Fondazione Eni Enrico Mattei

NOTA DI LAVORO

52.2015

Big Fish: Oil Markets and Speculation

Alessandro Cologni, Edison Trading S.p.A.
Elisa Scarpa, Edison Trading S.p.A., University of Milan, Polytechnic University of Milan
Francesco Giuseppe Sitzia, Edison Trading S.p.A.

Energy: Resources and Markets
Series Editor: Matteo Manera

Big Fish: Oil Markets and Speculation

By Alessandro Cologni, Edison Trading S.p.A.
Elisa Scarpa, Edison Trading S.p.A., University of Milan, Polytechnic University of Milan
Francesco Giuseppe Sitzia, Edison Trading S.p.A.

Summary
The role of speculators in the oil markets has been vastly investigated during the last few years. Several authors focused on the definition of speculation while others examined the relationship between oil prices and the behavior of trading actors. In this paper, we formulate a new theory able to describe "hedging needs" as well as the role of speculators in the crude oil market. According to our model, the different strategies of producers and consumers aimed at defending themselves against abrupt oil price changes can be satisfied only if speculators play a very active role. Due to the rapid growth in shale oil production, the importance of speculation in ensuring an equilibrium in the U.S. crude oil market has consequently grown noticeably. We estimate an econometric conditional Error Correction Model (ECM) applying Pesaran's bound tests, over the sample February 2000 November 2014, using WTI and CFTC data. Our theory is well supported by econometric evidence. In other words, our model is suitable to demonstrate how commercial operators act on the market. In addition, the increasing importance of future contracts (also known as financialisation of crude oil market) helps in reaching a level of prices close to the equilibrium one. Finally, we are able to find evidence of a positive impact of the action of speculators on the efficiency of oil markets as they help stabilizing prices.

Keywords: Oil Price Dynamics, Speculation and Fundamentals, Conditional Error, Correction Models, Pesaran Bounds Testing Approach
JEL Classification: Q4, Q41, D84, G15

The authors would like to thank Norberto Broggini, Bahattin Buyuksahin, Leo P. Drollas, Lutz Killian, Matteo Manera and Enrico Piccin, for insightful discussions, as well as participants at the International Workshop on "Oil and Commodity Price Dynamics" organized by Fondazione Eni Enrico Mattei (5-6 June 2014) for their helpful comments.

Address for correspondence
Elisa Scarpa
Edison Trading S.p.A.
Foro Buonaparte 31
20122 Milan
Italy
E-mail: elisa.scarpa@edison.it

Big Fish: Oil Markets and Speculation

Alessandro Cologni[a], Elisa Scarpa[a,b,c,*], Francesco Giuseppe Sitzia[a]

[a] *Edison Trading S.p.A.*
[b] *University of Milan*
[c] *Polytechnic University of Milan*

Abstract

The role of speculators in the oil markets has been vastly investigated during the last few years. Several authors focused on the definition of speculation while others examined the relationship between oil prices and the behavior of trading actors.
In this paper, we formulate a new theory able to describe "hedging needs" as well as the role of speculators in the crude oil market. According to our model, the different strategies of producers and consumers aimed at defending themselves against abrupt oil price changes can be satisfied only if speculators play a very active role. Due to the rapid growth in shale oil production, the importance of speculation in ensuring an equilibrium in the U.S. crude oil market has consequently grown noticeably.
We estimate an econometric conditional Error Correction Model (ECM) applying Pesaran's bound tests, over the sample February 2000 November 2014, using WTI and CFTC data. Our theory is well supported by econometric evidence. In other words, our model is suitable to demonstrate how commercial operators act on the market. In addition, the increasing importance of future contracts (also known as financialisation of crude oil market) helps in reaching a level of prices close to the equilibrium one. Finally, we are able to find evidence of a positive impact of the action of speculators on the efficiency of oil markets as they help stabilizing prices.

---

*Corresponding author
Email addresses: alessandro.cologni@edison.it (Alessandro Cologni), elisa.scarpa@edison.it (Elisa Scarpa), francesco.sitzia@edison.it (Francesco Giuseppe Sitzia)

Keywords: Oil price dynamics, speculation and fundamentals, conditional Error Correction models, Pesaran bounds testing approach.
JEL Classification Numbers: Q4, Q41, D84, G15.
Acknowledgments: The authors would like to thank Norberto Broggini, Bahattin Büyükşahin, Leo P. Drollas, Lutz Killian, Matteo Manera and Enrico Piccin, for insightful discussions, as well as participants at the International Workshop on "Oil and Commodity Price Dynamics" organized by Fondazione Eni Enrico Mattei (5-6 June 2014) for their helpful comments.

1. Introduction

Speculation is a very complex issue. In a broad sense, this activity involves buying and selling assets with high risk, in expectation of high returns. Speculation in financial markets finds very often a negative meaning, being responsible for driving up the price of financial assets. Especially in commodity markets speculators are pointed as guilty actors moving prices and favoring high volatility.

From an economic perspective speculators should be instead important agents for price discovery and for the creation of reliable markets as they are an important source of liquidity.

We do believe that financial markets have an impact on oil prices but this effect is often positive helping in reaching a long-run equilibrium price.

Measuring speculation is perhaps the hardest task and no consensus has been reached on the level of speculation which can be defined as extreme or excessive.

In this paper we have decided not to follow this stream but to develop a model able to describe and explain the whole complexity of oil market.

In general, we believe that the market is ruled by two categories of agents: commercial operators and speculators. On the one hand, the first ones are represented by oil companies, producing oil, or consumers, using it. This class of agents needs to hedge their positions ("hedging needs"). On the other hand, speculators are mainly driven by the objective of making money (what we call "money needs").

In depicting this theory we realized several evidences but two of them are the cornerstone of our paper: the financialisation of oil markets has been associated with the strong increase in U.S. crude oil production and with the evolution in the hedging needs of commercial operators; speculators act as price stabilizers, helping in reaching a level of price close to a long-run or equilibrium price.

Using a conditional Error Correction approach and bounds tests introduced by Pesaran et al. [29] and Pesaran and Smith [28] we are able to demonstrate the validity of our theory over the period February 2000 - November 2014.

Firstly hedging needs are described by fundamental variables such as supply-demand balance and storage dynamics. Secondly we demonstrate that oil price movements cannot be explained by operators positions but that the greater is the amount of speculation in the market the quicker the price approaches a long run equilibrium level.
Our paper is novel as it presents an econometric-based model that highlights the role played by speculators on international commodity markets.
The rest of this paper is organized as follows. Section 2 briefly reviews the existing literature on the increasing financialization of commodity markets. We also illustrate some of the major research works aimed at assessing the econometric relationship between speculation and oil future prices. A quick overview of the main speculative indexes introduced to measure the degree of speculation on financial markets is also presented. Section 3 presents the data. In particular, we describe how measures for speculative and hedging activities are calculated starting from open interest figures. In Section 4 the assumptions underlying our theoretical model are presented together with the main econometric results. We focus on the main relationships of interest by considering the two main building blocks of the model, i.e. description of (a) hedging needs of both producers and consumers; (b) crude oil price dynamics. Section 5 concludes.

2. Literature Review and Main Speculative Indexes

Numerous studies on international energy markets investigate the role of speculation and its effects on oil prices. Masters, in his 2008 testimony before the US Congress, argued that speculative positions taken by institutional investors had resulted in increases in futures and spot commodity prices (Masters [25]).
In this Section, following the main points outlined by Masters, recent published papers are divided in two main strands of the literature: (1) increased participation of institutional investors in commodities futures markets (*financialization* of commodity markets); (2) effects on spot and -futures prices of activity by speculators. Finally, a brief analysis of the major in-

dexes aimed at measuring the degree of speculative activity is also presented. Because of their relevance for our research we mainly concentrate on studies on crude oil markets.

2.1. On the financialization *of crude oil markets*

After 2003 financial investors entered the oil futures market in large numbers. This rapid inflow of investors/operators from outside the commodity industry into oil and other commodities was associated with a rapid growth of index funds (Fattouh [11]). Since index investors often buy simultaneously a portfolio made of different commodities, we should "expect to see a high degree of contemporaneous correlation in futures price movements through time" (Stoll and Whaley [35]).[1]

The effects of a broader participation of financial operators on the interdependence among the futures prices of different asset classes (e.g. oil prices, other commodity prices and stock prices) are studied empirically by several researchers (among others, Büyükşahin et al. [6], Büyükşahin et al. [7], Büyükşahin and Harris [4], Tang and Xiong [36], Büyükşahin and Robe [5] and Silvennoinen and Thorp [33]).[2]

According to Büyükşahin et al. [6] market participation by hedge funds and commodity index traders had surged since 2002. A higher presence of both commercial and non-commercial operators in commodity markets helps to explain how the statistical relationship among futures returns at different time maturities has increased since then.

Results obtained by employing dynamic conditional correlation and recursive cointegration techniques (Büyükşahin et al. [7]) show that the relation between prices (and returns) of investable commodity and US equity in-

[1]Nevertheless, early studies on the tendency by commodity prices to move together date back to the late 1980s (see, e.g., Pindyck and Rotemberg [30]). However, they advise that co-movements in commodity prices between 1960 and 1985 can not fully explained by common macroeconomic shocks.

[2]For a recent survey on the increased price co-movements between different commodities and the financialization of commodities markets see Fattouh et al. [12].

dexes had not changed noticeably in the preceding fifteen years. They also argue that there had not been a remarkable increase in correlations between the equity and commodity return series during periods characterized by extreme events[3].

The possibility that the interdependence among the prices of different asset classes had recently increased is also examined by Tang and Xiong [36]. They investigate the correlation of commodity prices with both stocks and the US dollar exchange rate returns as well the links between non-energy commodities and oil through formal regression analyzes. Their results seem to confirm that co-movements among different asset classes had significantly increased during recent years.

Drollas [9] notices that speculation activity on NYMEX (as measured by the CGES index of speculative intensity) [4] was relatively high from 2Q2008 till the summer of 2009. Conversely, activity by speculators declined significantly starting from the beginning of 2011. For the whole 2011, it stayed at rather low levels.

By using standard time-series econometric techniques[5], Büyükşahin and Harris [4] test the hypothesis of the increasing correlation among the prices of different asset classes (and, in particular, commodity futures and US stock index returns). Their findings provide additional empirical evidence that co-movements between stocks and fuel prices were often related to the entry of hedge funds taking positions in both equity and energy markets.

Other studies (e.g. Büyükşahin and Robe [5] and Silvennoinen and Thorp [33]) find that the conditional correlation between commodity futures returns and US stock index returns had increased especially in periods char-

[3]That is, periods in which the returns on the equity or stock indexes were significantly above the mean observed during a given period.

[4]This index measures the volume of pure speculative activity that is not offsetting opposite positions taken by hedgers.

