ASSESSING THE CROSS-COMMODITY RELATIONSHIPS IN ENERGY MARKETS

JULIEN CHEVALLIER
IPAG Business School, IPAG Lab, 184 Boulevard Saint-Germain, 75006 Paris, France

FLORIAN IELPO
Lombard Odier Asset Management , 6 avenue des Morgines, 1213 Petit-Lancy, Switzerland

ABSTRACT

This article reviews cointegration analyses in energy markets (e.g. WTI, Brent, Gasoil, Heating Oil, Natural Gas) during 1993-2011 using daily data. The main focus lies in the determination of long-term relationships between these specific energy commodities, with the inclusion of structural breaks. The results globally point out the existence of shared trends among energy markets. The crude oil price (with the WTI as its benchmark) seems to be the leader in the price discovery process, since most of the time it triggers the adjustement towards the long term stationary equilibrium between the variables in the cointegration system.

Key words
Energy Futures; Cointegration; Cross-Commodity Relationship; Structural Break

Q40; Q49; C32; C58

Corresponding Author:
IPAG Business School (IPAG Lab), 184 Boulevard Saint-Germain,
75006 Paris, France
Email: julien.chevallier04@univ-paris8.fr

1 INTRODUCTION

Cross-commodity relationships imply that two or several commodities share an equilibrium that links prices in the long run. These long-term connections – or inter-commodity equilibrium as denoted by Casassus et al. (2009) – include production relationships[1] where upstream commodity and downstream commodity are tied in a production process, and substitute/complementary relationships where two commodities serve as substitute[2] or complement.[3] The existence of inter-commodity equilibrium usually indicates long-term co-movement among commodity prices. Temporary deviations from this equilibrium (because of demand and supply imbalances caused by macroeconomic factors and inventory shocks, etc.) will be corrected over the long-run.

In this article, cointegration is used to analyze the long-term equilibrium relationships that may occur between energy markets. To illustrate the mechanisms at stake, consider the case where crude oil, unleaded gasoline, and heating oil futures prices are cointegrated. This suggests that the crack spreads[4] will not deviate without bounds, and will revert to their 'normal' levels. On the other hand, if the futures prices of crude oil, unleaded gasoline, and heating oil are not cointegrated, then the crack spreads can deviate without bounds, and using these spreads as risk management tool or as speculative vehicle will be questionable. Therefore, cointegration among these petroleum futures prices will make it possible to use statistical tools for determining extremes. The extremes in turn can then be used as a basis for trading strategies, and to explore risk arbitrage opportunities in crack spreads. Therefore, cointegration results will be used to identify relative mispricings that could be exploited.

As another illustrative example, we may refer to Gjolberg and Johnsen (1999), who argue that stable long run equilibrium relationships may exist between crude and oil product prices. Despite physical limits as to the relative amounts of different products that can be distilled from a barrel of crude, refiners do have some flexibility in their product mix. This flexibility can be utilized when shifts in relative prices occur, and can be enhanced by stock adjustments. Consequently, while relative prices may fluctuate, they are likely to gravitate back towards some long run equilibrium level. If such long run equilibrium price relationships do in fact exist,

---

[1] One commodity can be produced from another commodity when the former is the output of a production process that uses the other commodity as an input factor. For example, the petroleum refining process cracks crude oil into its constituent products, among which heating oil and gasoline are actively traded commodities on the NYMEX.

[2] A substitute relationship exists when two traded commodities are substitutes in consumption. Crude oil and natural gas are commonly viewed as substitute goods. Competition between natural gas and petroleum products occurs principally in the industrial and electric generation sectors. Similarly, ethanol and petroleum products are potentially competitive products.

[3] A complementary relationship exists when two commodities share a balanced supply or are complementary in consumption and/or production. in consumption and/or production.

[4] Defined as the spread created by purchasing oil futures and offsetting the position by selling gasoline and heating oil futures. The name of this investment strategy is derived from the fact that cracking oil produces gasoline and heating oil. Therefore, oil refiners are able to generate residual income by entering into these transactions. During the summer of 2005, the effects of hurricanes in the Southeastern United States created large volatility in the crack spread.

they may represent valuable information for risk management in integrated oil companies, which both produce and refine crude oil.[5] Finally, if the price process can be described as an error correction mechanism, this may be utilized for selective hedging, based on information about likely price movements back towards a long term equilibrium price relationship. In their empirical application, Gjolberg and Johnsen (1999) find indeed that for all refined products (possibly excluding heavy fuel oil) prices are cointegrated with the crude price. Hence, the current product-crude margin deviations from a long run equilibrium may contain significant information about the future changes in product prices and margins, which has implications in terms of product price forecasting and risk management.

Natural gas, coal and electricity prices are also characterized by equilibrium relationships in energy markets, especially from the perspective of power production. In the long run, Moutinho et al. (2011) underline that it is important to account for the electricity generating technologies, given that fuels compete on a cost basis in electricity production. In addition, fuel substitution capabilities within the electricity sector, either at plant- or grid-level, should contribute to the cointegrating relationship between energy prices. Hence, substitutability between crude oil, coal and gas products in the industrial sector, through direct use and cogeneration of electricity, can also influence the commodity price relationship. In order to understand why coal, crude oil and gas prices sometimes diverge from their long term equilibrium, it is also important to control for various short term factors that establish trends in the prices of electricity and other commodities.

Another central relationship in energy markets links oil and gas prices. Indeed, the hypothesis that oil and natural gas could be cointegrated comes from the fact that both fuels can be seen as substitutes in the production of many intermediary consumption goods. Hence, when their prices diverge, market agents would be able to arbitrate between both markets to adopt the cheapest energy source. But various idiosyncratic shocks can disrupt this relation, for instance the fact that Gas markets are more regional in essence than the crude oil market. Another example lies in the fact that oil is storable while storing and transporting gas is more difficult[6].

Many studies have documented empirically cointegrating relationships in energy markets. Girma and Paulson (1999) find a cointegration relationship in petroleum futures markets. Ai et al. (2006) document that the market-level indicators such as inventory and harvest size explain a strikingly large portion of commodity prices in the long run. Cortazar et al. (2008) have studied the statistical relationship among commodities in a multi-commodity framework using futures prices. Akram (2009) reveals that different pairs of the commodity prices may be cointegrated. However, none of these papers has provided a thorough analysis of cointegrating relationships within the specific group of energy commodities. Hence, this article aims at filling this gap in the literature. We choose to focus on the core of the long-run relationship within energy commodities based on the cointegration approach

---

[5] The net price risk for an integrated company will depend on the price variability of its portfolio of crude and refined products. Risk management should, therefore, take into account the covariance structure of all prices, and hedging should not be on a product-by-product basis. Furthermore, if product and crude prices are cointegrated, then the standard approach for establishing a risk-minimizing hedge may yield biased estimates.

[6] at least before the advent of shale gas, which is not discussed here.

(with/without structural breaks), which requires a strong economic rationale to attempt to relate two (or more) variables together overtime.

The rest of the article is structured as follows. Section 2 provides a literature review. Section 3 details the data used. Section 4 contains cointegration results for specific groups of energy markets. Section 5 conludes. The Appendix contains an overview of the cointegration methodology – with and without structural breaks.

2. LITERATURE REVIEW

Cointegration among energy markets has been extensively documented in the academic literature. We present below the main results related to three groups: *(i)* petroleum products, *(ii)* oil, gas and coal prices, and *(iii)* electricity and fuel prices.

2.1 Petroleum products

Chaudhuri (2001) tries to ascertain the role played by real oil prices in explaining the extremely volatile movements in real prices of primary commodities[7] by taking into account oil price shocks. The author shows that real commodity prices and real oil prices are cointegrated during 1973-1996, while the magnitude of oil price shocks could differ substantially among commodity markets. Additionally, the error-correction term stimulates the real commodity price adjustment (but not the real oil price adjustment).

Asche et al. (2003) investigate the relationship between Brent crude oil and refined product prices during 1992-2000. They find empirical evidence of a long run relationship between the prices of crude oil, gasoil, kerosene and naphtha. Evidence of a close relationship between the latter three refined products indicates that these markets are integrated. The crude oil price is found to be weakly exogenous, i.e. refined product prices are dependent on the crude oil price but not vice versa. No cointegration relationship can be found between crude and heavy fuel oil.

