the functioning of the model, a specific month and the decomposition implied by the estimated price model is discussed. In January 2006, the crude oil price jumped from 59.41 to 65.48 USD. Model (A1) in Table 4 calculates a value of 62.62 in December and 65.79 in January and model (E2) compute 61.23 and 65.77. Therefore, the observed price difference is 6.07 and the estimated price difference is 3.17 USD in model (A1) and 4.54 in model (E2). Given these specifications, the price differences are decomposed as shown in Table 6. This decomposition was conducted for the whole sample and the results can be interpreted as follows. If all effects for model (A1) are decomposed I find that ΔFUT_trade and the refining capacity variable mainly determine the WTI price change. Correspondingly, if large spot price jumps occurred, they were mainly caused by large movements in the futures variable. In particular, large price drops are strongly driven by ΔFUT_trade. The same analysis for model (E1) is more complicated due to the higher number of variables. However, if I summarize all effects of futures variables, i.e. ΔFUT_trade, ΔSPREAD1_2 and FUT1_Icont_6, and of all production variables, i.e. PROD_I97ref, ΔOPEC_capu, and ΔPROD_rig, then the same statements already mentioned for model (A1) apply again. Accordingly, all three futures variables contribute similarly to the development of the WTI spot series whereas the effect of production variables are mainly driven by the refining capacity variable.[40]

[40]The strong influence of PROD_rig in Table 6 was rather atypical.

Figure 8: Original WTI Spot Price and Estimated Series



Own Source: Corresponding regression results can be found in Table 4 and Table 5.

Table 6: Decomposition of Effects for January 2006

	Observations	Coefficients		Product	
Variable	Jan 2006	(A1)	(E2)	(A1)	(E2)
const	1	-0.908	0.318	-0.908	0.318
ΔQ_glob	-1.668	-0.466	-0.379	0.777	0.632
ΔFUT_trade	1.467	0.793	0.654	1.163	0.959
PROD_I97ref	1	2.132	1.545	2.132	1.545
I_event(+)	0	-2.249		0	
I_summer	0	1.906	1.779	0	0
ΔOPEC_capu	-2.004		0.279		-0.559
ΔPROD_rig	-43		-0.063		2.709
ΔSPREAD1_2	-0.020		-7.324		0.146
FUT1_Icont_6	1		-1.216		-1.216
Sum(=Price Diff.)				3.165	4.534

Own Source: The observations are the values observed for the corresponding variables in January 2006. The coefficients are the results reported in Table 4 and Table 5. Subsequently, the product of both observation values and coefficients is calculated to identify the most important effects.

Model Checking

I stressed that OLS is applicable in the price model. However, this statement is questionable if endogenous variables are used as covariates. There are two possible sources of simultaneity in model (A1). First, current quantity might be affected by current price movements. Second, FUT_trade might not only impinge on the oil price but might also be determined by the price series itself. All other variables in this system are exogenous since they are indicator variables capturing specific events or seasonal effects. To explain the variable ΔFUT_trade I use the following equation

$$\underset{}{\Delta\text{FUT\_trade}} = \underset{(1.275)}{0.317} + \underset{(3.687)}{0.437}\,\Delta WTI - \underset{(2.990)}{0.213}\,\Delta FUT1\_1 - \underset{(2.747)}{0.202}\,\Delta FUT1\_2 - \underset{(4.140)}{0.279}\,\Delta FUT1\_3 \quad (2)$$

$$N = 49, \quad R^2 = 0.596, \quad \overline{R}^2 = 0.559, \quad DW = 2.11 \quad |t - values|\ in\ ()$$

I estimate the reduced forms of equation (1) and (2) and use the residuals of both equations to perform a Hausman endogeneity test for model (A1) and (E2). Subsequently, an F-test clearly rejects the null of no endogeneity.[41] To correct for simultaneity in overidentified regression models I apply two stage least squares (2SLS) developed by Theil (1953) and Bassmann (1957). In Table 7 OLS and 2SLS regression results are compared. All exogenous variables of all equations are used as instruments in the 2SLS procedures. In both specifications the 2SLS coefficients are very similar to the OLS coefficients. However, in the 2SLS estimation process several variables become insignificant. The coefficients of the variables Q_glob, FUT_trade and PROD_I97ref are significant independent of the method applied. Therefore, the findings support the story set out above. Furthermore, the Breusch-Godfrey tests for autocorrelation usually failed to reject the null. Diagnostic checks such as Jarque-Bera tests on the residuals as well as Ramsey's (1969) specification error test always confirmed the validity of the chosen specifications. Finally, also multicollinearity is not an issue here. Hence, I am quite confident that the model is well-specified and captures the development of the crude oil price.