[5]The auto-regressive distributed lag - ARDL - cointegration approach proposed by Pesaran and Shin [27].

acterized by higher volatility.
Finally, rising financialization of commodity markets is usually associated with a breakdown of the statistical relationship between oil futures prices and inventories (Alquist and Kilian [1], Fattouh et al. [12] and Kilian and Murphy [22]). By reducing the price of risk implied in markets, greater financial market integration can reduce the cost of hedging and, consequently, determine increases in the levels of inventories. In this framework, higher oil inventories are responsible for higher spot prices. The possibility of further price increases in the future consequently shrinks.[6,7]
Fattouh et al. [12] suggest that "economic theory tells us that both spot prices and futures prices are determined simultaneously and respond to the same economic forces". They add that "[T]o the extent that global macroeconomic fundamentals have changed in recent years, for example, that fact could provide an alternative explanation of the observed comovement in spot and futures prices". So after 2003 "there is strong evidence that the spot and futures prices responded to the same economic fundamentals" (Kilian [20]). Finally, although there is some evidence of a stronger statistical relationship among the returns of several commodities, co-movements can also be found in markets in which "index funds do not operate and for which there are no futures exchanges".

2.2. On the effects of speculation on oil prices

Other authors examine how the large inflow of financial investors in oil futures markets had influenced recent movements of spot and futures crude oil prices. Several econometric analyzes are carried out with the pur-

[6] In other words, speculative activity may determine either purchases of oil futures contracts (given the existence and liquidity of oil futures markets, Gilbert [14]) or increases in oil inventories levels.

[7] Nevertheless, further research is required to give a more proper description of the relationship between oil futures and inventories. In fact, as Fattouh et al. [12] argue, "the absence or presence of speculative pressures in the oil market cannot be inferred from studying oil inventory data without a fully specified structural model".

pose of evaluating the relationship between the net long positions[8] held by non-commercial operators (*speculators*) and oil futures prices.

Nevertheless, research aimed at assessing the impact of speculation on commodity price changes has to address several econometric issues. A first important point to examine is related to the definition of "excessive speculation". Although it can be defined as "speculation that is not required for the oil market to function properly", providing a proper notion of "excessive speculation" from an operational point of view is particularly difficult. Consequently, many empirical works make no distinction between "socially desirable and undesirable speculation" (Kilian and Murphy [21]). Similarly, Fattouh et al. [12] argue that "the presence of speculators defined as noncommercial traders tells us nothing about whether speculation is excessive". They also add that, while commercial traders may take a position or decide not to hedge in the futures market despite having an exposure to the commodity, traders with a physical interest in the commodity (e.g. oil companies) can act as speculators on derivatives markets.

A second fundamental econometric issue researchers have to take into account is related to the endogeneity of net long positions held by both commercial and non-commercial traders. The causal relationship between positions held by operators and oil futures prices could, in fact, depend on the fact that traders' strategies often respond to the underlying fundamentals of the oil market. Hence, a fundamental question analysts and researchers should answer at this regard is the following: "are positions held by index funds exogenous? Or rather do they respond endogenously to other variables (such as economic incentives)?". Fattouh et al. [12] suggest that many empirical analyzes "raise more empirical questions than [they] answer".

Finally, as far as the econometric framework is concerned, the broader participation of financial investors and their effects on energy prices is often investigated by means of bivariate Granger causality tests. At this pur-

[8] Situation in which the number of long positions held by an operator in a given asset is higher than the number of short positions.

pose, either AutoRegressive Distributed Lag (ARDL) or Vector Autoregression (VAR) models are estimated by examining Commodity Futures Trading Commission (CFTC)'s Commitments of Traders (COT) data.

More recently, many researchers have underlined the importance to employ Structural Vector Autoregression (SVAR) models to identify the impact of speculation on crude oil price changes. In particular, SVAR techniques are designed to "disentangle demand and supply shocks in global oil markets" (Kilian and Murphy [22]). Hence, they can be used to test alternative explanations of the evolution of the real price of oil. More sophisticated econometric techniques (e.g. Generalized AutoRegressive Conditional Heteroskedasticity - GARCH - models) are employed by, among others, Du et al. [10], Cifarelli [8], Manera et al. [24].

Let us now turn our attention to the main results obtained by econometric research. A few empirical studies show that financial flows significantly affect crude oil price returns. For example, empirical results obtained by Singleton [34] display a statistically significant relationship between investor flows and futures prices. Results are also found to be robust to different specifications of the empirical framework. According to Singleton, informational frictions (i.e. heterogeneous beliefs about real economic activity) lead to speculative activity that may drift prices away from "fundamental" values.[9]

However, in general, econometric works (see, *inter alia*, Kilian and Murphy [21], Stoll and Whaley [35], Büyükşahin and Harris [4], Brunetti and Büyükşahin [3], Hamilton and Wu [17]) do reject the hypothesis that changes in positions held by financial traders cause movements in oil futures prices. Though they show strong evidence of speculation in 1979, 1986, 1990 and late 2002, Kilian and Murphy [21] conclude that speculative activity had not affected significantly real oil price movements during the 2003-mid 2008 period. Similar empirical findings are illustrated by, among others,

[9] A set of possible causes of the oil price shock of 2007-2008 is presented in Hamilton [15] and Hamilton [16]. They conclude that speculative activity partial contributed to the price increase observed during the summer of 2008.

Sanders et al. [32], Stoll and Whaley [35] and Büyükşahin and Harris [4].
In their analysis of the causality relationship between changes in net long positions (PNL) held by non-commercial firms and crude oil price changes, Sanders et al. [32] find that there is no consistent evidence that traders' percent net long positions Granger cause energy futures returns. Conversely, positive market returns significantly affect changes in the PNL positions of both noncommercial traders and commercial operators.
In Stoll and Whaley [35] bivariate Granger causality tests provide very little evidence that traders' positions help to forecast returns in crude oil futures markets. Büyükşahin and Harris [4] don't find empirical evidence that positions of hedge funds or other non-commercial investors help to predict changes in the futures price. On the contrary, they suggest that futures price changes often precede changes in the positions held by different operators.
By employing data representing individual positions taken by traders, Brunetti and Büyükşahin [3] suggest that movements in financial flows can not be used to forecast either the price level and volatility of oil futures prices.
In a similar way, in Hamilton and Wu [17] data since 2006 are studied in order to find a statistical significant relationship between market participation by index-fund investors and expected returns on futures contracts. However, they find little support to the hypothesis that index-fund investing in commodity markets had a sensible impact on commodity futures prices.
Up to now we have focused on the research that has studied the factors affecting crude oil futures prices. Other studies (for instance, Du et al. [10], Cifarelli [8] and Manera et al. [24]) address the determinants of oil price volatility.
In Du et al. [10] the role of several factors influencing the volatility of crude oil prices (such as speculation, *scalping*[10], and petroleum inventories) is

[10] That is, activities based upon opening and closing contract within a very short period of time in order to obtain small profits (particularly, from a temporary arbitrage possibility).

assessed through Bayesian Markov chain Monte Carlo methods. In this works, speculation is found to be an important factor in explaining oil price variation.

Cifarelli [8] studies the role of rational hedgers and informed oil traders in international crude oil markets. To test empirically the effects of changes in positions by these traders on price changes, non-linear smooth transition regime shift CCC (Constant Conditional Correlation)-GARCH techniques are employed. The author is able to define speculative futures price changes as those oil price movements due to "destabilizing hedgers' reactions to movements in the variability of the return of their covered cash position".

By using multivariate GARCH models, Manera et al. [24] document that financial speculation, measured by the Working's T index (Working [37]), does not affect significantly commodity returns. They also find that evidence of a negative correlation between agriculture and energy commodities. In other words, "high (low) volatilities in the agricultural markets correspond to low (high) volatility in the energy market".

Other empirical works show that there is no evidence that oil futures prices significantly improve the out-of-sample accuracy of forecasts of the spot price of oil (Alquist et al. [2]).

To conclude, following Alquist and Kilian [1] and Pirrong [31], we may suggest that, in order to understand real oil price changes, researchers should not focus solely on the role of speculation. In fact, in their opinion, prices also depend on changes in the equilibrium between demand and supply. In fact, after 2003, supply and demand forces have strongly affected crude oil prices and financial participation (Fattouh et al. [12] and Kilian and Murphy [21]). Similarly, since speculation is acknowledged to be "simply one component of the demand of oil", changes in speculative demand are related to changes in expectations due to future shifts in economic fundamentals.

In addition, the absence of evidence of an impact of speculation on energy futures prices does not mean that "the financialization of oil futures markets does not matter" (Fattouh et al. [12]). In other words, higher participation rates by institutional investors in crude oil markets not only help to explain the empirical evidence of structural instabilities in the term structure (affecting consequently the risk premium of oil futures prices). In addition (and in this paper we would like to stress this point) so-called speculators play an increasing important role in rapidly changing crude oil markets.

2.3. Main speculative indexes

Finally, an important issue researchers have to deal with concerns the measure to employ to quantify speculation. Table 1 report the expressions for some of the major indexes that empirical works have employed to describe the role of speculation on oil and other commodities markets.
Several authors (see, for instance, Büyükşahin and Harris [4]) employ in their analyses the Working's T index (Working [37]). This index measures the intensity of speculation relative to long (or short) hedging needs. In other words, it provides a measure of the adequacy of speculative positions to "absorb" (short and long) positions held by commercial traders (Peck [26]).
According to Working, the role of speculators has not to be interpreted in a negative way.[11] In fact, in his opinion, "what may be technically an 'excess' of speculation" could be "economically necessary for a well-functioning market". Accordingly, Kilian [20] suggests that "a high Working's T-index number does not necessarily indicate excessive speculation". Indeed, values for the index have to be compared with historical values for other (comparable) commodity markets.
Other measures often employed to assess the causality relationships between futures returns and positions held by traders are: (1) noncommercials' percent net long positions (PNL, difference between long and short positions divided by their sum); (2) the percent of total open interest held by each CFTC trader classification; (3) the Lakonishok, Shleifer and Vishny H(erding)-index (Lakonishok et al. [23], LSV). In particular, the index pre-

[11] As argued in the present research, this is also our viewpoint.

sented in LSV represents a measure of herding for a given stock in a given quarter. It is obtained by considering the number of money managers who increase/decrease their holdings in the stock in the quarter (net buyers/sellers) together with the expected proportion of money managers buying in that quarter relative to the number active. Measures of speculative and hedging activities calculated from volume and open interest data have often been employed to describe the role of speculation on affecting market prices. For instance, Garcia et al. [13] use, as proxies of the relative importance of the speculative behavior in a given contract, measures obtained by dividing the total volume of contracts traded in a period by either the size or the absolute change of open positions at the end of the period. Another measure they use is represented by the ratio between the change in open interest and overall volumes.

3. Data

In this Section we present the data used for our analysis. We focus on the US market selecting a sample over the period February 2000 - November 2014, composed by series of 771 weekly observations. This long span of data allows us to examine dynamics of oil markets in a period characterized by important shocks in markets fundamentals and prices.
We use both market and fundamental data: WTI prices from NYMEX, open interest data from the COT report published by the CFTC, supply-demand data from EIA and Saudi Arabia government balance data from the IMF WEO database, as described in more details in the following subsections.