Lanza et al. (2005) provide a comprehensive analysis of the price dynamics between 10 varieties of heavy crude oils[8] and product prices[9] in Europe and the USA during 1994-2002. They show that *(i)* product prices are statistically relevant in explaining short and long run adjustment in petroleum markets, and *(ii)* the long-run adjustment coefficients are sensitive to the gravity of the specific crude.[10]

Murat and Tokat (2009) analyze the crack spread[11] on the WTI crude oil by using weekly NYMEX futures during 2000-2008. They establish a causal impact of crack

---

[7] The commodities included in this analysis are: Aluminium, Bananas, Beef, Coal, Cocoa Beans, Coffee, Copper, Cotton, Groundnuts, Hides, Jute, Lamb, Lead, Maize, Manganese, Nickel, Plywood, Potash, Pulp, Rice, Rubber, Silver, Sugar, Tea, Tin, Tobacco, Wheat, Wool and Zinc.

[8] e.g. Brent, Urals, Iranian, Forcados, WTI, Maya, Boscan, Kern River, Thums.

[9] e.g. Gasoline, Gasoil, High Sulphur Fuel Oil (HSFO) and Low Sulphur Fuel Oil (LSFO).

[10] Prices of crude oils whose physical characteristics are more similar to the marker are likely to converge more rapidly to the long run equilibrium.

[11] Recall that the crack spread is a term used in the oil industry and futures trading for the differential between the price of crude oil and petroleum products extracted from it - that is, the profit margin that an oil refinery can expect to make by cracking crude oil (i.e. breaking its long-chain hydrocarbons into useful shorter-chain petroleum products). In the futures markets, the crack spread is a specific spread trade involving simultaneously buying and selling contracts in crude oil and one or more derivative products, typically gasoline and heating oil. Oil refineries may trade a crack spread to hedge the price risk of their operations, while speculators attempt to profit from a change in the oil/gasoline price differential.

spread futures on crude oil markets both in the long and short run after 2003 (where they detected a structural break). Westgaard et al. (2011) also examine the spread between gasoil and Brent crude oil futures prices on ICE Futures – i.e. the crack spread – during 1994-2009. A cointegration relationship is found for the 1- and 2-month futures contracts during 1994-2009. However, no cointegration relationships can be found during 2002-2009. The hurricane Katrina, the economic boom and the following financial crises might explain these results. In such volatile periods the spread between gasoil and crude oil is likely to deviate, and it might take several years until it reverts to its equilibrium value. For energy traders and hedgers, the authors suggest that exposures to the crack spread should therefore be treated with great care in such market environments.

2.2 Oil, gas and coal prices

Serletis and Herbert (1999) identify shared trends among the US Henry Hub (HH) natural gas price and fuel oil prices during 1996-1997 (using daily data). Interestingly, they also feature feedback relationships, which supposes the existence of effective arbitraging mechanisms across the two markets.

Bachmeier and Griffin (2006) evaluate the degree of market integration both within and between crude oil, coal, and natural gas markets during 1989-2004. They find that world crude oil prices[12] are cointegrated, but that the degree of market integration is much weaker for US coal prices.[13] Finally, they show that the crude oil, coal, and natural gas markets are only weakly integrated.

Panagiotidis and Rutledge (2007) find a cointegrating relationship among UK gas prices and Brent oil prices during 1996-2003. Despite the highly liberalized nature of the UK gas market, they show that gas and oil prices were still moving together in the long run. This latter result is highly debated in the literature. For instance, Villar and Joutz (2006) established earlier, based on a cointegration analysis, that the oil and natural gas prices may have appeared to decouple during 1989-2005. On the same topic, Brown and Yücel (2008) and Hartley et al. (2008) are able to identify a cointegration relationship between the WTI crude oil and HH natural gas prices during 1994-2007 and 1990-2006, respectively. While Brown and Yücel (2008) find that short-run deviations from the estimated long-run relationship could be explained by influences of weather, seasonality, natural gas storage and production in the Gulf of Mexico, Hartley et al. (2008) find that seasonal fluctuations and other factors such as weather shocks and changes in storage can have a significant influence on the short-run dynamic adjustment of prices.

Moutinho et al. (2011) reveal that the prices of Zeebrugge gas, API coal, fuel oil and Brent crude oil are cointegrated in Spain during 2002-2005. In addition, the prices of Brent tend to move to re-establish the price equilibrium. The suggested economic mechanism is the following: if there is an increase in demand, and taking into account a fixed production capacity, fuel and the raw material from which it is made, crude oil, becomes scarcer inherently making both commodities more expensive. Based on these characteristics, the authors predict that the tendency for

---

[12] e.g. West Texas Intermediate (WTI) traded at Cushing, Oklahoma, Brent crude from the UK sector of the North Sea, Dubai crude from the Middle East, Arun crude from Indonesia, and Alaskan North Slope (ANS) crude traded near Los Angeles.

[13] e.g. Colorado, Utah and Wyoming in the West, Kentucky and Ohio in the East.

crude oil and other fossil fuel prices (gas, coal and fuel oil) to move quickly and follow one another will strengthen, due to the substitutability of the four products in the heating and electricity markets.

2.3 Electricity and fuel prices

Asche et al. (2006) find that natural gas, crude oil and electricity prices are being cointegrated during 1995-1998 (i.e. after the deregulation of the UK gas market), with a leading role played by crude oil in the long term relationship.

Mjelde and Bessler (2009) find that the price series of natural gas, crude oil, coal and uranium are cointegrated with electricity prices from the US Pennsylvania - New Jersey - Maryland Interconnection (PJM) during 2001-2008. However, the authors are not able to detect one common trend, but that fuel prices tend to move electricity prices.

In his cointegration analysis, Mohammadi (2009) challenges this result by showing that coal, natural gas and crude oil do not affect electricity prices significantly during 1960-2007. Significant long-run relationships are found only between electricity and coal prices.

In a recent contribution, Bencivenga et al. (2011) analyze the relationships existing between crude oil, natural gas and electricity prices in the USA and Europe by using an error correction model (ECM) framework during 2001-2010. Their results illustrate that a long-run equilibrium exist between the various pairs of energy commodities in Europe and the North American market.

The main findings of cointegration analyses that can be found across energy markets are summarized in Table 1. Taken together, these studies provide overwhelming evidence in favor of a link between crude oil and other fuel prices in the long term. This result may be explained on solid economic grounds, given the indexation of many long term energy futures contracts on the price of oil, and the determination of other energy prices based on various qualities of oil products as an input to production. However, the link between oil and gas may have disappareared in the recent period (as investigated by Bachmeier and Griffin (2006) and Brown and Yücel (2008)) due to industrial changes in the production of natural gas at the regional level (especially in the US with the development of shale gas). Note that our database includes neither petroleum products, nor electricity prices. Hence, we are mostly interested in digging further the cointegration relationships between the WTI and Brent crude oil, Gasoil, Natural Gas and Heating Oil in our empirical application during 1993-2011.

Table 1: Energy Prices: Cointegrating Relationships

Authors	Period	Cointegration Relationship	SS	SB
Petroleum products				
Chaudhuri (2001)	1973-1996	Real Oil Prices ↔ Real Commodity Prices	No	No
Asche et al. (2003)	1992-2000	Crude Oil ↔ Gasoil, Kerosene, Naphta	No	No
		Crude Oil ∅ Heavy Fuel Oil		
Lanza et al. (2005)	1994-2002	Crude Oil ↔ Gasoline, Gasoil, HSFO, LSFO	No	No
Murat and Tokat (2009)	2000-2008	Crude Oil ↔ Gasoil (≈ Crack Spread)	Yes	Yes
Westgaard et al. (2011)	1994-2009	Brent Oil ↔ Gasoil (≈ Crack Spread)	Yes	No
Oil, gas and coal prices				
Serletis and Herbert (1999)	1996-1997	Fuel Oil ↔ Natural Gas	No	No
Bachmeier and Griffin (2006)	1989-2004	WTI ↔ Brent ↔ ANS ↔ Dubai ↔ Arun	No	No
		US Western Coal ↔ US Eastern Coal		
		WTI ∅ Wyoming Coal ∅ Natural Gas		
Villar and Joutz (2006)	1989-2005	WTI Oil ↔ Natural Gas	No	No
Panagiotidis and Rutledge (2007)	1996-2003	Brent Oil ↔ UK Natural Gas	No	No
Brown and Yücel (2008)	1994-2007	WTI Oil ↔ Natural Gas	No	No
Hartley et al. (2008)	1990-2006	WTI Oil ↔ Natural Gas	No	No
Moutinho et al. (2011)	2002-2005	Brent Oil ↔ Fuel Oil ↔ Gas ↔ Coal	No	No
Electricity and fuel prices				
Asche et al. (2006)	1995-1998	Crude Oil ↔ Natural Gas ↔ Electricity	No	No
Mjelde and Bessler (2009)	2001-2008	WTI ↔ Natural Gas ↔ Coal ↔ Uranium ↔ Elec	No	No
Mohammadi (2009)	1960-2007	Coal ↔ Electricity	No	No
		Crude Oil ∅ Gas ∅ Coal ∅ Electricity		
Bencivenga et al. (2011)	2001-2010	Crude Oil ↔ Natural Gas ↔ Electricity	No	No

Note: ↔ indicates the presence of a cointegration relationship. ∅ indicates the absence of a cointegration relationship. *SS* stands for 'Sub Sample' analysis in the paper considered. *SB* stands for 'Structural Break' analysis in the paper considered.