[41] Correspondingly, I can - as mentioned at the start of Section 4 - confirm that the demand variable seems to be exogenous whereas the endogeneity problem arises due to the futures variable.

Table 7: Inspection of Simultaneity

Dependent Variable: ΔWTI

(A1)	(A1)		(E2)	
Variable	OLS	2SLS	OLS	2SLS
const	-0.908^{+}	-0.915^{+}	0.318	0.105
	(1.83)	(1.73)	(0.57)	(0.14)
ΔQ_glob	-0.466*	-0.475*	-0.379*	-0.493*
	(2.19)	(2.35)	(2.36)	(2.37)
ΔFUT_trade	0.793**	0.794**	0.654**	0.704**
	(6.03)	(4.37)	(3.86)	(2.74)
PROD_I97ref	2.132**	2.143**	1.545**	1.684**
	(3.72)	(3.64)	(3.66)	(3.60)
I_event(+)	-2.249**	-2.273**		
	(2.84)	(3.01)		
I_summer	1.906**	1.916**	1.779**	1.903**
	(2.95)	(2.91)	(3.52)	(3.73)
ΔOPEC_capu			0.279*	0.342
			(2.40)	(0.68)
ΔPROD_rig			-0.063**	-0.054^{+}
			(3.08)	(1.76)
ΔSPREAD1_2			-7.324**	-7.585*
			(3.19)	(2.43)
FUT1_Icont_6			-1.216^{+}	-1.069
			(1.78)	(1.09)
N	49	49	49	49
R^2	0.601	0.601	0.724	0.717
$\overline{R}^2$	0.555	0.555	0.668	0.661
DW	2.182	2.188	2.297	2.373

Own Source: Q_glob is the traded quantity, FUT_trade is the difference between long and short contracts on the futures market in 10.000 contracts, PROD_I97ref is an indicator variable for the refining capacity, I_events(+) an indicator variable for positive events, and I_summer an indicator variable for summer months. OPEC_capu is the capacity utilisation of OPEC countries, PROD_rig counts the number of oil rigs, SPREAD1_2 is the two-month lagged spread between the WTI series and the one-month futures, and FUT1_Icont_6 is an indicator variable for contangos. N is the number of observations, $\overline{R}^2$ is the adjusted R^2 and DW is an abbreviation for the Durbin-Watson Statistic, absolute t-values in (). 1%, 5%, 10% significance levels are labelled by **, *, $^{+}$.

5 Conclusion

Unlike Kaufman et al. (2004), I cannot confirm a strong market power of OPEC in recent years. I only find a modest influence of OPEC's capacity utilisation on the WTI series. This finding possibly hints at the high interest in all production variables rather than at OPEC's market power. OPEC was more a passive observer than a price setter. In contrast, the estimation results imply that the upward trend at the spot market can be explained by an increasing crude oil demand of emerging markets. The additional demand meets the constrained refining capacity such that nearly no free refining capacities were available. Hence, an additional demand or a loss of refining capacity had caused an excess demand. Due to this potential threat investors steadily increased their demand for long WTI futures contracts and contemporaneously spot prices were pushed up.

Given these results, the question arises as to what caused this paucity of refining capacity. Moreover, it might be enlightening to understand incentives of refinery operators to extend or not extend the refining capacity. One explanation is that due to the 9/11 attacks the petroleum industry feared a declining world economy and postponed large investment projects. Then the economy recovered and the crude oil demand increased due to the rising demand of emerging countries. After this development continued for several months, investments in new refineries were made but these were not operable before the end of 2005. In accordance with this hypothesis, Asano (2002) showed for the US refining industry that a lag of four to seven years exists between economic conditions and the completion of new refinery capacities. However, incentives of refinery operators might also have contributed to this development. Refinery operators might prefer high oil prices and may take into account that possible production scarcities guarantee this aim. Since many refinery operators are also large crude oil producers this line of argument is appealing. Eicher (2006) argues that both fundamental factors, such as large investment costs and incentives, might have hindered a faster increase in global refining capacity.

6 Data Appendix

The Energy Information Administration provided WTI series, futures series, demand, and stock data. PROD_rig is data of Baker-Hughes from the website http://www.bakerhughes.com/investor/rig/index.htm and PROD_vlcc is available at http://www.pfbassoe.no/welcome.htm. Most data was observed each month. Otherwise, data was linearly interpolated. Finally, the refining capacity series was provided by EIA and the Oil&Gas Journal (OGJ), which published figures on the refining capacity in 2005. EIA figures which are based on the OGJ data inform on the annual average, whereas OGJ publishes end-of-year figures. For the end of the observation sample, the values of both sources are summarized in Table 8.