3.1. *Open interest as a measure for speculation and hedging*

We want to detect the impact on oil market of different operators behaviors/strategies so we employ CFTC data.
These data have been broadly used in literature and provide a universally recognized distinction between “commercial”, “non-commercial” and “non reporting” operators. The first category comprises operators joining

the derivatives market in order to reduce the risk of their natural position, in other words having "hedging needs". On the contrary, in the other two categories one can find actors such as funds and traders that are prone to risk and join the market playing according to what we call "money needs". Therefore, we define the ensemble of the two latter categories as "speculators" opposed to "commercials" ("hedgers"). Even if a precise market strategy is not declared for "non reporting" it is reasonable to include these operators into the "non commercial" because their positions are strongly positively correlated with those of non-commercial operators.

According to CFTC data, the Open Interest (OI) owned by operators is expressed in number of contracts ($1Lot = 1000Bbl$) divided in long, short and spreading positions.[12] Moreover, we decided not to consider positions in options, which clearly represent a significant part of financial activity, because the amount of open interest reported by the CFTC is calculated based on the option delta in order to represent the operator exposition to price level. Hence, it can change independently from operators' will as a consequence of price movements. We calculate three indexes based on OI data: (1) the percent net length, (2) the "speculation degree" in the market and (3) a modified version of the T-Index. The first represents our hedging measure while the other two are our speculation variables.

The PNL is a measure of how much a group of agents is biased towards long (positive) or short positions (negative) (see Table 1 for the PNL formula). The PNL positions of commercials (PNL_t^{Com}) and speculators (PNL_t^{Spec}) are strongly negatively correlated. However, as the total amount of OI by category evolves independently, this correlation is not perfect. In the last years the PNL held by speculators has strongly increased rising concerns about financial players being responsible for higher oil prices.

The first speculation index we calculate is simply the ratio between com-

[12]A spreading position consists of two positions with opposite sign at different maturities.

mercial ($Commercial\_OI_t$) and speculators OI ($Speculators\_OI_t$):

$$Spec\_Degree_t = \frac{Speculators\_OI_t}{Commercial\_OI_t}$$

This measure of speculation (that since 2012 has stably exceeded the level of one) simply assesses the quantity of speculation without envisioning whether it is excessive or adequate.

The T-Index assumes that the net commercial OI is the needed hedging that has to be provided by speculators, all the remaining speculators OI being excessive. Since we estimated the unobserved hedging needs (see Section 4), we can also introduce a modified T-Index defined as follows:

$$Mod\,T - INDEX_t = \frac{(Speculators\_OI_t - |Hedging\_Needs_t \cdot Commercial\_OI_t|)}{Commercial\_OI_t}$$

where $Hedging\_Needs_t$ denotes our measure of hedging needs estimated (as it will be shown in the next Subsections) taking into account the US supply-demand balance. Figure 1 shows the Open Interest for speculators and commercials together with our measure for speculative activity, $Spec\_Degree_t$ and $Mod\,T - INDEX_t$.

3.2. US crude oil data

We use weekly US data from EIA petroleum status report adjusted to match OI data by averaging the two weekly figures published in the dates near the CFTC report one. On the supply side, we consider the *crude oil field production* while on the demand side we use the *petroleum products supplied time series*. From these two variables we derive a supply-demand balance index that is calculated similarly to the PNL_t^{Com} (see Table 1).

$$Bal_t^{US} = \frac{Demand_t - Supply_t}{Demand_t + Supply_t}$$

Since the boom of the fracking technology this index has strongly decreased following the impressive crude supply expansion that interested the US mid-continent (see Figure 2).

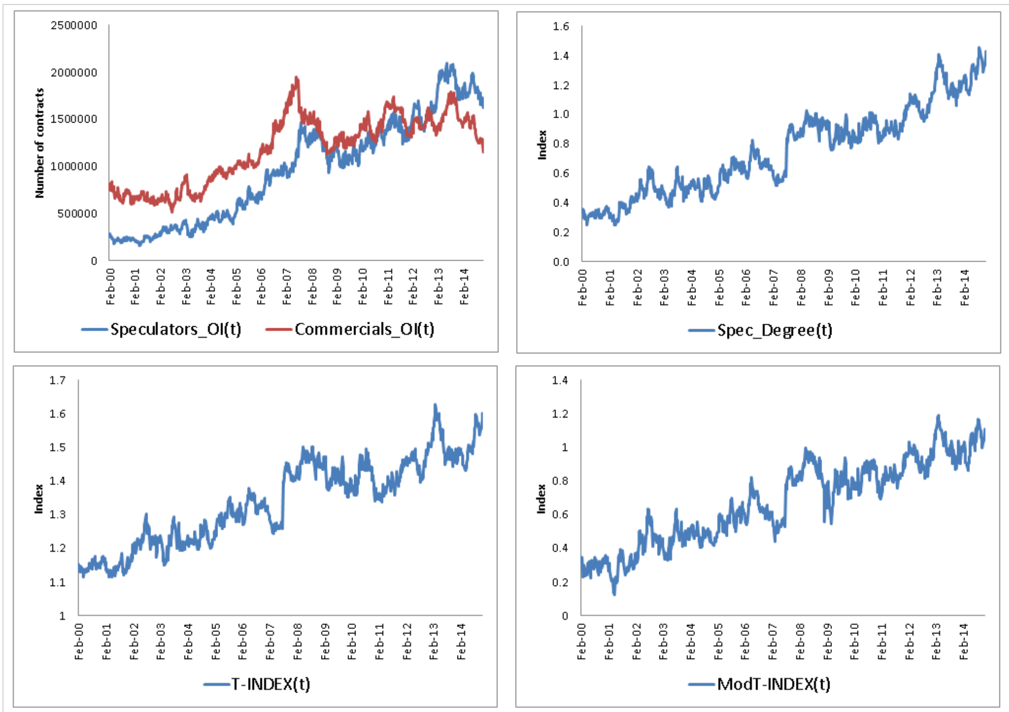

Figure 1: OI and speculation indexes

From NYMEX we collect the settlement prices for WTI futures contracts for the date in accordance to CFTC data. The weekly time series includes only the trading days for which the CFTC report is available, generally every Tuesday. The time-spread variable ($TS\_1\_4_t$) is expressed as a fraction of the first month contract price:

$$TS\_1\_4_t = \frac{WTI\_1_t - WTI\_4_t}{WTI\_1_t}$$

where WTI_1_t and WTI_4_t are the first and fourth month settlement prices, respectively. Figure 3 shows the relationship between the time-spread variable and WTI_1_t.

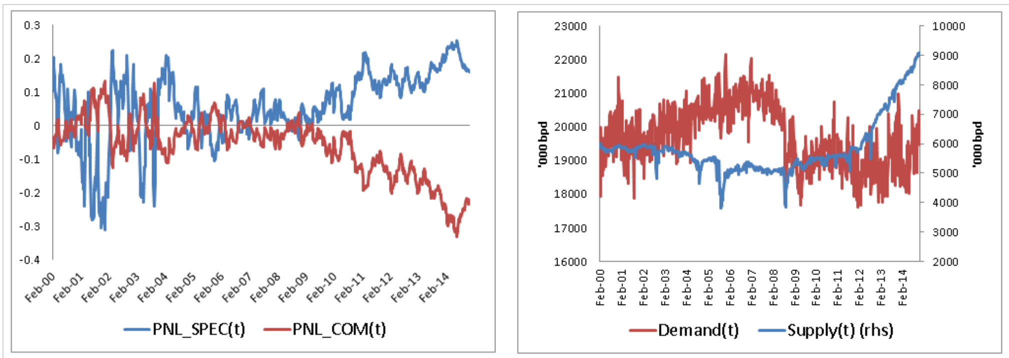

Figure 2: Net open interest and supply-demand data

3.3. An equilibrium price for oil

Our objective is to identify whether or not speculation and hedging affects the oil prices or vice-versa. In this subsection we want to define a long-run price for oil which allows us to detect whether or not the two classes of agents influence the WTI price. Does speculation cause high oil prices? Do high or low oil prices influence "hedging needs" or cause deviations from the hedging equilibrium?

We believe that the price should fluctuate around a long-run or equilibrium price. This equilibrium price is determined by many factors such as production costs, exploration costs, cost of capital, temporary supply and demand disruptions and so on and so forth. However, our aim is to define a long-run equilibrium price which, excluding shocks in fundamental variables, guarantee a stable evolution of the supply-demand balance. In this framework, the role of Saudi Arabia as one of the most important producing countries in the world should be taken into account. Saudi Arabia holds the greatest proven reserves (very recently surpassed by Venezuela) while ranks among the three biggest world oil producers (with Russia and the US). In addition, the Country has one of the lowest production cost in the world. The economy of Saudi Arabia is also almost entirely dependent on the oil market.

Since its foundation in 1960 OPEC has tried to protect the value of oil assets. Consequently, the Kingdom has often adjusted production in order to

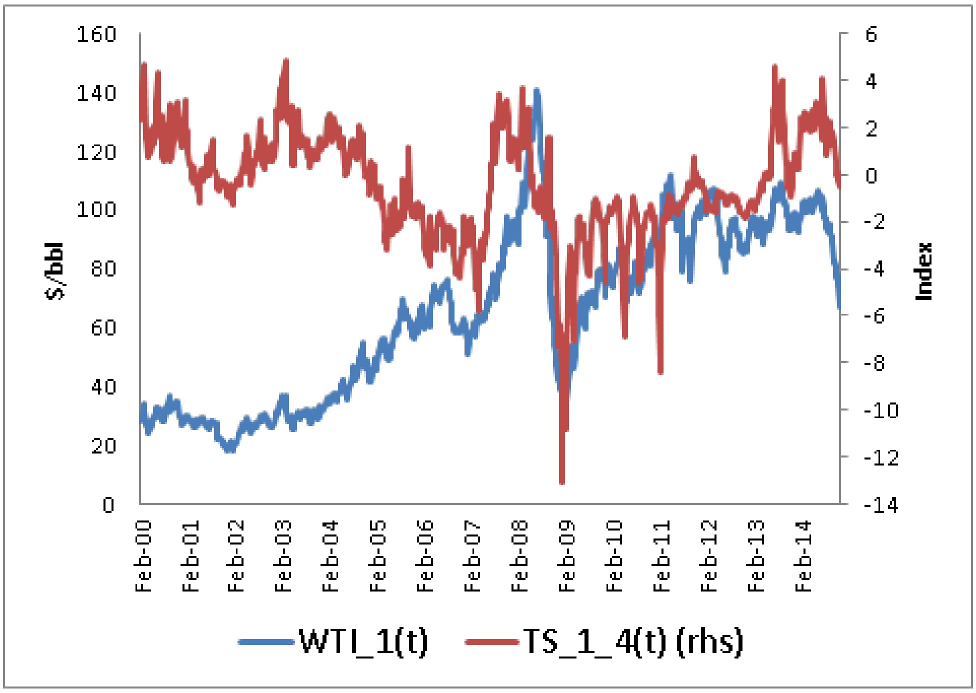

Figure 3: WTI price and time-spread series

stabilize oil prices. In other words, in the recent past, Saudi Arabia has often played the role of ”swing producer”. It decided unilaterally to reduce production in a context of weak demand and ample supplies or, conversely, to raise output in the presence of a relatively tight supply-demand balance. For these reasons, we employ Saudi Arabia’s breakeven price as our long-run equilibrium price of oil. This measure also highlights the fundamental role that one of the world’s biggest producer has on the crude oil market. More precisely, our measure of equilibrium price is constructed by simply assuming that Saudi government revenues (Gov_Rev_t) are represented as a function of Kingdom production ($Prod_t$) and oil price (P_t):

