Three Essays on Oil Markets and the Energy Consumption-Economic Growth Nexus

Elham Talebbeydokhti

Head of Ph.D. Program: Prof. Antonio Nicolo'
Supervisor: Prof. Fulvio Fontini

Department of Economics and Management
University of Padua
Cycle: XXX

ACKNOWLEDGEMENTS

Foremost, I would like to express my sincere gratitude to my advisor Professor Fulvio Fontini for the continuous support of my Ph.D study and research, for his patience, motivation, enthusiasm, and knowledge. His guidance helped me in all the time of research and writing of this thesis.

My sincere thanks also goes to Professor Massimiliano Caporin for his encouragement and insightful comments. I appreciate all his contributions of time and ideas to make my Ph.D. experience productive and stimulating.

I would also like to express my special thanks to Professor Muhammad Shahbaz, who provided me an opportunity to join his team as a visiting Ph.D. student in Montpellier Business School, and who gave me support to widen my research from various perspectives.

I thank my fiancé Francesco for the discussions, for the sleepless nights we were working together before deadlines, and for all the fun we have had in the last three years.

Last but not the least, I would like to thank my family: my father, mother and brother for supporting me spiritually throughout the writing of this thesis and my life in general. Without their love and support it would not be possible to conduct this research.

Contents

1 Introduction **5**

2 Chapter 1: Testing Persistence of WTI and Brent Long-run Relationship after the Shale oil Supply Shock **9**

2.1 Introduction 11

2.2 Methodology 15

2.2.1 Cointegration test with structural break 16

2.2.2 Vector error correction model 17

2.3 Data Analysis 19

2.4 Empirical Analysis 20

2.5 Conclusions 25

3 Chapter 2 : Assessing Income, Urbanisation and Industrialisation Impacts on Energy Consumption: Where's the EKC in an Open Economy **26**

3.1 Introduction 28

3.2 Literature Review 29

3.2.1 Industrialisation and Urbanisation 29

3.2.2 Income, Economic Growth 30

3.2.3 Trade openness 31

3.3 Data description and methodological framework 32

3.3.1 Cointegration methodology 33

3.3.2 ARDL bounds tests cointegration 33

3.3.3 The VECM Granger causality approach 34

3.4 Empirical results and discussion 35

3.5 Conclusion and policy implications 41

4 Chapter 3 : Resiliency and Asymmetric Reaction to Price Changes of Shale Oil Rig Counts 42

4.1 Introduction 44

4.2 Literature review 46

4.3 Data description 47

4.4 Research methodology 48

4.5 Research methodology 50

5 References 51

List of Figures

1 Figure 1: WTI, Brent monthly Prices and their Spread 60

2 Figure 2: US Crude oil and Shale oil Production 60

3 Figure 3: US Export of Crude oil 60

4 Figure 4: Impulse Response Function - Full-Sample 61

5 Figure 5: Impulse Response Function - First Sub-Sample 62

6 Figure 6: Impulse Response Function - Second Sub-Sample 63

7 Figure 7: Trends of the variables from 1960 to 2015 64

8 Figure 8: $CUSUM$ and $CUSUM_{SQ}$ test 65

9 Figure 9: Levels and the changes of the variables 66

List of Tables

1 Table 1: Descriptive analyses of monthly growth rates 67

2 Table 2: Perron unit root test 67

3 Table 3: Gregory - Hansen cointegration test results 67

4 Table 4: Full-sample cointegration estimation 68

5 Table 5: Selected sub-samples cointegration estimation 69

6 Table 6: Descriptive Statistics and Correlation Matrix 70

7 Table 7: Unit Root Analysis 70

8 Table 8: The Results of ARDL Cointegration Test 71

9 Table 9: Long Run Analysis 72

10 Table 10: Short Run Analysis 73

11 Table 11. VECM Granger Causality Analysis 74

1 Introduction

This thesis includes three essays which empirically deals with topics in energy economics. My contribution to papers regards the definition of the research question, the analysis of literature review, the design and implementation of the empirical strategy, the elaboration of data and the econometric estimates. The first paper, "Testing Persistence of WTI and Brent Long-run Relationship after the Shale oil Supply Shock" is a joint paper with Prof. Fontini and Prof. Caporin.

This paper mainly focus on impact of the fast-rising shale oil supplies on the spread between two main crude oil benchmarks. US oil market undergone a significant transformation with the unexpectedly strong rise in the US production of shale oil. The application of two technological innovations, horizontal drilling and hydraulic fracturing or fracking have enabled the US to grow dramatically the production of abundant shale oil resources. Shale oil, which is sometimes also known as tight oil or light tight oil (LTO) is petroleum that contains of light crude oil with low sulfur content, found in some rock formation deep below the earth's surface. Rapid production growth in shale oil had dramatic effects on US domestic oil price (WTI), which has decoupled from global indices (Brent). Historically, Brent and WTI crude oil prices tracked closely, with a price-difference per barrel between ±3 USD/bbl, with WTI usually priced higher. At the beginning of 2011, this historical relationship collapsed.From a global perspective, the rapid rise of US shale oil has been the main driver behind the increase in non-OPEC supply. The first chapter investigate what has been so far and what probably might be in the future the impact of shale oil production in the long-run WTI-Brent price relation- ship. More precisely, we investigate if we can statistically confirm that there was a long-run relationship in the WTI - Brent oil time series before the rise of shale oil production in the US market; if a structural break has occurred in the long-run relationship, and if so, when that has occurred and if it has coincided with the rise of the shale oil production; if, after the entrance in the market of abundant shale oil, a new long-run relationship has emerged between WTI and Brent prices, and if so of which kind. We also test the dynamics of the long-run relationships, namely, the

relative impact of the changes of one series on the other series; on the convergence to the equilibrium, if any; and by how much this has been influenced by the shale oil production rise. We use monthly data of WTI and Brent crude oil prices, as well as US shale oil quantities from January 2000 to November 2017. The empirical results of the cointegration test taking a structural break into account show that the structural break occurs in February 2011. We estimate a Vector Error Correction Model (VECM), considering the structural break suggested by the cointegration test results and the timing of the rise in shale oil production, and compare full-sample analysis with sub-sample estimates of the VECM models. Our analysis reveals that WTI and Brent crude oil prices have had a long-run relationship up to 2011 but there is no longer such a long-run relationship after the rise of the shale oil production, even though, on the basis of the available data, it is not unequivocally possible to assess whether the absence of a new long-run relationship is permanent or not in the new shale oil period. Also short-run dynamics are analysed through the use of Impulse Response Functions (IRF). Moreover, we assess the impact on WTI and Brent prices of the shale oil production.

The second paper, "Assessing Income, Urbanisation and Industrialisation Impacts on Energy Consumption: Where's the EKC in an Open Economy" is co-authored with Prof. Shahbaz and Prof. Md. Al Mamun.

The paper primarily focuses on the effect of energy price, economic growth, urbanisation, industrialisation and trade openness on energy consumption. Thus, the key dependent variable is energy consumption while other variables are the key explanatory or treatment variables. We investigate the environmental Kuznets curve (EKC) hypothesis between economic growth and energy consumption by including, energy prices, urbanisation, industrialisation and trade openness in energy consumption function for the United Kingdom annual data from1960 to 2015. In order to evaluate the EKC, the squared term of economic growth, urbanisation, industrialisation and trade openness is included in the set of explanatory variables. The EKC is said to exist for energy consumption if economic growth, urbanisation, industrialisation and trade openness is positively signed and economic growth, urbanisation, industrialisation and trade

openness squared has a negative coefficient. First, the traditional as well as structural breaks unit root tests are applied in order to examine the stationary properties of the variables. To validate the presence of cointegration between energy consumption and its determinants, we applied ARDL bounds testing approach to cointegration by accommodating structural breaks in the series. Finally, the VECM Granger causality approach has been employed to determine direction of causal relationship between the variables. The empirical results indicate existence of cointegration between the variables. Our results show that energy prices are negatively linked with energy consumption. Income is positively linked with energy consumption. Industrialisation and trade openness add in energy consumption but urbanisation declines it. The nexus between urbanisation- energy and trade-energy validate the presence of inverted U-shaped relationship i.e. EKC effect. The relationship between industrialisation- energy and income-energy consumption is U-shaped.

The research proposal, "Resiliency and Asymmetric Reaction to Price Changes of Shale Oil Rig Counts" is being developed together with Prof. Fontini, Prof. Caporin. We are still at the preliminary stage of the paper. We have a clear definition of research question and we already collect the necessary data for the empirical estimates. Even though, we define the baseline methodology, we need to run alternative tests with different methodologies, in order to increase the robustness of the analysis and investigate other model specification.

In this paper, we explore the relationship between shale oil rig count and US crude oil price. Since US shale revolution stimulated tremendous oil and gas production, the number of US rig count became widely publicised. A drilling rig is a machine that creates holes in the earth sub-surface in order to drill a new well to explore for, develop and produce oil. Recently, US oil market has been characterised by fluctuating WTI prices and this makes difficult for oil producers to determine the profitability in exploration and development. Thus, it is wise to take a closer look at the relationship between oil rig count and crude oil prices. From 2011 onward, the relative importance of shale oil production over the total US oil supply has been significantly increasing so it is reasonable to focus on the relationship between shale rig count

and crude oil prices in particular. Therefore, We split the total US rig count into shale rig count and non-shale rig count. We analyse how shale rig count and non-shale rig count and their production in US is affected by the changes in oil price while accounting for other determinants of this relationship. We also studied if this relationship is asymmetric for rises and drops of oil price. We test for the often claimed hysteresis hypothesis of the shale production in the case of crude oil price drops. The most relevant variables we use are weekly data on WTI price, the shale rig count, non-shale rig count, rig productivity and a set of potentially relevant economic and financial control variables from February 2011 up to October 2017 for total of 344 observations. This relationship is of significant interest to analysts, investors, oil companies, commercial banks, investment banks and policy makers.

Chapter 1

Testing Persistence of WTI and Brent Long-run Relationship after the Shale oil Supply Shock *

Massimiliano Caporin[†1], Fulvio Fontini[‡2], Elham Talebbeydokhti[§2]

[1]Department of Statistical Sciences, University of Padua, Italy
[2]Department of Economics and Management, University of Padua, Italy

Abstract

At the beginning of 2011, the spread between WTI and Brent crude oil prices, that was always contained in a range of roughly three dollars per barrel, rose dramatically and then felt again. This could indicate that something had intervened influencing a stable relationship between WTI and Brent oil price. It could be the result of the significant increase in the US shale oil production. This paper investigates if this was the case. More precisely, we test if there was a long-run relationship in the WTI - Brent oil time series; if a structural break has occurred in it; if this has been influenced by the rise in the US shale oil production, and if, after the boom of shale oil, a new long-run relationship has emerged and of which kind. We do so using monthly data of WTI and Brent crude oil prices, as well as US shale oil quantities from January 2000 to November 2017. The empirical results of the cointegration test taking a structural break into account show that the structural break occurs in February 2011. We estimate a Vector Error Correction Model (VECM), considering the structural break suggested by the cointegration test results and the timing of the rise in shale oil production, and compare full-sample analysis with sub-sample estimates of the VECM models. Our analysis reveals that WTI and Brent crude oil prices have had a long-run relationship up to 2011 but there is no longer such a long-run relationship after the rise of the shale oil production, even though, on the basis of the available data, it is not unequivocally possible to assess whether the absence of a new long-run relationship is permanent or not in the new shale oil period. Also short-run dynamics are analysed through the use of Impulse Response Functions (IRF). Moreover, we assess the impact on WTI and Brent prices of the shale oil production.

