CHARLES UNIVERSITY IN PRAGUE
FACULTY OF SOCIAL SCIENCES
Institute of Economic Studies

Pavel Šlechta

Speculative versus Fair Price of Crude Oil

Bachelor's Thesis

Prague 2012

Author: Pavel Šlechta
Supervisor: PhDr. Petr Teplý Ph.D.

Academic year: 2011/2012

Bibliografický záznam:

Šlechta, Pavel. Spekulativní versus fundamentální cena ropy, Praha, 2012. Bakalářská práce, Univerzita Karlova, Fakulta sociálních věd, Institut ekonomických studií. Vedoucí bakalářské práce: PhDr. Petr Teplý Ph.D.

Název práce: Spekulativní versus fundamentální cena ropy

Autor: Pavel Šlechta

Institut: Institut ekonomických studií

Vedoucí bakalářské práce: PhDr. Petr Teplý PhD.

E-mail vedoucího bakalářské práce: teply@fsv.cuni.cz

Abstrakt: Bakalářská práce se zabývá problematikou spekulace na trhu s ropou. Toto téma bylo často diskutováno v souvislosti s nárůsty ceny ropy v roce 2008, ale jeho aktualita je podtržena nedávnými opakovanými proraženími ceny 100 USD. V tezi zkoumáme vztah mezi open interest na burze s futures kontrakty New York Mercantile Exchange, poptávkovými a nabídkovými faktory na trhu s ropou a cenou ropy. Na základě analýzy vector error correction modelu s měsíčními pozorováními mezi lety 1994 a 2011 ukážeme, jak může zvýšení open interestu, který je již nyní tvořen z poloviny nekomerčními obchodníky, vést k persistentnímu zvýšení cen ropy. Věříme, že za dlouhodobou rovnováhou mezi cenou ropy a open interestem je riziková prémie, díky níž je možné částečně vysvětlit nárůsty ceny ropy, které nejsou zachyceny změnami poptávky a nabídky, jako například Hubbertův fenomén ropného vrcholu. Součástí práce je také analýza, zda může volatilita cen s ropou zvýšit open interest na trhu s ropou, což by znamenalo, že zvýšená volatilita přiláká více spekulativních obchodníků. Přes potvrzení Granger kauzality však nemůžeme říci, zda má volatilita pozitivní či negativní efekt na open interest.

Klíčová slova: Spekulace, trh s ropou, fundamentální cena ropy

Délka práce: 15 870 slov

Bibliographic entry:

Šlechta, Pavel. *Speculative versus Fair Price of Crude Oil*, Prague, 2012. Bachelor's thesis, Charles University, Faculty of Social Sciences, Institute of Economic Studies. Supervisor: PhDr. Petr Teplý, Ph.D.

Title: Speculative versus Fair Price of Crude Oil

Author: Pavel Šlechta

Department: Institute of Economic Studies

Supervisor: PhDr. Petr Teplý, Ph.D.

Supervisor's email address: teply@fsv.cuni.cz

Abstract: The thesis deals with the topic of speculation on the crude oil market. This topic has been frequently discussed in association with the price hikes in 2008, but since the oil price has recently repeatedly reached levels over USD 100, the topic is still very present. In our thesis we analyze the connection between the increasing open interest on the New York Mercantile Exchange crude oil futures market, the supply and demand factors for the crude oil and the crude oil price. Based on an error correction model analysis of monthly observations between 1994 and 2011, we show how an increase in the open interest, which is currently already comprised by the non-commercial traders by one half, can lead to a persistent increase in the crude oil prices. We believe it is the risk premium on the market which stands for the long-run equilibrium of the open interest and prices. Such a risk premium on the market of crude oil could explain the part of the increase in prices which could have not been captured by the simple supply and demand, as for example the concern about the Hubbert's peak oil. We also test whether the oil price volatility increases the open interest on the market, which would mean that the price volatility could attract more speculative traders. Although we find Granger causality, we cannot conclude a simply positive or negative effect.

Keywords: Speculation, crude oil market, fundamental price of oil

Length of the thesis: 15870 words

Declaration: I hereby declare that I have elaborated the Bachelor’s thesis on my own and I have used only listed sources and references. Furthermore, I have written this thesis only for the purpose of achieving the Bachelor’s degree at IES FSV UK. I acknowledge and agree with lending and publishing of the thesis for studying and research purposes.

In Prague, 18th May, 2012

Pavel Šlechta

Acknowledgments

I would like to thank my supervisor PhDr. Petr Teplý, Ph.D. for his support, valuable advices and comments on my thesis.

Table of Contents

Introduction 1

Part 1: Previous attempts to explain the oil price 2

1. 1 Introduction 2

1.2 The non-structural models, economic theory 3

1.2.1 Hotelling’s model 3

1.2.2 The inventory storage theory 4

1.2.3 The applicability of the models in the real world 5

1.3 Structural supply-demand models 6

1.3.1 Demand for oil 6

1.3.2 Supply 9

1.4 Informal models 14

1.4.1 Speculation 16

1.5 Summary 19

Part 2: Empirical part 20

2.1 Data 20

2.2 Model 22

2.3 Methodology 22

2.4 Estimation 23

2.5 Results 26

2.5.1 Impulse-response analysis 28

2.6 Is the crude oil price volatility attracting more speculative traders? 33

2.7 Forecasting the crude oil price 35

2.8 Summary of the empirical results 36

Conclusion 39

Bibliography 41

Appendix 47

List of Abbreviations:

ARCH – autoregressive conditional heteroskedasticity

boe – barrel of oil equivalent

EIA – U.S. Energy Information Administration

GDP – gross domestic product

I(0) – process integrated of order zero

I(1) – process integrated of order one

i.i.d – independent, identically distributed

IRF – impulse response function

NYMEX – New York Mercantile Exchange

OECD – Organization for Economic Co-operation and Development

OLS – ordinary least squares

OPEC – Organization of the Petroleum Exporting Countries

USA – United States of America

VAR – vector autoregression

VECM – vector error correction model

WTI – West Texas Intermediate

List of Tables

Table 1: Summary for the used data 21
Table 2: Estimated adjusting coefficients α 27
Table 3: Standard deviations in residuals 28
Table 4: Brief summary of recent findings in the research on the crude oil prices 38

List of Figures

Figure 1: The crude oil prices from 1984 to 2011 2
Figure 2: Petroleum consumption by organizations and states 7
Figure 3: Petroleum production by OPEC/non-OPEC and particular states 10
Figure 4: WTI open interest on NYMEX 17
Figure 5: Residuals of the estimated model 25
Figure 6: Linear prediction of the model in levels and differences 26
Figure 7: IRF, impulse: Crude oil price, response: Crude oil price 29
Figure 8: IRF, impulse: Liquid fuel consumption, response: Crude oil price 29
Figure 9: IRF, impulse: OPEC production, response: Crude oil price.................... 29
Figure 10: IRF, impulse: Crude oil production, response: Crude oil price 29
Figure 11: Impulse response analysis of OPEC and non-OPEC production to the liquid fuel consumption... 30
Figure 12: IRF, impulse: Open interest, response: Crude oil price 31
Figure 13: IRF, impulse: Crude oil price, response: Open Interest........................ 31
Figure 14: IRF of an unrestricted model, impulse: open interest, response: crude oil price 32
Figure 15: IRF of an unrestricted model, impulse: crude oil price, response: open interest.. 32
Figure 16: IRF, impulse: Liquid fuel consumption, response: Open interest.......... 33
Figure 17: IRF, impulse: OPEC crude oil production, response: Open interest 33
Figure 18: IRF, impulse: Non-OPEC crude oil production, response: Open interest 33
Figure 19: Out-of-sample forecasting for the crude oil price 35

Introduction

The West Texas Intermediate (WTI) crude oil price rocketed to the all-time high USD 145 in June 2008, only to plummet to the USD 34 at the end of 2008. Some researchers call it a clear speculative bubble, the others quote the rising demand from the developing countries and the tight supply. The true reasons behind the spike are not only an academic question, since the empirical studies generally agree on the negative effect of the oil prices on the world gross economy. Many argue that if the crude oil prices are truly driven up by speculation, then the speculation has a dangerous influence on the global economy and should be regulated. Also the most recent development on the crude oil market has shown that high oil prices are not a history, as we have witnessed WTI prices over USD 100 over relatively longer periods during the Libya war or the tension between Iran and the rest of the world. The discussions about whether it is the speculation which stands for at least part of the increases are thus continuing.

The topic of speculation on the crude oil market will be the key focus of this thesis. First, we will consider the existing theories and models explaining the behavior of the crude oil and outline the main obstacles in modeling the supply and demand factors, as for example the behavior of the OPEC cartel. Furthermore, we will cover the recent research on speculation on the crude oil market and show the difficulties in quantifying the speculation phenomena. Extending on the current stage of knowledge we will focus on the relationship between the increasing open interest[1] on the New York Mercantile Exchange (NYMEX), oil prices and the supply and demand factors. The open interest on the crude oil market has tripled over the last decade, mainly because of the increase in the non-commercial positions[2], which are often suspected to be speculative. We investigate whether such an increase could help explaining the increases in oil prices there where the traditional fundamental models fail, and whether an increase in the open interest can increase the oil prices without any change in the fundamental factors. We also examine whether an increase in oil price volatility can lead to an increase in the open interest on the market.

[1] "Open interest is the total of all futures and/or option contracts entered into and not yet offset by a transaction, by delivery, by exercise, etc." (CFTC 2012)
[2] Non-commercial positions are positions held by agents who are not physically involved with oil.

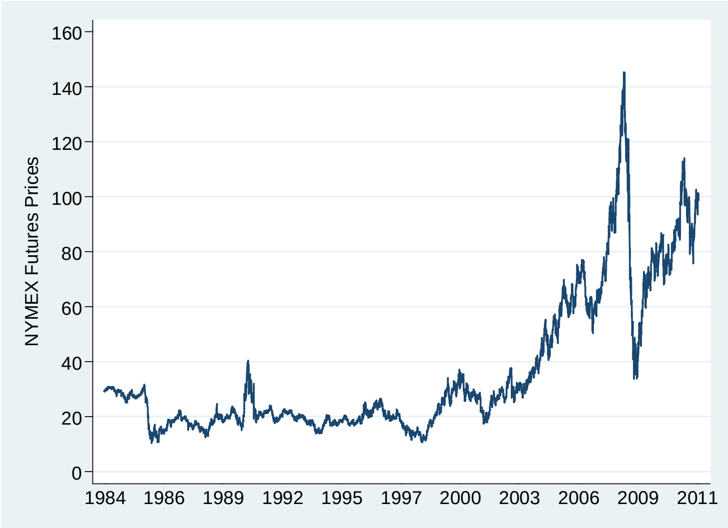


Figure 1: The crude oil prices from 1984 to 2011 (source: author based on the U.S. EIA)

Part 1: Previous attempts to explain the oil price

1. 1 Introduction

The first part of this thesis offers the predominantly used approaches in explaining the oil prices and discusses their weaknesses mainly with respect to the oil price spikes in recent years. The approaches can be divided into three categories, following the work by Fattouh (2007): the non-structural models, the structural models and informal approach models. The non-structural models offer an explanation based on mathematical optimization of oil producers' decisions with the supply and demand shocks usually represented by a random variable. The structural models, also called the supply-demand framework, try to capture the oil market by investigating the relationship of the oil supply and oil demand to the oil price, income and OPEC and non-OPEC production. These models proved the important property of oil – the demand and supply are unable to adjust to even big price changes in the short term. The informal approach includes other factors that are, at least in a short run, suspect to be able to influence the price of oil, as for example lack of spare production capacity or increasing number of speculation on the market. Whereas to prove the significance of production capacity or refinery utilization is statistically easy, the proof of speculation on the oil market can be a Herculean task.

On the one hand, the models based on fundamental parameters were unable to explain the behavior of oil price in the last years. There certainly are reasons to believe that the supply and demand factors were driving the oil price higher, the extent of the price increase has not been justified, though. On the other hand, a failure of fundamental models is not a proof of speculation, and the inefficiency on the oil market is hardly measurable. As none of the variables on its own proved to be able to drive the oil price so high, it is necessary to take into account all of the major variables and analyze their joint power. The first part of the thesis discusses the individual factors in the way they can influence the oil price and serves as the backbone for the second part of the thesis, where we build on the previous research and also use it to interpret our results.

1.2 The non-structural models, economic theory

This section discusses two major economic theories describing the price of oil; Hotelling model of exhaustible resources and the inventory storage theory, which has its origins in works by Kaldor (1939) and Working (1949) (Ederington et al. 2011).

1.2.1 Hotelling's model

The Hotelling's article The Economics of Exhaustible Resources is considered to be the starting point in the economic theory trying to explain the prices of exhaustible resources (Slade and Thille 2009, p. 3). An exhaustible resource is characterized by two features; it is not replaceable and its supply is limited relative to demand. This implies that the production and consumption in one period influence the consumption and production in following periods. Furthermore, the price of such a resource does not in microeconomic theory equal just the marginal cost, but includes also a scarcity rent. (Fattouh 2007)

Hotelling's producer of an exhaustible resource is maximizing the present value of future profits in the world of certainty. In other words, if the whole amount of non-renewable stock reserves is given and known, the model shows how much the producer should extract every period to maximize profit. The most cited implication of the model is that, assuming zero extraction costs, the price of the exhaustible resource should increase in the rate of free interest rate r[3].

Slade and Thille (2009) capture the current state of extensions to the theory, which allow for changes in cost of production depending on the current rate of extraction and remaining reserves, relax the assumption that all reserves are known in the period zero and allow for technical change in exploration or holding inventories. Gaudet (2007) tries to explain the sometimes declining oil prices with adding the market uncertainty, which implements a (sometimes negative) risk premium into the model.

1.2.2 The inventory storage theory

Another major theory on commodity prices is the inventory storage theory, which originates in work by Working (1949). Working focused on inter-temporal relations on the commodity market and investigated factors leading to differences between prices for two different dates of delivery. According to him, such a price difference is explained by the cost of carrying inventory (Working 1949, p. 1256). In the Hotelling's model, the producer was optimizing the production with respect to the whole given reserve of the resource. In the inventory storage model, the producer faces the decision of optimal storage when taking into account a potential profit of a price increase with respect to the storage costs. The key of the theory is therefore the net marginal cost of storage. The marginal net cost of storage can be divided into the marginal cost of physical storage, a marginal risk-aversion factor and the marginal "convenience yield" on stocks (Brennan 1958). The cost of physical storage represents the costs for storing the commodity in warehouses, the risk-aversion factor stands for the risk of a sudden fall in price, by which a company holding large inventories would be negatively affected, and the convenience yield is the yield for holding inventory in order to react on a possible positive demand shock. The producer optimizes the stored amount by putting the expected marginal revenue of stocks equal to the marginal costs of holding stocks.

The academicians using the model follow the logic that, if stocks are held by risk-neutral agents maximizing their profits, the expected price is tight to the current price through costs

---

[3] The result holds also for non-zero extraction cost, then the net price (market price minus extraction price) increases by r. (Hotelling 1931).

of holding stocks. Because of the existence of the convenience yield, the agents create serial dependence in the price even if there is no dependence in the underlying shocks to the supply and demand, which are in fact moderated by existence of inventories (Deaton and Laroque 1996).

1.2.3 The applicability of the models in the real world

Both theories gave birth to various non-structural econometrical models on commodity prices. Pindyck (1999) provides a study investigating the long term price evolution of several commodities, oil among them. Pindyck (1999) argues that structural models are more suitable for explaining the shorter term price evolution, whereas they are not appropriate for longer forecasting for the reason that it is difficult to forecast the underlying explanatory variables. In the long term the growth of the prices can be assumed to be stable, either because of the notion of depletion or because of the long term marginal cost. Using the basic Hotelling's model with constant marginal costs of extraction, Pindyck shows that oil prices are reverting to a long term trend line representing long-run total marginal cost, which itself is unobservable and is changing over time. The results on successfulness in forecasting the oil prices are mixed, though.

The basic version of Hotelling's model was used by many analysts to predict rising price of crude oil, even in the periods when the world witnessed sharply declining prices of oil (Fattouh 2007). Krautkraemer (1998, p. 2102) covers a variety of econometrical studies and concludes that there is strong empirical evidence that the basic Hotelling's model does not adequately explain the observed behavior of nonrenewable resource prices. According to him, even the models extended for the above mentioned features, despite increasing the performance of the Hotelling's model, do not completely apply to the observed data.

A model using the inventory storage theory is for example applied by Deaton and Laroque (1992) and Deaton and Laroque (1996), who try to explain the price of commodity by including speculative storage. They admit that predictions of their first model, where the shocks to harvest are assumed to be independent, identically distributed, are outperformed by a simple first-order linear autoregression model. In their latter study they allow for autoregressive behavior of the shocks, but they themselves describe the results as disappointing. Ederington et al. (2011) reach after a review of several papers a similar conclusion: models based solely on the theory of storage do not tend to do a very good job in predicting the price behavior.

The main criticism is aimed at the unrealistic assumptions of the models. That, however, does not necessarily mean the models are useless. It could also mean the models are just used in a different way than in which they were supposed to be used. As Fattouh (2007, p. 8) notes: “In my view, Hotelling’s model was not intended and did not provide a framework for predicting prices or analyzing the time series properties of prices of an exhaustible resource, aspects that the recent literature tends to emphasize.”