$$Gov\_Rev_t = f(x_t) = f(Prod_t \cdot P_t)$$

The oil price relevant for Saudi Arabia's finances (P_t) is calculated by considering a mix of the three main world oil benchmarks (WTI, Brent, Dubai). Taking a linear $f(x)$, the parameters of this linear regression are estimated by means of standard OLS econometric techniques. Saudi's breakeven price ($\widehat{P}_t$) is, hence, defined as a function of Saudi government expenditures (Gov_Exp_t):

$$\widehat{P}_t = \widehat{f}^{-1}(Gov\_Exp_t)$$

In other words, $\widehat{P}_t$ represents the yearly average price of oil that guarantees Saudi's balance of payments to be in equilibrium:

$$Gov\_Exp_t = Gov\_Rev_t$$

Finally, we determine the price variable to be used in the following Section (Dis_P_t) as the difference between the WTI first month price and the estimated break-even price divided by the latter:

$$Dis\_P_t = \frac{WTI\_1_t - \widehat{P}_t}{\widehat{P}_t}$$

4. Model. Empirical Framework and Results

In this Section we investigate the statistical properties of data used for the econometric analysis. Firstly, the order of integration is examined by applying Augmented Dickey Fuller (ADF) unit root tests. Then, we outline the reasons that led us to choose the Pesaran et al. [29]'s approach. Finally, the estimation is carried out in two steps. After examining the hedging needs of commercial operators, we introduce a model for the dynamics of oil prices.

4.1. Hedging model

We have already divided market operators in two broad categories, i.e. "commercials" having "hedging needs" and "speculators" - operators that, on the contrary, act in the name of "money needs".

Depending on market conditions, either the necessity to hedge short (producers dominate the market) or to hedge long (consumers dominate the market) could emerge. This mechanism is well represented by the percent net length, PNL_t^{Com}, that is the observed hedging position of commercial operators. It is straightforward to notice that without speculators this variable would always score 0.5.

In other words, commercial operators would have to balance themselves through price mechanisms by paying a premium or accepting a discount in order to find a counterpart. For this reason, prices could deviate from their "fair" or "equilibrium" value. In this framework, speculators are acknowledged to act in order to fulfill the needs of commercial operators. Accordingly, we can also add that, theoretically, the spot price approaches the equilibrium price (given a stable supply demand balance) when "hedging needs" are met via "money needs".

Therefore, which are the market conditions moving "hedging needs"? Our belief is that hedging needs should be mainly determined by fundamentals, e.g. supply-demand balance and storage dynamics.

In Section 3 we have presented the two time series Bal_t^{US} and $TS\_1\_4_t$. These two variables represent our measures for "fundamental" factors. A negative Bal_t^{US} value should be related to a negative PNL_t^{Com} since producers outnumber consumers. Under this assumption, the net position of commercials is "long commodity". Conversely (and consequently), their hedging needs become "negative" (i.e. "short") and vice-versa.

With regard to the storage dynamics, when $TS\_1\_4_t$ is negative (i.e. *contango*), there is incentive to stock the commodity. Thus, commercial operators should buy spot and sell forward increasing short positions in futures or, alternatively, reducing PNL_t^{Com}. On the contrary, a situation of *backwardation* (positive $TS\_1\_4_t$) prompts crude oil releases from storage. Hence, commercial operators should sell spot and buy forward in order to restore stocks. Both variables are assumed to have a positive impact on PNL_t^{Com}.

As outlined in more depth below, following Pesaran et al. [29], the set of variables Bal_t^{US}, PNL_t^{Com} and $TS\_1\_4_t$ are included in a conditional Error Correction Model (ECM) . This framework allows us to examine the relationships among the relevant variables without having preliminary knowledge on the order of integration of each time series. Figure 4 shows the behavior of the three variables PNL_t^{Com}, Bal_t^{US} and $TS\_1\_4_t$ together with Dis_P_t.

The determinants of crude oil prices will be examined independently in the final step of our empirical analysis once our measures of "hedging needs" (long run or equilibrium hedging requirements given the fundamental characteristics of crude oil markets) and "disequilibrium hedging" (difference between PNL_t^{Com} and these latter "hedging needs") are identified.

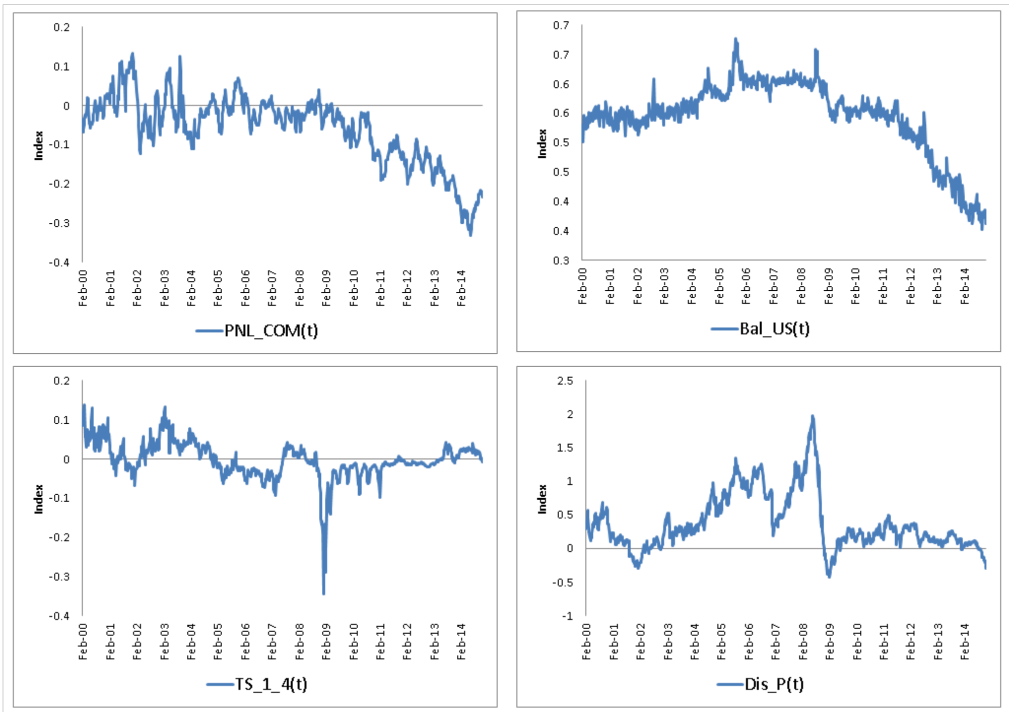

Figure 4: Main variables employed in the econometric analysis

As the estimation of a system of equations is influenced by the statistical properties of the variables, we start by investigating the order of integration of the four series: PNL_t^{Com}, Bal_t^{US}, $TS\_1\_4_t$ and Dis_P_t. Table 2 reports the results of Augmented Dickey Fuller (ADF) tests for variables expressed in

both levels or as first differences.
We can reject the null hypothesis of a unit root of $TS\_1\_4_t$ at a 5% confidence level, suggesting that this variable is $I(0)$. For PNL_t^{Com} the null hypothesis can be rejected only at a 10% test size. On the contrary, for both Bal_t^{US} and Dis_P_t, the null hypothesis can be rejected at any significance levels. Nevertheless, columns 4 and 5 of Table 2 show that the null hypothesis of stationarity can also be rejected for these latter three variables when they are expressed as first-differences. This would suggest that PNL_t^{Com}, Bal_t^{US} and Dis_P_t are variables integrated of order one (i.e., $I(1)$).
Our initial approach was to follow a standard econometric framework, that is the estimation of either Structural Vector Autoregressions (SVAR) or Vector Error Correction (VEC) models. However, as outlined above, our preliminary data analysis has showed the presence of both $I(0)$ and $I(1)$ variables. Accordingly, the standard Johansen procedure (Johansen [18], Johansen [19]) can not be employed since it would lead to unreliable results. The possibility to find long-run equilibrium relationships between stationary or a mixed set of $I(0)$ and $I(1)$ variables has been deeply investigated in the recent literature. In this particular case Pesaran et al. [29] provide a suitable bound test procedure that can be used to validate an econometric model.
Table 3 shows the results of the regression analyzes. According to the conditional ECM framework, the relevant equation to estimate is given by:

$$\Delta Y_t = \beta_0 + \theta_0 Y_{t-1} + \theta_1 X_{1,t-1} + \theta_2 X_{2,t-1} + \theta_3 X_{3,t-1} + \sum_{l=1}^{k_1} \beta_{1,l} \Delta Y_{1,t-l} + \\ + \sum_{l=1}^{k_1} \beta_{2,l} \Delta X_{2,t-l} + ... + \sum_{l=1}^{k_n} \beta_{n,l} \Delta X_{n,t-l} + \epsilon_t \quad (1)$$

where Y_t is our dependent variable, $X_{i,t}$ ($i = 1, 2, \dots, n$) is the set of n exogenous variables and ϵ_t is the vector of regression residuals. β and θ denote the set of parameters to estimate.
In a preliminary phase of our econometric procedure a *full* (that is, four-variable) conditional error ECM is estimated. In our stylized representation of the world crude oil market the relevant variables are defined as follows: (1) $Y_t : PNL_t^{Com}$; (2) $X_{1,t} : Bal_t^{US}$; (3) $X_{2,t} : TS\_1\_4_t$; (4) $X_{3,t} : Dis\_P_t$.

The number k_i of lags for each variable is selected by using the Schwarz information criterion. At this purpose, all possible combinations between 1 to 10 lags of the variables are considered. According to this selection procedure, we use in the conditional Error Correction model only one lag for each variable.
Nevertheless, the coefficient of the price variable in the long-run equation is not found to be statistical significant. The model is consequently reduced in order to include only the other three variables (*reduced* model).
Differently from the *full*, the *reduced* one does not show serial correlation in the residuals. Accordingly, we can employ the procedure proposed by Pesaran et al. [29]. This procedure consists in verifying whether the Wald-statistic (computed on the basis of the hypothesis that long-term coefficients are all equal to zero) falls outside the critical value bounds determined by Pesaran et al. [29]. If so, "a conclusive inference can be drawn without needing to know the integration/cointegration status of the underlying regressors".
For different cases (e.g. restrictions on either intercept or trend) and different numbers of variables $(n + 1)$, Pesaran et al. [29] provide the asymptotic (lower and upper) critical value bounds of the F-statistic, with a distribution which is totally non-standard. In each case, the lower bound values assume that all variables are $I(0)$, while on the contrary the upper bound is based upon the hypothesis that all of the variables are $I(1)$.
Table 4 shows the results of Pesaran et al. [29]'s bound tests. The results of Wald tests are reported together with the critical values that refer to the hypotheses of $I(0)$ or $I(1)$ variables.
Results show the computed F-statistic is greater than the upper bound presented in Pesaran et al. [29]. Thus, we can conclude that there is a long run equilibrium relationship among these three variables. This empirical evidence is also confirmed by the high value of the t-test of the coefficient of the lagged value of the dependent variable being lower than the $I(1)$ bound.
This preliminary step has allowed us to validate our model. The next phase of our empirical procedure consists in estimating the *full* structural VECM.