Keywords: WTI, Brent, Shale oil, Cointegration, Vector Error Correction Model, Impulse Response Function.

---

*This paper has been presented at the first AIEE Energy Symposium, Current and Future challenges to Energy Security Conference held in Milan, Italy, 30 November - 2 December 2016.

†massimiliano.caporin@unipd.it
‡fulvio.fontini@unipd.it
§elham.talebbeydokhti@phd.unipd.it

2.1 Introduction

Crude oil is not a homogenous product. There are various types of internationally traded crude oils with different qualities and characteristics. The quality of crude oil differentiates in terms of American Petroleum Institute (API) gravity formula which measures the oil's density (light to heavy) and its acidity and sulfur content (sweet referring to low-sulfur content and sour to high-sulfur content). The two dominant oil reference prices are the West Texas Intermediate (WTI) and the Brent crude oil. Brent is an European benchmark related to oil extracted from North Sea. Brent is composed of four crude blends: Brent, Forties, Oseberg, and Ekofisk (BFOE). The Brent and Forties blends are produced offshore in the waters of the United Kingdom, and the Ekofisk and Oseberg blends are mainly produced offshore in the waters of Norway (Energy Information Administration, EIA (2016)). Brent is traded on the Intercontinental Exchange (ICE) for delivery at Sullom Voe. In recent years, Brent has been used to price two third of the world's internationally traded crude oil supplies. WTI is the main oil benchmark in Americas (Fattouh, 2010). WTI refers to oil extracted from wells in the US and sent via pipeline to Cushing, Oklahoma. For over three decades, Cushing has been a major oil supply hub connecting oil suppliers to the Gulf Coast and therefore the price settlement point for WTI. WTI is traded on the New York Mercantile Exchange (NYMEX) for delivery at Cushing. WTI and Brent crude are classified as sweet light crude oil, sweet because of their low sulfur content and light because of their relatively low density, that making them ideal for the refining of diesel fuel, gasoline and other high demand products. However, Brent crude is not as sweet and light as WTI.

The difference in prices among the various types of oil can be supposed to depend on arbitrage rationale, that could be influenced by physical characteristics of the various crude oils, their short term variations in supply and demand, and market sentiment toward oil futures contracts. Many papers study crude oil price differentials, eg., Fattouh (2007, 2010), Kao and Wan (2012) and Borenstein *et al.* (2014). Historically, Brent and WTI crude oil prices tracked closely, with a price-difference per barrel between ±3 USD/bbl, with WTI usually priced higher

than Brent as show in Figure 1. This pattern of the price differential has probably denoted that arbitrage opportunities were kept at a minimum, with the price premium of WTI depending on the sweeter content of it and its rather stable behaviour denoting that WTI and Brent had been subject to a long-run equilibrium relationship. In early 2011, however, this longstanding relationship began to change, and Brent crude oil started to be priced much higher than WTI. Since WTI-priced stockpiles at Cushing could not easily be transported to the Gulf Coast, WTI crude was unable to be arbitraged in bringing the two prices back to parity and the price spread rose dramatically. A sign of the transportation difficulties of WTI was given by the evidence that the price of other US oil prices such as Louisiana Light Sweet (LLS) (priced at coastal areas of the US) were closer to Brent than WTI. In late 2014, the Brent-WTI spread has narrowed considerably. The spread, which was more than 27USD/bbl in Sep 2011, fell again to 3USD/bbl in February 2016. This massive volatility suggests something more than physical constrains in transportation have mattered. Several factors could have been the cause of this widening, such as oversupply of crude oil production in North America, dollar currency movements, variation in regional demand, slow economic rebound in North America and even politics. At the same time, Brent moved up in reaction to Libyan civil war, to civil unrest in Egypt and across the Middle East. Also the depletion of the North Sea oil fields might have helped increase the price of Brent.

During the same period the US oil industry was undergoing a major change. The application of two technological innovations, horizontal drilling and hydraulic fracturing or fracking have enabled the US to grow dramatically the production of abundant shale oil resources (See Figure 2). Shale oil, which is sometimes also known as tight oil or light tight oil (LTO) [1] is petroleum that contains of light crude oil with low sulfur content, found in some rock formation deep below the earth's surface. It is conventional oil trapped in unconventional formation of low

[1]We use here the U.S. Energy Information Administration (EIA) data. EIA " [...] has adopted the convention of using the term tight oil to refer to all resources, reserves, and production associated with low-permeability formations that produce oil, including that associated with shale formations." However, EIA acknowledges that "The oil and natural gas industry's colloquial use of the term tight oil is rather recent, and does not have a specific technical, scientific, or geologic definition." (Source: https://www.eia.gov/energy_in_brief/article/shale_in_the_united_states.cfm). For this reason and for coherence with the rest of the literature we shall use the term shale oil instead of tight oil.

permeability. The multi-stage hydraulic fracturing causes cracks in the rock formation that allow the crude oil, deposited within open spaces in the rock, to flow into well-bore. Most of this growth of the US shale oil has come from the following basins: Bakken in Montana and Dakota, Eagle Ford and Permian Basin in Texas. Rapid production growth in shale oil had dramatic effects on US domestic oil price, which has decoupled from global indices. The discrepancy between WTI and Brent prices might have resulted also from other factors, such as the rapid expansion of US shale oil production combined with limited pipeline capacity to move crude oil from production fields to storage locations, including land-locked delivery point of Cushing, Oklahoma to refining centres, and the longstanding US ban on exporting crude oil. Both might have put downward pressure on the price of WTI crude oil. The US oil export ban was signed into law since the 1975 Arab oil embargo, in an attempt to insulate the US from foreign oil price shock. Increase in US crude oil production mainly due to shale oil, rise of imports of heavy Western Canadian crude extracted from oil sands in Alberta and Saskatchewan, and US oil export ban might have determined an internal excess supply, pushing down domestic prices more than foreign ones. Of the three factors mentioned above, two have reduced their importance overtime. Since mid-2012, there has been considerable investment in expanding US crude oil infrastructure and new pipelines has been added at Cushing. In 2014, US allowed export of a type of minimally processed ultra-light oil as displayed in Figure 3. In December 2015, congress voted to lift the 40-year-old ban on crude oil export. However, the shale oil technology can be regarded as a permanent innovation, and shale oil production is still present. Indeed it has started to slightly decline only from late 2015, perhaps as a response to the dramatic oil price drop.

The purpose of this paper is to investigate what has been so far and what probably might be in the future the impact of shale oil production in the long-run WTI-Brent price relationship. More precisely, we investigate if we can statistically confirm that there was a long-run relationship in the WTI - Brent oil time series before the rise of shale oil production in the US market; if a structural break has occurred in the long-run relationship, and if so, when that has occurred and if it has coincided with the rise of the shale oil production; if, after the entrance

in the market of abundant shale oil, a new long-run relationship has emerged between WTI and Brent prices, and if so of which kind. We also test the dynamics of the long-run relationships, namely, the relative impact of the changes of one series on the other series; on the convergence to the equilibrium, if any; and by how much this has been influenced by the shale oil production rise.

Evidence from previous studies investigating whether WTI and Brent crude oil market is integrated is mixed. Reboredo (2011) suggest that crude oil prices are linked with the same intensity during bull and bear markets, thus supporting the hypothesis that oil market is "one great pool", in contrast with the hypothesis that states that oil market is regionalised. Gülen (1999) argues that market segmentation generates market inefficiencies and gives rise to arbitrage opportunities. Likewise, oil market co-movements have potentially important implications for portfolio allocation and hedging strategies involving spot and future oil contracts. Hammoudeh *et al.* (2008) and Wilmot (2013) supports the idea of the globalisation hypothesis. on the other hand, Wiener (1991) find out "the world oil market is far from completely united". Kim et al. (2013) have find that long-run relationships among WTI, Brent and Dubai crude oil prices hold during 1997M01 to 2012M07, even when the effects of the breaks are considered. An empirical study by Liu *et al.* (2015) found that the oil supply constraints at Cushing significantly affect the decreasing levels of cointegration between WTI and Brent. They also found that WTI dominates the price discovery process when WTI and Brent are cointegrated. Chen *et al.* (2015) use rigorous econometric methods to investigate the structural change in persistence of WTI - Brent spreads, i.e., a change from a stationary to a non-stationary time series. The CUSUM of the squares-based test of Leybourne *et al.* (2007), in which the breakpoint is not pre-specified, is undertaken and the estimated breakpoint is found to have occurred in 2010.

Aruga *et al.* (2015) investigate that WTI no longer have a long-run relationship with the Brent and Dubai crude oil markets. Büyükşahin *et al.* (2013) find structural break in the long-term relationship between WTI and Brent occur in 2008 and 2010. To the best of our knowledge,

there is no study in the literature that test statistically the long-run relationship between WTI and Brent considering shale oil quantity; at most, the relationship is conjectured in a narrative based on data patterns description. We aim at closing this gap, using cointegration and VECM methodologies to provide a statistically sound analysis of it.

The paper is structured as follows: Section 2 outlines the methodologies used for the cointegration analysis. Section 3 describes the data. Section 4 provides all the empirical results. Finally, section 5 contains the conclusions, and references follow.

2.2 Methodology

From a statistical viewpoint, energy commodities price time series behaves like time series of financial instruments' prices, and are usually non-stationary. Therefore, a simple OLS regression between two commodity prices, would potentially lead to spurious regression. Such an event do not realise if the series of interest share a common trend, and are linearly related by an equilibrium condition. In order to deal with this issue, we first assess whether the variables under consideration are non-stationary, or characterised by a unit root, or integrated of order one, $I(1)$. We might employ traditional unit root tests, like the Augmented Dickey-Fuller test, but we are aware that our time series might be contaminated by a structural break, due to the shale oil supply shock. Therefore, we need to take into account the possible existence of a break in each time series, either in the intercept, or in the trend or in both intercept and the trend. Consequently, we apply the Perron (1997) unit root test that allows for a break at an unknown location, and allow to evaluate both the null hypothesis of integration as well as to identify the break location.