1.3 Structural supply-demand models

The supply-demand framework to model the interactions on the market of oil gained in popularity after price spikes at the beginning of this century (Dées et al. 2007). The models show very low price elasticity of supply in the short run, which is in accord with the inventory storage theory and the Hotelling’s model from the previous chapter, where the producer was not considering the price of the commodity when he was making the optimal production decision. The low price elasticity of demand in the short run then justifies significant price movements in presence of demand shocks.

It is logical to model the price behavior of a commodity with the help of supply and demand, but modeling the market of oil is especially complex. Fattouh (2007) emphasizes the supply side, where there is hardly describable behavior of OPEC, reserves both proved and unproved and the depletion effect. Although these models generally fail to explain the spikes in oil prices in 2007, they are still capturing the essential variables and form the roots for models characterized by Fattouh (2007) as informal approach models.

1.3.1 Demand for oil

The EIA divides the sources of oil demand into two – OECD and non-OECD countries. Such a division distinguishes the two groups of countries with historically different roles. While the demand of OECD countries ruled the market for decades, it slightly declined between 2000 and 2010, whereas the demand from non-OECD countries has risen by 40 %, especially because of China, India and Saudi Arabia (EIA 2012b). The OECD countries still represent the majority of worldwide petroleum consumption (see the Figure 2), but its share declined to 54 % in the year 2010 from 62.5 % in the year 1990, with the steepest decline in recent years (data source: U.S. EIA 2012). In 2003, China took over Japan's second place in petroleum consumption and has been steadily growing so that in the year 2010 China consumed almost half as much petroleum as the USA. An unexpectedly strong growth from emerging economies is provided for example by Hicks and Kilian (2009) as a reason for the increasing oil prices between 2003 and 2008.

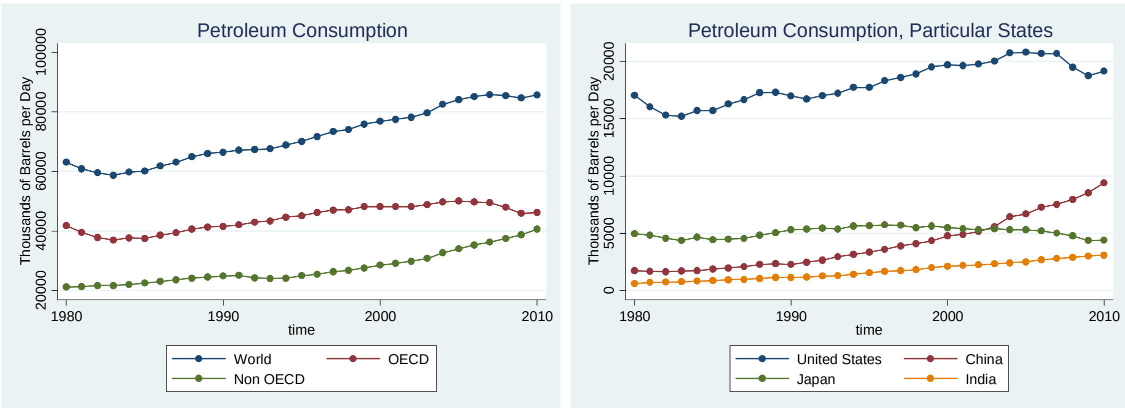


Figure 2: Petroleum consumption by organizations and states (source: author based on the U.S. EIA)

1.3.1.1 Economic activity

As oil is a commodity highly used in the industry, the demand for oil is depending on the growth or decline of the world economy. The researchers modeling the oil market therefore search for a suitable proxy for the world economic activity. Griffin and Schulman (2005) for example observe the retail prices of oil in OECD countries and are using OECD income per capita, Krichene (2006) uses in his analysis GDP of the G7 countries. Using such proxies raises concerns about their suitability, because, as already shown, the increase in the demand for oil comes nowadays mainly from non-OECD countries.

The problem of capturing a good proxy for the global real economic activity for the purpose of modeling the oil market is discussed by Kilian (2009). According to him, there are three main issues when constructing a proxy for the economic activity. First issue is the data availability – for many countries, measures of income are available neither on a monthly nor on a quarterly basis. Second, one must properly establish the constantly changing weights of particular countries to global economic activity, which is made even more difficult because of the continually changing exchange rates. Finally, the income measured by GDP can be distorted for our purpose as in industrialized countries the growth often relates to increasing share in services.

Kilian (2009) suggests an index of economic activity based on a global index of dry cargo single voyage freight rates. Under the assumption that the demand schedule for shipping services shifts with respect to economic activity, the price of shipping service can serve as a proxy for economic activity. Kilian then uses such an index in later analyses, see for example Baumeister and Kilian (2012); Hicks and Kilian (2009); Kilian and Murphy (2010). The index is also used by other researchers, for example by He et al. (2010). Kaufmann (2011), however, questions the validity of this index, showing that in opposite the high prices of oil might affect the freight rates. Moreover, Kilian (2009) is using the refiner acquisition cost of oil when calculating the index, which includes the transportation costs. Kaufmann (2011, p.107) further shows that "... the F test strongly rejects the null hypothesis that real oil prices do not 'Granger cause' SHIP [meaning the Kilian's index]."

1.3.1.2 Price and income elasticity

The demand models of oil market focus on the interactions between the oil demand, oil price and economic activity and result in the magnitude they can influence each other. Price elasticity shows the changes to the demand for oil with respect to the price of oil. In the short run, prices of oil have usually very small influence on the demand. Krichene (2006) for example reports the price elasticity for the G7 countries between -0.02 and -0.03, whereas Griffin and Schulman (2005, p. 11) estimate the short term price elasticity for 16 OECD countries at -0.11. The long term price elasticity enjoys much wider variety among researchers – Krichene (2006) reports the long term price elasticity between -0.03 and -0.08, Griffin and Schulman (2005) -0.36 and Gately and Huntington (2002) -0.64 for OECD countries and -0.18 for non-OECD countries. An explanation for such a difference can be that in non-OECD countries, the end price for the consumer is often influenced by various governmental subsidy programs, whereas the consumer in OECD countries feels the change immediately and can adjust to it, even if not in the short-term (EIA 2012c).

In general, very low price elasticity is assumed as well as the longer term price elasticity being higher than the short term one (Fattouh 2007). Overall, oil is in the short term highly price inelastic, in other words, changes in price do not cause changes in demand. In the long-term, results differ, but there is evidence of substitution effect and energy conservation taking place.

Income elasticity of demand stands for the magnitude of change in demand with rising income. As already described, researchers use different measures of economic activity, but mostly it is GDP or GDP per capita of the selected sample of countries. In the above mentioned studies, Krichene (2006) reaches the long-run income elasticity of 0.54 to 0.90, whereas Gately and Huntington (2002) report 0.56 for OECD and 0.53 for non-OECD countries. Fattouh (2007) after reviewing several other studies concludes that the demand for oil is more sensitive to changes in income than changes in price, there is a heterogeneity across countries and the responsiveness of oil demand to income has been declining for the OECD countries. A possible explanation for that is that the developing countries are more industry-intensive and hence require higher proportion of energy. Also, the transportation oil represents in the OECD countries just a small proportion of GDP, whereas it is more significant in the expanding economies (EIA 2012b).

1.3.2 Supply

Despite the described difficulties in modeling the demand side of oil market, the supply side has traditionally enjoyed more attention by researchers. Although there are some who approach it as a homogenous variable, as for example Kilian and Murphy (2010), most of them distinguish clearly between OPEC and non-OPEC production. OPEC's share of world production has been almost without a change roughly 40 % for the last 15 years (data source: U.S. EIA 2012). On the other hand, OPEC's oil exports represent 60 % of internationally traded oil (EIA 2012e) (see Figure 3). While it is believed that the non-OPEC producers behave competitively, the theories concerning the OPEC behavior differ (Fattouh 2007). Furthermore, the fear of oil depletion can push the prices higher as well.

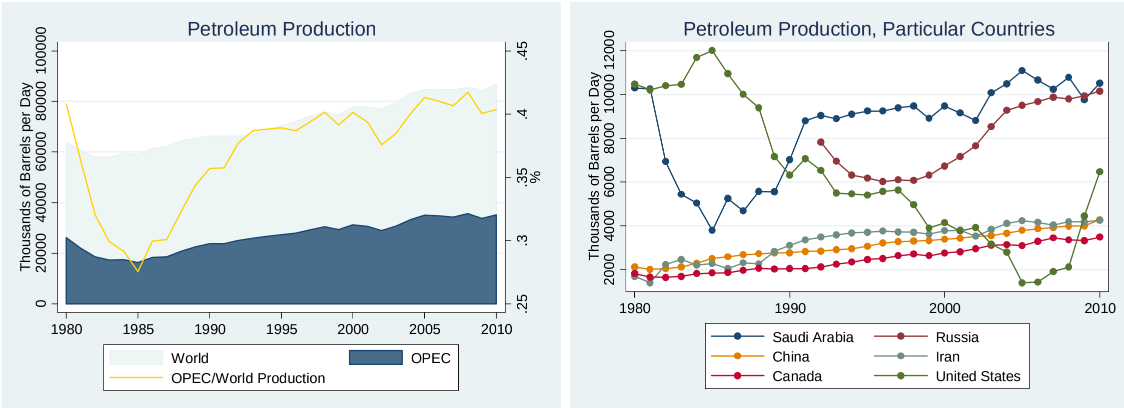


Figure 3: Petroleum production by OPEC/non-OPEC and particular states (source: author based on the U.S. EIA)

1.3.2.1 OPEC supply

"OPEC is a permanent intergovernmental organization of 12 oil-exporting developing nations that coordinates and unifies the petroleum policies of its Member Countries" (OPEC 2012). The theoretical debate about OPEC's role on the crude oil market, which serves as a base for econometrical research, is not centered on *whether* OPEC restricts output of oil, but on *why* it restricts it (Fattouh 2007, p. 21). Many researchers have tried to explain OPEC behavior on the oil market with different hypotheses. The two probably most popular theories are the cartel theory and the revenue target theory. The cartel explanation would mean that OPEC chooses price or quantities[4] such that the marginal revenue minus marginal cost increases at the rate of interest (Pindyck 1978). The target revenue theory, on the other hand, explains the OPEC supply decision as a response to the current budget requirements (Fattouh 2007).

Smith (2005, p. 6) provides a large literature overview to the topic and concludes that "... the empirical literature has failed to produce clear evidence regarding the nature of OPEC behavior, whether competitive or otherwise." A possible explanation according to him is that OPEC behavior simply varies over time depending on circumstances, which makes it impossible to capture it by one single hypothesis. Such a conclusion is in accord with Griffin and Neilson (1994), who observe historically relatively short period of 1983 – 1990 and find for OPEC's behavior consecutively two different fitting models. According to

---

[4] In 1986 OPEC abandoned the strategy of setting price in favor of the strategy of setting production quotas. (Fattouh 2007)

Smith (2005), OPEC is much more than a non-cooperative oligopoly, but much less than a frictionless cartel as well. He can neither confirm nor reject the hypothesis that Saudi Arabia behaves as a leader of the organization.

The different, often contradictory theoretical results hinder application of a simple econometrical model explaining the evolution of oil price according to OPEC's behavior. Dées et al. (2007) for example use OPEC's operable capacity and OPEC production, which depends on the world demand, production of the rest of the world, OECD's inventories and processing gains. According to their result, an increase in OPEC's capacity by 1 % has an immediate negative effect of 10 % on oil prices, in the long run the effect is 12 %. The reduction of prices leads to increase demand for OPEC's oil by 2 %. The overall estimated effect for OPEC is a decline in revenues by 8 %, which explains why OPEC is unwilling to increase the production capacity.

1.3.2.2 Non-OPEC supply

Non-OPEC producers represent roughly 60 % of the world oil production. The largest oil producing Non-OPEC country is Russia (10.1 ml. barrels per day) followed by the USA (9.7 ml. barrels per day) (data source: U.S. EIA).

The non-OPEC producers are usually regarded as price takers, since they are not able to influence the price by managing production. Hence, they are usually producing near their full capacity and they tend to have relatively higher production costs, as the cheaper oil is accessed by OPEC countries (EIA 2012d). Non-OPEC companies must therefore seek for new sources of oil from less profitable sources, as deepwater offshore or oil sands. As they are trying to lower the costs of drilling, they are also considered to be the technology leaders on the market (EIA 2012d). The EIA (2011) reports the upstream costs[5] for 28 reporting companies between 2006 and 2008 costs per barrel of oil of 41.49 USD for the USA (with decomposition of 37.32 for onshore and 74.20 for offshore), 38.75 for Canada and just 17.09 for the Middle East, the lowest amount among all the regions. Although the data describing the available oil supply in relation to the finding and lifting costs could

---

[5] The costs consist of finding costs and lifting costs, including the production taxes. The tax is usually fluctuating according to the price of oil on the markets and stands for a significant share of lifting costs – in 2008 the lifting costs for USA were 14.75 USD per barrel of oil equivalent (boe), including a tax 4.90 USD per boe. In 2009 the price of oil decreased and the tax slumped to 1.78 boe (EIA 2007, p. 21)

definitely help in explaining the supply side of oil, because of the low frequency of publishing they are just of a little use (Coleman 2012, p. 319-320).

Various researchers have tried to estimate the long term price elasticity of the supply of non-OPEC producers to quantify their response to price movements. In such models, price exogeneity is justified as the non-OPEC producers are assumed to be the price – takers. Despite this assumption, which makes the estimation easier than in the case of OPEC, the estimates vary largely across literature. Krichene (2006) estimates the long run price elasticity 0.08, which would mean that the oil supply barely responses to the changes in prices. Gately (2004) estimates the short – run price elasticity of supply between 0.03 and 0.05, whereas in the long – run it varies between 0.15 and 0.58. Such results show that even if the non-OPEC producers are positively motivated by higher prices, in the short run they are greatly limited by their production possibilities.

1.3.2.3 Reserves

There are many definitions of petroleum reserves, but mostly the definitions organize reserves according to the estimate of certainty of quantity that might be recovered (Oil and Gas Reserves Committee 2005). *Proved reserves* are usually denoted P90, which stands for 90 % probability that the recovered amount will be equal or greater than the estimate. Similar guideline is followed by the *unproved reserves* (P50) and *possible reserves* (P10) (Fattouh 2007, p. 17).

Due to the U.S. EIA data on proved petroleum reserves, the largest proved petroleum reserves were held by Saudi Arabia in 2011 (262.6 billion barrels), followed by Venezuela, which proved reserves almost doubled to around 237 billion barrels between 2010 and 2011. The third place is then attributed to Canada (175.2 billion barrels) with approximately 97 % of the reserves being unconventional, mainly from bitumen deposits (EIA 2012a). In 2009 the total proved reserves were almost 1342 billion barrels of crude oil. With the petroleum consumption 84.7 million barrels per day, the world had proved reserves for 43.4 years. This ratio has furthermore increased over time. In 1973, the world had proved oil reserves for 30 years, in 1983 for 35 years, in 1993 for 42 years and in 2003 for 41 years (Watkins 2006). This is mainly not because of new discoveries, but thanks to growth in preceding reserves, either because of the initially too conservative estimates or new technologies (Fattouh 2007).

The importance of oil reserves is derived from the well-known phenomena peak oil introduced by Hubbert (1949) and further developed in 1956, where Hubbert introduced the bell-shaped curve representing the complete cycle of production of any exhaustible resource. In this article he also successfully predicted the peak oil, in other words the moment, when the maximum extraction rate of oil is reached and starts to decline, for the U.S. oil production for approximately the year 1970. The idea of peak oil is now widely accepted (de Almeida and Silva 2009), the estimated years of reaching the peak oil vary mostly from 2010 to 2025 (Hirsch et al. 2006). Such an event would result in a huge liquid fuels problem (Hirsch et al. 2006), which would certainly justify the recent increases in the prices on the markets (de Almeida and Silva 2009).

1.3.3 The explanatory power of the structural models

All in all, the supply-demand models showed a small ability of both supply and demand to respond to oil prices in the short term. In the long term, the estimates of price elasticity of demand differ and some researchers even emphasize significant differences from the regional point of view, saying OECD countries can adjust their consumption according to changes in the price of oil, while the higher industrialized non-OECD countries are more dependent on oil and there is just a minor consumption adjustment with respect to price. The income elasticity of demand estimates and price elasticity of supply also differ significantly among available studies, which makes it difficult to draw any general conclusion from them.

Some simplified conclusion for the short term is possible, though. Let us assume the short term price elasticity of demand -0.05 and the price elasticity of supply 0.05, both of them well observed in the presented literature sample. Let us assume that OPEC behaves as a swing producer which fills the gaps between the world oil demand and the non-OPEC supply, but benefits from higher oil prices. Let us further assume a negative supply shock on the non-OPEC side, which decreases the world supplied quantity of oil by 1%. If OPEC does not take any corrective action, the equilibrium quantity on the market sinks by 0.5 % and the oil price rockets by 10 %. Even if this calculation simplifies the influence of OPEC on the supply side, it clearly illustrates the price sensitivity of the market to the supply and demand shocks, both of them witnessed in recent years.

Fattouh (2007) concludes that even if the supply demand framework provides useful insight into the oil market, it fails to succeed in projecting the prices, as it is highly sensitive to the assumptions about elasticities and OPEC behavior. If for example the assumed elasticities in the preceding paragraph had been 0.1 and -0.1, the price would have shown an increase only by 5 %. Furthermore, the framework lacks other context in which the price is determined besides the supply and demand.