This model is defined as follows:

$$\Delta X_{i,t} = C_{i01}\Delta X_{1,t-1} + C_{i02}\Delta X_{2,t-1} + C_{i03}\Delta X_{3,t-1} + \\ + C_{i06}\left(X_{1,t-1} - LC_1 - LC_2 X_{2,t-1} - LC_3 X_{3,t-1}\right) + \varepsilon_{it} \tag{2}$$

where $i = 1, 2, 3$, X_{1t}, X_{2t} and X_{3t} denote, respectively, PNL_t^{Com}, Bal_t^{US} and $TS\_1\_4_t$ The parameters of this model are estimated through maximum likelihood. This estimation technique also allows to compute impulse response functions in a proper manner.
Table 5 illustrates the results of the system of equations (2). The estimates of the long run equilibrium coefficients (LC_2 and LC_3) are consistent with the theory described above. Let us define the "hedging needs" of US producers and consumers as:

$$Hedging\_Needs_t = \widehat{LC}_1 + \widehat{LC}_2 \cdot Bal_t^{US} + \widehat{LC}_3 \cdot TS\_1\_4_t \tag{3}$$

The corresponding "disequilibrium hedging" (Dis_Hed_t) is given by:

$$Dis\_Hed_t = PNL_t^{Com} - Hedging\_Needs_t$$

The two variables $Hedging\_Needs_t$ and Dis_Hed_t are shown in Figure 5

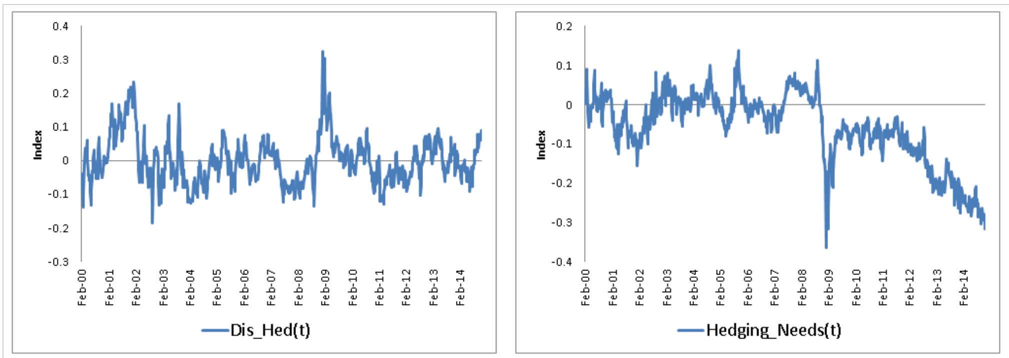

Figure 5: Hedging needs and disequilibrium hedge series

Econometric results empirically confirm that there exists a long-run relationship between the hedging needs of commercial operators ($Hedging\_Needs_t$),

US crude oil balance (Bal_t^{US}) and the 1-4 month WTI time spread.
As Bal_t^{US} decreases (producers *dominates* consumers, *ceteris paribus* long positions in futures increase noticeably), the hedging needs decrease as well. They consequently become more and more negative ($LC_2 > 0$). Conversely, LC_3 is found to be negative and highly significant: as $TS\_1\_4_t$ decreases (*contango* widens), incentives to stock the commodity increase. Short positions in futures go up signaling higher hedging needs by US crude oil operators.
With regard to the adjustment factors, Table 5 exhibits that the two coefficients C_{106} and C_{206} are strongly statistically significant. In other words, both Bal_t^{US} and PNL_t^{Com} tend to adjust to a disequilibrium in hedging needs.
As a last step of the empirical analysis aimed at describing the hedging needs of commercial operators, we have to test whether coefficients are constant over time. In order to check the possibility that parameters are structurally stable, the Vector ECM outlined above (system of equation (2)) is estimated through rolling regressions.
In the first panel of Figure 6, a sequence of rolling windows of fixed length of 371 observations (seven years of data or - approximately - half the full estimation period) is used to investigate whether coefficients of the long-run equilibrium equation (3) (parameters $\widehat{LC}_1$, $\widehat{LC}_2$ and $\widehat{LC}_3$) do not vary across samples. The second panel of Figure 6 exhibits the results we obtain when the structural properties of the reversion coefficients C_{101}, C_{201} and C_{301} are examined.
As shown in Figure 6 empirical evidence is definitely acceptable and in line with insights from our theoretical framework. The estimated coefficient $\widehat{LC}_2$ is found to be significant different from zero for almost all rolling regressions. Apart from a small number of samples (those that include the dramatic collapse of crude oil prices during the second part of 2008), results suggest that there is a strong economic relationship between Bal_t^{US} and the percent net long positions of commercial operators. Similar findings can be highlighted when the structural stability of coefficient $\widehat{LC}_3$ is analyzed. The effects of the 1st-4th month time spread on the long-run equilibrium

hedging needs of commercial operators are found to be statistically significant for most of the rolling regressions. However, even in the case of the impact of $TS\_1\_4_t$ on $Hedging\_Needs_t$, we identified some forms of misspecification in the econometric model between 2008 and 2009. During the worst months of the Great Recession, crude oil prices rapidly dropped from $150\$/Bbl$ to $40\$/Bbl$ and the time-spread weakened extremely as a consequence of the sudden oversupply on the market. This situation of very negative time-spread (i.e. super-contango) persisted for several months. This situation coincided with extremely high storage levels implying a huge amount of floating storage to balance the market.
Finally, the structural stability of the reversion coefficients C_{101} and C_{201} is clearly evident from the second panel of Figure 6. The mechanisms of adjustment of PNL_t^{Com} to the potential emergence of long-run hedging needs are found to be quite rapid and statistically relevant. Similarly, in most of all the samples, the US supply-demand balance tends to adjust rapidly to the disequilibrium described by the long-run relationship among PNL_t^{Com}, $TS\_1\_4_t$ and Bal_t^{US}. Similarly to the case of coefficients $\widehat{LC}_2$ and $\widehat{LC}_3$, this relationship loses part of its statistically relevance during the second part of 2008.
In other words, econometric analyzes confirm one of the major predictions of our economic theory: that is, the impact of US supply-demand balance and storage dynamics on hedging needs.

4.2. Price model

The second (and last) phase of our econometric analysis involves the estimation of a simple regression model that describes the short-term dynamics of the disequilibrium price variable. At this purpose, the following equation is estimated through Ordinary Least Square (OLS) econometric techniques:

$$\Delta Dis\_P_t = \delta_0 + \delta_1 \Delta Dis\_P_{t-1} + \delta_2 \Delta PNL_{t-1}^{Com} + \delta_3 \Delta Bal_{t-1}^{US} + \quad (4)$$
$$+\delta_4 \Delta TS\_1\_4_{t-1} + \delta_5 \Delta Dis\_Hed_{t-1} + \omega_0 Dis\_P_{t-1} + \omega_1 \cdot S_{t-1} + \eta_t$$

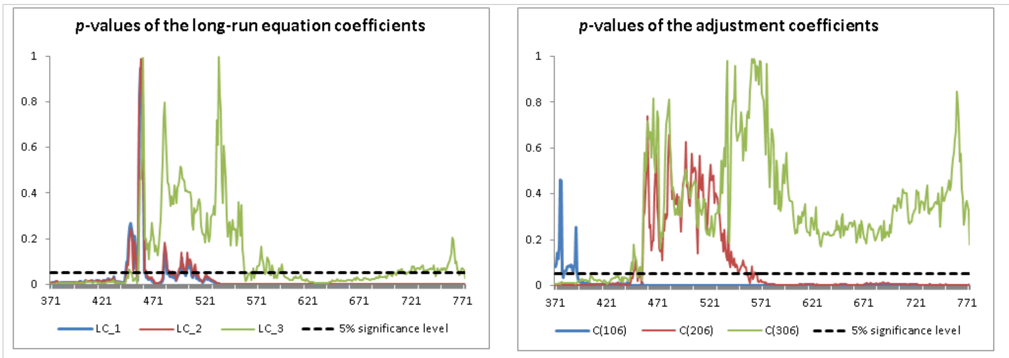

Figure 6: Hedging needs model: 7-year rolling regressions

here S_t denotes a variable obtained by multiplying Dis_P_t by a measure of speculative activity on the market; that is, $S_t = Dis\_P_t \cdot Spec\_var_t$ where $Spec\_var_t$ is either $Spec\_Degree_t$ or $Mod\,T - INDEX_t$.
The crude oil price (expressed as ratio to the equilibrium price)[13] is modeled through a mean reverting process with reversion speed ω_1. Our idea is to test whether the amount of speculative positions in the market increases, $\omega_1 < 0$ (or decreases, $\omega_1 > 0$) the speed of adjustment to the equilibrium price.
In the *full* version of the model, the two variables Dis_P_{t-1} and S_{t-1} are both included in our specification of disequilibrium price. As shown by equation (4), the evolution of the dependent variable, ΔDis_P_t, can also depend on the lagged variables describing the hedging and speculation activity, i.e. "hedging needs" and "disequilibrium hedging".
We have also estimated a *reduced* form of equation (4). This reduced expression for ΔDis_P_t is obtained by using as explanatory variable only one variable between Dis_P_{t-1} and S_t. In other words, either ω_0 or ω_1 is constrained to be zero.
As outlined in Section 3, we believe that the role of speculation in the crude oil market can be properly described by the relative importance of positions

[13] Variable that, as we have seen, is characterized by a unit root, i.e. Dis_P_t is $I(1)$.