If the series are integrated, despite possibly affected by a structural break, we are allowed to test for the existence of cointegration, that is, the existence of a long-run relationship between the series of interest. If the series are cointegrated there exist a linear combination of the

series which is stationary, or not characterised by unit root, or integrated of order zero, $I(0)$. We follow here the Gregory and Hansen (1996) cointegration analysis as again, we have to deal with the possible presence of structural breaks. If the test show evidences in favour of cointegration, we can specify a Vector Error Correction Model (VECM) to identify both the long-run relationships between the variables of interest and, at the same time, accounting for the structural break. The model also allows to verify if the long-run disequilibrium lead to an adjustment process to the prices. Impulse Response Function (IRF) has also been generated to explain the response to shale oil supply shock amongst the oil price variables. We now provide a few details on the approaches we follow.

2.2.1 Cointegration test with structural break

Gregory and Hansen (1996) introduced a cointegration test that allows for possible structural breaks. Four models can be applied in order to test cointegration according to the type of structural change that might realise:

Model 1: Standard cointegration

$$y_{1t} = \mu + \alpha^T y_{2t} + e_t, \quad t = 1, ..., n \tag{1}$$

where y_{1t} and y_{2t} are $I(1)$ and e_t is $I(0)$. This is the baseline model, representing the case where we do not have structural breaks. To introduce the latter, that is to allow the model to take into account a regime shift either in μ or α, we start by defining the dummy variable:

$$\phi_{t\tau} = \begin{cases} 0 & if \quad t \leq [n\tau] \\ 1 & if \quad t > [n\tau] \end{cases}$$

where the unknown parameter $\tau \in (0, 1)$ denotes the timing of the structural change. We then introduce the three models with structural change.

Model 2: Level shift (C)

$$y_{1t} = \mu_1 + \mu_2 \phi_{t\tau} + \alpha^T y_{2t} + e_t, \quad t = 1, ..., n \tag{2}$$

where μ_1 is the intercept before the shift, μ_2 is the change in the intercept at the time of the shift.

Model 3: Level shift with trend (C/T)

$$y_{1t} = \mu_1 + \mu_2 \phi_{t\tau} + \beta t + \alpha^T y_{2t} + e_t, \quad t = 1, ..., n \tag{3}$$

where we introduce a time trend into the level shift model.

Model 4: Regime shift (C/S)

$$y_{1t} = \mu_1 + \mu_2 \phi_{t\tau} + \alpha^T y_{2t} + \alpha_2 y_{2t} \phi_{t\tau} + e_t, \quad t = 1, ..., n \tag{4}$$

here we allow for a structural break in the cointegrating vector α as well as in the intercept μ.

In all four models, the test of the null hypothesis of no cointegration is residual-based. In other words, we can regard y_{1t} and y_{2t} as being cointegrated by examining whether the residuals e_t does not have a unit root. Gregory and Hansen (1996) constructed three statistics for those test: Z_a^*, Z_t^* and ADF^*. The test statistics Z_a^* and Z_t^* build upon the Philips-Perron test statistics, while ADF^* is based on the augmented Dickey-Fuller (ADF) statistics. The null hypothesis is rejected if the statistic, either ADF^*, Z_a^* or Z_t^*, is smaller than the corresponding critical value. In all cases, the critical values are tabulated as the test statistics do not follow a standard density.

2.2.2 Vector error correction model

A Vector Error Correction Model (VECM) is a restricted Vector Auto Regressive (VAR) model designed for use with non-stationary series that are known to be cointegrated. VECM

can lead to a better understanding of the nature of long-run and short-run relationships among the modelled variables. In our case, the model allows evaluating the existence of long-run relationships across three variables, namely, WTI and Brent price and shale oil quantity. In fact, we resort to the tests of Johansen (1988, 1995) and determine the number of cointegrating relations. In its most general representation, the VECM model takes the following form:

$$\Delta X_t = \Pi X_{t-1} + \sum_{j=1}^{p} \phi_j \Delta X_{t-j} + \delta D_t + \epsilon_t, \tag{5}$$

where X_t is the vector of the modelled variables (in logs), Δ identifies the first difference of the variables (i.e. the growth rates), the summation monitors the short-run dynamics of the series growth rates, and, finally, D_t contains a set of deterministic variables, namely a constant and a linear trend.

In the presence of cointegrating relationships, as identified by the Johansen (1988, 1995) Trace and Maximum Eigenvalue tests, the matrix Π might be decomposed as $\Pi = \alpha\beta'$, where the product $\beta' X_{t-1}$ contains disequilibrium errors, the matrix β provides the cointegration coefficient, i.e. the coefficient of the long-run relationship between variables, and the vector α contains the adjustment coefficients to past disequilibrium. In other words, the matrix β indicates the coefficients that link the variables of interest in the so-called long-run relation, while the parameters α show the impact of disequilibrium on the short-run dynamic, and provides also a measure of the speed of adjustment back to equilibrium.

In our setting, we expect the existence of a unique long-run (cointegrating) relationship. This single equilibrium relation would link the two prices and the shale oil quantity would act as the driver of the diverging behaviour among them. In fact, assuming a stable, over time, the link between Brent and WTI prices would read as:

$$WTI_t - \mu - \beta_1 Brent_t - \beta_2 ShaleQ_t = \epsilon_t, \tag{6}$$

where WTI_t denotes the WTI real log-price, $Brent_t$ is the Brent real log- price, $ShaleQ_t$ is

the shale oil total quantity (in logs), and ϵ_t is the stationary error term. If the two oil prices have a stable and long-run link, we expect the coefficient β_1 to be statistically significant and close to one, suggesting that movements in the Brent (WTI) price are replicated in the WTI (Brent) price. Moreover, shocks to one of the two prices might lead to adjustments toward the equilibrium by means of the VECM structure, as governed by the adjustment coefficients and the lag structure. With the surge of shale oil production, even the shale oil quantity might play a role in the long-run equilibrium relation.

Once the VECM has been estimated, short-run dynamics and the adjustment due to disequilibrium in the long-run relationship can be examined by considering the impulse response functions (IRF). These functions measure the time profile of the effect of a shock, or impulse on the (endogenous) variables of interest. A crucial element is given by the approach we might choose to orthogonalize the shocks. We stick to the most common approach, i.e. the use of a Cholesky decomposition of the innovation's covariance of the VECM model. In that case the variables ordering becomes fundamental. We thus set the variable order as follows: Brent price, WTI price and shale oil quantity. We set Brent as the first variable because Brent has more liquid market than WTI and might be safely taken as a reference price. Furthermore, we are interested in verifying the effect on the shale oil quantity of price disequilibriums. By setting the shale oil quantity as the last variable, we do not allow for structural shocks on the shale oil quantity to directly impact on the price equations. The effect of shale oil quantities shocks will go through the error correction term and the dynamic interaction among the growth rates of the variables (i.e. the lags). In the empirical analysis we focus on IRF up to 24 lags (two years) and compute bootstrap confidence intervals.

2.3 Data Analysis

We use the monthly Real Spot Prices of WTI and Brent Crude Oil (Dollars per barrel) and monthly Tight oil quantities and US crude oil production (Thousand barrels per day) from US Energy Information Administration (EIA). The observation period ranges from January 2000 to

November 2017, which yield a sample size of 215 observations. Table 1 gives summary statistics for the monthly growth rates of crude oil prices and quantities. Considering the significant rise of shale oil production which starts at February 2011, we divide the full-sample period into two sub-samples. The growth rate of the two prices are, on average, negative in the second sub-sample, while the US crude oil production (WTI quantity) increases from February 2011 onward, mostly due to the increase in shale oil production.

We then assess the stationarity property of WTI, Brent prices and shale oil quantity before performing the analysis of cointegration. The proper break dates can be identified by looking at the result of the Perron test for each model and each time series that has the lowest test statistics. The results display in Table 2 shows that the three series of WTI, Brent prices and shale oil quantity are non-stationary and they have break. The lowest test statistics correspond always to the model with break in intercept. Interestingly enough, the break in shale oil quantity occurs in February 2011 when the production of shale oil started to rise. We check also that all time series are stationary at their first differences (data is not reported in the paper).

2.4 Empirical Analysis

Since the unit root test results suggest that the WTI, Brent prices and shale oil quantity are $I(1)$, first, we evaluate the possible presence of cointegration between WTI and Brent prices in full-sample using Gregory and Hansen (1996) methodology. Then, we replicate the test including also shale oil quantity. Assuming that the timing of a structural change is unknown, Gregory and Hansen (1996) suggest a cointegration test that allows a structural change in the cointegration vector. In the case of conflicting results, the conclusions are based on the Z_t^* statistic which is described by Gregory and Hansen (1996) as the most powerful of all three statistics considered. We estimate all three models that assess type of structural change (C, C/T, C/S), as explained in section 2.1. The optimal lag is selected based on estimates of Akaike Information Criterion (AIC). The test results are presented in Table 3. When we consider cointegration between WTI and Brent prices, statistics point out to the existence of

a cointegration relationship, with a structural break occurring in 2010M10 and 2011M02. Also in the case of WTI, Brent prices and shale oil quantity, the three statistics defiantly reject the null hypothesis of no cointegration in the models. In this case, the break occur in 2011M02 and 2013M04. Overall, the results of the tests indicate that structural changes have occurred between 2010M10 and 2013M04. This confirms the time identification of the break as coinciding with the rise in the shale oil production, from early 2011 onward. The discrepancies in the time lags can be attributed to a slow impact of the shale oil on the cointegration relationships or to anticipation effect of the expectation of the shale oil rise.

Having identified the timing of the structural break, we can test for the existence of a long-run relationship between the two prices as well as among the three considered variables. We test for the presence of cointegration following the approach of Johansen *et al.* (2000). We introduce a break in the form of a step dummy assuming value 1 from February 2011, and interacting the dummy with the cointegration equation intercept and trend. We stress that the specification of the deterministic terms for both the WTI and Brent relationship and the WTI, Brent and shale oil quantity one, when using the full-sample of data, is intercept (no trend) in cointegrating equation and intercepts in VAR. This is the most common specification for trending series like the ones we used here. Furthermore, the appropriate maximum lag length has been chosen base on the Schwarz Information Criterion (SIC), which ensure that the VAR is well-specified. In the presence of serial correlation in the residuals, we increase the lag until any autocorrelation issue is resolved.