1.4 Informal models

After the surge in price and volatility of oil prices in 2008, the crude oil enjoyed increased attention from researchers. Khan (2009) discusses three possible causes of the spike. The first possible explanation is the start of the peak oil as described in the previous section, which, however, was at variance with further rising oil production. The second explanation can be called fundamentalists explanation, as the proponents of the theory believe the surge in the oil prices was caused by fundamental factors, like GDP growth on one hand and capacity constraints on the other hand. Finally, there are researchers believing the fundamental factors were not powerful enough to account for such a price increase and that the price was manipulated by speculation on further increases. Unfortunately, speculation is, unlike GDP, oil barrels or the number of OPEC members, hard to measure (Cifarelli and Paladino 2010, p. 371; Kaufmann 2011, p. 109; Khan 2009, p. 4; Parsons 2009, p. 32).

Time showed that the traditional basic structural models were unable to capture the observed evolution on the market. Researchers have therefore tried to capture other variables that could have an influence on the price in the short-term, as the lack of spare capacity, geopolitical situation, the role of inventories or speculation on the markets. Some of the factors have already been described earlier, we will just present some further ones and focus on the main topic of speculation more in-depth.

Breitenfellner and Cuaresma (2008) for example examine the relationship of oil prices and the EUD/EUR exchange rate, using it as a proxy for nominal effective exchange rates, as it represents a significant share of oil imports from OPEC countries. They identify and discuss five channels in which this relationship can take place; the purchasing power channel, the local price channel, the investment channel, the monetary policy channel and the currency market channel. Although the direction of causality between USD/EUR and oil prices is unclear, exchange rate information is significantly improving the oil prices forecasts (Breitenfellner and Cuaresma 2008, p. 118).

Kaufmann et al. (2008) find a negative relationship between the crude oil prices and the refinery capacity utilization. They interpret this relationship in the way that in case of higher refinery utilization, the refineries are accepting crude oils with other properties, for example another API gravity index, than are the ones the refinery is most suitable for[6]. Such crude oils increase the operating costs for the refineries, which are willing to pay lower prices for them.

Finally, the important role of inventories has been confirmed by multiple analyses, for example Chevillon and Rifflart (2009); Dées et al. (2007); Kilian and Murphy (2010). Dées et al. (2007) perceive the effect of holding stock as a positive externality on the market for crude oil. The agents on the market usually hold stocks for reasons described in the inventory storage model earlier in this thesis, mostly to avoid the risk of disruption. According to Dées et al. (2007) higher OECD oil stocks suppress the oil prices and they suggest that policy makers develop instruments to enhance the willingness of agents to hold stocks.

As the speculation on the crude oil market is the key topic of this thesis, it will be given more space in the next separate chapter.

[6] Light sweet crude oils tend to be produced first, as they are because of their properties more expensive and generate more revenues for both producers and refineries. Most of the refineries are therefore set for these crudes and higher refinery utilization is usually connected to refining heavy and sour crudes (Kaufmann et al. 2008, p. 2615).

1.4.1 Speculation

> "It is hard to explain current oil prices in terms of fundamentals alone. The recent surge in the oil price (from USD 80 to USD 100 a barrel) seems to go well beyond what would be indicated by the growth of the world economy. Producers and many analysts say it is speculative activity that is pushing up oil prices now. Producers in particular argue that fundamentals would yield an oil price of about USD 80 a barrel, with the rest being the result of speculative activity." (IMF 2008, p. 27)

One can find dozens of such quotes across news articles and many even in academic articles. What one cannot find is any statistical evidence that the price spike was caused by speculation.

Among the proponents of the hypothesis that the oil spike in 2008 was not entirely caused by fundamentals there are for example Parsons (2009), Chevillon and Rifflart (2009), Cifarelli and Paladino (2010), Coleman (2012), Kaufmann (2011), Kaufmann and Ullman (2009), Khan (2009) and Ellen and Zwinkels (2010).

A major influence of speculation on crude oil markets is rejected by the Interagency Task Force on Commodity Markets (2008), Hamilton (2009b) or Kilian (2009), who argues that the price increase was led by higher demand especially from China.

How to prove that oil prices are driven up by speculators? One would say that an easy solution would be to regress the amount of speculative capital on the market on oil prices and test the significance of the influence. Distinguishing the speculative from the non-speculative capital proved quite difficult, though. Speculation can be defined as an act of purchasing (selling) an asset with the intent to sell it (re-purchase) at a later date with a motive of expectation of a change in price (Kaldor 1939, p. 1). The only data which at least roughly fit the definition are provided by the U.S. Commodity Futures Trading Comission (CFTC).

CFTC distinguishes between two types of agents on the futures markets; commercial and non-commercial traders. The classification of a commercial trader is achieved when the trader files a statement with CFTC confirming his hedging intention on the futures or option market (Interagency Task Force on Commodity Markets 2008, p.19). Usually one assumes that the speculation on the market is represented by the non-commercial group of traders, but because the motivation to take the hedge of a commercial trader is not investigated and one therefore cannot say whether the hedge is done with an intention to lower risk or it is a market bet, the speculation could arise also within the commercial group.

The increase in open interest because of the inflow on non-commercial traders is obvious from the figure Figure 4. If the global oil production reached on average 86,838 thousands of barrels per day in 2010, then just the volume of open positions on NYMEX futures market exceeded this number 15.4 times. This fact certainly can lead to a concern whether such an inflow of speculative money could push the oil prices higher.

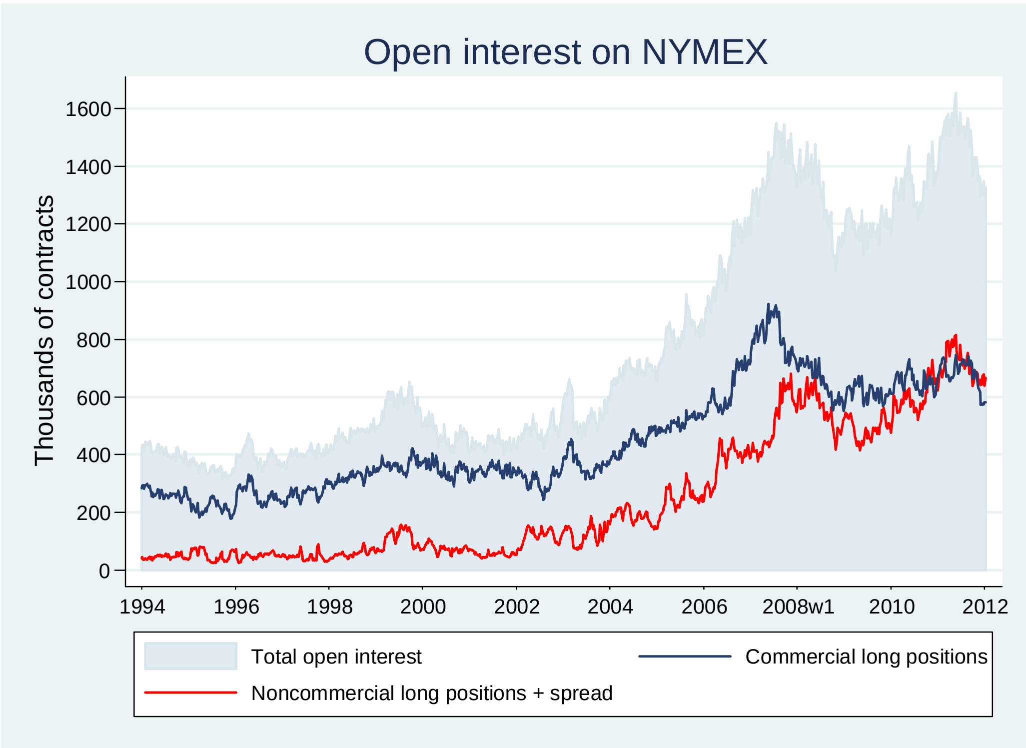


Figure 4: WTI open interest on NYMEX (source: author based on CFTC)

A special report of CFTF, a Commitment of Traders report (COT) further distinguishes between commercial producers, manufacturers, commercial dealers, swap dealers (all these are part of the commercial traders group), hedge funds and floor brokers & traders (these are part of the non-commercial traders group (Interagency Task Force on Commodity Markets 2008, p. 21). These categories accounted in 2008 for roughly 80 % of open interest in the crude oil futures market. The groups suspect of speculation are then non-commercial traders in total, the hedge funds, swap dealers and non-commercial traders combined with swap dealers (Interagency Task Force on Commodity Markets 2008, p.27). The report tests the Granger-causality relating daily position changes to price changes in the NYMEX WTI crude oil futures contract in years 2000 to 2008, but does not find any evidence that any group's change in the position precedes the change in the price. It finds, however, a reverse Granger-causality for some of the groups, which means that the increase in the speculative capital occurred after the increase in price. The report concludes that the results are consistent with the view of crude oil being driven up by fundamentals (Interagency Task Force on Commodity Markets 2008, p. 29).

Buyuksahin and Harris (2011) also fail to find that traders on NYMEX, who can be considered speculators, lead the price changes, based on the Granger causality tests, and find a reverse relationship indicating that speculators and swap dealers are trend followers.

What indicators do then lead to the opinion that the crude oil markets are driven by speculation? Kaufmann (2011) discusses three indicators of speculation. First, he occupies himself with the role of inventories. In the inventory storage theory, described earlier in this thesis, if a producer expects a price rise, it is a reason to increase the oil inventories. Kaufmann shows and statistically tests for a reversal in a 22 years downward sloping trend starting at 1982, when the inventories start to rise in 2004. Although such a test cannot be considered to be a proof of speculation, Kaufmann regards it more as a "smell test".
A so called "Law of One Price" is Kaufmann's second indicator of speculation. The Law of One Price is a widely tested and confirmed belief that the world crude oil market is unified, which implies that the spread between two crude oils on different markets is given by their physical measures of quality, as for example the sulfur content or API gravity index, and arbitrage transaction costs (Fattouh 2010; Kaufmann 2011, p. 111). The spread between prices of two crude oils should therefore be stationary, unless speculative expectations play an important role in price formation. Kaufmann indeed identifies several extended periods when the cointegrating relationship breaks down.

Finally, Kaufmann points out an inability of fundamental models to capture the price changes in the beginning of the millennium. On an example of model proposed by Kaufmann et al. (2008) he shows relatively good performance of the model based on fundamental measures until the year 2007. After 2007, when the model under-predicts the real values when there is a price increase and over-predicts when there is a price drop, both consistent with the behavior in presence of speculation.

1.5 Summary

In the first part of the thesis we have covered the major approaches trying to explain the oil prices. The Hotelling's theory and the inventory storage theory present an alternative to the classic microeconomic theory, where the price on the market would equal to the marginal cost of production. In case of a commodity which can be depleted, we need to take into account also its scarcity rent. Both of the models fail in predicting the oil prices even in the longer period, even if some argue that it was never their goal.

Furthermore, the fundamental variables on the crude oil market were considered. We described the tightness between the supply and demand, which has been steadily growing because of the developing countries. The supply side, on the other hand, is complicated because of the presence of the OPEC cartel, whose market share is also slowly increasing. Another very important factor is the threat of the peak oil, which would have far-reaching consequences.

At the end, we briefly introduced further factors which were proven to be important for explaining the crude oil prices, for example the OECD inventories. We dedicated more space to the topic of speculation, which is suspected to be responsible for the large changes in oil price in the recent years. We discussed the possible connection between the increasing amount of the open interest on the crude oil market and the speculation and noted that so far, no causality has been found from the increases in open interest to the increases in the oil prices.

Part 2: Empirical part

In the second part of the thesis we will try to answer the two following questions:

1) Is there any long term relationship between the crude oil prices and the open interest on the futures market?

2) Can an increase in oil price volatility lead an increase in the open interest?

The first question relates to whether simply the changing in the investor perception of the market without any change in the supply and demand factor can lead to higher prices, the second question then whether an increase in the price volatility can cause such a change in the perception.

First, we will describe the data we will use for our analysis and suggest a model which could help us answer the first question. Then, we have a closer look at the reasoning for the methodology used to the estimation model and the estimation itself. In the fifth part of the chapter presents the result of our analysis. Part six is dedicated to the answer to our second research question. Part seven briefly discusses the topic of the crude oil price forecasting. Finally, we summarize our results and compare them to the result of the recent research.

2.1 Data

For the price of oil, the front month future of West Texas Intermediate (WTI) light sweet crude oil with delivery in Cushing in Oklahoma is used, as the WTI price together with Brent and Dubai oil serve as international pricing benchmarks (Fattouh 2010). We used the data on the non-OPEC crude oil production provided by the EIA on a monthly basis and treated the crude oil produced by OPEC as a separable variable. The data on the world crude oil consumption are unfortunately not provided on a monthly basis by the EIA for the non-OECD countries. Since we have already pointed out the importance of the demand growth coming from these countries, we decided to use the data on the world liquid fuel consumption, which is assumed to be a good proxy for crude oil consumption. We also included the petroleum stocks in OECD countries, as the data on the inventories of the non-OECD countries are unfortunately not available. The financial market is represented by the data on open interest on the WTI futures market provided by the U.S. Commodity Futures Trading Commission.

All the variables are monthly, as that is the lowest frequency the EIA is releasing the fundamental data. The first month the data on the liquidity fuel consumption are available is the January 1994 and the last data used are from the October 2011, the sample therefore consists out of 214 observations. We would of course prefer a larger data sample, but that would lead us to the need to drop some of our variables. Such a relatively small sample prevents us from including other variables that could be interesting for modeling the crude oil market, as for example the strategic reserves, OECD crude oil imports, OPEC spare capacity, the global proven reserves, refinery utilization etc.

In the second analysis in this thesis we use weekly data on the open interest on the WTI, Cushing Oklahoma provided by the U.S. Commodity Futures Trading Commission, and the weekly standard deviation of the WTI front month futures contracts, calculated for the last five days every Tuesday, for when the data on the open interest are published as well. We are working with the time period 1994 to 2011, which gives us a sample over 900 observations.

The table I shows the basic summary of the used data. We can observe the great variance in the data – for example the crude oil price, where the highest value is more than ten times the lowest one. The standard deviation of the OPEC production being higher than the non-OPEC one is also notable.

Name of the variable	**Units**	**Mean**	**Standard deviation**	**Minimum**	**Maximum**
Crude oil price	USD	44	29	11.3	134
OECD inventories	bl. barrels	2910	1343	710	6940
Liquid fuel consumption	bl. barrels per day	79	6	66	89
Non-OPEC crude oil production	bl. barrels per day	40	1.96	35.7	43
OPEC crude oil production	bl. barrels per day	29	2.47	24.6	33.1
Open interest of WTI	'000 of contracts	750	390	315	1554
Weekly standard deviation of the crude oil prices	USD	0.8723	0.776	0.0234	7.83

Table 1: Summary for the used data (source: author based on EIA, CFTC)

2.2 Model

We hypothesized the existence of two equilibria on the crude oil market, the first equilibrium on the physical oil market, where the fundamental variables influence each other, and the equilibrium on the speculative financial market, which is represented by the size of the market. Even if the U.S. Commodity Futures Trading Commission provides a distinction between commercial and non-commercial traders and one could search for the speculation on the crude oil market strictly on the non-commercial side, the connection between speculation and the crude oil inventories has been often suggested (Hamilton 2009a; Kaufmann 2011). In such a case, one could also seek for the speculation within the commercial traders. Our suggested model is hence as follows:

$$price_t = \alpha_1 + \beta_1 * lfc\_w_t - \beta_2 * cop\_rest_t - \beta_3 * opec_t - \beta_4 oecd\_inv_t \quad (1)$$

$$price_t = \alpha_2 + \beta_5 * open\_interest_t + \beta_6 * oecd\_inv_t \quad (2)$$

Price stands for the crude oil futures price, *lfc_w* stands for world liquid fuel consumption, *cop_rest* for non-OPEC crude oil production, *opec* for OPEC crude oil production, *oecd_inv* for OECD petroleum inventories and *open_interest* for the open interest on the futures market. This model will be estimated through the error correction method suggested by Johansen (1995).

2.3 Methodology

To decide which method to use for estimating a model, it is essential do find out whether our data are stationary or non-stationary. According to the definition used by Kočenda and Černý (2007, p. 15), a time series is stationary, if "[...] any shock that occurs in time *t* has a diminishing effect over time and finally disappears in time *t* + *s* as *s* -> ∞." Whereas a stationary process tends to fluctuate around its mean and has a finite variance, a non-stationary one demonstrates a changing mean and an increasing variance with the sample size (Harris 1995, p. 15). Such processes are also called unit root processes, or processes integrated of order one, I(1), or of a higher order. The main problem with I(1) processes is that they usually violate the assumptions of the ordinary least squares method (Wooldridge 2003). Moreover, a simple regression of two I(1) series often results in significant *t* statistic, pointing out a relationship that does not exist (Wooldridge 2003). This anomaly has been tackled by Granger and Newbold (1973), who proved that such a regression is simply mis-specified and the presented R^2 is misleading. An I(1) process is described by Wooldridge (2003, p. 797) as "a time series process that needs to be first-differenced in order to produce a I(0) process," whereas a process integrated of order zero, I(0), is defined as "a stationary, weakly dependent time series process [...]" that fulfills the standard OLS assumptions.