held by speculators. This means that, in our first attempts, the $Spec\_var_t$ variable is represented by the $Spec\_Degree_t$ variable (ratio between the open interest held by speculators and positions held by commercials).
As an exercise of robustness check, we employ as speculative variable also our modified T-Index, $Mod\,T - INDEX_t$. This measure has the clear advantage to take into account the objective of commercial operators to protect themselves against abrupt changes in oil market conditions.
Figures 7 and 8 show the estimated coefficients of ω_0 and ω_1 and as well as their statistical significance when equation (4) is estimated by reducing the sample size on the basis of the quantile distribution of Dis_P_t. In other words, according to this quantile-regression approach, we estimate equation (4) by removing a percentage α of observations according to the distribution of Dis_P_t. On the x-axis, we show the percentage of observations of the dependent variable not used for the analysis (α). Estimates of coefficients ω_0 and ω_1 are reported on the y-scale together with their p-values.
Both specifications of the $Spec\_var_t$ variable (i.e., $Spec\_Degree_t$ and $Mod\,T - INDEX_t$) are employed. In addition results refer to the estimation of either the *full* or the *reduced* form of equation (4).
Tables 6 and 7 report the results obtained by estimating equation (4) using only the observations of Dis_P_t that satisfy $q(0.2) < Dis\_P_t < q(0.8)$. That means that, according to our estimation procedure we do not consider the quantiles corresponding to the lower and upper 20% of the distribution of Dis_P_t (see also discussion above).
Our results suggest that ω_0 and ω_1 are jointly negative and statistically significant for both our speculation measures, $Spec\_Degree_t$ (Table 6) and $Modified\_T\_Index$ (Table 7).
Similarly to the majority of the previous literature, we find that the hedging process in futures markets does not lead price movements. Neither the PNL measure or the disequilibrium in hedging are found to be statistically significant.
To conclude, by adopting this procedure we find empirical evidence that, when the price is not too far from the long-run equilibrium level, the speed of adjustment is greater and more statistically significant. The statistical

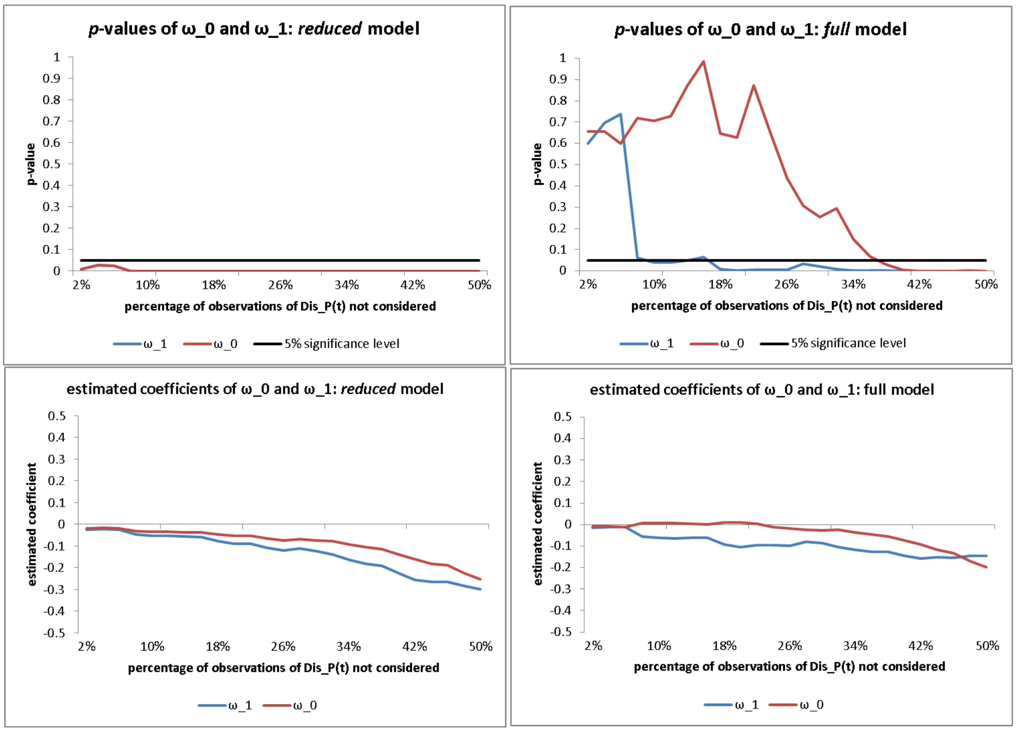

Figure 7: Oil price model: estimation by quantiles of Dis_P_t ($Spec\_var_t = Spec\_Degree_t$)

relevance of the speculation variable strongly diminishes when the price is very far from equilibrium (that is, when crude oil markets are characterized by important shocks in fundamental factors).

5. Concluding Remarks

During the last few years several theoretical and empirical analyzes have studied the role of speculators in crude oil markets. Often, this research has described their activities in a negative way. Speculative positions taken by institutional investors have been often associated with important (and unjustified) increases in both futures and spot commodity prices.

In this paper, we highlight a different point of view. We present a theoreti-

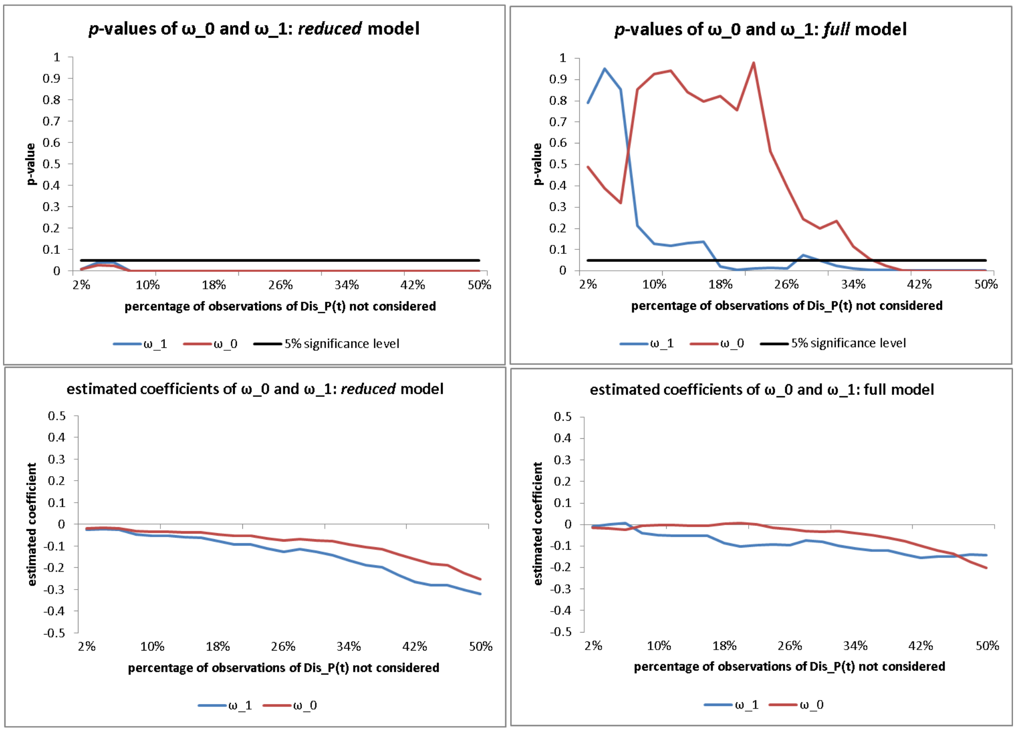

Figure 8: Oil price model: estimation by quantiles of Dis_P_t ($Spec\_var_t = Mod\ T - INDEX_t$)

cal framework based upon the relationships between two classes of agents: commercial operators and speculators.

The first category of agents comprises both oil producing companies and oil consumers. These operators have - what we called - “hedging needs”. That is, commercials use financial markets to hedge the risk associated to ”adverse” oil price movements. On the contrary, speculators operate as they have “money needs”. They act as counterparts of commercial operators. In doing these activities, their objective is to “make money”.

Our paper is novel in many aspects. In particular, the role played by the different classes of agents on crude oil markets and the effects of these relationships on prices is examined through state-of-art time-series econometric techniques.

Financial markets are argued to have strong effects on oil prices. Nevertheless, in our opinion, this impact has to be evaluated in a positive manner. In fact, the interconnections between commercials and speculators help to stabilize crude oil prices. As a result, prices approach a level close to their long-run equilibrium.

Accordingly, we set up an econometric procedure aimed at describing hedging needs as well as short-term dynamics of the disequilibrium oil price.

As a preliminary analysis the statistical properties of the series are studied by employing the procedure proposed by Pesaran et al. [29]. This framework allows us to examine the relationships among the relevant variables without having preliminary knowledge on the order of integration of each time series.

The core of the econometric analysis is based upon two steps. In the first one, a structural VECM is estimated using, as endogenous variables, measures of US supply-demand balance, storage dynamics and positions held by commercials. The system of equations is augmented to include a long-run relationship representing the hedging needs of US producers and consumers.

In the second part of the econometric analysis, a long-run measure for the crude oil price is introduced. Short-term deviations from this equilibrium level are modeled by including as explanatory variables proxies of speculative activity.

In accordance with our theory, empirical evidence shows a strong long-run relationship between the hedging needs of commercial operators, US crude oil balance and 1-4 month WTI time spread. Results also suggest that endogenous variables tend to adjust rapidly to a disequilibrium in hedging needs.

Estimation results of the dynamics of the disequilibrium price are consistent with a statistically significant impact of speculative activity on the adjustment mechanism of prices. Increases in the amount of speculative positions in the market are associated with a more rapid process of prices towards their long-run equilibrium.

Finally, our empirical findings are found to be robust across different sample periods. In fact, our research show that the econometric model is valid over a recent and long sample of data, period that was characterized by important oil price shocks.

Table 1: **Measured employed to assess the impact of speculation on oil prices**

Working's T-Index:

$$T-Index_t = \begin{cases} 1+\frac{SS_t}{HS_t+HL_t} & \text{if } HS_t \geq HL_t \\ 1+\frac{SL_t}{HS_t+HL_t} & \text{if } HS_t < HL_t \end{cases}$$

Percent of total open interest held by each CFTC trader classification:

$$\text{Reporting Commercials' percent of } TOI_t = \frac{CL_t + CS_t}{2 \cdot TOI_t}$$

Percent net long (PNL) position:

$$PNL_t^{Com} = \frac{CL_t - CS_t}{CL_t + CS_t}$$

$$PNL_t^{Spec} = \frac{SL_t - SS_t}{SL_t + SS_t + 2 \cdot SP_t}$$

Lakonishok, Shleifer and Vishny (1992)'s H-Index:

$$H_t = \left| \frac{B_t}{B_t + S_t - p_t} \right| - AF_t$$

Measured related to volume and open interest data:

$$\text{(a) } R1_t = \frac{V_t}{TOI_t}, \text{ (b) } R2_t = \frac{V_t}{|\Delta TOI_t|} \text{(c) } R3_t = \frac{\Delta TOI_t}{V_t}$$

Legenda: SS_t (resp. SL_t): speculation short (resp. long); HS_t (resp. HL_t): hedging short (resp. long); CS_t (resp. CL_t): reporting commercials' net short (resp. long) positions; SP_t: amount of spreading positions; p_t: expected proportion of money managers buying relative to the number of active; B_t (resp. S_t): net buyers (resp. sellers); AF_t: adjustment factor; V_t: total volume of contracts traded and TOI_t: market's total open interest.