Results are reported in Table 4. The upper part of the table describes the Johansen (1988, 1995) Trace test and Max. Eigenvalue. The first column in the table refer to the WTI-Brent relationship and the second one refer WTI, Brent and shale oil. For both cases, tests indicates the presence of cointegration when accounting for a break in the linear trend between the two series at the 5% significant level. The lower part of the table reports the results of the VECM model. Evidence claims for the existence of a strong long-run relationship between WTI and Brent prices, in both cases with and without shale oil. The Brent price coefficient is almost

equal to 1 and highly significant. This means that a given percentage change in Brent price is coupled with an almost equivalent to the percentage change in WTI price. When we look at the first column, the equation adjustment coefficient for WTI price is negative, -0.52, and significant, while the Brent price equation adjustment coefficient is not statistically significant. This indicates that when there is some external force that drives the two series away, it is WTI price that react closing the gap and bringing down the price time series to converge to their long-run equilibrium with the Brent one, while Brent price does not react to disequilibrium. This is as expected, given that Brent markets is much more liquid than WTI one and Brent is priced worldwide while WTI in US only. In the second column, empirical evidence again suggests the presents of a single cointegration relationship among them, i.e the presence of a long-run equilibrium. We notice that the Brent price coefficient remains almost unchanged and highly significant compared to the previous one. The shale oil quantity is positive and slightly significant at 10% level. The sign of the coefficient is coherent with the economic rationale of WTI market: a rise in shale oil quantity implies an increase in WTI supply and therefore reduces its price. The low level of significance will be explained further below. The equation adjustment coefficient for WTI price is negative and significant, and the Brent one is not significant, as before, thus confirming the same behaviour. The adjustment coefficient of the shale oil quantity is not significant, showing that the production of shale oil in the full-sample period did not respond to changes in the long-run relationship between WTI and Brent prices.

We then move to the analysis of the impulse response functions. Figure 4 reports the latter for full-sample. The response of the two crude oil prices to a shock to shale oil supply has the expected negative sign. The effectiveness of shale oil supply shock is significant, persistently negative, and decreasing up to 8 periods. Afterwards, both crude oil prices reach their new long-run equilibrium. Moreover, we observe that a shock on the Brent price has an effect of comparable size on both the Brent and WTI prices, as expected given the cointegration. We note that a shock on the WTI price is not impacting on the Brent. This is a consequence of the Cholesky ordering adopted and of the limited dynamic interdependence among the two prices. Notably, a positive shock on the WTI price impacts on the shale oil quantity, leading to an

increase in the production.

We then take into account the existence of a break in the structural relation, as suggested by the Gregory-Hansen test results and the abrupt increase of shale oil production in 2011, in order to assess the role played by the rise of shale oil production before and after the structural break. Therefore, we replicate the analysis run for the full-sample period splitting it into two sub-samples. We test for the existence of cointegration relationships in the first sub-sample from January 2000 up to January 2011, then for the second sub-sample from February 2011 up to November 2017, between WTI and Brent prices, and also between WTI, Brent prices and shale oil quantity. Through this process, we can ascertain whether the long-run relationship between WTI and Brent prices has changed after a structural break. We tested several choices of the deterministic parameters and selected the one that provided the best results. For the case of WTI and Brent prices only in both sub-samples, we allow for the presence of a intercepts in both the cointegrating equation and the VAR and exclude the trend in the cointegration equation. On the contrary, in order to test the existence of a long-run relationship between WTI, Brent and shale oil quantity analysis, we introduce a linear trend in the cointegrating equation and the intercepts in both the cointegrating equation and the VAR.

Table 5 displays the results. When looking at the first sub-sample between WTI and Brent prices and between WTI, Brent prices and shale oil quantity, first column and third column in Table 5, Trace and Max. Eigenvalues indicate that cointegration exists among the variables. The result also confirm the existence of long-run relationship between variables. The Brent price coefficient is close to the coefficient of the full-sample analyses and highly significant. The shale oil quantity is positive and significant. The figure of the β_1 for the two models with and without shale oil are almost equivalent, which is as expected, given the limited quantity of shale oil before 2011. The equation adjustment coefficient for WTI price is negative and significant in both cases, while the Brent one is not. Note that now the shale oil adjustment coefficient is positive and significant; even if shale oil quantity for the first sub period was very limited and therefore its impact on long run relationship is modest, nevertheless the fact that

it is positive in sign implies it had for the first sub-sample the potentiality to push away the time series of WTI and Brent from their long run equilibrium. Effectively, this is what we observed in the second sub-sample. The boom of shale oil quantity has been so strong to break up the long run relationship: second and forth column in Table 5 show that there is no longer cointegration. Thus, the sub-sample evidence confirms that the relationship between WTI and Brent prices has been affected by rise of shale oil production, in the sense that there is no long-run equilibrium in the second sub-sample.

Figure 5 displays the impulse responses of all variables in first sub-sample used in VECM. The positive shale oil supply shock have negative response for the crude oil prices, however, the effects are not statistically significant. We interpret it as the shale oil production was very limited in first period thus we do not expect a big impact on the crude oil prices. The other impulse response functions are comparable to those in Figure 4, confirming the strong relation between the two prices and the impact of WTI shocks on the shale oil quantity. For the second sub-sample shown in Figure 6, prior to estimation, we take a first-difference of all three variables. We choose VAR in first differences because all the variables of interest in second sub-sample are non-stationary and are not cointegrated. Initially, the response of both crude oil prices to shale oil supply shock is significant and decreasing. After one period this phenomenon increase but its effect fades away after two periods. Furthermore, we note that we observe a reaction of WTI price to Brent price shocks but not the opposite, coherently with the first sub-sample and the full-sample analysis. WTI price react to Brent price due to potential arbitrage opportunities for crude oil market participants. Interestingly, the shale oil production is not reacting to crude oil prices shocks. That could be caused by the short sub-sample size. It could also indicates a high resilience of the shale oil industry to temporary price shocks. We leave the investigation on the reaction of the shale industry to price changes to further studies.

2.5 Conclusions

This study analyses the long-run relationship between WTI and Brent before and after the rise in shale oil production. We analyse the monthly spot prices of WTI and Brent, which are benchmarks for the North America, Europe and Eurasia, respectively, from January 2000 to November 2017. We apply the Perron unit root test in the presence of a structural break, and show that all the considered time series are non-stationary at their level and that increase in shale oil production has determined a structural break in 2011. To investigate whether the time series are integrated with a structural break, we apply the Gregory-Hansen test. We find that there exist cointegration relationships between WTI and Brent as well as between WTI, Brent and shale oil, with a structural break which occurs at February 2011. We then test the long-run behaviour applying VECM. First, we demonstrate the existence of a positive long-run relationship between WTI and Brent prices in the full-sample, with or without shale oil quantity. Then, we replicate the analysis by splitting the data into two sub-samples, based on Gregory-Hansen test result and before and after increase in shale oil production. The empirical results demonstrate that in the period, January 2000 - January 2011, the long-run relationship between WTI and Brent still exists and there is a unit value. This means that a given percentage change in Brent price correspond to the same percentage change in WTI price. Moreover, the analysis of the adjustment coefficients indicates that it is WTI price reacts after an external shock and bringing back the two series to their long run relationship and not the Brent price. In the period, February 2011 - November 2017, however, there is no longer a long-run relationship between time series. Lack of cointegration when focusing on the most recent years seems to suggest that the instability created by the rise of shale oil production has not yet been fully recovered by the market. No new long-run relationship has emerged. However, on the basis of this analysis it cannot be assessed if the lack of the cointegration relationship is a permanent feature of the WTI - Brent relationship or if on the contrary it is just due to a yet too short time period. In order to see if a new long-term equilibrium is reached by the market, further research should proceed on longer samples.

Chapter 2

Assessing Income, Urbanisation and Industrialisation Impacts on Energy Consumption: Where's the EKC in an Open Economy

Elham Talebbeydokhti*[1], Muhammad Shahbaz†[2], Md. Al Mamun‡[3]

[1]Department of Economics and Management, University of Padua, Italy
[2]Energy and Sustainable Development, Montpellier Business School, France
[3]Department of Economics and Finance, La Trobe University, Australia

Abstract

This paper investigates the environmental Kuznets curve (EKC) hypothesis between economic growth and energy consumption by including, energy prices, urbanisation, industrialisation and trade openness in energy consumption function for the United Kingdom. In doing so, traditional as well as structural breaks unit root tests are applied in order to examine the stationary properties of the variables. To validate the presence of cointegration between energy consumption and its determinants, we applied ARDL bounds testing approach to cointegration by accommodating structural breaks in the series. Finally, the VECM Granger causality approach has been applied to determine direction of causal relationship between the variables. The empirical results indicate existence of cointegration between the variables. Our results show that energy prices are negatively linked with energy consumption. Income is positively linked with energy consumption. Industrialisation and trade openness add in energy consumption but urbanisation declines it. The nexus between urbanisation-energy and trade-energy validate the presence of inverted U-shaped relationship i.e. EKC effect. The relationship between industrialisation-energy and income-energy consumption is U-shaped.

Keywords: Income, Urbanisation, Industrialisation, Energy consumption, Trade Openness, ARDL, EKC hypothesis.

*elham.talebbeydokhti@phd.unipd.it
†m.shahbaz@montpellier-bs.com
‡mamunacademic@gmail.com

3.1 Introduction

Industrialisation leads to urbanisation by constructing economic growth as well as job opportunities that draw population shift from rural to urban areas and from agricultural to industrial employment. This structural transformation of the economy causes many fundamental changes in energy uses. As urbanisation and industrialisation are growing rapidly, there are questions as to how these two attributes of modernisation will impact energy consumption. The history of developed countries shows that economic growth is a process of urbanisation and industrialisation. Most studies argue that urbanisation and industrialisation increase energy consumption (Jones (1991), Zhang and Lin (2012), Solarin and Shahbaz (2013)). On the contrary, some researchers find the relationship is negative (Wang (2014), Sathaye *et al.* (1994)). This inconsistent and conflicting results is partly due to differences in methods and sample data, but also reflecting different impacts of urbanisation and industrialisation on energy consumption at different development stages. Rapid industrialisation, urbanisation population growth and trade expansion have increased the demand for energy consumption.

An essential component of economic growth is trade openness and increase in international trade that increases the economic activities and the energy demand (Sadorsky (2012)). Trade openness enables developing countries to import of goods, services and advanced technologies from developed countries. Shahbaz *et al.* (2014b) explores the relationship between trade openness and energy consumption using data of 91 high, middle and low income countries for period of 1980 - 2010. They found an inverted U-shaped in high income countries but U-shaped in middle and low income countries between the trade openness and energy consumption. Furthermore, empirical research on energy had placed little attention on the existence of an U-shaped relationship between income and energy consumption. This study aims to investigate the effects of urbanisation, industrialisation and economic growth on energy consumption. We considered economic activity, industrialisation and financial development as additional determinants in understanding the nexus between variables and energy consumption. The paper also seeks to validate the existence of cointegration relationship between energy consumption and its deter-

minants. Finally, the paper examines the causal relationship between variables.

The structure of the rest of the article organised as follows. Section 2 presents the literature review in detail. Section 3 provides the empirical model, data and methodology. Section 4 reports and discuss empirical results. Finally, Section 5 concludes the article and provides policy implications.