Since the time series used in our analysis are highly suspected of containing a unit root, one option would be to first-difference them to obtain I(0) processes. All of the six variables are presumed to be endogenous in the model, as the literature review showed that the variables respond to each other. Such a constraint further complicates the use of a standard OLS method, as including an endogenous variable in an OLS regression can cause bias and inefficiency (Wooldridge 2003, p.83-84). A solution to this would be to use a simultaneous equations model (SEM) or a vector autoregression model (VAR) for the differenced data. Engle and Granger (1987), however, showed that it is possible that for two time series x_1 and x_2, both I(1), there exists such a linear combination $x_1 + \beta x_2$ that is stationary. In that case we say that x_1 and x_2 are cointegrated of order (1,1). This linear combination defines a long-term equilibrium relationship, driven by some economic forces, toward which the variables will be forced by an error-correction mechanism. A simple vector autoregression model in differences omits such an error correction term (Engle and Granger 1987, p. 275).

2.4 Estimation

First, we tested whether our time series were non-stationary. Since the unit root tests have in general low power properties (Harris 1995), we applied multiple tests to decide on the I(1) of the process. The basic Dickey-Fuller test indicated some of the variables could be trend-stationary, which led us to use a trend in our later analysis of the model. We also applied a modified Dickey-Fuller test developed by Elliott et al. (1996), which is supposed to have a greater power in the presence of a trend, and KPSS test developed by Kwiatkowski et al. (1992). Based on the results of the tests we decided to assume that all of our variables are at least I(1). Finally, the first differences of the time series were also tested for a presence of a unit root, to find out whether some of the variables were I(2), where the null hypothesis of a unit root has been rejected in all cases. This indicates that all of the variables are I(1) and invalidates the usage of the OLS method. Instead we decided to use the error correction model developed by Johansen (1995).

To decide on the number of cointegrating vectors within our six variables, we used Johansen's trace statistics method with three lags, based on Akaike's information criterion, and a deterministic trend in the levels of the variables. We included a linear trend in the levels as we believe the supply and demand for crude oil could be trending, based on the advice given by Johansen (1995, p. 82) that one should, when deciding on the inclusion of the time trend, rely on the economic insight. The oil price can also be suspected of a trend, as we are using the nominal oil price. The results show that we can reject the null hypothesis of more than two cointegration factors on 1 % and 5 % level of significance. In our further analysis we will therefore assume the hypothesized 2 cointegrating relationships (see Part I in the Appendix).

Finally, knowing the cointegration rank of the covariance matrix, we used a vector-error correction model estimation technique implemented in the Stata program. We allowed for a linear trend both in the levels of the variables and the cointegrating equations and four lags, as the results indicated the presence of serial correlation in residuals in case of fewer lags. As we have two cointegrating vectors, Stata imposes automatically four restrictions on the resulting cointegrating relationships. Harris (1995, p. 117) stresses, however, that it is necessary to impose some restrictions based on economic arguments rather than forming the economic arguments based on the restrictions imposed automatically. We therefore imposed the restrictions based on the two suggested equilibria previously described. For the sake of the later interpretation we set the coefficient on the crude oil price equal to one. We omitted the variable *open_interest_all* in the equation (1) and the demand and supply variables in the equation (2). Even though six restrictions were implied, the test of the overidentifying restrictions does not reject the null hypothesis that the restriction is valid. We also imposed restrictions on the short-run adjustment parameters α of the *OECD inventory*, which represent the speed of adjustment to the equilibrium in the short run. Both of the α parameters of *OECD inventory* seemed insignificant, which we confirmed by a joint test. We therefore concluded the weak exogeneity of *OECD inventory* in the model, which should improve the overall stochastic properties of the model (Harris 1995). Such a restriction means that the OECD inventories do not adjust to the disequilibrium on the market in the short run. We have not restricted the α coefficients of the *OPEC production*, even though the coefficients also seemed insignificant. Such a restriction would go against the theory of OPEC being a swing producer.

The post-estimation tests suggest that the regression fulfills the stability condition and there is no serial correlation in the residuals. The test for normality strongly rejects the normality in residuals for most of the variables as well as no skewness and kurtosis. When we plot the residuals of the oil price (see the Figure 5), we can easily observe growing variance in the residuals in the latter periods, when the price of oil also tends to fluctuate more. This is probably not so surprising, as the evidence of conditional heteroskedasticity is present in many estimated dynamic regression models in finance, especially for monthly, weekly and daily data (Gonçalves and Kilian 2004, p. 1). Adding additional lags does not solve the non-normality in residuals, therefore we assume that there are important variables missing in the model. This, again, is not surprising, as the financial markets are extremely complex and trying to capture all the relevant relationships would mean adding many additional variables, but, as discussed before, there are not enough observations for such an analysis. Despite the non-normality of the residuals, the analysis is not completely invalid, as Johansen (1995, p. 29) states: “The methods derived are based upon the Gaussian likelihood but the asymptotic properties of the method only depend on the i.i.d. assumption of the errors. Thus the normality assumption is not so serious for the conclusion, [...].”

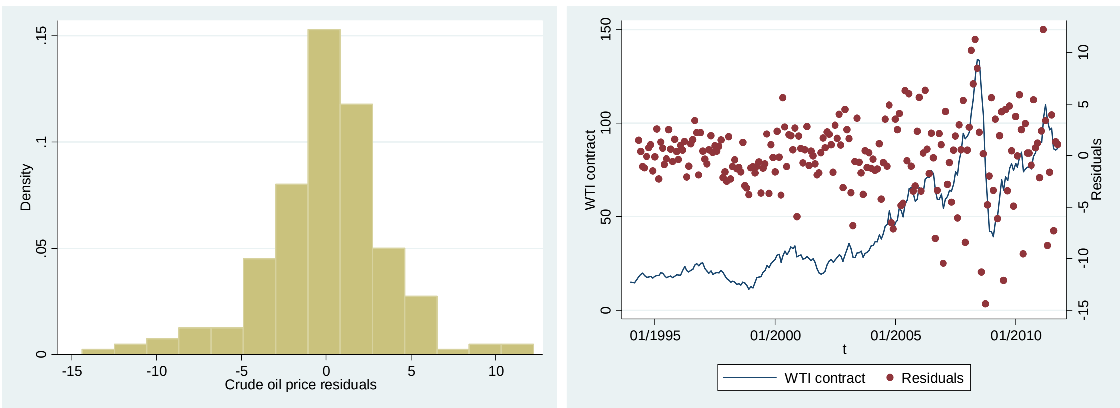


Figure 5: Residuals of the estimated model

2.5 Results

We estimated the previously described model of two long run equilibria depicted by the equations (1) and (2) on the market for crude oil using the Johansen vector error correction method, as follows:

$$price_t = 223 + 0.24 * t + 8.8 * lfc\_w_t - 18.4 * cop\_rest_t - 5.2 * opec_t - 5.2 * oecd\_inv_t \quad (3)$$

$$price_t = 12 + 0.1 * t + 0.05 * open\_interest_t - 3.88 * oecd\_inv_t \; , \quad (4)$$

where *price* stands for the crude oil price, *lfc_w* for world liquid fuel consumption, *cop_rest* for the crude oil production in non-OPEC countries, *opec* for OPEC crude oil production, *oecd_inv* for OECD petroleum inventories and *open_interest* for the total open interest of the closest future contract on NYMEX. All of the coefficient are highly significant, even if the statistical inferences can be biased by the heteroskedastic residuals (for the fit of the model see the Figure 6, for more details on the parameters see Part I in the appendix).

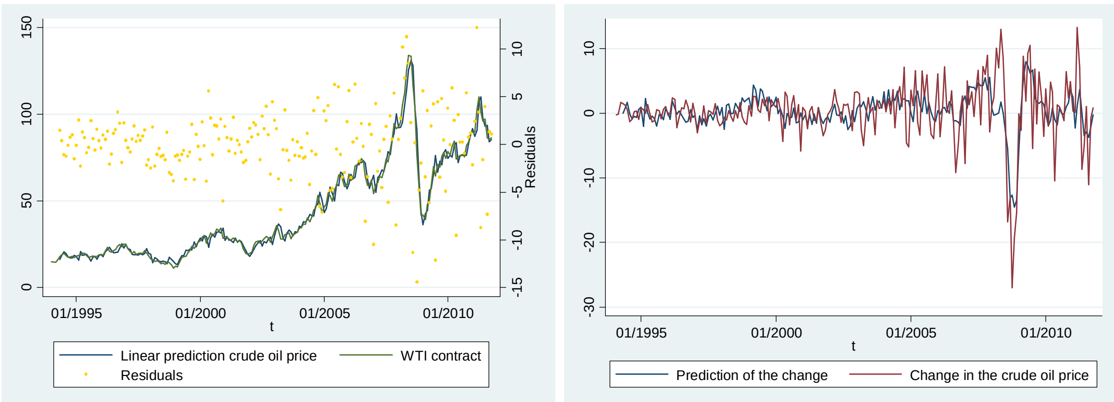


Figure 6: Linear prediction of the model in levels and differences

The directions of the coefficients in the equation (3) keep with the economic theory; an increased demand has a positive effect on the crude oil price, an increasing supply a negative effect. One can see that the coefficient on the non-OPEC crude oil production is higher than on the OPEC crude oil production. Even if the interpretation that increasing the non-OPEC production has a stronger negative effect on the crude oil prices is tempting, one cannot interpret the cointegrating relationship that easily, as one cannot simply count on the *ceteris paribus* condition. We will elaborate on the results in the next section. The effect of increased OPEC inventories is also negative. Although we would like to provide a similar interpretation for the equation (4) as well, meaning that the increase in the open interest on the crude oil market would lead to higher prices, since there are is no underlying economic theory, we will rather rely on the impulse-response analysis, as it is for such a case suggested by Lütkepohl and Reimers (1992).

The adjusting coefficients of the variables α, which measure the speed they return to the equilibrium, are shown in Table 2. Most of the coefficients are significant and negative, showing that the variables tend to return to the equilibrium. The coefficient α of the crude oil showing its response to the first equilibrium is not statistically significant, which would mean that the crude oil price could in the short run fundamentally overvalued.

alpha	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
D_wtif						
_ce1 L1.	-.0542816	.0391837	-1.39	0.166	-.1310802	.022517
_ce2 L1.	-.1505146	.0429408	-3.51	0.000	-.234677	-.0663522
D_lfc_w						
_ce1 L1.	.0728971	.013147	5.54	0.000	.0471293	.0986648
_ce2 L1.	-.0712989	.0144076	-4.95	0.000	-.0995373	-.0430605
D_oecdinv						
_ce1 L1.	(omitted)					
_ce2 L1.	(omitted)					
D_cop_rest						
_ce1 L1.	-.0109495	.0032897	-3.33	0.001	-.0173972	-.0045018
_ce2 L1.	.0084958	.0036051	2.36	0.018	.0014299	.0155618
D_cop_o						
_ce1 L1.	-.0006924	.0028802	-0.24	0.810	-.0063374	.0049526
_ce2 L1.	-.0029026	.0031563	-0.92	0.358	-.0090889	.0032837
D_oia						
_ce1 L1.	.1168255	.4156819	0.28	0.779	-.697896	.931547
_ce2 L1.	-.7552431	.4555392	-1.66	0.097	-1.648083	.1375972

Table 2: Estimated adjusting coefficients α

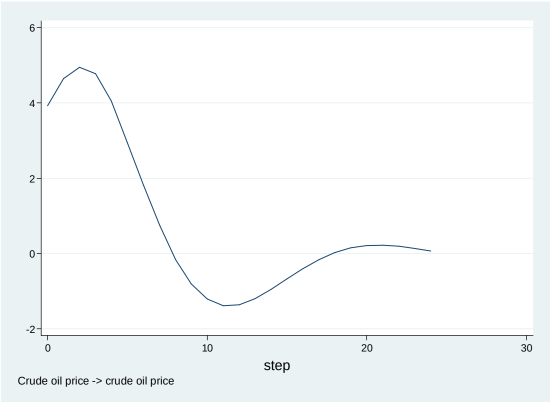
2.5.1 Impulse-response analysis

The impulse-response analysis allows for an analysis of a shock of one endogenous variable on itself or on another endogenous variable (StataCorp 2009). Such an analysis can provide an especially valuable insight for the vector-error correction model, as it is often difficult to interpret the innovations between the variables directly (Lütkepohl and Reimers 1992, p. 54). We used the orthogonalized impulse response function with Cholesky decomposition, which, unlike the ordinary impulse response, allows that a shock to one variable is accompanied by shocks to other variables (StataCorp 2009, p. 173). The y-axis represents the shock to the response variable in the size of one standard deviation (to find out the size of the shock see the Table 3) of the residuals of the impulse variable, the x-axis represents the time periods, which are in our case months. Unfortunately, Stata does not provide the confidence intervals for the IRF analysis of a VECM model.

Variable	**Standard deviation of the residuals, units**
Crude oil price	3.72 USD
OECD inventories	420 bl. barrels
Open interest on NYMEX	39.4 thousands of contracts
Non-OPEC production	0.31 bl. barrels per day
OPEC production	0.43 bl. barrels per day
Liquid fuels consumption	1.25 bl. barrels per day

Table 3: Standard deviations in residuals

The orthogonalized impulse-response function (OIRF) of the crude oil price on the crude oil price shows that such a shock is transitory. The initial increase in the oil price is followed by additional increases within the first two months, and then the effect fades away. Such a development could be interpreted as a hoard behavior on the financial markets.



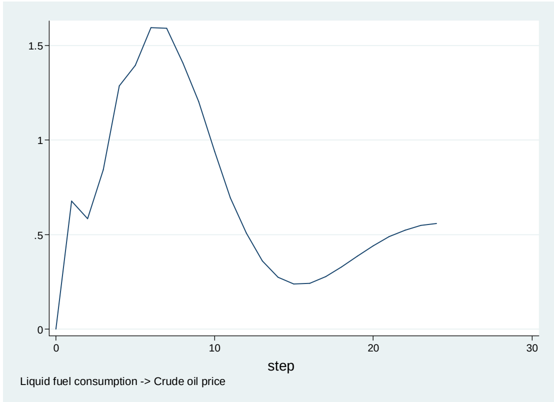
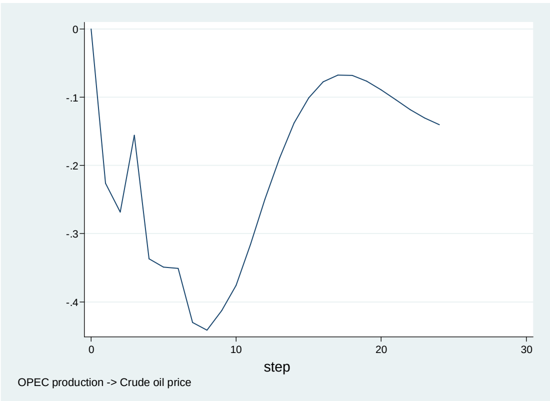


Figure 7: IRF, impulse: Crude oil price, response: Crude oil price

Figure 8: IRF, impulse: Liquid fuel consumption, response: Crude oil price



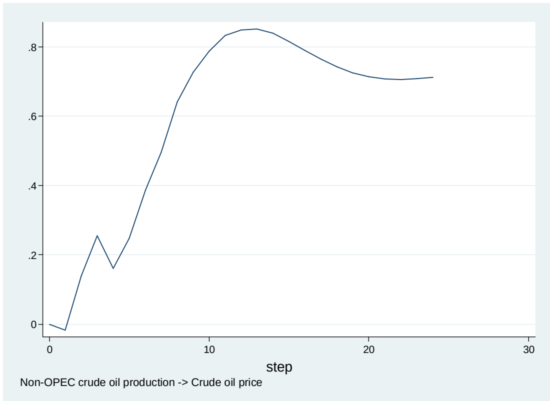


Figure 9: IRF, impulse: OPEC production, response: Crude oil price

Figure 10: IRF, impulse: Crude oil production, response: Crude oil price

An innovation to the liquid fuel consumption has an expected positive effect on the crude oil price, the effect seems to be transitory. A similar development can be observed if the OPEC crude oil production is increased by one standard deviation (430 ml. barrels per day). A bit surprising is a positive shock to non-OPEC crude oil production. Such a shock has according to our analysis a persistent positive effect on the crude oil prices. Why would an increase in the OPEC production decrease the crude oil prices and an increase in the non-OPEC crude oil production increase them? A possible explanation is depicted by the Figure 11. An increase in both of the productions has a positive permanent effect on the liquid fuel consumption, which is significantly smaller for the OPEC production (even after considering the different sizes of the shock). The increase in the non-OPEC production transforms immediately into an increased consumption, which promptly pushes the price of oil up. Moreover, the role of OPEC cannot be really modeled under the market equilibrium assumption – OPEC usually does not increase the production if it is not needed as low prices probably are not in its interest. Furthermore, the fear of insufficient spare capacity of OPEC probably plays an important role as well.