Table 2: **ADF tests for unit root**

Regressor	Level		First Difference		Result
	t-statistics	Prob.	t-statistics	Prob.	
PNL_t^{Com}	-2.667	0.080*	-23.547	0.000***	$I(1)$
Bal_t^{US}	-0.049	0.953	-23.385	0.000***	$I(1)$
$TS\_1\_4_t$	-2.895	0.046**	-30.838	0.000***	$I(0)$
Dis_P_t	-2.125	0.235	-28.176	0.000***	$I(1)$

Notes: This table reports the unit root tests using the ADF tests for the set of variables considered by the ARDL analysis. ***, ** and * denote rejection of the null hypothesis of unit root at the 1%, 5% and 10% significance level, respectively.

Table 3: **ARDL Model with Dependent variable** ΔPNL_t^{Com}

Regressor	full model		reduced model	
	value	prob	value	prob
constant	-0.043	0.000***	-0.035	0.000***
PNL_{t-1}^{Com}	-0.052	0.000***	-0.048	0.000***
Bal_{t-1}^{US}	0.072	0.000***	0.059	0.000***
$TS\_1\_4_{t-1}$	0.010	0.003***	0.028	0.023**
Dis_P_{t-1}	-0.001	0.457		
ΔPNL_{t-1}^{Com}	0.130	0.001***	0.181	0.000***
ΔBal_{t-1}^{US}	-0.002	0.956	0.002	0.966
$\Delta TS\_1\_4_{t-1}$	-0.017	0.321	-0.046	0.160
ΔDis_P_{t-1}	-0.028	0.006***		
F-statistic	8.437	0.000***	8.191	0.000***
Akaike info criterion	-5.514		-5.496	
Schwarz criterion	-5.459		-5.454	
Durbin-Watson stat	2.016		1.989	
Breusch-Godfrey Serial Correlation LM Test	2.238	0.107	0.530	0.589

Notes: This table shows the estimated coefficients and diagnostic properties of residuals of the AutoRegressive Distributed Lag Model (1). ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively.

Table 4: **Pesaran et al. [29]'s bound tests**

Wald Test			**Critical bounds ($k=2$)**		
Null hypothesis $H_0: \theta_0 = \theta_1 = \theta_2 = 0$			**Unrestricted intercept and no trend**		
Test Statistic	*df*	Value	Prob.	$I(0)$	$I(1)$
F-Statistic	$(3, 764)$	8.61***	0.01	5.15	6.36
Wald Test			**Critical bounds ($k=2$)**		
$H_0: \theta_0 = 0$, $H_1: \theta_0 < 0$			**Unrestricted intercept and no trend**		
Test-statistic	*df*	Value	Prob.	$I(0)$	$I(1)$
t-Statistic	(770)	-4.821***	0.01	-3.43	-4.10

Notes: This table shows the results obtained by following the ARDL approach presented in Pesaran et al. [29]. The estimated coefficients refer upon the conditional ECM given by eq. (1). ***, ** and * denote rejection of the null hypothesis at the 1%, 5% and 10% significance level, respectively.

Table 5: **VECM Estimates: Model for Hedging Needs**

Coefficient	**Value**	**Std. Error**	**$t-$Statistic**	**Prob.**
C_{101}	0.176	0.036	4.946	0.000***
C_{102}	0.003	0.043	0.081	0.936
C_{103}	-0.052	0.032	-1.590	0.112
C_{106}	-0.039	0.009	-4.202	0.000***
LC_1	-0.808	0.108	-7.509	0.000***
LC_2	1.381	0.197	7.006	0.000***
LC_3	1.043	0.308	3.385	0.001***
C_{201}	0.003	0.029	0.110	0.913
C_{202}	-0.313	0.034	-9.106	0.000***
C_{203}	0.001	0.026	0.045	0.964
C_{206}	0.014	0.007	2.074	0.038**
C_{301}	0.050	0.039	1.276	0.202
C_{302}	0.084	0.047	1.784	0.075*
C_{303}	-0.163	0.036	-4.560	0.000***
C_{306}	0.030	0.010	3.139	0.002***
Included observations: 771				
Total system (balanced) observations 2313				

Notes: This table shows the results of the VECM estimates (system of equations (2)). ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively.

Table 6: **Oil price model ($Spec\_var_t = Spec\_Degree_t$): Full Model**

Regressor	**Coefficient**	**Std. Error**	$t-$**Statistic**	**Prob.**
c	0.049	0.006	7.894	0.000***
ΔDis_P_{t-1}	-0.001	0.053	-0.026	0.979
ΔPNL_{t-1}^{Com}	0.262	0.204	1.287	0.199
ΔBal_{t-1}^{US}	-0.072	0.226	-0.318	0.751
$\Delta TS\_1\_4_{t-1}$	0.157	0.243	0.645	0.519
ΔDis_Hed_{t-1}	-0.064	0.057	-1.119	0.264
S_{t-1}	-0.146	0.040	-3.626	0.000***
Dis_P_{t-1}	-0.073	0.026	-2.859	0.004***
R-squared	0.146	Mean dependent var		0.001
Adjusted R-squared	0.133	S.D. dependent var		0.068
S.E. of regression	0.063	Akaike info criterion		-2.668
Sum squared resid	1.808	Schwarz criterion		-2.597
Log likelihood	623.065	Hannan-Quinn criter.		-2.640
F-statistic	11.100	Durbin-Watson stat		1.833
Prob(F-statistic)	0.000***			

Dependent Variable: ΔDis_P_t
Method: Least Squares
Sample: $q(0.2) < Dis\_P_t < q(0.8)$
Included observations: 461

Notes: This table shows the results obtained by estimating equation (4). ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively.

Table 7: **Oil price model ($Spec\_var_t = Mod\,T - INDEX_t$): Full Model**

Regressor	**Coefficient**	**Std. Error**	**$t-$Statistic**	**Prob.**
c	0.047	0.006	7.655	0.000***
ΔDis_P_{t-1}	-0.002	0.054	-0.044	0.965
ΔPNL_{t-1}^{Com}	0.258	0.205	1.260	0.208
ΔBal_{t-1}^{US}	-0.058	0.226	-0.256	0.798
$\Delta TS\_1\_4_{t-1}$	0.157	0.244	0.645	0.520
ΔDis_Hed_{t-1}	-0.058	0.058	-1.015	0.311
S_{t-1}	-0.138	0.045	-3.095	0.002***
Dis_P_{t-1}	-0.078	0.027	-2.930	0.004***
R-squared	0.140	Mean dependent var		0.001
Adjusted R-squared	0.127	S.D. dependent var		0.068
S.E. of regression	0.063	Akaike info criterion		-2.661
Sum squared resid	1.822	Schwarz criterion		-2.589
Log likelihood	621.293	Hannan-Quinn criter.		-2.632
F-statistic	10.519	Durbin-Watson stat		1.832
Prob(F-statistic)	0.000***			
Dependent Variable: ΔDis_P_t				
Method: Least Squares				
Sample: $q(0.2) < Dis\_P_t < q(0.8)$				
Included observations: 543				
Notes: See notes to Table 6.				

References

[1] Ron Alquist and Lutz Kilian. What do we learn from the price of crude oil futures? *Journal of Applied Econometrics*, 25(4):539–573, 2010.

[2] Ron Alquist, Lutz Kilian, and Robert J. Vigfusson. Forecasting the price of oil. Technical report, Bank of Canada Working Paper, 2011.

[3] Celso Brunetti and Bahattin Büyükşahin. Is speculation destabilizing? *SSRN eLibrary*, 2009.

[4] Bahattin Büyükşahin and Jeffrey H. Harris. Do speculators drive crude oil futures prices? *The Energy Journal*, 32(2):167–202, 2011.

[5] Bahattin Büyükşahin and Michel A. Robe. Speculators, commodities and cross-market linkages. *Journal of International Money and Finance*, 42:38–70, 2014.

[6] Bahattin Büyükşahin, Michael S. Haigh, Jeffrey H. Harris, James A. Overdahl, and Michel A. Robe. Fundamentals, trader activity and derivative pricing. *CFTC Working Paper*, 2009.

[7] Bahattin Büyükşahin, Michael S. Haigh, and Michel A Robe. *Commodities and Equities: A "Market of One"?* Nova Science, 2009.

[8] Giulio Cifarelli. Smooth transition regime shifts and oil price dynamics. *Energy Economics*, 38:160–167, 2013.

[9] Leon P. Drollas. Price determination in the oil market. FEEM Conference, Milan, 1213 January, 2012.

[10] Xiaodong Du, Cindy L. Yu, and Dermot J. Hayes. Speculation and volatility spillover in the crude oil and agricultural commodity markets: A Bayesian analysis. *Energy Economics*, 33(3):497–503, 2011.

[11] Bassam Fattouh. *Oil market dynamics through the lens of the 2002-2009 price cycle*. Oxford Institute for Energy Studies, 2010.

[12] Bassam Fattouh, Lutz Kilian, and Lavan Mahadeva. The role of speculation in oil markets: What have we learned so far? *The Energy Journal*, 34(3):7–33, 2013.

[13] Philip Garcia, Raymond M Leuthold, and Hector Zapata. Lead-lag relationships between trading volume and price variability: New evidence. *Journal of Futures Markets*, 6(1):1–10, 1986.

[14] Christopher L. Gilbert. Commodity speculation and commodity investment. 2010.

[15] James D. Hamilton. Causes and consequences of the oil shock of 2007-08. *Brookings Papers on Economic Activity*, 2009.

[16] James D. Hamilton. Understanding crude oil prices. *Energy Journal*, 30 (2), 2009.

[17] James D Hamilton and Jing Cynthia Wu. Effects of index-fund investing on commodity futures prices. Technical report, National Bureau of Economic Research, 2014.

[18] Søren Johansen. Statistical analysis of cointegration vectors. *Journal of economic dynamics and control*, 12(2):231–254, 1988.

[19] Søren Johansen. Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica: Journal of the Econometric Society*, pages 1551–1580, 1991.

[20] Lutz Kilian. Speculation in oil markets? What have we learned? *VoxEU.org*, 2012.

[21] Lutz Kilian and Daniel P. Murphy. Why agnostic sign restrictions are not enough: understanding the dynamics of oil market VAR models. *Journal of the European Economic Association*, 10(5):1166–1188, 2012.

[22] Lutz Kilian and Daniel P Murphy. The role of inventories and speculative trading in the global market for crude oil. *Journal of Applied Econometrics*, 29(3):454–478, 2014.

[23] Josef Lakonishok, Andrei Shleifer, and Robert W. Vishny. The impact of institutional trading on stock prices. *Journal of financial economics*, 32 (1):23–43, 1992.

[24] Matteo Manera, Marcella Nicolini, and Ilaria Vignati. Financial speculation in energy and agriculture futures markets: A multivariate garch approach. *The Energy Journal*, 34(3):55–81, 2013.

[25] Michael W. Masters. Testimony before the committee on homeland security and governmental affairs. *US Senate, Washington, May*, 20, 2008.

[26] Anne E. Peck. The role of economic analysis in futures market regulation. *American Journal of Agricultural Economics*, 62(5):1037–1043, 1980.

[27] M. Hashem Pesaran and Yongcheol Shin. An autoregressive distributed-lag modelling approach to cointegration analysis. *Econometric Society Monographs*, 31:371–413, 1998.