3. DATASET AND UNIT ROOT TESTS

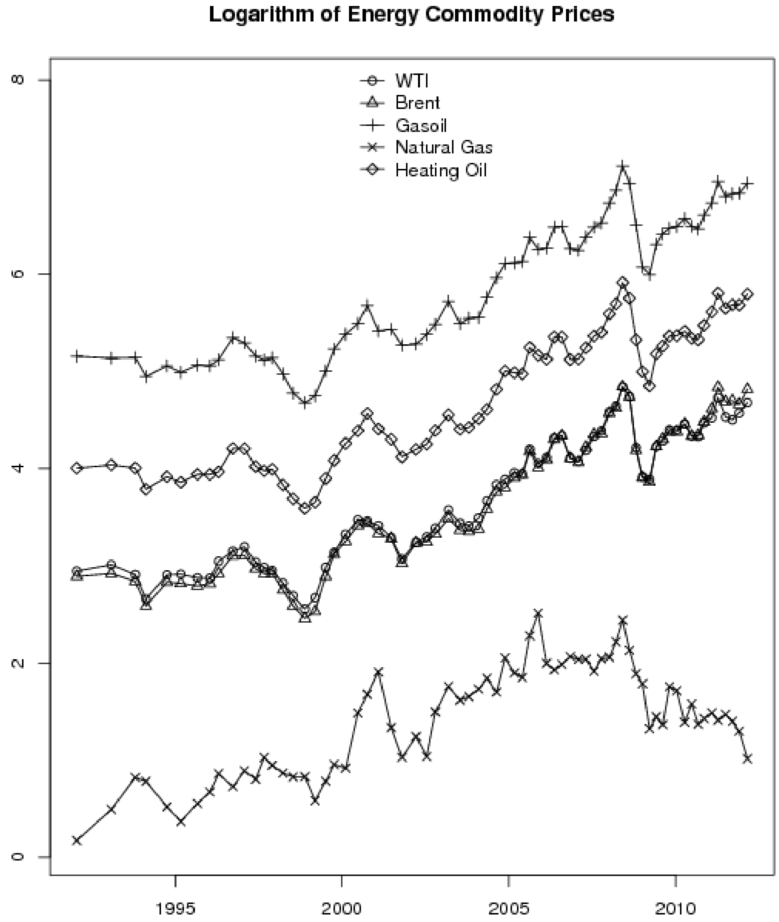
Table 2: Descriptive statistics for energy prices [1993-2011]

	Min	Max	Mean	Std. Dev.	Skew.	Kurt.	JB
WTI	2.3997	4.9826	3.7166	0.6623	-0.0490	4.5999	217.7409
Brent	2.2773	4.9877	3.6916	0.7118	-0.0020	4.4380	212.5636
Gasoil	4.5136	7.1894	5.8695	0.6970	0.0109	3.9951	216.4479
Natural Gas	0.1266	2.7361	1.4199	0.5385	0.0168	6.8089	93.4769
Heating Oil	3.4049	6.0190	4.7411	0.6827	-0.0068	3.1474	216.5013

Note: The number of observations is equal to 2,757. Std. Dev. stands for Standard Deviation, Kurt. for Kurtosis, Skew. for Skewness, and JB for the Jarque Bera test statistic.

Table 2 presents descriptive statistics on the energy prices used in this article, i.e. WTI, Brent, Gasoil, Natural Gas, and Heating Oil. All the data comes from Bloomberg in daily frequency. The dataset starts in 1993 and ends in 2011. The start of the dataset corresponds to the maximum historical that could be accessed from Bloomberg for all the time series included in this article. More precisely, concerning the specific time series retained in our article, WTI is the New York Mercantile Exchange (NYMEX) WTI Light Sweet Crude Oil futures price traded in US Dollar per barrel, Brent is the InterContinental Exchange (ICE) Brent Crude futures price traded in US Dollar per barrel, Gasoil is the Reformulated Regular Gasoline Blendstock (RBOB) Gasoline futures price traded in US Dollar per gallon, Natural Gas is the Henry Hub Natural Gas futures price traded in US Dollar per million British thermal units (mmBtu), and Heating Oil is the NYMEX Heating Oil futures price traded in US Dollar per gallon. The descriptive statistics concern the variables transformed to log-returns. We can notice departure from normality with excess kurtosis and a skewness coefficient different from 3. This comment is further confirmed by the value of the Jarque Bera test statistic.

Figure 1: Logarithm of Time Series for Energy Commodities



The time series are shown in logarithm form in Figure 1. We notice visually that energy markets display an homogeneous behavior across the time period, with similar price movements during periods of expansion / recession (except perhaps for natural gas). Hence, by means of this preliminary visual inspection, we validate intuitively the need to resort to cointegration to analyze in more details the behavior of the respective groups of commodities included in this article.

Table 3: Unit Root Test Results for Energy Prices

	ADF None	ADF Drift	ADF Trend	PP Constant	PP Trend	KPSS
WTI	-38.6010	-38.6222	-38.6159	-54.7841	-54.7748	0.0434
Brent	-38.0714	-38.1020	-38.0977	-55.5461	-55.5393	0.0465
Gasoil	-36.7820	-36.8087	-36.8045	-52.6444	-52.6380	0.0542
Natural Gas	-37.0949	-37.0898	-37.1216	-54.3157	-54.3509	0.0222
Heating Oil	-37.6437	-37.6683	-37.6633	-54.6559	-54.6487	0.0482

Note: Test statistics are given. ADF stands for the Augmented Dickey-Fuller unit root test, PP for the Phillips-Perron unit roor test, and KPSS for the Kwiatkowski Phillips Schmidt Shin unit root test. Corresponding critical values (at 5% level) can be found in Greene (2011): -1.9409 for ADF None, -2.8623 for ADF Drift, -3.4114 for ADF Trend, -2.8623 for PP Constant, -3.4114 for PP Trend, and 0.4630 for KPSS.

Table 3 reports the usual unit root tests (ADF, PP, KPSS) results for the prices of energy markets. These tests are meant to check formally the stationarity of the time series under consideration, in addition to the preliminary investigation of the plot for each time series. All tests have been conducted on raw data, and then on log-returns. The results reproduced concern the log-returns, which are all shown to be stationary. Indeed, the test statistics reproduced for the ADF and PP tests are far smaller than the corresponding critical values (which can be found in Greene (2011) for instance, and mentioned at the bottom of the table). The conclusion is that we can safely reject the null hypothesis of unit root for these two tests. For the KPSS test, the test statistics are smaller than the critical values (see again Greene (2011)), which leads us to accept the null hypothesis of stationarity. We verify that all series are integrated of the same order (*I(1)*), which is a pre-condition for cointegration.

4. COINTEGRATION ANALYSES FOR ENERGY MARKETS

In the context of energy markets, we are mainly interested in testing the hypothesis of cointegration in two categories: *(i)* Petroleum products, and *(ii)* Oil and gas prices. Note that the analysis cannot be performed on the entire spectrum of fuel prices, i.e. including coal. Indeed, coal futures contracts are mostly traded on a very long term basis (i.e. several years ahead). Therefore, the availability of long time series is limited and not currently included in our Bloomberg database. Hence, we cannot fully replicate the results from previous literature dealing with coal data.

Electricity prices are not included in our review for three main reasons. First, electricity prices are mostly studied by resorting to spreads, for instance documenting the cost to produce one MWh of electricity from a coal- vs. gas-fired power plant. This would lead us to many technicalities which are dealt with in specialized books or articles. Second, the use of raw electricity prices is most often performed by using hourly data (24-hour) which is not compatible with the daily frequency of the data used in this article. Third, the discussion about peak vs. base electricity prices complicates the treatment of the information in the cointegration framework that we develop.