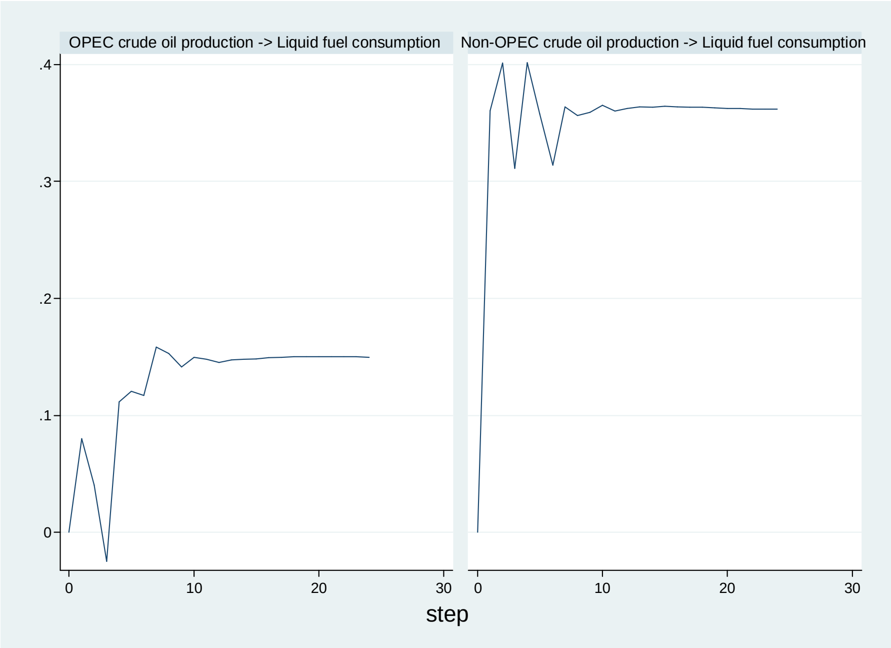


Figure 11: Impulse response analysis of OPEC and non-OPEC production to the liquid fuel consumption

Finally, we will focus on the influence of the open interest on the crude oil price. Our analysis has already shown there is a long-run equilibrium between the number of futures contracts and the crude oil price. The following figures model what happens if one of the variables is pushed out of the equilibrium. We can see both of the shocks are permanent. In case of a shock of one standard deviation to the oil price (USD 3.7), there is an immediate inflow of investors on the market. Our analysis suggests there is 10 000 more long positions (and of course the short positions as well). In case the shock does not fade away, the investors start leaving the market, as one might well consider the price to be overvalued. The analysis suggests that in case of a pure price shock, the open interest finally decreases under its initial value, but we cannot confirm that without confidential intervals.

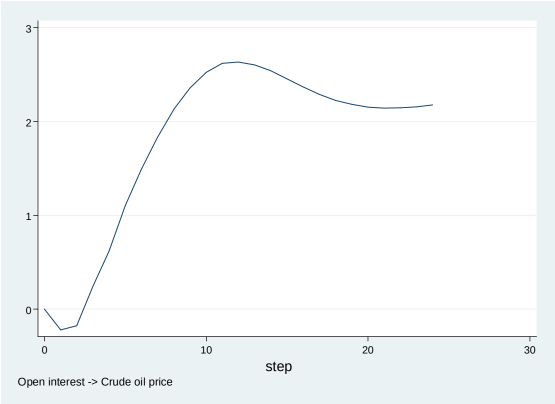


Figure 12: IRF, impulse: Open interest, response: Crude oil price

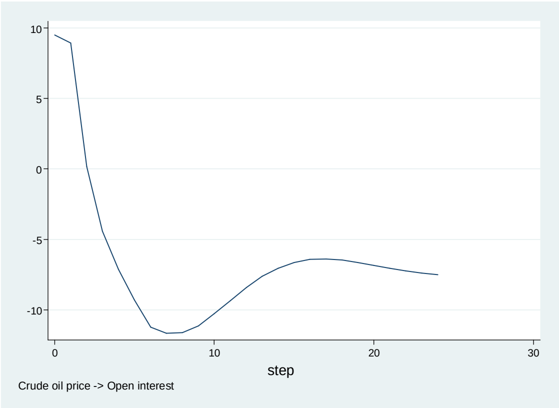


Figure 13: IRF, impulse: Crude oil price, response: Open Interest

In case of a shock on the open interest side, we assume a creation of 40 thousand new long and short contracts. The figure suggests that at the presence of new investments, the price shock is a positive and permanent one. That would mean that not only fundamental factors, but also the amount of investment on the crude oil price would be able to move the price.

To provide a better robustness for our estimates, we re-estimated the model in the statistical software Gretl without imposing any restrictions, because Gretl allows for including the bootstrap confidence intervals. The analysis showed in the two following figures shows not only that the shock in the unrestricted model looks the same (which justifies the restrictions we imposed), but also that the 95 % confidence band of the impulse of the open interest on the oil price is above zero. The shock to the crude oil price on the open interest, on the other way, does not need to be significant.

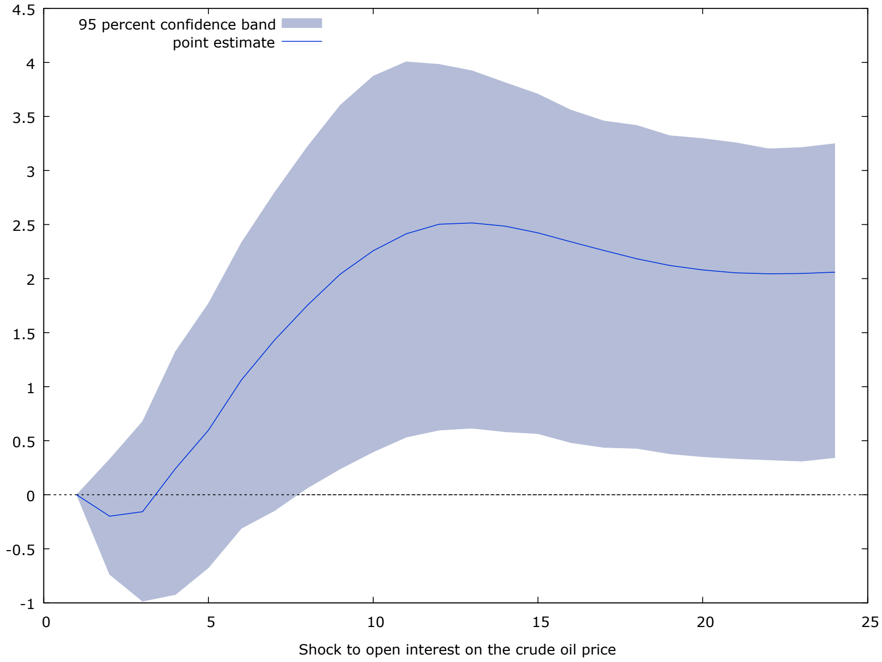


Figure 14: IRF of an unrestricted model, impulse: open interest, response: crude oil price

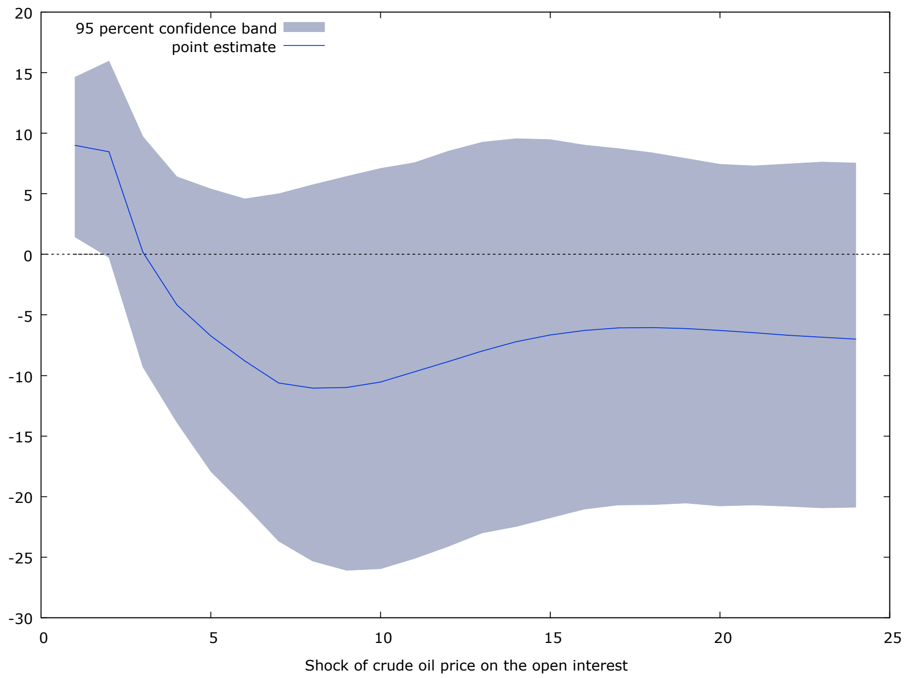


Figure 15: IRF of an unrestricted model, impulse: crude oil price, response: open interest

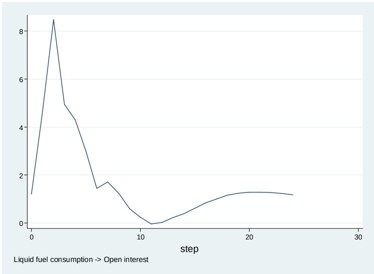


Figure 16: IRF, impulse: Liquid fuel consumption, response: Open interest

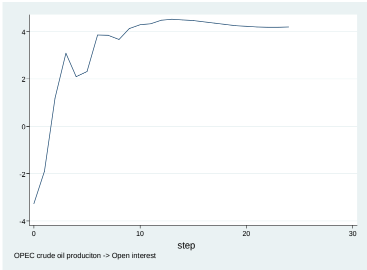


Figure 17: IRF, impulse: OPEC crude oil production, response: Open interest

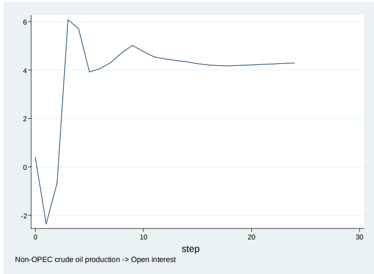


Figure 18: IRF, impulse: Non-OPEC crude oil production, response: Open interest

As we can see in the Figures 16 to 18, basically any shock to a fundamental variable on the crude oil prices leads to an increase of the open interest on the market. We have discussed the uncertainty regarding supply and demand caused by many unknowns already, for example the peak oil. Because of that, basically any price on the market is justifiable by some economic reasoning. Most of the participants on the crude oil market are now the non-commercial ones, which means they trade the futures contracts only to earn profit on selling them (or buying them in case of a short position). An unexpected inflow of such participants can lead the oil price to a higher level even without changes in the underlying fundamentals.

2.6 Is the crude oil price volatility attracting more speculative traders?

After showing how an inflow in the open interest on the crude oil market can cause a persistent positive shock to crude oil prices, we will take a closer look at whether an increased volatility in the crude oil prices can attract more speculative capital to the market. We will test this hypothesis with the concept of Granger non-causality introduced by Granger (1969), which simply tests whether the lagged values of one variable increase the explanatory power for the current value of another variable (Kočenda and Černý 2007, p. 158). In our case, such an analysis becomes more complicated because the open interest is assumed to be I(1). We will therefore rely on the method suggested by Toda and Yamamoto (1995), who proposed overidentifying the model by adding *d* extra lags to the true number of lags *k*, where *d* is the maximal suspected order of integration in the VAR model, and testing for joint significance of the first *k* lags. The expected inefficiency added to the model is supposed to be small for models with a generous lag length and a small number or variables (Toda and Yamamoto 1995, p. 246). This method has been further investigated by Hacker and Hatemi-J (2006), who concluded that there is an improved power of the test after using a bootstrapped distribution, especially in presence of the autoregressive conditional heteroskedasticity (ARCH).

As the bootstrapping option is not allowed while estimating a VAR model under the current version of Stata, we estimated two restricted models using the OLS method.

$$open\_interest_t = c_1 + \sum_{i=1}^{n} open\_interest_{t-i} + \sum_{i=1}^{n} variation\_in\_prices_{t-i} + \varepsilon_1 \qquad (5)$$

$$variation\_in\_prices_t = c_2 + \sum_{i=1}^{n} open\_interest_{t-i} + \sum_{i=1}^{n} variation\_in\_prices_{t-i} + \varepsilon_2 \quad (6)$$

where *open_interest* stands for the open interest on the crude oil futures market NYMEX and *variation_in_prices* for the changes in the crude oil futures on NYMEX measured by the standard deviation of the prices within the week.

We got different results in estimating the lag length using the Akaike's information criterion, Schwartz's Bayesian information criterion and the Hannah and Quinn information criterion. Being aware that the suggested method might be distorted by having too few lags (Hacker and Hatemi-J 2006, p. 1499), we decided to rely on the Hannah and Quinn information criterion's suggestion of 15 lags in the model.

Finally, we estimated the equations (5) and (6) using 16 lags, one extra lag to follow the Toda and Yamamoto method. The results indicated a strong presence of ARCH in the residuals, therefore we followed the example of Hacker and Hatemi-J (2006, p. 1498) and used the bootstrapping method. Finally, we applied the Wald test for the joint significance of the first 15 lagged variables. The test for the joint significance of the lagged variables of the open interest in the equation (6) shows that we cannot reject the null hypothesis that the *open interest* does not Granger cause the *variation in prices*. The test for the joint significance of the lagged variables of the *variation in prices* rejects the null hypothesis of the variables being jointly zero on the 5 % level of significance, in other words, the test rejects the null hypothesis that *variation in prices* does not Granger cause the *open interest*. We therefore conclude that there is small evidence that using past variation of prices for explaining the current values of the open interest improves the explanatory power of the model. Because of the large number of lags and the different signs on the lags it is difficult to state whether the increased volatility leads the increases in the open interest or not, though.

2.7 Forecasting the crude oil price

After showing the behavior of our two equilibria model in the sample, we want to have a brief look at the behavior of the model for out-of-sample forecasting. The following figure confirms that the out-of sample forecasting is extremely difficult and usually unreliable.

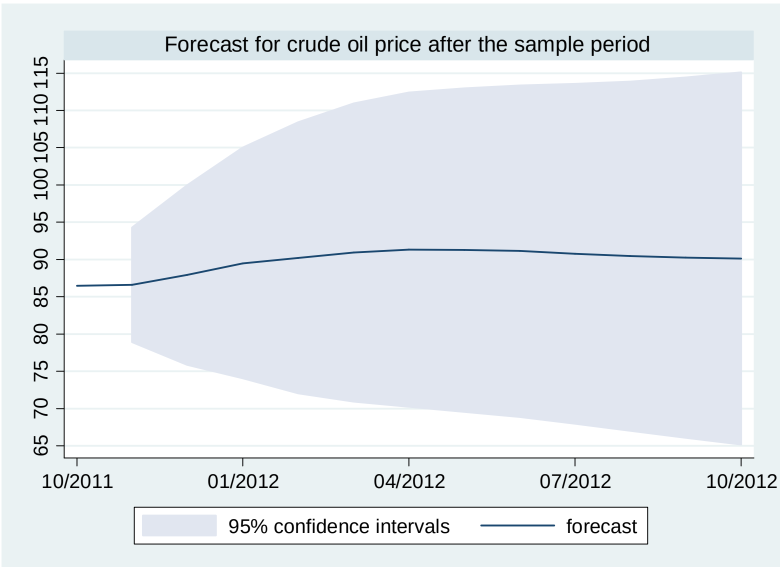


Figure 19: Out-of-sample forecasting for the crude oil price

In the figure we can easily observe the low predictive power of the model. Already in the first forecasted month the 95 % confidence intervals show the possible range of approximately USD 15, and since we are dealing with a VECM model, the forecasted confidence intervals are further widening with the time, so that within just one year, we have a range from USD 65 to USD 115. As the second value is almost twice as big as the first one, we can imagine the decision makers basing their decisions on the crude oil forecasts would like to use a better forecast. The failure of our model to provide a better forecast is by no means a sporadic one. Peter Davies, a former chief economist of British Petroleum, said for example at the Annual Conference of the International Association for

Energy Economics in 2007 that “we cannot forecast oil prices with any degree of accuracy over any period whether short or long” (Davies 2007 in Alquist et al. 2011, p. 26). Alquist et al. (2011) provide a detailed comparison of forecasting powers of different models. They note that many financial institutions, for example the International Monetary Fund or some central banks, use the price of NYMEX oil futures as a proxy for the expected spot price, but conclude that such a forecast is neither more accurate than a no-change forecast represented by a random walk nor more successful in predicting the sign of the change in the nominal price. At the one or three months horizon, such a forecast is according to them actually inferior to tossing a coin (Alquist et al. 2011, p. 22). Alquist et al. furthermore investigated the forecasts provided by professional macroeconomic forecasters and concluded that the forecasts are mostly lower than the oil price at the time of the forecast, which resulted into large and persistently negative forecast errors especially during large oil fluctuation periods, for example 2003-2009. The authors provide a VAR model which tends to have better predictive powers than a no-change estimate, but only for up to six months period. The conclusion is therefore similar to ours – it is impossible to provide a reliable long-term oil price forecasts.