3.2 Literature Review

The theory of energy use comes from the hypothesis of dematerialisation in which it is in line with reduction in material and energy consumption along the path of economic growth. Recalde and Ramos-Martin (2012) reveals that dematerialisation supports the Environmental Kuznets Curve (EKC) which assumes the existence of an inverted-U shaped relationship between energy growth and environmental degradation. The EKC hypothesis claims that environmental degradation increases with economic activity up to a turning point after which income increases tend to improve environmental quality. Therefore, EKC hypothesis is based on the concept of intensity of energy use, implying that energy intensity can be explained by economic growth (income). Here, we emphasise some other factors that can prominently influence intensity.

3.2.1 Industrialisation and Urbanisation

The industrial revolution had a significant effect on urbanisation. Industrialisation creates economic growth that increase the demand for the improved education and public works agencies that are characteristic of urban areas. Since businesses are looking for new technology in order to increase their productivity, they require an educated workforce. This attract skilled workers to the area for pleasant living conditions and seeking economic opportunities. While this process of modernisation is associated with increases in income and living standards, there are also important impacts on energy uses. Increased industrial activity uses more energy than traditional agriculture does. An interesting number of studies investigate the impact of urban-

isation and industrialisation on energy consumption. Typically, empirical literatures found a positive relationship between urbanisation, industrialisation and energy demand. Poumanyvong and Kaneko (2010) shows a positive relationship but insignificant between urbanisation and energy use in a balanced panel dataset for 99 countries by using the STIRPAT (Stochastic Impacts by Regression on Population, Affluence, and Technology) model, while industrialisation positively affects energy use. York (2007) establish a positive relationship between variables for 14 EU countries. Cole and Neumayer (2004) found a positively associated and a U-shaped relationship between urbanisation and energy demand. Although many studies show that higher urbanisation rate might lead to higher energy use, some empirical researchers finds the relationship is negative. Mishra *et al.* (2009) stated a negative impact on energy consumption in New Caledonia, but a positive relation in Fiji, French Polynesia, Samoa and Tonga. l**?** also found that it might exert a negative influence on energy consumption of 17 developing countries over the period 1960 - 2005.

3.2.2 Income, Economic Growth

A considerable number of studies have been conducted to support a direct association between energy consumption and economic growth. Kraft and Kraft (1978) examined the causal link between energy consumption and economic growth for the USA. He confirmed a unidirectional causality relation from GNP to energy consumption for the period 1947 - 1974. Akarca and Long (1980), Yu and Choi (1985), Abosedra and Baghestani (1989) found a positive relationship between energy uses and economic growth. Contrarily, some studies found a conflicting results about the direction of causality (Noor and Siddiqi (2010), Asghar (2008) and Amirat and Bouri (2010)). The direction of causality in energy-growth nexus plays an important role in policy designation issues such as energy production, consumption, distribution, economic growth, trade and many other factors. Yu and Choi (1985) used cross-country analysis to examine the casual relationship between national income and different forms of energy consumption. Parikh and Shukla (1995) used a pooled data set on developed and

developing countries from 1965 - 1987 to investigate the influence of economic growth on energy consumption. The authors found some evidence of a non-linear relationship between income and energy consumption for the Fixed Effect specification but insignificance evidence of this relationship in the Random Effects specification. Kahsai *et al.* (2012) investigated the long-run relationship between total energy consumption and economic growth for a panel of 19 African countries. First, they use Choi (2001), Levin and Lin (1993) and Hadri (2000) panel unit root tests in order to test the integrating properties of real GDP and total energy consumption. Second, by using Pedroni (1999) panel cointegration approach indicating that both variables are cointegrated for long-run relationship. Also they confirm that economic growth is the cause of energy consumption in the long run as well as in short run.

3.2.3 Trade openness

Trade is a basic economic concept involving the export and imports of goods and services. The crucial factors that increase world trade are energy prices, trade liberalisation, and improvement in transportation technology. Rapid increase in trade leads to productivity gain for domestic firms due to the presence of foreign firms. Accordingly, this productivity gain influence energy consumption. There are few studies investigating the relationship between trade openness and energy consumption. For instance, Cole (2006) studied the relationship between trade liberalisation and energy consumption over 32 countries. He obtained trade liberalisation build up economic growth in which benefits energy demand. Moreover, he found that greater openness to trade in these developing countries had increased energy consumption during the period 1975 and 1995. Sadorsky (2011) empirically investigated the casual relationship between energy consumption, trade openness and economic growth in 8 Middle Eastern countries. The empirical evidence reported an existence of a long-run relationship between the variables of interests. Moreover, he founds that exports Granger cause energy consumption and in short-run a reaction is found between imports and energy consumption. Sadorsky (2012) confirms the long-run relationship between trade and energy consumption by considering a sample of 7

South American countries. Subsequently, Fisher-Vanden *et al.* (2004) examined that rising in energy prices devoted significantly to the decline of firm-level energy consumption. Similarly, Shahbaz (2012) demonstrate the impact of trade openness on income in long-run by using cointegration, Granger causality tests, and the innovative accounting approach. This conclusion is in line with that of Shahbaz *et al.* (2014a), Narayan and Smyth (2009), Dedeoğlu and Kaya (2013) and Raza *et al.* (2015).

3.3 Data description and methodological framework

Let E_t, P_t, Y_t, U_t, I_t and O_t represent energy consumption (kg f oil equivalent), oil price, real GDP, urban population, industry value added, trade openness respectively after natural logarithmic transformation. Data used are annual from 1960 to 2015, taken from World Development Indicators; each in per capita terms. (Figure 7)

The paper primarily focuses on the effect of energy price, economic growth, urbanisation, industrialisation and trade openness on energy consumption. Thus, the key dependent variable is energy consumption while other variables are the key explanatory or treatment variables. The long run relationship between variables is tested using the following linear logarithmic functional form:

$$\ln E_t = \alpha_0 + \beta_{01} \ln P_t + \beta_{02} \ln Y_t + \beta_{03} \ln U_t + \beta_{04} \ln I_t + \beta_{05} \ln O_t + \epsilon_t \qquad t = 1, 2, ..., n$$

where, t is the time period and ϵ_t is the error term and assumed to be normally distributed. In order to evaluate the EKC, the squared term of economic growth, urbanisation, industrialisation and trade openness is included in the set of explanatory variables. The EKC is said to exist for energy consumption if economic growth, urbanisation, industrialisation and trade openness is positively signed and economic growth, urbanisation, industrialisation and trade openness squared has a negative coefficient.

3.3.1 Cointegration methodology

In the past two decades, after the seminal work of Zivot and Andrews (2002) cointegration techniques, Engle and Granger (1987) and Johansen and Juselius (1990) have been extensively used in empirical research to examine the long run relationship of economic variables in a bivariate or multivariate framework. However, the empirical exercise to investigate cointegration between the variables in these models become invalid if any variable is integrated at $I(0)$ in the VAR system or mixed order of integration of the variables. Pesaran *et al.* (2001) introduced Autoregressive-Distributed lag (ARDL) bounds tests approach for cointegration. One of the main benefits of ARDL bounds tests procedure is that it can be applied regardless of whether the underlying variables are stationary i.e. $I(0)$, integrated of order one i.e. $I(1)$ or fractionally integrated i.e. $I(0)/I(1)$. Second, the long-run and short-run parameters of the model in question can be estimated simultaneously. Third, ARDL bounds testing approach to cointegration performs better than all conventional cointegration approaches for small sample data while investigating the cointegration between the series. The critical values are easily available for small data to compare with our calculated F-statistics.

3.3.2 ARDL bounds tests cointegration

An ARDL model is a general dynamic specification, which uses the lags of the dependent variable and the lagged and contemporaneous values of the independent variables, through which the short-run effects can be directly estimated, and the long-run equilibrium relationship can be indirectly estimated. ARDL technique involves estimating unrestricted error correction model. An ARDL representation of equation is given as follows:

$$\Delta \ln E_t = \alpha_0 + \alpha_D D + \Sigma_{i=1} b_{1i} \Delta \ln E_{t-1} + \Sigma_{i=1} b_{2i} \Delta \ln P_{t-1} + \Sigma_{i=1} b_{3i} \Delta \ln Y_{t-1} + \Sigma_{i=1} b_{4i} \Delta \ln U_{t-1} + \Sigma_{i=1} b_{5i} \Delta \ln I_{t-1} + \Sigma_{i=1} b_{6i} \Delta \ln O_{t-1} + \sigma_1 \ln E_{t-1} + \sigma_2 \ln P_{t-1} + \sigma_3 \ln Y_{t-1} + \sigma_4 \ln U_{t-1} + \sigma_5 \ln I_{t-1} + \sigma_6 \ln O_{t-1} + \epsilon_{1t}$$

where, Δ is difference operator and D indicates the structural break point based on findings

of ADF unit root test with structural break. F-test is used to examine whether a cointegrating relationship exists among the variables. The null hypothesis of no cointegration among the variables in equation (2), $H_0 : \sigma_1 = \sigma_2 = \sigma_3 = \sigma_4 = \sigma_5 = \sigma_6 = 0$ against the alternative $H_1 : \sigma_1 \neq \sigma_2 \neq \sigma_3 \neq \sigma_4 \neq \sigma_5 \neq \sigma_6 \neq 0$, which is denoted as $F_E(E_t, P_t, Y_t, U_t, I_t, O_t)$. If computed F-statistics exceeds the upper critical bound (UCB) then the series are cointegrated. If the computed F-statistic lies below the lower critical value (LCB), no cointegration exist. If the computed F-statistics falls between the UCB and LCB, the test is uncertain. Moreover, the parameter stability is checked by applying the $CUSUM$ and $CUSUM_{SQ}$ tests proposed by Brown *et al.* (1975).