Our main results regarding energy prices are contained in Table 4. The Appendix contains a review of the cointegration methodology with/without structural breaks. Let us start our investigation with the first category labelled as petroleum products.

**Table 4: Cointegration Analyses of Energy Prices:
Summary of the Main Results**

Period	Cointegration Relationship	SB
Petroleum products		
-2011	WTI ↔ Gasoil	No
-2000	WTI ↔ Gasoil	No
-2011	WTI ↔ Gasoil	No
-2011	WTI ↔ Gasoil	Yes
-2011	Brent ↔ Gasoil	No
-2000	Brent ↔ Gasoil	No
-2011	Brent ↔ Gasoil	No
-2011	Brent ↔ Gasoil	Yes
-2011	WTI ↔ Heating Oil	No
-2000	WTI ↔ Heating Oil	No
-2011	WTI ↔ Heating Oil	No
-2011	WTI ↔ Heating Oil	Yes
-2011	Brent ↔ Heating Oil	No
-2000	Brent ↔ Heating Oil	No
-2011	Brent ↔ Heating Oil	No
-2011	Brent ↔ Heating Oil	Yes
Oil and gas prices		
-2011	Heating Oil ∅ Natural Gas	No
-2000	Heating Oil ↔ Natural Gas	No
-2011	Heating Oil ∅ Natural Gas	No
-2011	Heating Oil ↔ Natural Gas	Yes
-2011	WTI ↔ Natural Gas	No
-2000	WTI ↔ Natural Gas	No
-2011	WTI ∅ Natural Gas	No
-2011	WTI ↔ Natural Gas	Yes

Note: ↔ indicates the presence of a cointegration relationship.
∅ indicates the absence of a cointegration relationship.
SB stands for 'Structural Break' analysis.

4.1 Petroleum products

4.1.1 WTI and Gasoil

Studying cointegration between the WTI crude oil futures price and Gasoil consists by definition in testing the stationarity of the crack spread. In Table 4, we verify that this hypothesis consistently holds during the full period (with or without break), as well as during the corresponding sub-periods.[14]

---

[14] Note that the choice of the two sub-periods 1993-2000 and 2000-2011 were chosen ex-ante in order to leave approximately the same number of datapoints in each sub-periods. Further exploration of statistical structural break tests goes beyond the scope of the present article.

Table 5: Cointegration Test for the 1993-2000 Sub Period, *Without* Structural Break, between WTI and Gasoil

1993-2000	Max. Eigen.	10%	5%	1%
r ≤ 1	5.76	10.49	12.25	16.26
r = 0	36.85	16.85	18.96	23.65

In Table 5, we only reproduce the results obtained for the first sub-period. This sub period was chosen among other possible results, since it provides us with the most satisfactory output concerning the estimation of the VECM. The remaining results can be obtained upon request to conserve space.[15] Table 5 clearly illustrates the fact that there exists a cointegration relationship between WTI and Gasoil during 1993-2000 (at the 1% level).

Table 6: VECM Results for the 1993-2000 Sub Period, W*ithout* Structural Break, between WTI and Gasoil

Error Correction Term		
WTI	1	
Gasoil	-0.900	
Trend	-0.001	
VECM	Δ WTI	Δ Gasoil
ECT	-0.028	0.038
(t.stat)	(-1.72)	(2.54)
Intercept	-0.046	0.064
(t.stat)	(-1.7)	(2.56)
Δ WTI(-1)	-0.034	0.294
(t.stat)	(-0.84)	(7.96)
Δ Gasoil(-1)	-0.004	-0.238
(t.stat)	(-0.09)	(-6.02)

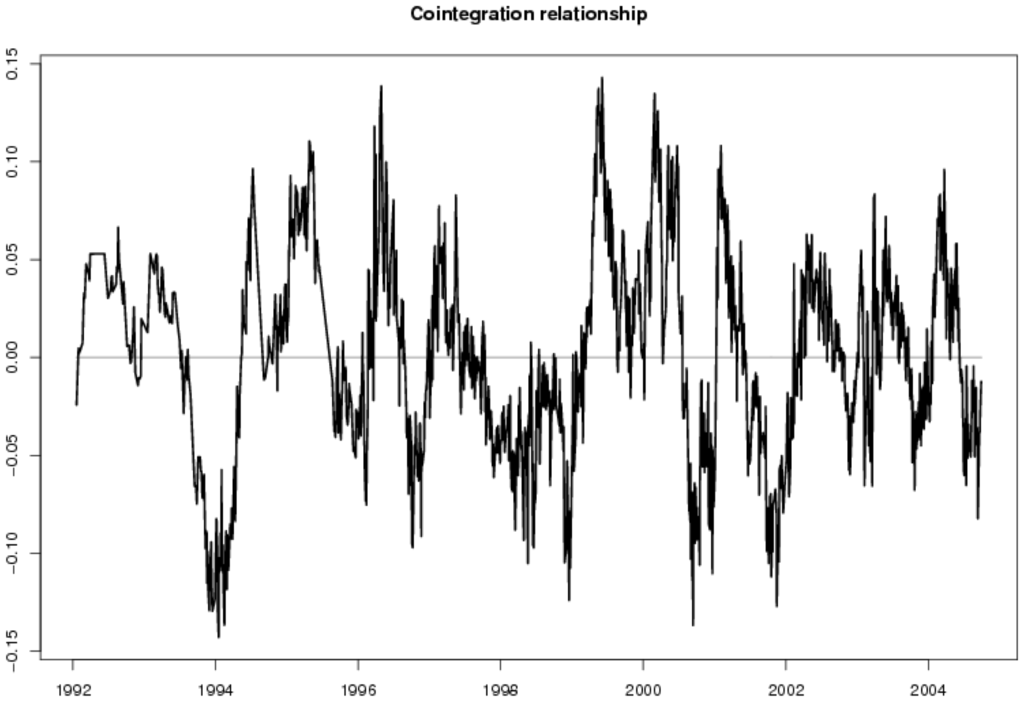
The corresponding VECM estimates can be found in Table 6. While both error correction terms are statistically significant, only the sign for the WTI variable is negative. Hence, the stationarity of the crack spread is made possible through the feedback effects coming from the WTI crude oil futures price in this system.

As a final diagnostic check, we can observe in Figure 2 that the cointegrating relationship between WTI and Gasoil is stable during 1993-2000. The same comments apply for the other graphs that could be produced out of this cointegration exercise between the two time series. Our results in this first category of petroleum products are line with the findings by Murat and Tokat (2009), who analyzed previously the crack spread over the period 2000-2008. Note that we were able to identify one structural break in this relationship on January 4, 2011.

---

[15] This comment applies in the remainder of this article.

Figure 2: Cointegration Relationship for the 1993-2000 Sub Period, *Without* Structural Break, between WTI and Gasoil



4.1.2 Brent and Gasoil

Similar results can be obtained when performing the cointegration analysis between Brent and Gasoil. Recall that Brent is mostly produced out of North European shores, while the WTI price is a world benchmark crude oil price delivered in Cushing, Oklahoma (USA). Apart from these geographical differences, the stationarity of the crack spread – defined as the difference between Brent and Gasoil here – should still hold. The inspection of Table 4 indeed allows us to validate this statement, and we can only underline the remarkable stability of this relationship across all specifications designed to test for cointegration in this article.

Table 7: Cointegration Test for the 1993-2000 Sub Period, *Without* structural break, between Brent and Gasoil

1993-2000	Max. Eigen.	10%	5%	1%
$r \leq 1$	5.82	10.49	12.25	16.26
$r = 0$	31.52	16.85	18.96	23.65

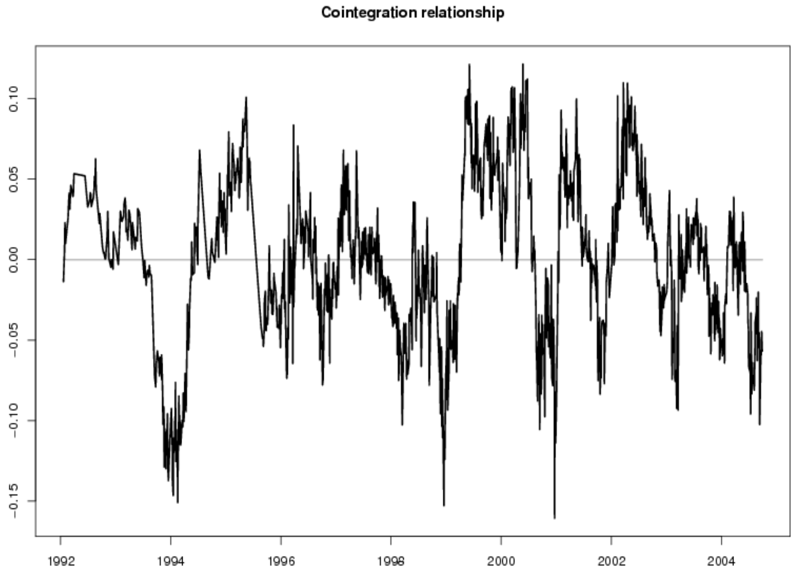
In Table 7, we report again the results for the first sub-period (which bring the best fit to the data when looking at the VECM). The rank of the cointegration r is at least equal to 1 (at the 1% significance level).