Table 8: Comparison of Refining Capacity Data

Data Source	Average 2004	Begin 2005	Average 2005	Begin 2006
EIA	82.260		82.800	
OGJ		82.409		85.127

Own Source: All figures are measured in mbd. OGJ values represent the situation on 1st January, whereas the REF variables are averages of mid-month values from December and January (or averages of June and July figures).

References

Adelman, M. A. (1993). "Modelling world oil supply." The Energy Journal, 14(1): 1-32.

Allen, M. (2005). "Oil Market Developments and Issues." International Monetary Fund, Prepared by the Policy Development and Review Department, March 1 2005.

Asano, H. (2002). "An empirical analysis of lumpy investment: the case of US petroleum refining industry." Energy Economics 24(6): 629-645.

Asche, F., O. Gjolberg, and T. Völker (2003). "Price relationships in the petroleum market: an analysis of crude oil and refined product prices." Energy Economics 25(3): 289-301.

Bacon, R. and S. Tordo (2004). "Crude Oil Prices - Predicting Price Differentials Based on Quality." The World Bank Group, October 2004, Note Number 275.

Bahram, A., A. Chatrath, F. Song, F. Szidarovszky (2006), Petroleum spreads and the term structure of futures prices, Applied Economics, 38, 16, 1917-1929.

Balabanoff, S. (1995). "Oil futures prices and stock management: a cointegration analysis." Energy Economics 17(3): 205-210.

Basmann, R.L. (1957). "A Generalized Classical Method of Linear Estimation of Coefficients in a Structural Equation." Econometrica 25: 77-83.

Birol, F. and W., Davie (2001). "Oil supply costs and enhanced productivity." in: IEA (Hrsg.). Energy Prices and Taxes, 4. Quartal 2001, S. XVII - XXII, Paris.

Box, G. E. P. and G. M. Jenkins (1970). "Time Series Analysis." Forecasting and Control. Holden Day.

Brook, A.-M., R. Price, D. Sutherland, N. Westerlund, and C. André (2004). "Oil Price Developments: Drivers, Economic Consequences and Policy Responses." Economics Department Working Paper, No. 412, http://www.oecd.org/eco/.

Campbell, C. J., Laherrere, J. (1998). "The end of cheap oil." Scientific American 278(3): 7883.

Campbell, C. J. (2004). "Essence of Oil & Gas Depletion Collected Papers and Excerpts." Multi Science Pub Co Ltd.

Clarida, R., J. Gali and M. Gertler (2000). "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory." The Quarterly Journal of Economics 115(1): 147-180.

Considine, T.J. and D.F. Larson (2001), Uncertainity and the convenience yield in crude oil price backwardations, Energy Economics 23, 533-548.

Curlee, T.R., A.F. Turhollow, and S. Das (1988). "Oil supply disruptions and modelling methodologies." Energy Economics 10(2): 147-154.

Dees, S., P. Karadeloglou, R. K. Kaufmann, and M. Sánchez (2005). "Modelling the World Oil Market - Assessment of a Quarterly Econometric Model." Working Paper, European Central Bank.

de La Granville, O. (1980). "Capital Theory, Optimal Growth, and Efficiency Conditions with Exhaustible Resources." Econometrica 48(7): 1763-1776.

De Santis, R. A. (2003). "Crude oil price fluctuations and Saudi Arabia's behaviour." Energy

Economics 25: 155-173.

Eicher, C. (2006). "Refining Capacity and the Cost of Refined Oil Products." 3 November 2006, mimeo.

Forbes, K. F. and E. M. Zampelli (2000). "Technology and the Exploratory Success Rate in the U.S. Offshore." The Energy Journal, 21(1): 109-120.

Foster, A. J. (1996). "Price discovery in oil markets: a time varying analysis of the 1990-91 Gulf conflict." Energy Economics 18: 231-246.

Franses, P.H. and B. Hobijn (1990). "Testing for seasonal unit roots in monthly data." Econometric Institute Report 9032A, Erasmus University, Rotterdam.

Franses, P.H. and B. Hobijn (1997). "Critical values for unit root tests in seasonal time series." Journal of Applied Statistics 24(1): 25-47.

Gjolberg, O. and T. Johnsen (1999). "Risk management in the oil industry: can information on long-run equilibrium prices be utilized?" Energy Economics 21(6): 517-527.