2.8 Summary of the empirical results

In our vector error correction model, we were able to find two cointegrating relationships indicating two long run equilibria between our variables on the crude oil market. The first equilibrium was described as an equilibrium on the physical market for the crude oil, where the supply and demand variables play role, while the second equilibrium as an equilibrium on the speculative market, represented by the futures crude oil price, the open interest on the NYMEX market and the physical inventories, which speculative role has been suggested since the work by Working (1949). The presence of inventories allows even the commercial traders to take part in speculation, when they hold larger inventories in the expectation of higher prices. Therefore we did not distinguish in our analysis between commercial and non-commercial traders, as it was the case for example for Buyuksahin and Harris (2011) or the Interagency Task Force on Commodity Markets (2008). We based our analysis on a monthly basis, as we believed the monthly basis allowed us to not only snap the market development from both fundamental and speculative point of view, but also provided us with more observations, as many of the studies are based on quarterly data with not more than 70 observations. We were also searching for a longer-term relationship, which, as we believe, cannot be provided by the daily data.

The directions of our coefficients on the fundamental variables are corresponding to the economic theory and the previous literature (for a brief summary see the table IV), the signs of the OPEC and non-OPEC productions are negative and the sign on the liquid fuel consumption is positive. We also find a significant negative effect of holding stocks on the crude oil price, which corresponds to the studies by Chevillon and Rifflart (2009); Dées et al. (2007) or Kilian and Murphy (2010). Finally, we find a long-term positive relationship between the crude oil prices and the open interest on the market. Using the impulse-response analysis, we show that when the variables are in an equilibrium, an increase in the open interest can lead to a persistent increase in the crude oil price. A shock to the crude oil prices, on the other end, leads probably just to a transitory increase in the open interest.

Such a result is in a contradiction to the results obtained by Buyuksahin and Harris (2011) or the Interagency Task Force on Commodity Markets (2008), who found Granger causality from the crude oil prices to the open interest, but not the other way round. This can be caused by the usage of different intervals in the analysis, as both of the studies are based on a daily period analysis. Both of the studies therefore focus on a very short term relationship, trying to find out, whether the daily changes in the open interest can lead to the changes in the crude oil price. We believe that such a process would be a longer-term one, as also our analysis suggests.

We also tested the hypothesis whether the oil price volatility Granger causes the open interest on the market. We rejected the hypothesis that the price volatility does not Granger cause the open interest, although the p-value was only slightly lower than the 5 % and the direction of the effect was unclear. On the other hand, it seems that the open interest does not Granger cause the volatility in prices on the crude oil market.

The author, year	Period, data	Methodology	Results
Dées et al. (2007)	1984Q1-2002Q1, quarterly	Dynamic OLS	very low price elasticity of supply and demand, higher supply elasticity for the USA (5.5 %) OPEC capacity can cause permanent changes in crude oil price significant influence of OECD stocks on the crude oil price collapse of OPEC cartel would lead to a decrease in crude oil price, but also an increase in price volatility
Kaufmann et al. (2008)	1986-2000, quarterly	Dynamic OLS	higher U.S. refinery utilization rates depress the crude oil price non-linear effect of OPEC operating capacity higher OECD stocks lower the crude oil prices backwardation or contango[7] represent a concern of the future prices and can also influence the current prices
Interagency Task Force on Commodity Markets (2008)	January 2003 – June 2008, daily	Granger non-causality tests	Granger causality from changes in prices to some of the groups in the open interest on the crude oil market, not the other way round
Breitenfellner and Cuaresma (2008)	1983-2007, monthly	VAR, VECM	using the EUR/USD exchange rate improves the oil price forecast
Chevillon and Rifflart (2009)	1989-2006, quarterly	VECM	based the model strictly on the fundamental variables, the rest is supposed to be the risk premium, this premium accounts for half of the increase in the oil price between 2000 and 2006 the demand and supply weakly exogenous in the short run, does not need to hold for longer periods
Kilian and Murphy (2010)	1973.2-2009.8, monthly	Structural VAR	emphasize the power of OECD inventories on the crude oil price, reject the role of speculation in the 2003-2008 price hike regard the low demand price elasticity estimates to be non-credible, their estimate substantially higher
Buyuksahin and Harris (2011)	July 2000- June 2004, July 2004- March 2009, daily	Granger non-causality tests based on bivariate VAR	Granger causality from changes in prices to changes in traders' position, not the other way round
Šlechta (2012)	1994 – 2011, monthly	VECM, Granger causality	a long term relationship between the crude oil prices and the open interest on NYMEX an increase in the open interest can lead to a persistent increase in the prices the oil price volatility Granger causes the open interest, even if the direction is unclear

Table 4: Brief summary of recent findings in the research on the crude oil prices

[7] backwardation – the price of the near month future contract is higher than the far-month one, contango – the price of the near month future contract is lower than the far month one (Kaufmann et al. 2008)

Conclusion

In our analysis, we found a long term relationship between the crude oil prices and the open interest on the NYMEX. Furthermore, we showed that when both variables are in an equilibrium and there is an unexpected positive shock to the open interest, there is also a persistent increase in the oil prices. We believe it is the risk factor which stands for the economic force behind the cointegrating relationship of the price and the open interest. Such a risk premium has been suggested for example by Chevillon and Rifflart (2009), who attributed to the risk premium half of the increase in the crude oil prices between the years 2000 and 2006. There are plenty of reasons why the crude oil became a risky asset. Starting from the phenomenon of the peak oil, which possible consequences are described by Hirsch et al. (2006, p. 7) as "[...] pervasive and long lasting.", to the fact that the shortages in the crude oil supply are usually compensated by an oil cartel, whose remaining reserves are questionable as well. These problems have been further catalyzed by the growing consumption coming from the developing countries. Well consistent with the Hotelling's theory, the commodity price is influenced by its scarcity rent, but in a world of uncertainty, no one knows how high the scarcity rent should be. Such an uncertainty provides a playground for speculation on the market and explains, why the models based just on fundamental variables generally failed to explain the oil price increases after 2003, even if they used to provided satisfactory results for earlier periods (see for example Kaufmann 2011). A change of the risk awareness could be a foundation for such a shock to the open interest we simulated.

The field for further research is extremely wide for the crude oil prices, as we have not found any model which power to explain the oil prices we would find completely satisfactory. We restricted our analysis to the factors that we believed were the major ones, since we did not want to lose more observations by adding further variables. Other factors have been proven significant for the crude oil market, though, and should be therefore included into the models as well. Kaufmann et al. (2008) for example emphasize the effect of refinery utilization, for which they because of the lack of better data use the U.S. refinery utilization rates. The same authors find a non-linear effect in the OPEC operating capacity. Breitenfellner and Cuaresma (2008) show an increased predictive power of a model when using the USD/EUR exchange rate. The EIA offers a weighted basket of currencies against the dollar with the weights equal to the share of world consumption, but at the time of writing this thesis, the index has been only computed for the period after 1995. The EIA also offers the world gross domestic product weighted by the countries' consumption of the crude oil, which is only available for the period after 2002. We believe that there are also further factors which are important for the analysis, as for example the global proven oil reserves, which are provided by the IEA only on the annual basis, which makes the use of the data limited. Even if the reserves have been growing in the past, the drilling became more expensive as more of the crude oil is from unconventional sources. Utilization of such data would probably also help in explaining the crude oil prices. One must, however, find the ideal balance between the number of variables used in the econometrical model and the number of observations. On the one hand, more variables are provided quarterly than monthly, on the other hand, the analysis of quarterly data does not provide the researcher with enough observations to utilize the various variables. Furthermore, as we have already outlined, there are just limited sources for the data of the developing countries, although these countries have played the major roles in the recent years.

We believe it is almost impossible to gather all the relevant fundamental data for the crude oil price analysis, and therefore the risk factor will keep on playing an important role on the crude oil market, providing space for speculation on the prices. This obstacle, together with the impossibility to predict various macroeconomic and political events around the world, will further prevent the researchers, policy-makers or investors from being able to forecast the crude oil prices for the longer-run more precisely then up to now.

Bibliography

Alquist R./ Kilian, L./ Vigfusson, R. J. (2011): *Forecasting the Price of Oil*, Board of Governors of the Federal Reserve System International Finance Discussion Papers Number 1022.

Baumeister C./ Kilian, L. (2012): *Real-Time Analysis of Oil Price Risks Using Forecast Scenarios*, Bank of Canada Working Paper 2012-1.

Breitenfellner A./ Cuaresma, J. C. (2008): Crude Oil Prices and the USD/EUR Exchange Rate, in: *Monetary Policy & the Economy*, Q4/08, p. 102-121.

Brennan M. J. (1958): The Supply of Storage, in: *The American Economic Review*, 48(1), p. 50-72.

Buyuksahin B./ Harris, J. H. (2011): Do Speculators Drive Crude Oil Futures Prices?, in: *The Energy Journal*, 32(2), p. 167-202.

Chevillon G./ Rifflart, C. (2009): Physical Market Determinants of the Price of Crude Oil and the Market Premium, in: *Energy Economics*, 31(4), p. 537-549.

Cifarelli G./ Paladino, G. (2010): Oil Price Dynamics and Speculation: A Multivariate Financial Approach, in: *Energy Economics*, 32(2), p. 363-372.

Coleman L. (2012): Explaining Crude Oil Prices Using Fundamental Measures, in: *Energy Policy*, 40(1), p. 318-324.

Oil and Gas Reserves Committee (2005): *"Mapping" Subcommittee Final Report: Comparison of Selected Reserves and Resource Classiifcations and Associated Definitions*, Society of Petroleum Engineers.

de Almeida P./ Silva, P. D. (2009): The Peak of Oil Production—Timings and Market Recognition, in: *Energy Policy*, 37(4), p. 1267-1276.

Deaton A./ Laroque, G. (1996): Competitive Storage and Commodity Price Dynamics, in: *Journal of Political Economy*, 104(5), p. 896-923.

Dées S./ Karadeloglou, P./ Kaufmann, R. K./ Sánchez, M. (2007): Modelling the World Oil Market: Assessment of a Quarterly Econometric Model, in: *Energy Policy*, 35(1), p. 178-191.

Ederington, L./ Fernando, C./ Lee, T./ Linn, S./ May, A. (2011): *Factors Influencing Oil Prices: A Survey of the Current State of Knowledge in the Context of the 2007-2008 Oil Price Volatility*, U.S. Energy Information Administration Federal Government of the United States, Washington D.C.

Ellen S. t./ Zwinkels, R. C. J. (2010): Oil Price Dynamics: A Behavioral Finance Approach with Heterogeneous Agents, in: *Energy Economics*, 32(6), p. 1427-1434.

Elliott G./ Rothenberg, T. J./ Stock, J. H. (1996): Efficient Tests for an Autoregressive Unit Root, in: *Econometrica*, 64(4), p. 813-836.

Engle R. F./ Granger, C. W. J. (1987): Co-Integration and Error Correction: Representation, Estimation, and Testing, in: *Econometrica*, 55(2), p. 251-276.

Fattouh B. (2007): *The Drivers of Oil Prices: The Usefulness and Limitations of Non-Structural Model, the Demand-Supply Framework and Informal Approaches*, Oxford Institute for Energy Studies WPM 32.

Fattouh B. (2010): The Dynamics of Crude Oil Price Differentials, in: *Energy Economics*, 32(2), p. 334-342.

Gately D. (2004): OPEC's Incentives for Faster Output Growth, in: *Energy Journal*, 25(2), p. 75-96.

Gately D./ Huntington, H. G. (2002): The Asymmetric Effects of Changes in Price and Income on Energy and Oil Demand, in: *Energy Journal*, 23(1), p. 19.

Gonçalves S. l./ Kilian, L. (2004): Bootstrapping Autoregressions with Conditional Heteroskedasticity of Unknown Form, in: *Journal of Econometrics*, 123(1), p. 89-120.

Granger C. W. J. (1969): Investigating Causal Relations by Econometric Models and Cross-spectral Methods, in: *Econometrica*, 37(3), p. 424-438.

Granger C. W. J./ Newbold, P. (1973): Spurious Regressions in Econometrics, in: *Journal of Econometrics*, 2, p. 111-120.

Griffin J. M./ Neilson, W. S. (1994): The 1985-86 Oil Price Collapse and Afterwards: What Does Game Theory Add?, in: *Economic Inquiry*, 32(4), p. 543-561.

Griffin J. M./ Schulman, C. T. (2005): Price Asymmetry in Energy Demand Models: A Proxy for Energy-Saving Technical Change?, in: *Energy Journal*, 26(2), p. 1-21.

Hacker R. S./ Hatemi-J, A. (2006): Tests for Causality Between Integrated Variables Using Asymptotic and Bootstrap Distributions: Theory and Application, in: *Applied Econometrics*, 38, p. 1489-1500.

Hamilton J. D. (2009a): Causes and Consequences of the Oil Shock of 2007-08, in: *Brookings Papers on Economic Activity*, 1(215-261), p.

Hamilton J. D. (2009b): Understanding Crude Oil Prices, in: *The Energy Journal*, 30(2), p. 179-206.

Harris R. I. D. (1995): *Using Cointegration Analysis in Econometric Modelling*: Harvester Wheatsheaf, Prentice Hall.

He Y./ Wang, S./ Lai, K. K. (2010): Global Economic Activity and Crude Oil Prices: A Cointegration Analysis, in: *Energy Economics*, 32(4), p. 868-876.

Hicks B./ Kilian, L. (2009): *Did Unexpectedly Strong Economic Growth Cause the Oil Price Shock of 2003-2008?*, CEPR Discussion Paper No. DP7265.

Hirsch R. L./ Bezdek, R./ Wendling, R. (2006): Peaking of World Oil Production and Its Mitigation, in: *AIChE Journal*, 52(1), p. 2-8.

Hotelling H. (1931): The Economics of Exhaustible Resources, in: *Journal of Political Economy*, 39(2), p. 137-175.

Hubbert M. K. (1949): Energy from Fossil Fuels, in: *Science*, 109, p. 103-109.

Hubbert M. K. (1956): Nuclear Energy and Fossil Fuels, Publication Number 95, Shell Development Company, Houston, Texas, Presented on the Spring Meeting on Drilling and Production Practise, March 7-9, American Petroleum Institute, San Antonio, Texas.

IMF (2008): *Regional Economic Outlook: Middle East and Central Asia*, Washington, D.C.

Johansen S. (1995): *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, New York: Oxford University Press.

Kaldor N. (1939): Speculation and Economic Stability, in: *The Review of Economic Studies*, 7(1), p. 1-27.

Kaufmann K. K. (2011): The Role of Market Fundamentals and Speculation in Recent Price Changes for Crude Oil, in: *Energy Policy*, 39(1), p. 105-115.

Kaufmann R. K./ Dees, S./ Gasteuil, A./ Mann, M. (2008): Oil prices: The Role of Refinery Utilization, Futures Markets and Non-linearities, in: *Energy Economics*, 30(5), p. 2609-2622.

Kaufmann R. K./ Ullman, B. (2009): Oil Prices, Speculation, and Fundamentals: Interpreting Causal Relations among Spot and Futures Prices, in: *Energy Economics*, 31(4), p. 550-558.

Khan M. S. (2009): *The 2008 Oil Price "Bubble"*, Peterson Institute for International Economics PB09-19.

Kilian L. (2009): Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market, in: *American Economic Review*, 99(3), p. 1053-1069.

Kilian L./ Murphy, D. P. (2010): *The Role of Inventories and Speculative Trading in the Global Market for Crude Oil*, University of Michigan Working Paper.

Kočenda E./ Černý, A. (2007): *Elements of Time Series Econometrics: An Applied Approach*, Prague: Karolinum Press.

Krautkraemer J. A. (1998): Nonrenewable Resource Scarcity, in: *Journal of Economic Literature*, XXXVI, p. 2065-2107.

Krichene N. (2006): *World Crude Oil Markets: Monetary Policy and the Recent Oil Shock*, International Moneraty Fund Working Paper No. 06/62.

Kwiatkowski D./ Phillips, P. C. B./ Schmidt, P./ Shin, Y. (1992): Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root?, in: *Journal of Econometrics*, 54(1–3), p. 159-178.

Lütkepohl H./ Reimers, H.-E. (1992): Impulse Response Analysis of Cointegrated Systems, in: *Journal of Economic Dynamics and Control*, 16(1), p. 53-78.

OPEC (2012): *About us*, [online], available: http://www.opec.org/opec_web/en/17.htm [Retrieved 15/05/2012].

Parsons J. E. (2009): *Black Gold & Fool's Gold: Speculation in the Oil Futures Market*, CEEPR 09-013.

Pindyck R. S. (1978): Gains to Producers from the Cartelization of Exhaustible Resources, in: *The Review of Economics and Statistics*, 60(2), p. 238-251.

Pindyck R. S. (1999): The Long-Run Evolutions of Energy Prices, in: *The Energy Journal*, 20(2), p. 1-28.

Slade M. E./ Thille, H. (2009): Whither Hotelling: Tests of the Theory of Exhaustible Resources, in: *Annual Review of Resource Economics*, 1(1), p. 239-260.

Smith J. L. (2005): Instructable OPEC? Behavioral Tests of the Cartel Hypothesis, in: *The Energy Journal*, 26 (1), p. 51-82.

StataCorp (2009): *Stata: Release 11*, Statistical Software, College Station, TX: StataCorp LP

Toda H. Y./ Yamamoto, T. (1995): Statistical Inference in Vector Autoregressions with Possibly Integrated Processes, in: *Journal of Econometrics*, 66, p. 225-250.

U.S. CFTC (2012): *Explanatory Notes*, [online], available: http://www.cftc.gov/MarketReports/CommitmentsofTraders/ExplanatoryNotes/index.htm [Retrieved 15/05/2012].

U.S. EIA (2011): *Performance Profiles of Major Energy Producers 2009*, U.S. Department of Energy, Washington, DC.