Table 8: VECM results for the 1993-2000 sub period, *without* structural break, between Brent and Gasoil

Error Correction Term		
Brent	1	
Gasoil	-0.933	
Trend	-9.976	
VECM	Δ Brent	Δ Gasoil
ECT	-0.015	0.044
(t.stat)	(-0.93)	(2.85)
Intercept	-0.028	0.084
(t.stat)	(-0.90)	(2.87)
Δ Brent(-1)	-0.045	0.332
(t.stat)	(-1.11)	(8.48)
Δ Gasoil(-1)	0.009	-0.261
(t.stat)	(0.22)	(-6.51)

In the VECM displayed in Table 8, only the error correction term for Gasoil is significant, but it is not positive. What concerns the ECT for Brent, we record as expected a negative sign, but it is not significant. Therefore, we cannot entirely validate this specification (despite the strong indication in favor of cointegration between Brent and Gasoil) due to this lack of significance coming from the error correction model. In that particular case, it is not guaranteed that the Brent price will be able to restore the long term equilibrium should deviations occur from either of the two time series in the short term.

Figure 3: Cointegration Relationship for the 1993-2000 Sub Period, W*ithout* structural break, between Brent and Gasoil



Finally, we can investigate the stationarity of the crack spread as pictured in Figure 3. The relationship appears very stable around the mean. However, we have discarded this model previously following the analysis of the VECM estimates. These results differ from the most recent analysis by Westgaard et al. (2011), who concluded in favor of the stationarity of the crack spread (defined as the difference between Brent and Gasoil) during 1994-2009. Note that, in another specification, we have identified a structural break on September 24, 2003.

4.1.3 WTI and Heating Oil

The third specification tested in the context of petroleum products consists of WTI and Heating Oil. According to Table 4, the same results apply here, as the validity of the cointegration model is accepted in all the cases considered.

Table 9: Cointegration Test for the 1993-2000 Sub Period, W*ithout* Structural Break, between WTI and Heating Oil

1993-2000	Max. Eigen.	10%	5%	1%
r ≤ 1	6.14	10.49	12.25	16.26
r = 0	36.38	16.85	18.96	23.65

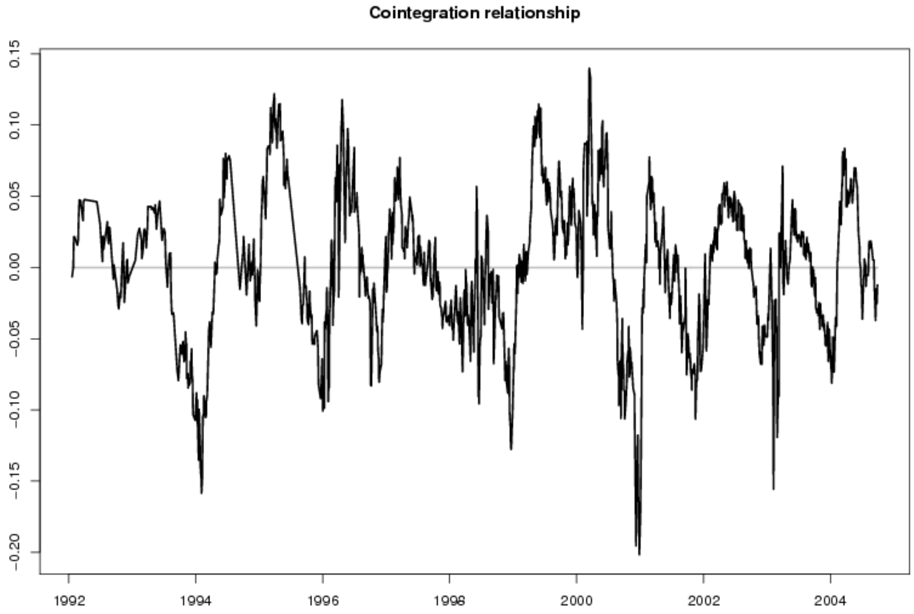
In Table 9, we verify readily the presence of one cointegrating relationship between WTI and Heating Oil (at the 1% level).

Table 10: VECM Results for the 1993-2000 Sub Period, W*ithout* Structural Break, between WTI and Heating Oil

Error Correction Term		
WTI	1	
Heating Oil	-0.928	
Trend	-8.161	
VECM	Δ WTI	Δ Heating Oil
ECT	-0.022	0.033
(t.stat)	(-1.40)	(2.06)
Intercept	-0.016	0.026
(t.stat)	(-1.34)	(2.11)
Δ WTI(-1)	-0.035	0.041
(t.stat)	(-0.72)	(0.82)
Δ Heating Oil (-1)	0.001	-0.075
(t.stat)	(-0.01)	(-1.52)

However, in Table 10, the VECM model estimated is not entirely satisfactory. Indeed, we observe that the error correction term for Heating Oil is significant but positive, while the ECT for WTI is negative but insignificant (i.e. it does not reach the 10% level). Therefore, we must reject the validity of the VECM model based on our observations.

Figure 4: Cointegration relationship for the 1993-2000 sub period, W*ithout* structural break, between WTI and Heating Oil



In Figure 4, we could not detect any instability in the cointegration relationship displayed between WTI and Heating Oil. Similarly to our previous case with Brent and Gasoil, we must dismiss the validity of cointegration relationship between WTI and Heating purely based on the results from the VECM. We agree on this point with the conclusions by Asche et al. (2003), who rejected the cointegration between Crude Oil and Heavy Fuel Oil (whose Heating Oil is one component) during the study period 1992-2000. Finally, note that a structural break could be detected here on January 4, 2011.

4.1.4 Brent and Heating Oil

According to Table 4, similar conclusions can be reached when replacing WTI with Brent in the cointegration exercise. Indeed, the hypothesis of cointegration between Brent and Heating Oil can be accepted in all the specifications tested here.

Table 11: Cointegration Test for the 1993-2000 Sub Period, W*ithout* structural break, between Brent and Heating Oil

1993-2000	Max. Eigen.	10%	5%	1%
$r \leq 1$	5.98	10.49	12.25	16.26
$r = 0$	30.84	16.85	18.96	23.65

Table 11 reflects the results obtained during the first sub-period, where the rank of the cointegration r is found to be equal to at least 1 (at the 1% level).

Table 12: VECM results for the 1993-2000 sub period, W*ithout* structural break, between Brent and Heating Oil

Error Correction Term		
Brent	1	
Heating Oil	-0.964	
Trend	-6.963	
VECM	Δ Brent	Δ Heating Oil
ECT	-0.011	0.037
(t.stat)	(-0.79)	(2.42)
Intercept	-0.010	0.037
(t.stat)	(-0.74)	(2.46)
Δ Brent(-1)	-0.053	0.043
(t.stat)	(-1.10)	(0.85)
Δ Heating Oil (-1)	0.017	-0.076
(t.stat)	(0.37)	(-1.57)

According to Table 12, we cannot accept anymore the validity of the cointegration hypothesis by looking at the VECM estimates. Indeed, the error correction term for Heating Oil is significant but positive. What concerns the Brent variable, the ECT is negative but not statistically significant. Hence, we cannot find evidence of any feedback mechanism to correct the deviations from the long run equilibrium in this system.