3.3.3 The VECM Granger causality approach

The next step is to determine the direction of casual relationship after the validation of cointegration between the variables. If the cointegration exist between series of integrated order one, $I(1)$, Granger (1969) suggest that there must be causality (in Granger sense) relation at least running from one side. Moreover, Granger (1969) argued that the presence of cointegration between the variables leads us to determine the short run as well as long run causal relationship. For instance, the concept of Grager causality reveals that Granger causality from E to Y if and only if, the changes in Y are predicted by the past values of E and similarly, Y Granger cause E if and only, the past values of Y predict the deviation in E. It is exposed by Granger (1969) to apply the Vector Error Correction Model (VECM) if the series are integrated at $I(1)$. The empirical equation of the VECM Granger causality is modelled as follows:

$$(1-L)\begin{bmatrix} \ln\ E_t \\ \ln\ P_t \\ \ln\ Y_t \\ \ln\ U_t \\ \ln\ I_t \\ \ln\ O_t \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \end{bmatrix} + \sum_{i=1}^{P}(1-L)\begin{bmatrix} b_{11i}\ b_{12i}\ b_{13i}\ b_{14i}\ b_{15i}\ b_{16i} \\ b_{21i}\ b_{22i}\ b_{23i}\ b_{24i}\ b_{25i}\ b_{26i} \\ b_{31i}\ b_{32i}\ b_{33i}\ b_{34i}\ b_{35i}\ b_{36i} \\ b_{41i}\ b_{42i}\ b_{43i}\ b_{44i}\ b_{45i}\ b_{46i} \\ b_{51i}\ b_{52i}\ b_{53i}\ b_{54i}\ b_{55i}\ b_{56i} \\ b_{61i}\ b_{62i}\ b_{63i}\ b_{64i}\ b_{65i}\ b_{66i} \end{bmatrix}$$

$$\times \begin{bmatrix} \ln\ E_{t-1} \\ \ln\ P_{t-1} \\ \ln\ Y_{t-1} \\ \ln\ U_{t-1} \\ \ln\ I_{t-1} \\ \ln\ O_{t-1} \end{bmatrix} + \begin{bmatrix} \alpha \\ \beta \\ \gamma \\ \delta \\ \theta \\ \vartheta \end{bmatrix} ECM_{t-1} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \\ \varepsilon_{4t} \\ \varepsilon_{5t} \\ \varepsilon_{6t} \end{bmatrix}$$

where, $(1-L)$ denotes the difference operator. The lagged error term i.e. EMC_{t-1} is obtain by using the long run ARDL relationship. The error terms, $\varepsilon_{1t}, \varepsilon_{2t}, \varepsilon_{3t}, \varepsilon_{4t}, \varepsilon_{5t}$ and ε_{6t} are assumed to have normal distributions with zero mean and constant variance i.e. $N(0, \sigma)$. The statistical significance of the coefficient of lagged error term ECM_{t-1} suggests the long run causal relationship between the variables. The statistical significance of the first differences of the series confirms the nature of the short run causal relationship. Additionally, joint long and short runs causal relationship can be estimated by joint significance of both ECM_{t-1} and the estimate of lagged independent variables. For instance, $b_{12,i} \neq 0\ \forall_i$ means that energy use Granger-cause energy price, while causality runs from energy price to energy use is indicated by $b_{21,i} \neq 0\ \forall_i$.

3.4 Empirical results and discussion

Table 6 deals with the explanation of descriptive statistics and pair-wise correlation. The Jarque-Bera test reveals the normal distribution of data. This support us for further analysis

to investigate the relationship between variables. The pair-wise correlation analysis reports that economic growth, industrialisation and trade openness have positive correlation with energy consumption but the correlation of energy price and urbanisation is negative with energy consumption. Economic growth, industrialisation, urbanisation and trade openness are positively correlated with energy price. The correlation of industrialisation, urbanisation and trade openness are positive with economic growth. Urbanisation and trade openness are positively correlated with industrialisation. Lastly, a positive correlation exists between trade openness and energy urbanisation.

In order to examine unit root properties of the variables, we applied traditional ADF (Augmented Dickey-Fuller) unit root test and ADF unit root test with structural break in the series introduced by Kim and Perron (2009). Although the bounds test for cointegration does not require that all variables be integrated of order 1, $I(1)$, it is essential to conduct the stationarity tests in order to ensure that the variables are not integrated of order 2, $I(2)$. Otherwise, the F-test would be spurious in the presence of $I(2)$ because both the critical values of the F-statistics computed by Pesaran *et al.* (2001) and Narayan (2005) are based on the assumption that the variables are $I(0)$ or $I(1)$. The null hypothesis of both the unit root tests is that the series in question has a unit root against the alternative of stationarity. The result of unit root tests is presented in Table 7. The empirical results of ADF test reveals that all the variables are non-stationary in their level data. However, we noted that all the variables are stationary after first difference. This shows that all the variables are integrated at $I(1)$. This empirical results reported by ADF may ambiguous due to its low explanatory power. Furthermore, ADF unit root test does not accommodate information of structural break stemming in the series which may be cause of unit root problem. This issue is solved by applying ADF unit root test with structural break and results are reported in lower part of Table 7. We find that all the variables contain unit root problem in the presence of structural breaks at level with intercept and time trend. The structural break periods are $1963, 1973, 2007, 2008, 1969$ and 1972 for energy consumption, oil prices, economic growth, industrialisation, urbanisation and trade openness. The presence of structural breaks in the variables is outcome of economic, energy and trade

policies. After first difference, all the variables are found stationary. This shows that energy consumption, oil prices, economic growth, industrialisation, urbanisation and trade openness unique order of integration i.e. $I(1)$.

The unique order of integration of the variables leads us for applying cointegration between energy consumption and its determinants. In doing so, we choose to apply the autoregressive distributive lag modelling (ARDL) or bounds testing approach to cointegration developed by Pesaran *et al.* (2001). The dummy variable capturing structural break based on Kim and Perron (2009) unit root test empirical findings is included while investigating ARDL F-statistic proposed by Pesaran *et al.* (2001). Before proceeding to apply the ARDL bounds testing in order to examine cointegration between energy consumption and its determinants, we choose appropriate lag length of the variables by applying unrestricted vector auto-regressive (VAR) model. The ARDL F-test is sensitive with lag length. The empirical results will be misleading if we choose inappropriate lag length of the variables. The lag length of the variables is chosen following Akiake information criteria (AIC). The AIC is superior compared to other criterion due to its explanatory power. The appropriate lag length based on AIC is reported in second column of Table 8. The results of ARDL bounds testing approach to cointegration are reported in Table 8. It is noted that ARDL F-statistic exceeds upper critical bound as we treated energy consumption, oil prices, economic growth, industrialisation and trade openness as dependent variables. This shows the presence of five cointegrating vectors stemming in energy demand function. This confirms the existence of cointegration relationship between the variables. We may conclude that energy consumption, oil prices, economic growth, industrialisation, urbanisation and trade openness have long-run relationship for UK over the period of $1960 - 2015$.

The existence of cointegration relationship between the variables leads us to examine the long run impact on oil price, economic growth, urbanisation, industrialisation and trade openness on energy consumption. The long-run results are reported in Table 9 and we find that oil prices are inversely linked with energy demand i.e. rise in oil prices decline energy consump-

tion. A 1% increase in oil prices will decline energy consumption by 0.5435%. The relationship between economic growth and energy consumption is positive and statistically significant. This shows that economic growth has positive effect on energy consumption. Keeping other things constant, a 0.5437% increase in energy consumption is linked with 1% increase in economic growth. The association between urbanisation and energy demand is negative and statistically significant at 1% level. This reveals that urbanisation declines energy demand. A 1% increase in urbanisation declines energy consumption by 0.6263% by keeping all else constant. Industrialisation affects energy demand positively and significantly. This shows that rise in industrialisation increases energy consumption. Keeping other factors constant, a 0.5502% increase in energy consumption is linked with 1% increase in industrialisation. The relationship between trade openness and energy consumption is positive and it is statistically significant at 1% level. It implies that trade openness increases energy consumption. A 1% rise in trade openness leads energy consumption by 0.4908% by keeping other factors constant. The dummy variable captures the implementation of the Electricity and Gas Act 1963 for UK in energy demand function. The effect of dummy variable is positively and statistically significant on energy consumption. The squared terms of economic growth, industrialisation, urbanisation and trade openness are included in energy demand function in order to examine whether relationship between the variables is inverted-U shaped or U-shaped. The results are reported in Table 9. We find that linear and squared terms of economic growth have negative and positive effect on energy consumption. This confirms the presence of U-shaped relationship between economic growth and energy consumption. This U-shaped association between economic growth and energy consumption reveals that energy consumption decreases initially and starts to increase with an increase in economic growth after a threshold level. The relationship between urbanisation and energy consumption is inverted-U shaped. This shows that energy consumption increases with urbanisation initially but after threshold level of urbanisation, energy consumption starts to decline. The relationship between industrialisation and energy consumption is U-shaped and statistically significant at 1% level. This implies that initially, industrialisation is negatively linked with energy consumption but energy consumption starts to rise with an

increase in industrialisation after a threshold level. The association between trade openness and energy demand is inverted-U shaped as linear and squared terms of trade openness have positive and negative effect on energy consumption which is statistically significant at 1% level. This empirical evidence reveals that energy consumption initially increases with trade openness and after threshold level of trade, energy consumption starts to decline. Overall, linear and squared energy demand functions are statistically significant confirmed by F-statistics. The linear and squared energy demand function are 94.29% and 98.01% explained by independent variables. There is an absence of autocorrelation in both models. The stability analysis shows the absence of serial correlation, auto-regressive conditional heteroscedasticity and white heteroscedasticity. The functional forms of linear and squared energy demand functions are well formulated confirmed by Ramsey reset test.

Table 10 presents the short-run empirical results. We find that oil price inversely and significantly linked with energy demand. Economic growth is positively and significantly linked with energy consumption. Urbanisation adds to energy demand insignificantly. The association between industrialisation and energy consumption is negative but statistically insignificant. Trade openness increases energy consumption significantly. In squared energy demand function, the association between economic growth and energy demand is inverted-U shaped but insignificant. A significant U-shaped relationship exists between urbanisation and energy consumption. The relationship between industrialisation and energy consumption is U-shaped but statistically insignificant. An inverted-U shape linkage is found between trade openness and energy consumption but it is statistically insignificant. The coefficients of ECM_{t-i} of linear and quadratic energy demand functions are negative and statistical significant at 1% and 5% levels, respectively. This confirms that an established long-run relationship between energy demand and its determinants is corroborated. The negative and significant estimate of ECM_{t-i} shows the speed of adjustment from short-run disequilibrium towards long-run equilibrium path. This shows that short-run deviations are corrected by 7.48% and 24.88% each year for linear and quadratic energy demand functions for UK. Linear and quadratic energy demand functions are 72.18% and 77.25% explained by oil prices, economic growth, urbanisation, industrialisation

and trade openness in short-run. The F-statistics confirm the overall significance of linear and quadratic energy demand function at 1% level. There is no empirical evidence of autocorrelation. The diagnostic analysis reports the absence of serial correlation. The residual term is normally distributed. An absence of autoregressive conditional heteroscedasticity and white heteroscedasticity is confirmed. The functional form is well formulated confirmed by Ramsey reset test. The $CUSUM$ and $CUSUM_{SQ}$ test are also applied to examine the reliability of ARDL long-run and short-run estimates. Figure 8 show empirical results of linear energy demand (Figure 1, Figure 2) and quadratic energy demand function (Figure 3, Figure 4). We find that $CUSUM$ and $CUSUM_{SQ}$ test are between critical bounds at 5% level. This confirms that ARDL estimates are reliable and consistent.