U.S. EIA (2012a): *Canada*, [online], available: http://www.eia.gov/countries/cab.cfm?fips=CA [Retrieved 15/05/2012].

U.S. EIA (2012b): *Demand: Non-OECD*, [online], available: http://205.254.135.7/finance/markets/demand-nonoecd.cfm [Retrieved 15/05/2012].

U.S. EIA (2012c): *Demand: OECD*, [online], available: http://205.254.135.7/finance/markets/demand-oecd.cfm [Retrieved 15/05/2012].

U.S. EIA (2012d): *Supply: Non-OPEC*, [online], available: http://205.254.135.7/finance/markets/supply-nonopec.cfm [Retrieved 15/05/2012].

U.S. EIA (2012e): *Supply: OPEC*, [online], available: http://205.254.135.7/finance/markets/supply-opec.cfm [Retrieved 15/05/2012].

Watkins G. C. (2006): Oil Scarcity: What Have the Past Three Decades Revealed?, in: *Energy Policy*, 34(5), p. 508-514.

Wooldridge J. M. (2003): *Introductory Econometrics: A Modern Approach* Vol. 2e, Mason, Ohio: South-Western College Publ.

Working H. (1949): The Theory of Price of Storage, in: *The American Economic Review*, 39(6), p. 1254-1262.

Sources of the data:

U.S. EIA	
Total Oil Supply	http://205.254.135.7/cfapps/ipdbproject/IEDIndex3.cfm?tid=50&pid=53&aid=1
Total Petroleum Consumption	http://205.254.135.7/cfapps/ipdbproject/IEDIndex3.cfm?tid=50&pid=54&aid=2
Total Petroleum Stocks	http://205.254.135.7/cfapps/ipdbproject/IEDIndex3.cfm?tid=50&pid=5&aid=5
Crude Oil Proved Reserves	http://205.254.135.7/cfapps/ipdbproject/IEDIndex3.cfm?tid=5&pid=57&aid=6
NYMEX Futures Price	http://205.254.135.7/dnav/pet/pet_pri_fut_s1_d.htm
Short-Term Energy Outlook	http://205.254.135.7/forecasts/steo/query/
U.S. CFTC	
Open Interest on NYMEX	http://www.cftc.gov/MarketReports/CommitmentsofTraders/HistoricalCompressed/index.htm

Data last accessed on May 15th, 2012

Appendix

Part I: The VECM estimation

The rank of cointegration

```
. vecrank  wtif lfc_w oecdinv cop_rest cop_o oia, trend(constant) lags(3) levela
```

Johansen tests for cointegration

Trend: constant — Number of obs = 211

Sample: 04/1994 - 10/2011 — Lags = 3

maximum rank	parms	LL	eigenvalue	trace statistic	5% critical value	1% critical value
0	78	-2234.2993		132.6953	94.15	103.18
1	89	-2206.9296	0.22851	77.9560	68.52	76.07
2	98	-2185.685	0.18239	35.4667*1*5	47.21	54.46
3	105	-2177.3122	0.07630	18.7212	29.68	35.65
4	110	-2172.228	0.04705	8.5527	15.41	20.04
5	113	-2168.6154	0.03366	1.3276	3.76	6.65
6	114	-2167.9516	0.00627			

The test for weak exogeneity

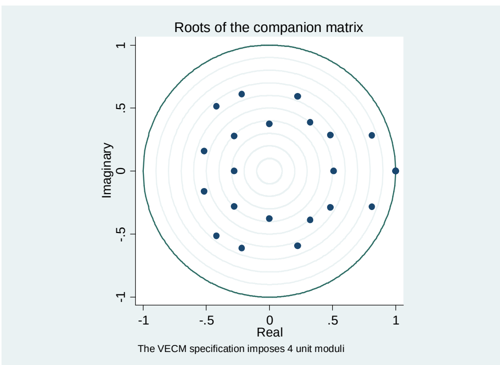
```
. test [D_oecdinv]L1._ce1 [D_oecdinv]L1._ce2

 ( 1)  [D_oecdinv]L._ce1 = 0
 ( 2)  [D_oecdinv]L._ce2 = 0

           chi2(  2) =    1.22
         Prob > chi2 =    0.5436
```

The post-estimation tests

Lagrange-multiplier test

lag	chi2	df	Prob > chi2
1	40.2370	36	0.28811
2	28.0622	36	0.82498

H0: no autocorrelation at lag order



Eigenvalue stability condition

Eigenvalue	Modulus
1	1
1	1
1	1
1	1
.814068 + .2832488*i*	.861938
.814068 - .2832488*i*	.861938
-.4198115 + .5135629*i*	.663316
-.4198115 - .5135629*i*	.663316
-.2181851 + .6113981*i*	.649163
-.2181851 - .6113981*i*	.649163
.2233643 + .5926453*i*	.63334
.2233643 - .5926453*i*	.63334
.4828501 + .2868241*i*	.561616
.4828501 - .2868241*i*	.561616
-.518614 + .1600109*i*	.542737
-.518614 - .1600109*i*	.542737
.5077588	.507759
.324741 + .3870242*i*	.505217
.324741 - .3870242*i*	.505217
-.2812878 + .2785707*i*	.395884
-.2812878 - .2785707*i*	.395884
-.00113247 + .3749025*i*	.374904
-.00113247 - .3749025*i*	.374904
-.2806498	.28065

The VECM specification imposes 4 unit moduli.

Jarque-Bera test

Equation	chi2	df	Prob > chi2
D_wtif	49.930	2	0.00000
D_lfc_w	2.631	2	0.26832
D_oecdinv	63.978	2	0.00000
D_cop_rest	2.348	2	0.30912
D_cop_o	1016.152	2	0.00000
D_oia	19.169	2	0.00007
ALL	1154.209	12	0.00000

Skewness test

Equation	Skewness	chi2	df	Prob > chi2
D_wtif	-.44844	7.038	1	0.00798
D_lfc_w	-.25874	2.343	1	0.12583
D_oecdinv	.3348	3.923	1	0.04763
D_cop_rest	-.25183	2.220	1	0.13626
D_cop_o	-.89125	27.801	1	0.00000
D_oia	.29581	3.063	1	0.08012
ALL		46.388	6	0.00000

Kurtosis test

Equation	Kurtosis	chi2	df	Prob > chi2
D_wtif	5.214	42.892	1	0.00000
D_lfc_w	2.8186	0.288	1	0.59152
D_oecdinv	5.6198	60.055	1	0.00000
D_cop_rest	3.1211	0.128	1	0.72010
D_cop_o	13.628	988.351	1	0.00000
D_oia	4.3568	16.107	1	0.00006
ALL		1107.820	6	0.00000

The results of the VECM model

Vector error-correction model

Sample: 05/1994 - 10/2011	No. of obs	=	210
	AIC	=	21.675
Log likelihood = -2143.875	HQIC	=	22.52553
Det(Sigma_ml) = 29.68365	SBIC	=	23.7789

Equation	Parms	RMSE	R-sq	chi2	P>chi2
D_wtif	21	3.71094	0.3666	122.5584	0.0000
D_lfc_w	21	1.24262	0.3450	110.6988	0.0000
D_oecdinv	19	.414844	0.1349	32.74831	0.0257
D_cop_rest	21	.311259	0.1965	51.52189	0.0002
D_cop_o	21	.429261	0.1586	42.3876	0.0038
D_oia	21	39.3006	0.1443	35.31047	0.0261

(1) [D_oecdinv]L._ce1 = 0
(2) [D_oecdinv]L._ce2 = 0

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
D_wtif						
_ce1						
L1.	-.0542816	.0391837	-1.39	0.166	-.1310802	.022517
_ce2						
L1.	-.1505146	.0429408	-3.51	0.000	-.234677	-.0663522
wtif						
LD.	.3878692	.0687826	5.64	0.000	.2530577	.5226806
L2D.	.2486873	.0742156	3.35	0.001	.1032274	.3941472
L3D.	.0981519	.0770195	1.27	0.203	-.0528034	.2491073
lfc_w						
LD.	.023755	.3333906	0.07	0.943	-.6296787	.6771886
L2D.	-.3491955	.295506	-1.18	0.237	-.9283766	.2299857
L3D.	-.2910112	.2400319	-1.21	0.225	-.7614651	.1794427
oecdinv						
LD.	-.6710134	1.00103	-0.67	0.503	-2.632995	1.290968
L2D.	1.316095	.9843429	1.34	0.181	-.6131813	3.245372
L3D.	.9970289	.9863175	1.01	0.312	-.9361178	2.930176
cop_rest						
LD.	.7699733	.9542771	0.81	0.420	-1.100375	2.640322
L2D.	.5064523	.9139016	0.55	0.579	-1.284762	2.297666
L3D.	.903498	.8797435	1.03	0.304	-.8207676	2.627764
cop_o						
LD.	-.4848684	.9659897	-0.50	0.616	-2.378173	1.408437
L2D.	.3909956	.9451106	0.41	0.679	-1.461387	2.243378
L3D.	.8115865	.9483006	0.86	0.392	-1.047048	2.670221
oia						
LD.	-.0124395	.0074441	-1.67	0.095	-.0270296	.0021506
L2D.	-.0053638	.0072489	-0.74	0.459	-.0195714	.0088438
L3D.	.0023967	.0072533	0.33	0.741	-.0118195	.0166129
_cons	-.4711109	.4761967	-0.99	0.323	-1.404439	.4622176

D_lfc_w						
_ce1						
L1.	.0728971	.013147	5.54	0.000	.0471293	.0986648
_ce2						
L1.	-.0712989	.0144076	-4.95	0.000	-.0995373	-.0430605
wtif						
LD.	-.0103517	.0230325	-0.45	0.653	-.0554946	.0347912
L2D.	-.0249525	.024855	-1.00	0.315	-.0736674	.0237624
L3D.	.0372115	.0257959	1.44	0.149	-.0133474	.0877704
lfc_w						
LD.	-.0188503	.111773	-0.17	0.866	-.2379213	.2002207
L2D.	.1372472	.0990271	1.39	0.166	-.0568423	.3313368
L3D.	.2307208	.0804081	2.87	0.004	.0731239	.3883177
oecdinv						
LD.	-.0996449	.3351966	-0.30	0.766	-.7566182	.5573285
L2D.	-.2004468	.3296109	-0.61	0.543	-.8464723	.4455788
L3D.	-.1759505	.3302703	-0.53	0.594	-.8232684	.4713675
cop_rest						
LD.	-.1283008	.3196456	-0.40	0.688	-.7547947	.4981932
L2D.	.0254659	.3060611	0.08	0.934	-.5744029	.6253347
L3D.	-.1609057	.2945875	-0.55	0.585	-.7382865	.4164751
cop_o						
LD.	-.1212588	.323476	-0.37	0.708	-.7552601	.5127425
L2D.	-.3751699	.3164791	-1.19	0.236	-.9954575	.2451176
L3D.	-.5747763	.3175428	-1.81	0.070	-1.197149	.0475962
oia						
LD.	-.0033158	.0024932	-1.33	0.184	-.0082023	.0015707
L2D.	-.0041226	.0024278	-1.70	0.089	-.008881	.0006359
L3D.	-.0060436	.0024293	-2.49	0.013	-.0108049	-.0012822
_cons	-.5677836	.1596588	-3.56	0.000	-.8807091	-.2548582

D_oecdinv						
_ce1						
L1.	4.07e-19	1.07e-18	0.38	0.703	-1.69e-18	2.50e-18
_ce2						
L1.	-4.07e-19	1.07e-18	-0.38	0.703	-2.50e-18	1.69e-18
wtif						
LD.	-.0022687	.0076438	-0.30	0.767	-.0172503	.012713
L2D.	-.0178026	.0079761	-2.23	0.026	-.0334354	-.0021697
L3D.	-.0083404	.0081183	-1.03	0.304	-.024252	.0075712
lfc_w						
LD.	-.0024316	.0232802	-0.10	0.917	-.0480598	.0431967
L2D.	-.0582429	.0259076	-2.25	0.025	-.1090209	-.0074649
L3D.	-.0182576	.0239353	-0.76	0.446	-.0651699	.0286548
oecdinv						
LD.	.1027298	.1118992	0.92	0.359	-.1165886	.3220483
L2D.	.0177958	.1098838	0.16	0.871	-.1975725	.2331641
L3D.	.0823293	.1102502	0.75	0.455	-.1337571	.2984157
cop_rest						
LD.	-.0613263	.0976228	-0.63	0.530	-.2526634	.1300109
L2D.	.2522903	.0988234	2.55	0.011	.0585999	.4459807
L3D.	-.0217252	.0980363	-0.22	0.825	-.2138728	.1704223
cop_o						
LD.	.1705024	.1069361	1.59	0.111	-.0390884	.3800933
L2D.	.0085913	.1050676	0.08	0.935	-.1973375	.2145201
L3D.	.19283	.1057924	1.82	0.068	-.0145193	.4001793
oia						
LD.	.0003621	.0007896	0.46	0.647	-.0011856	.0019097
L2D.	-.0003146	.0007662	-0.41	0.681	-.0018163	.0011871
L3D.	.0003925	.0007643	0.51	0.608	-.0011054	.0018905
_cons	-.0039722	.0321067	-0.12	0.902	-.0669002	.0589559

D_cop_rest						
_ce1						
L1.	-.0109495	.0032897	-3.33	0.001	-.0173972	-.0045018
_ce2						
L1.	.0084958	.0036051	2.36	0.018	.0014299	.0155618
wtif						
LD.	.0006002	.0057693	0.10	0.917	-.0107074	.0119077
L2D.	-.0122977	.0062254	-1.98	0.048	-.0244991	-.0000962
L3D.	-.0032827	.0064608	-0.51	0.611	-.0159456	.0093801
lfc_w						
LD.	-.0535566	.0279797	-1.91	0.056	-.1083959	.0012826
L2D.	-.025993	.0247949	-1.05	0.294	-.0745902	.0226042
L3D.	.0019679	.0201368	0.10	0.922	-.0374995	.0414354
oecdinv						
LD.	.0938188	.0839623	1.12	0.264	-.0707443	.2583818
L2D.	-.0381998	.0825629	-0.46	0.644	-.2000201	.1236205
L3D.	.1142283	.0827283	1.38	0.167	-.0479162	.2763728
cop_rest						
LD.	-.1530844	.0800533	-1.91	0.056	-.309986	.0038171
L2D.	-.1817618	.076659	-2.37	0.018	-.3320107	-.0315128
L3D.	-.137839	.0737897	-1.87	0.062	-.2824643	.0067862
cop_o						
LD.	.0972002	.0810248	1.20	0.230	-.0616054	.2560059
L2D.	.0293741	.0792729	0.37	0.711	-.1259979	.184746
L3D.	.0833229	.0795399	1.05	0.295	-.0725724	.2392183
oia						
LD.	.0003906	.0006244	0.63	0.532	-.0008332	.0016145
L2D.	.0003121	.0006081	0.51	0.608	-.0008797	.0015039
L3D.	.0003629	.0006084	0.60	0.551	-.0008297	.0015554
_cons	.1444101	.0399657	3.61	0.000	.0660788	.2227414
D_cop_o						
_ce1						
L1.	-.0006924	.0028802	-0.24	0.810	-.0063374	.0049525
_ce2						
L1.	-.0029026	.0031563	-0.92	0.358	-.0090889	.0032837
wtif						
LD.	.0051229	.0079285	0.65	0.518	-.0104166	.0206624
L2D.	.0183697	.0083887	2.19	0.029	.0019281	.0348113
L3D.	.0179284	.0086095	2.08	0.037	.0010541	.0348026
lfc_w						
LD.	.0248601	.0307652	0.81	0.419	-.0354387	.0851588
L2D.	.061293	.0300046	2.04	0.041	.002485	.120101
L3D.	.0433613	.0260195	1.67	0.096	-.0076359	.0943584
oecdinv						
LD.	-.0573107	.1157902	-0.49	0.621	-.2842554	.169634
L2D.	-.0976561	.1137675	-0.86	0.391	-.3206363	.125324
L3D.	.086267	.1140857	0.76	0.450	-.1373369	.3098709
cop_rest						
LD.	.060229	.1048997	0.57	0.566	-.1453705	.2658286
L2D.	-.2081502	.1036677	-2.01	0.045	-.4113351	-.0049653
L3D.	.0915886	.1015728	0.90	0.367	-.1074905	.2906677
cop_o						
LD.	-.1361366	.111093	-1.23	0.220	-.3538747	.0816016
L2D.	-.2132572	.1089642	-1.96	0.050	-.426823	.0003087
L3D.	-.0542885	.1095599	-0.50	0.620	-.269022	.160445
oia						
LD.	-.0004624	.0008351	-0.55	0.580	-.0020992	.0011744
L2D.	-.0003334	.0008116	-0.41	0.681	-.0019241	.0012572
L3D.	-.0011659	.0008106	-1.44	0.150	-.0027548	.0004229
_cons	.019409	.0433962	0.45	0.655	-.0656459	.104464