Figure 5: Cointegration Relationship for the 1993-2000 Sub Period, W*ithout* Structural Break, between Brent and Heating Oil

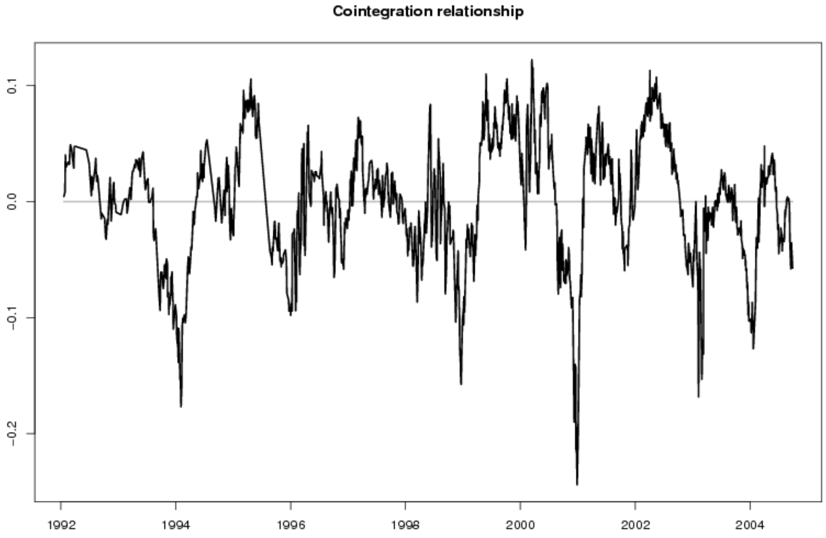


Figure 5 tells us that the relationship is quite stable over time. But the cointegration hypothesis must still be rejected, given our comments at the stage of the VECM. Our results are opposite to the recent findings by Moutinho et al. (2011), who concluded in favor of the cointegration hypothesis between Brent and Fuel Oil (among other fuels such as Gas and Coal). In another specification during the full period, note that we were able to isolate a structural break on July 15, 2008.

4.2 Oil and gas prices

4.2.1 Heating Oil and Natural Gas

We now focus our attention on Heating Oil and Natural Gas. Overall, the results contained in Table 4 are less satisfactory than for the previous category, since we are able to detect a cointegrating relationship in 2 of the 4 specifications tested only, i.e. during the first sub-period and during the full period with the modeling of one structural break.

Table 13: Cointegration Test for the 1993-2011 Full Period, with Structural Bbreak, between Heating Oil and Natural Gas

1993-2011	Max. Eigen.	10%	5%	1%
r ≤ 1	13.21	5.42	6.79	10.04
r = 0	29.38	13.78	15.83	19.85

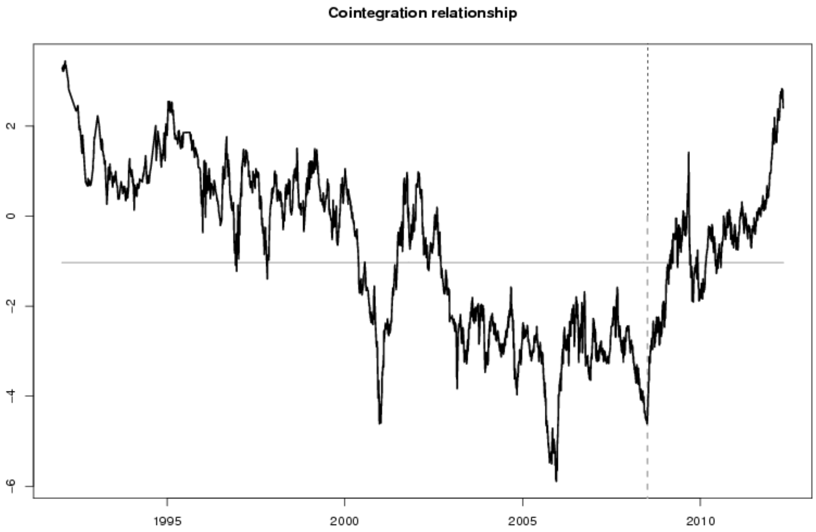
In Table 13, we have chosen to reproduce the results relative to the full period with one structural break. The null hypothesis of no cointegration ($r=0$) between Heating Oil and Natural Gas can be safely rejected at the 1% level.

Table 14: VECM Results for the 1993-2011 Full Period, with Structural Break, between Heating Oil and Natural Gas

Error Correction Term		
Heating Oil	1	
Natural Gas	-4.066	
VECM	Δ Heating Oil	Δ Natural Gas
ECT	0.001	0.003
(t.stat)	(2.75)	(3.77)
Δ Heating Oil (-1)	-0.060	-0.020
(t.stat)	(-2.91)	(-0.57)
Δ Natural Gas (-1)	0.037	-0.028
(t.stat)	(3.06)	(-1.38)

As shown in Table 14, this specification is not valid however given that both error correction terms are significant but positive.

Figure 6: Cointegration Relationship for the 1993-2011 Full Period, With Structural Break, between Heating Oil and Natural Gas



In Figure 6, we notice indeed that the stability of the cointegration relationship cannot be granted as easily as in the former category. There remains some areas for instability in the cointegration relationship between Heating Oil and Natural Gas before and even after the structural break detected on July 7, 2008. Overall, our results contradicts the previous findings by Serletis and Herbert (1999) who studied this relationship over the period 1996-1997.

4.2.2 WTI and Natural Gas

The next relationship under scrutiny consists of WTI and Natural Gas. In Table 4, our results point to the fact that the hypothesis of cointegration between Oil and Gas holds for most specifications, except during the second sub-period (2000-2011).

Table 15: Cointegration Test for the 1993-2011 Full Period, *Without* Structural Break, between WTI and Natural Gas

1993-2011	Max. Eigen.	10%	5%	1%
r ≤ 1	5.26	10.49	12.25	16.26
r = 0	17.53	16.85	18.96	23.65

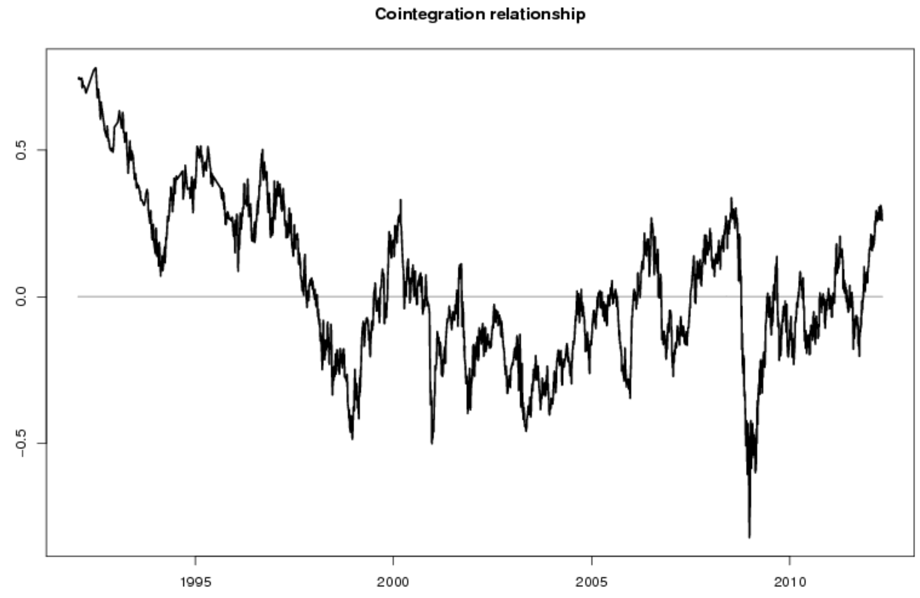
The results reproduced in Table 15 concern the 1993-2011 full period, without structural break. We infer from this Table that the rank of the cointegration r between WTI and Natural Gas is equal to at least 1 (at the 10% level).

The VECM provides better results in Table 16 than previously for Heating Oil and Natural Gas. Indeed, both error correction terms are significant. Moreover, the sign of the ECT for WTI is negative, which allows us to validate the model. The main implication of this result is that if deviations occur between Oil and Gas, they will be corrected by the fluctuations of the WTI variable in order to ensure the stationarity of the system.

Table 16: VECM Results for the 1993-2011 Full Period, W*ithout* structural break, between WTI and Natural Gas

Error Correction Term		
WTI	1	
Natural Gas	-0.454	
Trend	-0.001	
VECM	Δ WTI	Δ Natural Gas
ECT	-0.006	0.010
(t.stat)	(-2.45)	(2.61)
Intercept	0.013	- 0.020
(t.stat)	(2.56)	(-2.57)
Δ WTI(-1)	-0.050	- 0.033
(t.stat)	(-2.54)	(-1.08)
Δ Natural Gas(-1)	0.022	-0.030
(t.stat)	(1.74)	(-1.54)

Figure 7: Cointegration Relationship for the 1993-2011 Full Period, W*ithout* Structural Break, between WTI and Natural Gas



In Figure 7, we can detect some forms of non-stationarity at the beginning of the period (1993-1998), but the behavior of the cointegrating relationship is globally stable past that date. Based on this interpretation of the VECM, we confirm the validity of the cointegration hypothesis between WTI and Natural Gas, as identified previously by many authors (among others: Villar and Joutz (2006), Brown and Yücel (2008), Hartley et al. (2008)). In another specification, note that we have identified the presence of one structural break on July 7, 2008 (i.e. at the period of time corresponding to strong fluctuations in the WTI price series).