The results of VECM Granger causality are reported in Table 11. In long run, we find the bidirectional causality between oil prices and energy consumption. This shows that oil prices and energy consumption are complementary. Contrarily, Bekhet and Yusop (2009) reported that oil prices lead energy consumption for Malaysian economy. Economic growth causes energy consumption and in resulting, energy consumption causes economic growth in Granger sense. This empirical evidence is not consistent with Balcilar *et al.* (2010) and Tiwari and Mutascu (2011) who reported neutral effect between energy consumption and economic growth in UK. The unidirectional causality exists running from urbanisation to energy consumption and oil prices. Industrialisation is Granger cause of urbanisation. This empirical evidence is consistent with Shahbaz (2012) who reported that urbanisation leads industrialisation. Urbanisation Granger causes trade openness. This empirical evidence is consistent with Shahbaz *et al.* (2015) who reported that urbanisation leads trade openness. Energy consumption is cause of industrialisation and in resulting, industrialisation is cause of energy consumption in Granger sense. This empirical finding is similar with Shahbaz (2012) who noted the feedback effect between industrialisation and economy consumption. The bidirectional causality exists between trade openness and energy consumption. Similarly, Shahbaz *et al.* (2015) reported that trade openness causes energy consumption and in resulting, energy consumption causes trade openness. The feedback effect also exists between oil prices and industrialisation (trade

openness). In short run, we find the bidirectional causality between economic growth and energy consumption. Energy consumption is Granger cause of trade openness. The unidirectional causal relationship exists running from urbanisation and trade openness to oil prices. Economic growth causes industrialisation but opposite is not true. Trade openness is cause of economic growth in Granger sense.

3.5 Conclusion and policy implications

This paper explore the relationship between economic growth and energy consumption by adding oil prices, industrialisation, urbanisation and trade openness in energy demand function. In doing so, we have applied ADF unit root test without and with structural breaks in order to examine stationary properties of the variables. For examining cointegration between energy consumption and its determinants, the ARDL bounds testing approach to cointegration is applied by accommodating structural breaks in the series. The VECM Granger causality is employed for examining direction of causal relationship between the variables. The empirical evidence validates the existence of cointegration between energy consumption and its determinants in the presence of structural breaks stemming in the series. Further, oil prices rise declines energy demand. Economic growth is positive linked with energy consumption. Urbanisation reduces energy consumption. Energy demand is increased with an increase in industrialisation. Trade openness adds in energy consumption. The causality analysis reveals the presence of feedback effect between economic growth and energy consumption. Energy consumption causes oil prices and in resulting, oil prices cause energy consumption in Granger sense. Industrialisation causes energy consumption and energy consumption causes industrialisation as well. The feedback effect also exists between trade openness and energy consumption. Urbanisation Granger causes energy consumption, oil prices, economic growth, industrialisation and trade openness. The government of UK should revisit its technological policy and implement energy efficiency in individual sector to enhance domestic production for maintaining sustainable economic development in long run.

Chapter 3

Resiliency and Asymmetric Reaction to Price Changes of Shale Oil Rig Counts

Massimiliano Caporin*[1], Fulvio Fontini†[2], Elham Talebbeydokhti‡[2]

[1]Department of Statistical Sciences, University of Padua, Italy
[2]Department of Economics and Management, University of Padua, Italy

Abstract

Since US shale revolution stimulated tremendous oil production, the number of US rig count became widely publicised. A drilling rig is a machine that creates holes in the earth sub-surface in order to drill a new well to explore for, develop and produce oil. Recently, US oil market has been characterised by fluctuating WTI prices and this makes difficult for oil producers to determine the profitability in exploration and development. Thus, it is wise to take a closer look at the relationship between oil rig count and crude oil prices. From 2011 onward, the relative importance of shale oil production over the total US oil supply has been significantly increasing so it is reasonable to focus on the relationship between shale rig count and crude oil prices in particular. Therefore, We split the total US rig count into shale rig count and non-shale rig count. We analyse how shale rig count and non-shale rig count and their production in US is affected by the changes in energy prices while accounting for other determinants of this relationship. We also studied if this relationship is asymmetric for rises and drops of crude oil prices. We test for the often claimed hysteresis hypothesis of the shale oil production in the case of crude oil prices drops. The most relevant variables we use are weekly data on WTI price, the shale rig count, non-shale rig count, rig productivity and a set of potentially relevant economic and financial control variables from February 2011 up to October 2017 for total of 344 observations. This relationship is of significant interest to analysts, investors, oil companies, commercial banks, investment banks and policy makers.

Keywords: WTI Price, Shale oil, Rig Count.

*massimiliano.caporin@unipd.it
†fulvio.fontini@unipd.it
‡elham.talebbeydokhti@phd.unipd.it

4.1 Introduction

A drilling rig is a machine that creates holes in the earth sub-surface in order to drill a new well to explore for, develop and produce oil. Advanced in horizontal drilling and hydraulic fracturing have recently expanded the US supply of petroleum. The US development of shale reserves has already shifted the global balance of supply and demand for oil, reduced OPEC market power, and brought oil prices to a new, lower equilibrium level. Since US shale revolution stimulated tremendous oil production, the number of US rig count became widely publicised. The long-run impact of shale oil will depend on how much can be produced, therefore, the rig count can be consider as a good indicator to evaluate this point. However, increasing rig efficiency has been helping US oil companies produce more crude oil with fewer rigs. The changes in the price of crude oil are often considered an important factor for understanding fluctuations of crude oil production and rig count. The oil price have plummeted from the peak of $108/bbl in June 2014 to a low of $28/bbl in February 2016. Drilling activity responded swiftly and the shale oil rig count dropped from a high of 1127 in September 2014 to a low of 232 in May 2016, a 79.41 percent decline. The non-shale rig count dropped from a high of 485 in October 2014 to a low of 75 in June 2016, a 84.54 percent decline. Figure 9 plots the levels and the changes of our interest variables. The oil prices may therefore affect oil rig count to exploring new areas and developing new fields. Consequently, oil price changes influence shale oil production mainly through their impact on investments in exploration activity and field development. The number of active rigs could rise if oil price trade firm. However, rising rigs could bring more supply to the market, which would in turn pressure oil price. Although it seems obvious that higher oil price eventually bring about higher investment in oil fields but the recent relationship between changes in oil price and oil rig counts may not be that obvious and direct because of the presence of the lagged response between these variables Black and LaFrance (1998). According to research by Morgan Stanely, over the past ten years, the lows and highs of the oil rig count and crude oil have been three to four months apart. The rig counts are expected to be related to the oil price. Yet in the literature the oil price - rig count nexus has hardly been

quantitatively explored. An exception is Khalifa *et al.* (2017), who empirically show that there is a delayed and non-linear impact of changes in oil prices on oil rig counts. This paper however neither distinguishes across different types of oil production, nor consider the causality of the oil price - rig counts relationship. A technical innovation of oil cultivation, namely, horizontal drilling and fracking, has allow extracting oil entrapped in shale rocks formation, named shale oil. In the US, shale oil has become the largest component of total oil production and has boomed the US oil supply. The US total crude oil production increased from 5392 thousand barrels per day in February 2011 to 9203 thousand barrels per day in August 2017 mostly due to rise in shale oil production which has 51% share of total US crude oil production. This has fostered the US oil exports (Langer *et al.* (2016)) and contributed (at least partially) to the fall of oil price (Aguilera and Radetzki (2015), Baumeister and Kilian (2016), Dale (2016), Khan (2017)). Shale industry has shown to be extremely resilient to oil price drops, witnessing a surprisingly high ability to resist to oil price reduction (OPEC (2016), Ansari (2017)), even though business reports cast doubt on its possibility to last overtime (Nussbaum and Wethe (2017)). In this paper, we empirically test if the oil rig counts - oil price nexus is affected by the nature of the oil extraction, distinguishing between shale oil rigs and non shale oil rigs. We also investigate if there is an asymmetry in the relationship between shale oil rig counts and oil price when considering positive and negative oil prices; if such a relationship differ across types of rig counts, namely, shale and non shale oil rig counts; if and how much more resilient to reduction in oil prices is the shale industry compared to the non shale one. We do so by estimating and testing the difference in the delayed response of oil rig counts to increase and falls of oil prices distinguishing between the two types of oil rigs. Given that the number of shale oil rigs can be influenced by the oil price change, but also the opposite effect can take place since oil rig counts affect oil supply and thus might impact on oil price, we also study the revers causality and feedback hypothesis of the oil price - shale oil rig nexus, controlling for the difference between shale and non shale oil rigs. The paper proceed as follows: Section 2 provides a brief review of the literature. Section 3 describes the data. Section 4 presents the research methodology. Section 5 explain the empirical results. Section 6 concludes.

4.2 Literature review

In this section we discuss relevant studies and position our paper in the literature. It seems like the literature analysing recent energy prices and shale rig count is fairly limited.

Some studies generally focus on the economics of well production. Gülen *et al.* (2013) specifically focus on the sensitivity and economic viability of drilling new natural gas wells in Texas' Barnett Shale to changes in natural gas prices. Additionally, a number of studies focus on costs and revenues and thus, the economic impact of oil and gas industry (Howard and Harp Jr (2009); Ewing *et al.* (2014); Snead (2005); Henriques and Sadorsky (2011)).

Osmundsen *et al.* (2010) analyses the development in drilling productivity in exploration wells at the Norwegian continental shelf. Osmundsen *et al.* (2012) pioneered on change in drilling speed and oil prices. They test impact of experience on drilling speed. The authors find that congestion externalities and depletion effects on average dominate learning effects. These effects may not be identified at the aggregate level because of averaging out.

Osmundsen *et al.* (2010) empirically evaluate the importance of relationship-specific learning, using a detailed data from the onshore oil and gas drilling in Texas. He show that the productivity of an oil production company and its drilling contractor increases with their joint experience. Newell and Raimi (2014) estimate the elasticity of shale gas supplies in Texas. They find that prices provide a significant incentive to drilling by regressing changes in drilling activity against changes in expected revenues from production. Smith and Lee (2017) present a new method of shale oil development to assess how much of the US resources base is likely to be economically viable at various price level, and what share of potential drilling sites are likely to be exploited. They find the volume of reserves to be highly inelastic with respect to price.

Chen (2011) examine the impact changes in futures price have on drilling activity, as measured by changes in the number of rigs actively employed. Their study emphasizes a global comparison of the US versus the rest of the world. Ringlund *et al.* (2008) investigates how

oil rig count in different non-OPEC regions is affected by crude oil price by using dynamic regression models augmented with latent components capturing trend and seasonality. Their result show a positive relationship between oil rig activity and crude oil price and depending on the oil industry structure and the reaction of the oil rig count to changes of oil prices, the persistence of this relationship differs across regions. Toews and Naumov (2015) formulate a VAR for the oil and gas upstream industry and annual data to examine the dynamic effects of the oil price and drilling activity. These authors find a directional relationship that runs from changes in real oil prices to rig counts with a one-year lag. Khalifa *et al.* (2017) shows that the changes in oil prices on oil rig counts relationship has lag up to one quarter and it is non-linear by using both the quantile regression and quantile-on-quantile approach. Within the same line of research, Black and LaFrance (1998) question the lag based on an empirical investigation on oil fields in Montana. Apergis *et al.* (2016) examine the dynamic relationship among oil production, rig count and crude oil prices for six US oil producing regions. They find a long-run equilibrium relationship exists in each of the six regions with the coefficient on rig count being the largest for the Permian region. However, in their study, the rig count data is mix of oil rig count and gas rig count.