D_oia						
_ce1						
L1.	.1168255	.4156819	0.28	0.779	-.697896	.931547
_ce2						
L1.	-.7552431	.4555392	-1.66	0.097	-1.648083	.1375972
wtif						
LD.	.7376604	.7284545	1.01	0.311	-.6900842	2.165405
L2D.	-1.295678	.7860795	-1.65	0.099	-2.836365	.2450096
L3D.	-.1110895	.8158266	-0.14	0.892	-1.71008	1.487901
lfc_w						
LD.	4.691377	3.534438	1.33	0.184	-2.235995	11.61875
L2D.	6.107383	3.131603	1.95	0.051	-.0304462	12.24521
L3D.	.6985402	2.542937	0.27	0.784	-4.285524	5.682604
oecdinv						
LD.	7.265288	10.60137	0.69	0.493	-13.51302	28.04359
L2D.	17.65853	10.4247	1.69	0.090	-2.773511	38.09056
L3D.	4.966366	10.44556	0.48	0.634	-15.50656	25.4393
cop_rest						
LD.	-12.18679	10.10904	-1.21	0.228	-32.00015	7.626573
L2D.	-5.651483	9.679707	-0.58	0.559	-24.62336	13.32039
L3D.	9.791845	9.316994	1.05	0.293	-8.469128	28.05282
cop_o						
LD.	3.670313	10.23062	0.36	0.720	-16.38133	23.72196
L2D.	9.283253	10.00935	0.93	0.354	-10.33471	28.90122
L3D.	4.36228	10.04302	0.43	0.664	-15.32167	24.04623
oia						
LD.	-.1115632	.0788494	-1.41	0.157	-.2661053	.0429788
L2D.	-.1064182	.0767831	-1.39	0.166	-.2569103	.0440739
L3D.	-.0607367	.0768304	-0.79	0.429	-.2113216	.0898482
_cons	.1490413	5.048623	0.03	0.976	-9.746078	10.04416

Cointegrating equations

Equation	Parms	chi2	P>chi2
_ce1	4	134.4822	0.0000
_ce2	2	75.57157	0.0000

Identification: beta is overidentified

(1) [_ce1]wtif = 1
(2) [_ce2]wtif = 1
(3) [_ce1]oia = 0
(4) [_ce2]lfc_w = 0
(5) [_ce2]cop_rest = 0
(6) [_ce2]cop_o = 0

beta	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
_ce1						
wtif	1	.	.	.	.	.
lfc_w	-9.035065	.9464847	-9.55	0.000	-10.89014	-7.179989
oecdinv	4.821967	1.244983	3.87	0.000	2.381844	7.262089
cop_rest	18.51262	1.81517	10.20	0.000	14.95495	22.07028
cop_o	5.453178	1.156849	4.71	0.000	3.185795	7.72056
oia	(omitted)					
_trend	-.2336901	.0693511	-3.37	0.001	-.3696156	-.0977645
_cons	-215.8173	.	.	.	.	.
_ce2						
wtif	1	.	.	.	.	.
lfc_w	(omitted)					
oecdinv	3.444691	1.071664	3.21	0.001	1.344269	5.545113
cop_rest	(omitted)					
cop_o	(omitted)					
oia	-.0503604	.0065559	-7.68	0.000	-.0632097	-.0375111
_trend	-.1026704	.0434317	-2.36	0.018	-.1877949	-.0175459
_cons	-11.11678	.	.	.	.	.

LR test of identifying restrictions: chi2(4) = 1.465 Prob > chi2 = 0.833

Part II: The tests for causality

Estimating the lag length

lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
0	-13964.1				5.6e+10	24.747	24.747	24.747
1	-11521.4	4885.3	4	0.000	2.8e+08	19.4341	19.4421	19.4551
2	-11483.3	76.166	4	0.000	2.6e+08	19.3598	19.3758	19.4018
3	-11408.2	150.31	4	0.000	2.2e+08	19.2048	19.2288	19.2678
4	-11358.3	99.827	4	0.000	2.0e+08	19.1047	19.1368	19.1888
5	-11320.4	75.646	4	0.000	1.8e+08	19.0311	19.0712	19.1361
6	-11312.3	16.321	4	0.003	1.8e+08	19.022	19.0701	19.1481
7	-11295.7	33.14	4	0.000	1.8e+08	18.9946	19.0507	19.1417
8	-11282	27.436	4	0.000	1.7e+08	18.9734	19.0376	19.1415
9	-11277.8	8.286	4	0.082	1.7e+08	18.9731	19.0453	19.1622
10	-11225.8	104.11	4	0.000	1.6e+08	18.8684	18.9486	19.0785*
11	-11214.7	22.251	4	0.000	1.5e+08	18.8529	18.9411	19.084
12	-11203.9	21.467	4	0.000	1.5e+08	18.8382	18.9345	19.0904
13	-11190.5	26.828	4	0.000	1.5e+08	18.8177	18.922	19.0909
14	-11181.4	18.182	4	0.001	1.5e+08	18.8066	18.9189	19.1008
15	-11170.9	21.075	4	0.000	1.5e+08	18.7924	18.9127*	19.1076
16	-11165.7	10.396	4	0.034	1.5e+08	18.7898	18.9181	19.126
17	-11162.7	6.0532	4	0.195	1.5e+08	18.7919	18.9282	19.1491
18	-11150.6	24.109	4	0.000	1.4e+08	18.7744	18.9187	19.1526
19	-11145.6	10.027*	4	0.040	1.4e+08*	18.7721*	18.9245	19.1714
20	-11142.8	5.6517	4	0.227	1.4e+08	18.7747	18.9351	19.1949

OLS of lagged *open interest* and *standard deviation* on *standard deviation*

```
. regress std oia_1 oia_2 oia_3 oia_4 oia_5 oia_6 oia_7 oia_8 oia_9 oia_10 oia_11 oia_
> 12 oia_13 oia_14 oia_15 oia_16 std_1 std_2 std_3 std_4 std_5 std_6 std_7 std_8 std_9
>  std_10 std_11 std_12 std_13 std_14 std_15 std_16, vce(bootstrap, reps(1000))
(running regress on estimation sample)

Bootstrap replications (1000)
----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
..................................................    50
..................................................   100
```

output omitted

```
..................................................   950
..................................................  1000

Linear regression                               Number of obs      =       922
                                                Replications       =      1000
                                                Wald chi2(32)      =    796.30
                                                Prob > chi2        =    0.0000
                                                R-squared          =    0.6276
                                                Adj R-squared      =    0.6142
                                                Root MSE           =    0.4837
```

std	Observed Coef.	Bootstrap Std. Err.	z	P>|z|	Normal-based [95% Conf.	Interval]
oia_1	1.15e-06	7.54e-07	1.53	0.125	-3.22e-07	2.63e-06
oia_2	-2.39e-06	1.20e-06	-1.99	0.047	-4.75e-06	-3.53e-08
oia_3	2.02e-06	1.22e-06	1.66	0.098	-3.72e-07	4.41e-06
oia_4	-4.52e-07	1.13e-06	-0.40	0.688	-2.66e-06	1.75e-06
oia_5	1.15e-06	1.10e-06	1.05	0.294	-9.98e-07	3.30e-06
oia_6	-1.66e-06	1.05e-06	-1.57	0.116	-3.72e-06	4.09e-07
oia_7	-5.66e-07	1.03e-06	-0.55	0.582	-2.58e-06	1.45e-06
oia_8	1.47e-06	1.22e-06	1.21	0.228	-9.23e-07	3.87e-06
oia_9	-1.27e-06	1.30e-06	-0.97	0.331	-3.82e-06	1.29e-06
oia_10	1.40e-06	1.32e-06	1.06	0.290	-1.19e-06	4.00e-06
oia_11	-1.50e-06	1.10e-06	-1.36	0.175	-3.66e-06	6.67e-07
oia_12	-7.15e-08	1.07e-06	-0.07	0.947	-2.17e-06	2.03e-06
oia_13	2.39e-07	1.09e-06	0.22	0.827	-1.90e-06	2.38e-06
oia_14	-1.46e-06	1.26e-06	-1.16	0.246	-3.92e-06	1.01e-06
oia_15	3.18e-06	1.59e-06	2.00	0.046	6.22e-08	6.30e-06
oia_16	-9.69e-07	9.37e-07	-1.03	0.302	-2.81e-06	8.69e-07
std_1	.1242312	.0654109	1.90	0.058	-.0039718	.2524342
std_2	.0941667	.0414022	2.27	0.023	.0130199	.1753136
std_3	.1455191	.0597518	2.44	0.015	.0284079	.2626304
std_4	.0764381	.0566008	1.35	0.177	-.0344974	.1873736
std_5	-.0421359	.0717162	-0.59	0.557	-.182697	.0984252
std_6	.0039488	.0465074	0.08	0.932	-.0872041	.0951017
std_7	.0167936	.0457189	0.37	0.713	-.0728138	.1064009
std_8	.0422598	.0502612	0.84	0.400	-.0562504	.14077
std_9	.0315801	.0512057	0.62	0.537	-.0687813	.1319415
std_10	.0671344	.0631634	1.06	0.288	-.0566636	.1909325
std_11	.0761234	.0634359	1.20	0.230	-.0482086	.2004554
std_12	.050366	.048165	1.05	0.296	-.0440357	.1447677
std_13	.0472005	.0591717	0.80	0.425	-.0687739	.1631749
std_14	.0211822	.0702429	0.30	0.763	-.1164914	.1588559
std_15	.0499591	.0828277	0.60	0.546	-.1123802	.2122984
std_16	.0053325	.0534463	0.10	0.921	-.0994204	.1100853
_cons	-.0427067	.0326127	-1.31	0.190	-.1066264	.021213

Test for joint significance of the lagged *open interest*

```
. test(oia_1 oia_2 oia_3 oia_4 oia_5 oia_6 oia_7 oia_8 oia_9 oia_10 oia_11 oia_12 oia_
> 13 oia_14 oia_15)

 ( 1)  oia_1 = 0
 ( 2)  oia_2 = 0
 ( 3)  oia_3 = 0
 ( 4)  oia_4 = 0
 ( 5)  oia_5 = 0
 ( 6)  oia_6 = 0
 ( 7)  oia_7 = 0
 ( 8)  oia_8 = 0
 ( 9)  oia_9 = 0
 (10)  oia_10 = 0
 (11)  oia_11 = 0
 (12)  oia_12 = 0
 (13)  oia_13 = 0
 (14)  oia_14 = 0
 (15)  oia_15 = 0

           chi2( 15) =   12.21
         Prob > chi2 =    0.6632
```

OLS of lagged *open interest* and *standard deviation* on *open interest*

```
. regress oia oia_1 oia_2 oia_3 oia_4 oia_5 oia_6 oia_7 oia_8 oia_9 oia_10 oia_11 oia_
> 12 oia_13 oia_14 oia_15 oia_16 std_1 std_2 std_3 std_4 std_5 std_6 std_7 std_8 std_9
>  std_10 std_11 std_12 std_13 std_14 std_15 std_16, vce(bootstrap, reps(1000))
(running regress on estimation sample)

Bootstrap replications (1000)
----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
..................................................    50                              ..
..................................................   100
```

output omitted

```
..................................................   950
..................................................  1000

Linear regression                               Number of obs      =       922
                                                Replications       =      1000
                                                Wald chi2(32)      = 195907.70
                                                Prob > chi2        =    0.0000
                                                R-squared          =    0.9965
                                                Adj R-squared      =    0.9964
                                                Root MSE           =  2.41e+04
```

oia	Observed Coef.	Bootstrap Std. Err.	z	P>|z|	Normal-based [95% Conf.	Interval]
oia_1	.909369	.0391301	23.24	0.000	.8326755	.9860625
oia_2	.0058247	.0502101	0.12	0.908	-.0925852	.1042346
oia_3	-.0622794	.0580263	-1.07	0.283	-.1760088	.05145
oia_4	.2999718	.05538	5.42	0.000	.1914289	.4085146
oia_5	-.1883786	.0560898	-3.36	0.001	-.2983126	-.0784447
oia_6	-.0385355	.0543437	-0.71	0.478	-.1450472	.0679762
oia_7	.0178957	.0508631	0.35	0.725	-.0817942	.1175857
oia_8	.0411466	.0524014	0.79	0.432	-.0615583	.1438516
oia_9	.2482009	.0542844	4.57	0.000	.1418055	.3545963
oia_10	-.3385292	.054594	-6.20	0.000	-.4455314	-.231527
oia_11	-.0124041	.0563846	-0.22	0.826	-.1229159	.0981078
oia_12	-.0103775	.0544073	-0.19	0.849	-.1170139	.0962588
oia_13	.2570267	.0585174	4.39	0.000	.1423348	.3717186
oia_14	-.1782285	.0616442	-2.89	0.004	-.2990489	-.057408
oia_15	.0110072	.0582047	0.19	0.850	-.1030718	.1250863
oia_16	.0454458	.0445236	1.02	0.307	-.041819	.1327105
std_1	-5686.615	2444.635	-2.33	0.020	-10478.01	-895.2172
std_2	-628.8697	2196.281	-0.29	0.775	-4933.501	3675.762
std_3	-1998.655	1824.769	-1.10	0.273	-5575.137	1577.827
std_4	-3993.434	2056.764	-1.94	0.052	-8024.618	37.75044
std_5	218.861	2649.227	0.08	0.934	-4973.528	5411.25
std_6	1791.474	2143.908	0.84	0.403	-2410.509	5993.457
std_7	-1247.765	2038.616	-0.61	0.540	-5243.378	2747.849
std_8	712.3746	2016.416	0.35	0.724	-3239.728	4664.477
std_9	-957.9416	2089.861	-0.46	0.647	-5053.994	3138.111
std_10	2737.936	1916.194	1.43	0.153	-1017.734	6493.606
std_11	256.3713	2078.747	0.12	0.902	-3817.899	4330.642
std_12	-3099.929	2192.24	-1.41	0.157	-7396.641	1196.784
std_13	1861.137	2207.907	0.84	0.399	-2466.28	6188.554
std_14	-740.451	2336.858	-0.32	0.751	-5320.609	3839.707
std_15	1299.954	2516.912	0.52	0.606	-3633.102	6233.01
std_16	4099.617	1953.052	2.10	0.036	271.7064	7927.529
_cons	874.3735	1695.121	0.52	0.606	-2448.003	4196.75

Test for joint significance of the lagged *standard deviation*

```
. test(std_1 std_2 std_3 std_4 std_5 std_6 std_7 std_8 std_9 std_10 std_11 std_12 std_
> 13 std_14 std_15)

 ( 1)  std_1 = 0
 ( 2)  std_2 = 0
 ( 3)  std_3 = 0
 ( 4)  std_4 = 0
 ( 5)  std_5 = 0
 ( 6)  std_6 = 0
 ( 7)  std_7 = 0
 ( 8)  std_8 = 0
 ( 9)  std_9 = 0
 (10)  std_10 = 0
 (11)  std_11 = 0
 (12)  std_12 = 0
 (13)  std_13 = 0
 (14)  std_14 = 0
 (15)  std_15 = 0

           chi2( 15) =   25.76
         Prob > chi2 =    0.0406
```

UNIVERSITAS CAROLINA PRAGENSIS
založena 1348

Univerzita Karlova v Praze
Fakulta sociálních věd
Institut ekonomických studií

Opletalova 26
110 00 Praha 1
TEL: 222 112 330,305
TEL/FAX: 222 112 304
E-mail: ies@mbox.fsv.cuni.cz
http://ies.fsv.cuni.cz

Akademický rok 2010/2011

TEZE BAKALÁŘSKÉ PRÁCE

Student:	Pavel Šlechta
Obor:	Ekonomie
Konzultant:	PhDr. Petr Teplý, Ph.D.

Garant studijního programu Vám dle zákona č. 111/1998 Sb. o vysokých školách a Studijního a zkušebního řádu UK v Praze určuje následující bakalářskou práci

Expected name of the bachelor's thesis:

Speculative versus Fair Price of Crude Oil

Preliminary scope of work:

The main focus of the bachelor thesis is the crude oil market and the influence of the speculation on the crude oil prices. Employing the vector error correction estimation, it will be examined whether an increase in speculative trading can cause an increase in the oil prices. The main fundamental factors that influence the oil prices will be considered, for example the demand and the supply, where it will be distinguished between the OPEC and non-OPEC supply, and the open interest on the market will be used as a proxy for the risk or speculative factor.

References:

Fattouh, B. (2007): The Drivers of Oil Prices: The Usefulness and Limitations of Non-Structural Model, the Demand-Supply Framework and Informal Approaches, Oxford Institute for Energy Studies WPM 32.
Hamilton, J. D. (2009b): Understanding Crude Oil Prices, in: The Energy Journal 30(2), p. 179-206.
Johansen, S. (1995): Likelihood-Based Inference in Cointegrated Vector Autoregressive Models, New York: Oxford University Press.
Kaufmann, K. K. (2011): The role of market fundamentals and speculation in recent price changes for crude oil, in: Energy Policy 39(1), p. 105-115.
Kaufmann, R. K./ Dees, S./ Gasteuil, A./ Mann, M. (2008): Oil prices: The role of refinery utilization, futures markets and non-linearities, in: Energy Economics 30(5), p. 2609-2622.
Kočenda, E./ Černý, A. (2007): Elements of Time Series Econometrics: An Applied Approach, Prague: Karolinum Press.
Lütkepohl, H./ Reimers, H.-E. (1992): Impulse response analysis of cointegrated systems, in: Journal of Economic Dynamics and Control 16(1), p. 53-78.

Date of registering:	Květen 2011
Date of submission:	Květen 2012

Podpisy konzultanta a studenta:

V Praze dne 3. 6. 2011