Note that we will not test the specifications by Bachmeier and Griffin (2006) who used several coal price series absent from our dataset, and who also focused on the issue of the Law of One Price between various qualities of crudes (which is deliberately left out of our research in the present article). Note also that, in the spirit of Panagiotidis and Rutledge (2007), the investigation of the cointegration relationship between Brent and Natural Gas cannot be performed in our setting. While these authors relied on the UK time series for Natural Gas (which makes sense to be compared with the European Brent time series), we have gathered in our dataset a US time series for Natural Gas (labelled Henry Hub).

5. CONCLUSION

This article examines the relationships between the prices of different forms of energy based on the cointegration tool with/without structural breaks. This topic has been studied in the previous academic literature, but there lacks a central and up-to-date body of knowledge to synthetize this information. Based on daily data from 1993 to 2011, we have divided our database of commodities into two groups: *(i)* petroleum products, and *(ii)* oil and gas prices.

Compared to the systematic reproduction of the results from previous literature, we document that there are more cross-commodity linkages at hand in energy markets than it is usually agreed upon among market practitioners. Even if the relationships at hand are not always stable, they do exist among the various groups of energy commodities investigated in our study. Besides, the interest of resorting to the cointegration technique with the explicit modeling of one structural break has been clearly underlined, compared to linear cointegration tools or sub-period decomposition only.

Among the variety of cointegration analyses developed in this article, it is noteworthy to remark that energy markets seem strongly inter-related. Indeed, 21 of the 24 specifications tested confirmed the presence of cointegration either between petroleum products, or between oil and gas prices. The core of the results implies that energy markets share common trends overtime, such as the worldwide demand for energy which has encountered an unprecedent growth over the last decades. Besides, energy sources offer the advantage of being substitutable technologically speaking, which allows producers and/or consumers to switch between their fuel inputs depending on the cheapest energy source. Our results globally point out the existence of shared trends among energy markets, despite the fact that idiosyncratic shocks (e.g. excess demand or supply, geopolitical events, etc.) can still affect one market in particular. Of all energy markets, the crude oil price (with the WTI as its benchmark) seems to be the leader in the price discovery process, since most of the time it triggers the adjustement towards the long term stationary equilibrium between the variables in the cointegration system.

In terms of economic implications, when cointegration is detected, then the long run stationary combination of the time series considered implies that idiosyncratic shocks will be corrected by feedback effects. The forces at stake in the error correction mechanisms can be related to substitutes and/or complementary relationships between pairs of energy prices. There are several reasons to explain why one might expect asymmetric responses from different markets in the short run. One is that the markets may have different access times to the information being delivered. Another is that the information may be interpreted differently initially. However, because the commodities trade on common trends, arbitrage opportunities between the markets would eventually result in a multi-market consensus concerning the value of new information. These characteristics are of primary importance for the consumers and producers of the commodity, but also to investment managers and traders who would lose valuable information by ignoring them.

REFERENCES

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

Asche, F., Osmundsen, P., Sandsmark, M. 2006. The UK market for natural gas, oil and electricity: Are the prices decoupled? *Energy Journal* 27(2):27-40

Ai, C., Chatrath, A., Song, F. 2006. On the comovement of commodity prices. *American Journal of Agricultural Economics* 88:574-588

Akram, QF. 2009. Commodity prices, interest rates and the dollar. *Energy Economics* 31:838-851

Bachmeier, LJ., Griffin, JM. 2006. Testing for Market Integration: Crude Oil, Coal, and Natural Gas. *The Energy Journal* 27(2):55-71

Bencivenga, C., Sargenti, G., D'Ecclesia, R. 2011. Integration of energy commodity markets in Europe and the USA. *Journal of Risk Management in Financial Institutions* 4(3):301-313

Brown, SPA., Yücel, MK. 2008. What drives natural gas prices? *Energy Journal* 29:45-60

Casassus, J., Liu, P., Tang, K. 2009. Commodity Prices in the Presence of Inter-commodity Equilibrium Relationships. *Working Paper SSRN* #1306022, Social Science Research Network, USA

Chaudhuri, K. 2001. Long-run prices of primary commodities and oil prices. *Applied Economics* 33:531-538

Cortazar, G., Milla, C., Severino, F. 2008. A multicommodity model of futures prices: Using futures prices of one commodity to estimate the stochastic process of another. *Journal of Futures Markets* 28:537-560

Dickey, DA., Fuller, WA. 1979. Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association* 74:427-431

Dickey, DA., Fuller, WA. 1981. Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root. *Econometrica* 49:1057-1072

Girma, PB., Paulson, AS. 1999. Risk arbitrage opportunities in petroleum futures spreads. *Journal of Futures Markets* 19:931-955

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

Greene, WH. 2011. *Econometric Analysis*. Prentice Hall, 7 edition, USA

Hamilton, JD. 1996. *Time Series Analysis*. Princeton University Press, USA

Hartley, PR., Medlock III, KB., Rosthal, J. 2008. The relationship of natural gas to oil prices. *Energy Journal* 29(3):47-65

Johansen, S., Juselius, K. 1990. Maximum likelihood estimation and the demand for money inference on cointegration with application. *Oxford Bulletin of Economics and Statistics* 52:169-210

Johansen, S. 1988. Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control* 12:231-254

Johansen, S. 1991. Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica* 59:1551-1580

Kwiatkowski, DP., Phillips, PCB., Schmidt, P., Shin, Y. 1992. Testing the Null Hypothesis of Stationnarity Against the Alternative of the Unit Root: How Sure are we that Economic Time Series are Non Stationary? *Journal of Econometrics* 54:159-178

Lanza, A., Manera, M., Giovannini, M. 2005. Modeling and forecasting cointegrated relationships among heavy oil and product prices. *Energy Economics* 27:831-848

Lütkepohl, H., Saikkonen, P., Trenkler, C. 2004. Testing for the cointegrating rank of a VAR with level shift an unknown time. *Econometrica* 72:647-662

Mjelde, JW., Bessler, DA. 2009. Market integration among electricity markets and their major fuel source markets. *Energy Economics* 31(3):482-491

Mohammadi, H. 2009. Electricity prices and fuel costs: Long-run relations and short-run dynamics. *Energy Economics* 31(3):503-509

Moutinho, V., Vieira, J., Carrizo-Moreira, A. 2011. The crucial relationship among energy commodity prices: Evidence from the Spanish electricity market. *Energy Policy* 39:5898-5908

Murat, A., Tokat, E. 2009. Forecasting oil price movements with crack spread futures. *Energy Economics* 31:85-90

Panagiotidis, T., Rutledge, E. 2007. Oil and gas markets in the UK: Evidence from a cointegrating approach. *Energy Economics* 29:329-347

Pfaff, B. 2008. *Analysis of Integrated and Cointegrated Time Series with* R. Springer, New York Dordrecht Heidelberg London, UK

Phillips, PCB., Peron, P. 1988. Testing for a unit root in time series regression. *Biometrika* 75:335-346

Serletis, A., Herbert, J. 1999. The message in North American energy prices. *Energy Economics* 21(5):471-483

Trenkler, C. 2003. A new set of critical values for systems cointegration tests with a prior adjustment for deterministic terms. *Economics Bulletin* 3:1-9

Villar, JA., Joutz, FL. 2006. The relationship between crude oil and natural gas prices. *Technical Report*, *Energy Information Administration*, Office of Oil and Natural Gas, Washington, DC, USA

Westgaard, S., Estenstad, M., Seim, M., Frydenberg, S. 2011. Co-integration of ICE Gas oil and Crude oil futures. *Energy Economics* 33:311-320

APPENDIX: COINTEGRATION METHODOLOGY

Cointegration without Structural Breaks

Cointegration can be seen as a useful econometric tool to decompose the long term trend between pairs (or groups) of variables, and the short-term departures from the trend. In the context of commodity markets, a cointegration relationship will tell us whether a pair (or a group) of individual commodities are tied together in the long run (which means that there exists a strong economic rationale to link these variables in the economic analysis), and to which extent exogenous perturbations from this equilibrium can occur.