[28] M Hashem Pesaran and Ron P Smith. Structural analysis of cointegrating vars. *Journal of Economic Surveys*, 12(5):471–505, 1998.

[29] M Hashem Pesaran, Yongcheol Shin, and Richard J Smith. Bounds testing approaches to the analysis of level relationships. *Journal of applied econometrics*, 16(3):289–326, 2001.

[30] Robert S. Pindyck and Julio J. Rotemberg. The excess co-movement of commodity prices. *The Economic Journal*, 100(403):1173–1189, 1990.

[31] Craig Pirrong. Stochastic fundamental volatility, speculation, and commodity storage. *Speculation, and Commodity Storage (October 17, 2008)*, 2008.

[32] Dwight R Sanders, Keith Boris, and Mark Manfredo. Hedgers, funds, and small speculators in the energy futures markets: an analysis of the cftc's commitments of traders reports. *Energy Economics*, 26(3):425–445, 2004.

[33] Annastiina Silvennoinen and Susan Thorp. Financialization, crisis and commodity correlation dynamics. *Journal of International Financial Markets, Institutions and Money*, 24:42–65, 2013.

[34] Kenneth J. Singleton. Investor flows and the 2008 boom/bust in oil prices. *Management Science*, 60(2):300–318, 2013.

[35] Hans R. Stoll and Robert E. Whaley. Commodity index investing and commodity futures prices. *Journal of Applied Finance*, 20(1):7–46, 2010.

[36] Ke Tang and Wei Xiong. Index investment and the financialization of commodities. *Financial Analysts Journal*, 68(6), 2012.

[37] Holbrook Working. Speculation on hedging markets. *Food Research Institute Studies*, (02), 1960.

NOTE DI LAVORO DELLA FONDAZIONE ENI ENRICO MATTEI

Fondazione Eni Enrico Mattei Working Paper Series

Our Note di Lavoro are available on the Internet at the following addresses:

http://www.feem.it/getpage.aspx?id=73&sez=Publications&padre=20&tab=1
http://papers.ssrn.com/sol3/JELJOUR_Results.cfm?form_name=journalbrowse&journal_id=266659
http://ideas.repec.org/s/fem/femwpa.html
http://www.econis.eu/LNG=EN/FAM?PPN=505954494
http://ageconsearch.umn.edu/handle/35978
http://www.bepress.com/feem/

NOTE DI LAVORO PUBLISHED IN 2015

ERM	1.2015	Elena Verdolini, Laura Diaz Anadon, Jiaqi Lu and Gregory F. Nemet: The Effects of Expert Selection, Elicitation Design, and R&D Assumptions on Experts' Estimates of the Future Costs of Photovoltaics
CCSD	2.2015	James Lennox and Ramiro Parrado: Capital-embodied Technologies in CGE Models
CCSD	3.2015	Claire Gavard and Djamel Kirat: Flexibility in the Market for International Carbon Credits and Price Dynamics Difference with European Allowances
CCSD	4.2015	Claire Gavard: Carbon Price and Wind Power Support in Denmark
CCSD	5.2015	Gunnar Luderer, Christoph Bertram, Katherine Calvin, Enrica De Cian and Elmar Kriegler: Implications of Weak Near-term Climate Policies on Long-term Mitigation Pathways
CCSD	6.2015	Francisco J. André and Luis M. de Castro: Incentives for Price Manipulation in Emission Permit Markets with Stackelberg Competition
CCSD	7.2015	C. Dionisio Pérez Blanco and Thomas Thaler: Water Flows in the Economy. An Input-output Framework to Assess Water Productivity in the Castile and León Region (Spain)
CCSD	8.2015	Carlos M. Gómez and C. Dionisio Pérez-Blanco: Simple Myths and Basic Maths about Greening Irrigation
CCSD	9.2015	Elorri Igos, Benedetto Rugani, Sameer Rege, Enrico Benetto, Laurent Drouet, Dan Zachary and Tom Haas: Integrated Environmental Assessment of Future Energy Scenarios Based on Economic Equilibrium Models
ERM	10.2015	Beatriz Martínez and Hipòlit Torró: European Natural Gas Seasonal Effects on Futures Hedging
CCSD	11.2015	Inge van den Bijgaart: The Unilateral Implementation of a Sustainable Growth Path with Directed Technical Change
CCSD	12.2015	Emanuele Massetti, Robert Mendelsohn and Shun Chonabayashi: Using Degree Days to Value Farmland
CCSD	13.2015	Stergios Athanassoglou: Revisiting Worst-case DEA for Composite Indicators
CCSD	14.2015	Francesco Silvestri and Stefano Ghinoi : Municipal Waste Selection and Disposal: Evidences from Lombardy
CCSD	15.2015	Loïc Berger: The Impact of Ambiguity Prudence on Insurance and Prevention
CCSD	16.2015	Vladimir Otrachshenko and Francesco Bosello: Identifying the Link Between Coastal Tourism and Marine Ecosystems in the Baltic, North Sea, and Mediterranean Countries
ERM	17.2015	Charles F. Mason, Lucija A. Muehlenbachs and Sheila M. Olmstead: The Economics of Shale Gas Development
ERM	18.2015	Anna Alberini and Charles Towe: Information v. Energy Efficiency Incentives: Evidence from Residential Electricity Consumption in Maryland
CCSD	19.2015	ZhongXiang Zhang: Crossing the River by Feeling the Stones: The Case of Carbon Trading in China
CCSD	20.2015	Petterson Molina Vale: The Conservation versus Production Trade-off: Does Livestock Intensification Increase Deforestation? The Case of the Brazilian Amazon
CCSD	21.2015	Valentina Bosetti, Melanie Heugues and Alessandro Tavoni: Luring Others into Climate Action: Coalition Formation Games with Threshold and Spillover Effects
CCSD	22.2015	Francesco Bosello, Elisa Delpiazzo, and Fabio Eboli: Macro-economic Impact Assessment of Future Changes in European Marine Ecosystem Services
CCSD	23.2015	Maryse Labriet, Laurent Drouet, Marc Vielle, Richard Loulou, Amit Kanudia and Alain Haurie: Assessment of the Effectiveness of Global Climate Policies Using Coupled Bottom-up and Top-down Models
CCSD	24.2015	Wei Jin and ZhongXiang Zhang: On the Mechanism of International Technology Diffusion for Energy Technological Progress
CCSD	25.2015	Benjamin Michallet, Giuseppe Lucio Gaeta and François Facchini: Greening Up or Not? The Determinants Political Parties' Environmental Concern: An Empirical Analysis Based on European Data (1970-2008)
CCSD	26.2015	Daniel Bodansky, Seth Hoedl, Gilbert Metcalf and Robert Stavins: Facilitating Linkage of Heterogeneous Regional, National, and Sub-National Climate Policies Through a Future International Agreement
CCSD	27.2015	Giannis Vardas and Anastasios Xepapadeas: Time Scale Externalities and the Management of Renewable Resources
CCSD	28.2015	Todd D. Gerarden, Richard G. Newell, Robert N. Stavins and Robert C. Stowe: An Assessment of the Energy-Efficiency Gap and Its Implications for Climate Change Policy
CCSD	29.2015	Cristina Cattaneo and Emanuele Massetti: Migration and Climate Change in Rural Africa
ERM	30.2015	Simone Tagliapietra: The Future of Renewable Energy in the Mediterranean. Translating Potential into Reality
CCSD	31.2015	Jan Siegmeier, Linus Mattauch, Max Franks, David Klenert, Anselm Schultes and Ottmar Edenhofer: A Public Finance Perspective on Climate Policy: Six Interactions That May Enhance Welfare
CCSD	32.2015	Reyer Gerlagh, Inge van den Bijgaart, Hans Nijland and Thomas Michielsen: Fiscal Policy and CO2 Emissions of New Passenger Cars in the EU
CCSD	33.2015	Marie-Laure Nauleau, Louis-Gaëtan Giraudet and Philippe Quirion: Energy Efficiency Policy with Price-quality Discrimination

CCSD	34.2015	Eftichios S. Sartzetakis, Anastasios Xepapadeas and Athanasios Yannacopoulos: Regulating the Environmental Consequences of Preferences for Social Status within an Evolutionary Framework
CCSD	35.2015	Todd D. Gerarden, Richard G. Newell and Robert N. Stavins: Assessing the Energy-efficiency Gap
CCSD	36.2015	Lorenza Campagnolo and Fabio Eboli: Implications of the 2030 EU Resource Efficiency Target on Sustainable Development
CCSD	37.2015	Max Franks, Ottmar Edenhofer and Kai Lessmann: Why Finance Ministers Favor Carbon Taxes, Even if They Do not Take Climate Change into Account
CCSD	38.2015	ZhongXiang Zhang: Carbon Emissions Trading in China: The Evolution from Pilots to a Nationwide Scheme
CCSD	39.2015	David García-León: Weather and Income: Lessons from the Main European Regions
CCSD	40.2015	Jaroslav Mysiak and C. D. Pérez-Blanco: Partnerships for Affordable and Equitable Disaster Insurance
CCSD	41.2015	S. Surminski, J.C.J.H. Aerts, W.J.W. Botzen, P. Hudson, J. Mysiak and C. D. Pérez-Blanco: Reflections on the Current Debate on How to Link Flood Insurance and Disaster Risk Reduction in the European Union
CCSD	42.2015	Erin Baker, Olaitan Olaleye and Lara Aleluia Reis: Decision Frameworks and the Investment in R&D
CCSD	43.2015	C. D. Pérez-Blanco and C. M. Gómez: Revealing the Willingness to Pay for Income Insurance in Agriculture
CCSD	44.2015	Banchongsan Charoensook: On the Interaction between Player Heterogeneity and Partner Heterogeneity in Two-way Flow Strict Nash Networks
CCSD	45.2015	Erin Baker, Valentina Bosetti, Laura Diaz Anadon, Max Henrion and Lara Aleluia Reis: Future Costs of Key Low-Carbon Energy Technologies: Harmonization and Aggregation of Energy Technology Expert Elicitation Data
CCSD	46.2015	Sushanta Kumar Mahapatra and Keshab Chandra Ratha: Sovereign States and Surging Water: Brahmaputra River between China and India
CCSD	47.2015	Thomas Longden: CO_2 Intensity and the Importance of Country Level Differences: An Analysis of the Relationship Between per Capita Emissions and Population Density
CCSD	48.2015	Jussi Lintunen and Olli-Pekka Kuusela: Optimal Management of Markets for Bankable Emission Permits
CCSD	49.2015	Johannes Emmerling: Uncertainty and Natural Resources - Prudence Facing Doomsday
ERM	50.2015	Manfred Hafner and Simone Tagliapietra: Turkish Stream: What Strategy for Europe?
ERM	51.2015	Thomas Sattich, Inga Ydersbond and Daniel Scholten: Can EU's Decarbonisation Agenda Break the State-Company Axis in the Power Sector?
ERM	52.2015	Alessandro Cologni, Elisa Scarpa and Francesco Giuseppe Sitzia: Big Fish: Oil Markets and Speculation