Gray, L. C. (1914). "Rent under the Assumption of Exhaustibility." The Quarterly Journal of Economics 28(3): 466-489.

Greene, D. L., J. L. Hopson, J. Li (2006). "Have we run out of oil yet? Oil peaking analysis from an optimist's perspective." Economic Policy 34: 515-531.

Griffin, J. M. (1985). "OPEC Behaviour: A Test of Alternative Hypotheses." American Economic Review 75: 954-963.

Haider, G. M. (2000). "World Oil Reserves: Problems In Definition And Estimation." OPEC Review 24(4): 305-327.

Hamilton, J.D. and A.M. Herrera (2004). "Oil Shocks and Aggregate Macroeconomic Behavior: The Role of the Monetary Policy." Journal of Money, Credit, and Banking, 36(2): 265-286.

Hammoudeh, S. and H. Li (2004). "The impact of the Asian crisis on the behavior of US and international petroleum prices." Energy Economics 26: 135-160.

Hornsell, P. and R. Mabro (1993). "Oil Markets and Prices: The Brent Market and the Formation of World Oil Prices." Oxford University Press, Oxford.

Hotelling, H. (1931). "The Economics of Exhaustible Resources." Journal of Political Economy 39(2): 135-79.

Hylleberg, S., R.F. Engle, C.W.J. Granger, B.S. Yoo (1990). "Seasoanl integrationg and cointegration." Journal of Econometrics 44: 215-238.

Jabir, I. (2001). "The shift in US oil demand and its impact on OPEC's market share." Energy Economics 23(6): 659-666.

Johansen, S. (1995). Likelihood-based Infenerence in Cointegrated Vector Autoregressive Models. Oxford: Oxford Univesity Press.

Jones, C. T. (1990). "OPEC behaviour under falling prices: Implications for cartel stability." The Energy Journal 90(11): 117-129.

Kaufmann, R. K. (1995). "A model of the world oil market for Project LINK: Integrating economics, geology, and politics." Economic Modelling 12(2): 165-178.

Kaufmann, R. K., S. Dees, P. Karadeloglou, and M. Sánchez (2004). "Does OPEC Matter? An

Econometric Analysis of Oil Prices." The Energy Journal 25(4): 67-90.

Krichene, N. (2002). "World crude oil and natural gas: a demand and supply model." Energy Economics 24(6): 557-576.

Litzenberger, R.H. and N. Rabinowitz (1995). "Backwardation in the Oil Futures Market: Theory and Empirical Evidence." Journal of Finance 50(5): 1517-1545.

Matthies, K. (2005). "Oil Prices Out of Control?" Intereconomics - Review of European Economic Policy 40(4): 178-179.

Energy Information Administration, US Department of Energy (2006). *Monthly Energy Review.* http://www.eia.doe.gov/emeu/mer/pdf/mer.pdf (download January 3rd, 2007).

Nakamura, D. (2005). "Refiners add 2.7 million b/d of crude refining capcity in 2005." Oil & Gas Journal, Week of Dec. 19: 60-64.

Pindyck, R. S. (2001). "The Dynamics of Commodity Spot and Futures Markets: A Primer." The Energy Journal 22(3): 1-29.

Ramcharran, H. (2002). "Oil production responses to price changes: an empirical application of the competitive model to OPEC and non-OPEC countries." Energy Economics 24(2): 97-106.

Ramsey, J. B. (1969). "Tests for Specification Errors in Classical Linear Least Squares Regression Analysis." Journal of the Royal Statistical Society B 31: 350-371.

Rotemberg, J.J., and M. Woodford (1996). "Imperfect Competion and the Effects of Energy Price Increases on the Economic Activity." Journal of Money, Credit and Banking 28: 549-577.

Saikkonen, P. and Lütkepohl, H. (2000). "Testing for the cointegrating rank of a VAR process with an intercept." Econometric Theory 16: 373-406.

Theil, H. (1953). "Repeated Least-Squares Applied to Complete Equation Systems." The Hague: The Central Planning Bureau, The Netherlands (mimeographed).

Tvedt, J. (2002). "The effect of uncertainty and aggregate investments on crude oil price dynamics." Energy Economics 24(6): 615-628.

Wang, Z. (2003). "Hedonic prices for crude oil." Applied Economic Letters 10: 857-861.

Watkins, G.C.(2006). "Oil scarcity: What have the past three decades revealed?" Economic Policy 34: 508-514.

ISSN: 1438-2733