Preliminary Conditions

As a pre-requisite condition for cointegration, the time series need to be integrated of the same order. For instance, the econometrician can check, based on standard stationarity tests, that the prices of the raw time series considered are non stationary and integrated of order one ($I(1)$). This amounts to checking that they are difference stationary[16].

In practice, the Augmented-Dickey-Fuller (1981, ADF) or Phillips-Peron (1988, PP) tests are used. Extensions of these stationarity tests were also developed by Kwiatkowski et al. (1992, KPSS). We apply these three tests in our article.

Johansen cointegration tests

To keep the notations parsimonious, let us consider here the cointegration setting with only two variables[17]. As is standard in a linear cointegration exercise, the econometrican needs to check first if the variables are cointegrated, i.e. if β exists such that $R_t = X_t^e - \beta X_t^{e'}$ is stationary. This can be done by performing an OLS regression of X_t^e on $X_t^{e'}$, or more rigorously by using the Johansen cointegration test (Johansen and Juselius (1990), Johansen (1991)).

Let X_t be a vector of N variables, all $I(1)$:

$$X_t = \Phi_1 X_{t-1} + \ldots + \Phi_p X_{t-p} + \varepsilon_t \quad (1)$$

with $\varepsilon_t : WGN(0,\Omega)$, WGN denotes the White Gaussian Noise, Ω denotes the variance covariance matrix, and Φ_i $(i = 1,\ldots,p)$ are parameter matrices of size $(N \times N)$.

Under the null H_0, there exists r cointegration relationships between N variables, *i.e.* X_t is cointegrated with rank r.

---

[16] Stationarity is a central concern in time series analysis, which implies that the mean of the variable shall be time invariant (in the weak sense of stationarity). See Hamilton (1996) for further reference.

[17] Note however that the Johansen cointegration framework can be generalized to k variables.

Note that the Johansen cointegration tests can be performed on the logarithmic transformation of the time series under consideration.

For a financial modeling viewpoint, if we find that commodities are cointegrated, i.e. that there exists a stationary combination of these variables in the long term, the direct implication would be that they share at least one common risk factor in the long term. Hence, their joint analysis can bring fruitful results to the econometrician.

Error-correction Model

The next step of the cointegration model consists in describing the dynamics of the variables in terms of the residuals of the long-term relation (Johansen (1988)). We want to introduce an error-correction mechanism on the levels and on the slopes between the variables e and e':

$$\begin{pmatrix}\Delta X_t^e \\ \Delta X_t^{e'}\end{pmatrix} = \begin{pmatrix}\mu_e \\ \mu_{e'}\end{pmatrix} + \sum_{k=1}^{p}\Gamma_k \begin{pmatrix}\Delta X_{t-k}^e \\ \Delta X_{t-k}^{e'}\end{pmatrix} + \begin{pmatrix}\Pi_e \\ \Pi_{e'}\end{pmatrix} R_t + \begin{pmatrix}\varepsilon_t^e \\ \varepsilon_t^{e'}\end{pmatrix} \qquad (2)$$

where

- e stands for the first variable, and e' stands for the second variable;
- X_t^e is the log price of variable e at time t;
- the 2×1 vector process $\Delta Z_t = \left(\Delta X_t^e = X_{t+1}^e - X_t^e, \Delta X_t^{e'} = X_{t+1}^{e'} - X_t^{e'}\right)$ is the vector of the variables price returns;
- $\mu = (\mu_{X,e}, \mu_{X,e'})$ is the 1×2 vector composed of the constant part of the drifts;
- Γ_k are 2×2 matrices of real valued parameters expressing dependence on lagged returns;
- $(R_t = X_t^e - \beta X_t^{e'})$ is the process composed of the deviations to the long-term relation between the variables log prices;
- Π is a 2×1 vector matrix expressing the sensitivity to the deviations to the long-term relation between the variables prices;
- the residual shocks $(\varepsilon_t^e, \varepsilon_t^{e'})$ are assumed to be i.i.d with a centered bi-variate normal distribution $N(0, \sum)$.

However, by considering a purely linear model, it is possible that the econometrician will either misspecify the model, or ignore a valid cointegration relationship. That is why we detail below the cointegration methodology with an unknown structural break.

Cointegration with Structural Breaks

In this section, we explore the possibility of wrongly accepting a cointegration relationship, when some of the underlying time series are contaminated by a structural break. For instance, sharp deviations from the long-term trend can occur between a group of commodities, which would imply that the cointegration relationship is not valid anymore during specific sub-samples. The structural breakpoint detection allows to take into account these events in the cointegration analysis, instead of simply ignoring them.

We present the procedure for estimating a vector error-correction model (VECM) with a structural shift in the level of the process, as developed by Lütkepohl et al. (2004). By doing so, we draw on the notations by Pfaff (2008).

Framework

Let $\vec{y}_t$ be a $K \times 1$ vector process generated by a constant, a linear trend, and level shift terms[18]:

$$\vec{y}_t = \vec{\mu}_0 + \vec{\mu}_1 t + \vec{\delta} d_{t\tau} + \vec{x}_t \quad (3)$$

with $d_{t\tau}$ a dummy variable which takes the value of one when $t \geq \tau$, and zero otherwise. The shift point τ is unknown, and is expressed as a fixed fraction of the sample size:

$$\tau = [T\lambda], \quad 0 < \underline{\lambda} \leq \lambda \leq \overline{\lambda} < 1 \quad (4)$$

where $\underline{\lambda}$ and $\overline{\lambda}$ define real numbers, and $[\cdot]$ the integer part. Therefore, the shift cannot occur at the very beginning or the very end of the sample. The estimation of the structural shift is based on the regressions:

$$\vec{y}_t = \vec{v}_0 + \vec{v}_1 t + \vec{\delta} d_{t\tau} + \vec{A}_1 \vec{y}_{t-1} + \ldots + \vec{A}_1 \vec{y}_{t-p} + \varepsilon_{t\tau}, \quad t = p+1, \ldots, T \quad (5)$$

with $\vec{A}_i, \quad i = 1, \ldots, p$ the $K \times K$ coefficient matrices, and ε_t the white noise K-dimensional error process. The estimator for the breakpoint is defined as:

$$\hat{\tau} = \arg \min_{\tau \in \mathrm{T}} \det \left(\sum_{t=p+1}^{T} \hat{\vec{\varepsilon}}_{t\tau} \hat{\vec{\varepsilon}}'_{t\tau} \right) (6)$$

---

[18] Note that Lütkepohl et al. (2004) develop their analysis in the context where $\vec{x}_t$ can be represented as a VAR(p), whose components are at most $I(1)$ and cointegrated with rank r .

with $\mathrm{T} = [T\underline{\lambda}, T\overline{\lambda}]$, and $\hat{\vec{\varepsilon}}_{t\tau}$ the least squares residuals of Eq. (5). Once the breakpoint $\hat{\tau}$ has been estimated, the data are adjusted as follows:

$$\hat{\vec{x}}_t = \vec{y}_t - \hat{\vec{\mu}}_0 - \hat{\vec{\mu}}_1 t + \hat{\vec{\delta}} d_{t\hat{\tau}} \quad (7)$$

The test statistic writes:

$$LR(r) = T \sum_{j=r+1}^{N} \ln(1 + \hat{\lambda}_j) \quad (8)$$

with corresponding critical values found in Trenkler (2003).

Estimation of the VECM

The error-correction model (ECM) writes:

$$\Delta X_t = \Pi_1 \Delta X_{t-1} + \ldots + \Pi_{p-1} \Delta X_{t-p+1} + \Pi_p X_{t-p} + \varepsilon_t \ (9)$$

where the matrices Π_i $(i = 1, \ldots, p)$ are of size $(N \times N)$. All variables are $I(0)$, except X_{t-p} which is $I(1)$. For all variables to be $I(0)$, $\Pi_p X_{t-p}$ needs to be $I(0)$ as well.

Let $\Pi_p = -\beta\alpha^{'}$, where $\alpha^{'}$ is an (r, N) matrix which contains r cointegration vectors, and β is an (N, r) matrix which contains the weights associated with each vector. If there exists r cointegration relationships, then $Rk(\Pi_p) = r$. Johansen's cointegration tests are based on this condition. We can thus rewrite Eq.(9):

$$\Delta X_t = \Pi_1 \Delta X_{t-1} + \ldots + \Pi_{p-1} \Delta X_{t-p+1} - \beta\alpha^{'} X_{t-p} + \varepsilon_t \quad (10)$$

The estimation of the corresponding vector error-correction model (VECM) is performed through maximum likelihood methods (Johansen and Juselius (1990), Johansen (1991)).