Our study extends the existing literature. We examine the time series empirical relationship between shale/non-shale production and shale/non-shale rig count, accounting for the energy prices for oil and gas separately.

4.3 Data description

The weekly data for active rig count have been collected from the Baker Hughes Rig Counts by basins. In order to distinguish between shale and non shale oil rigs, we reconstruct the time series of weekly US oil rig counts data attributing the oil rigs data to shale and non-shale rigs on the basis of the geographical distribution of the basins. The basins consider as shale basins are: Ardmore Woodford, Arkoma Woodford, Barnett, Cana Woodford, DJ-Niobrara, Eagle Ford, Fayetteville, Garnite Wash, Haynesville, Marcellus, Mississippian, Permian, Utica,

Williston. We also use data series for WTI crude oil spot prices, US shale oil production as well as data for several economic and financial control variables from February 2011 to October 2017 for total of 344 observations.

4.4 Research methodology

Recently, US oil market has been characterised by fluctuating WTI prices and this makes difficult for oil producers to determine the profitability in exploration and development. Thus, it is wise to take a closer look at the relationship between oil rig count and crude oil prices. From 2011 onward, the relative importance of shale oil production over the total US oil supply has been significantly increasing so it is reasonable to focus on the relationship between shale rig count and crude oil prices in particular.

To verify the existence of interdependence between oil rigs and oil prices, following the work of Khalifa *et al.* (2017), we start by considering cross-correlation and exceedance correlations. We thus evaluate the correlation between the changes in oil rig counts, total, shale and non-shale, and lagged relative changes of oil prices cumulated over increasing periods. This sheds light on the time needed for oil rigs to react to changes in prices. Furthermore, we repeat the same analysis pointing at exceedance correlations, thus separating the oil relative changes depending on their sign. This allows evaluating if negative or positive price movements have equal effects on the change in rig counts. We repeat the same cross-correlation analysis between the relative changes in the oil prices and lagged cumulated changes in rig counts. This allows focusing on the impact of changes in production on the oil price.

Given these preliminary estimates, we proceed to the evaluation of the dynamic interdependence between the relative changes in oil prices and changes in rig counts by focusing on a VAR model. We exclude a priori the possible presence of contemporaneous effects between the two variables (an assumption that is validated by the previous analyses) and focus on the dynamic. The lag structure is specified both by building on the cross-correlation analysis as

well as by focusing on economic reasoning. The VAR model is augmented by control variables and extended by introducing a distinction between positive and negative price changes. By specifying a VAR model, we can construct the impulse response functions that are used to analyse the impact and resilience of oil price shocks on the rig counts, by conditioning on the type of rigs and the sign of the shocks. Moreover, we evaluate the effect on oil prices of shocks on the production side.

The dependent variable y_t is a set of two variables, changes in shale oil rig count (ΔSOR_t) and changes in non-shale rig count ($\Delta NSOR_t$).

$$y_t = \begin{bmatrix} \Delta SOR_t \\ \Delta NSOR_t \end{bmatrix}$$

Model 1: standard VAR model

$$y_t = \Phi_0 + \sum_{j=1}^{p} \Phi_j y_{t-j} + \varepsilon_t \tag{7}$$

where the lag length is chosen with information criteria using a maximum lag of 13 weeks (i.e. 3 months).

Model 2: restricted VAR

$$y_t = \Phi_0 + \Phi_1 y_{t-1} + \Phi_m y_{t-1|t-4} + \Phi_q y_{t-1|t-13} + \varepsilon_t \tag{8}$$

where $y_{t-1|t-j} = \sum_{i=1}^{j} y_{t-i}$. Here, we construct a VAR model where estimation with considering different lags. The idea is to verify if there is an impact of monthly and quarterly movements rather than all lags.

Model 3: restricted VARX

$$y_t = \Phi_0 + \Phi_1 y_{t-1} + \Phi_m y_{t-1|t-4} + \Phi_q y_{t-1|t-13} + \Psi_1 x_{t-1} + \Psi_m x_{t-1|t-4} + \Psi_q x_{t-1|t-13} + \delta F_{t-1}\varepsilon_t \quad (9)$$

where, x_t is log difference of oil price ($x_t = \Delta OP_t$) and F_t is a collection of economic/financial variables. Here, we introduce a model allows for prices to impact on rig counts.

Model 4: restricted VARX

$$\begin{aligned} y_t &= \Phi_0 + \Phi_1 y_{t-1} + \Phi_m y_{t-1|t-4} + \Phi_q y_{t-1|t-13} + \Psi_1^+ x_{t-1}^+ + \Psi_m^+ x_{t-1|t-4}^+ + \Psi_q^+ x_{t-1|t-13}^+ \quad (10) \\ &\quad + \Psi_1^- x_{t-1}^- + \Psi_m^- x_{t-1|t-4}^- + \Psi_q^- x_{t-1|t-13}^- + \delta F_{t-1}\varepsilon_t \end{aligned}$$

where $x_{t-1|t-j}^- = \sum_{i=1}^{j} x_{t-i} I\left(x_{t-i} < 0\right)$ and $x_{t-1|t-j}^+ = \sum_{i=1}^{j} x_{t-i} I\left(x_{t-i} \geq 0\right)$. We also studied if this relationship is asymmetric for rises and drops of oil price.

4.5 Research methodology

Preliminary results confirm the literature finding of a positive relationship between oil rigs and oil price, delayed up to 3 months. When distinguishing between shale and non shale oil rigs, we see that the impact of the delayed oil price on shale rigs is higher than on non shale ones. We then separate positive and negative changes in prices and see that when considering the overall number of rigs the positive oil price changes have positive impacts on rigs, while the negative ones reduces the rigs, as expected. The effect of the latter is smaller and more delayed than the impact of positive price changes on rigs. Disaggregating between shale and non-shale rigs, we see that for the shale rigs the relationships are stronger and more significant than for non-shale oil rig counts. For the shale oil rigs, positive price changes impact more than the negative ones, with the expected sign, showing resilience to price drops. There is a lower resilience of the non-shale industry. Causality impacts and feedback is still under investigation.

5 References

References

Abosedra, S. and Baghestani, H. (1989). New evidence on the causal relationship between united states energy consumption and gross national product. *The Journal of Energy and Development*, pp. 285–292. 3.2.2

Aguilera, R. F. and Radetzki, M. (2015). *The price of oil*. Cambridge University Press. 4.1

Akarca, A. T. and Long, T. V. (1980). On the relationship between energy and gnp: a reexamination. *The Journal of Energy and Development*, **5** (2), 326–331. 3.2.2

Amirat, A. and Bouri, A. (2010). Energy and economic growth: the algerian case. *Accessed*, **24** (1), 2012. 3.2.2

Ansari, D. (2017). Opec, saudi arabia, and the shale revolution: Insights from equilibrium modelling and oil politics. 4.1

Apergis, N., Ewing, B. T. and Payne, J. E. (2016). A time series analysis of oil production, rig count and crude oil price: Evidence from six us oil producing regions. *Energy*, **97**, 339–349. 4.2

Aruga, K. *et al.* (2015). Testing the international crude oil market integration with structural breaks. *Economics Bulletin*, **35**, 641–649. 2.1

Asghar, Z. (2008). Energy-gdp relationship: a causal analysis for the five countries of south asia. 3.2.2

Balcilar, M., Ozdemir, Z. A. and Arslanturk, Y. (2010). Economic growth and energy consumption causal nexus viewed through a bootstrap rolling window. *Energy Economics*, **32** (6), 1398–1410. 3.4

Baumeister, C. and Kilian, L. (2016). Understanding the decline in the price of oil since june 2014. *Journal of the Association of Environmental and Resource Economists*, **3** (1), 131–158. 4.1

Bekhet, H. A. and Yusop, N. Y. M. (2009). Assessing the relationship between oil prices, energy consumption and macroeconomic performance in malaysia: co-integration and vector error correction model (vecm) approach. *International Business Research*, **2** (3), 152. 3.4

Black, G. and LaFrance, J. T. (1998). Is hotelling's rule relevant to domestic oil production? *Journal of Environmental Economics and Management*, **36** (2), 149–169. 4.1, 4.2

Borenstein, S., Kellogg, R. *et al.* (2014). The incidence of an oil glut: Who benefits from cheap crude oil in the midwest? *Energy Journal*, **35** (1), 15–33. 2.1

Brown, R. L., Durbin, J. and Evans, J. M. (1975). Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 149–192. 3.3.2

Büyükşahin, B., Lee, T., Moser, J. and Robe, M. (2013). Physical market conditions, paper market activity, and the wti–brent spread. *Energy Journal*, **34** (3). 2.1

Caporin, M. and Fontini, F. (2016). The long-run oil–natural gas price relationship and the shale gas revolution. *Energy Economics*, http://dx.doi.org/10.1016/j.eneco.2016.07.024.

Chen, F. (2011). Energy futures prices and drilling activity. 4.2

Chen, W., Huang, Z. and Yi, Y. (2015). Is there a structural change in the persistence of wti–brent oil price spreads in the post-2010 period? *Economic Modelling*, **50**, 64–71. 2.1

Choi, I. (2001). Unit root tests for panel data. *Journal of international money and Finance*, **20** (2), 249–272. 3.2.2

Cole, M. A. (2006). Does trade liberalization increase national energy use? *Economics Letters*, **92** (1), 108–112. 3.2.3

— and Neumayer, E. (2004). Examining the impact of demographic factors on air pollution. *Population & environment*, **26** (1), 5–21. 3.2.1

Dale, S. (2016). New economics of oil. oil and gas. *Natural Resources, and Energy Journal*, **1** (3). 4.1

Dedeoğlu, D. and Kaya, H. (2013). Energy use, exports, imports and gdp: New evidence from the oecd countries. *Energy Policy*, **57**, 469–476. 3.2.3

EIA (2016). Short-term energy outlook: Market prices and uncertainty report. *US Energy Information Administration*. 2.1

Engle, R. F. and Granger, C. W. (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica: journal of the Econometric Society*, pp. 251–276. 3.3.1

Ewing, B., Watson, M., McInturff, R. and McInturff, R. (2014). The economic impact of the permian basin?s oil and gas industry. *Report Prepared for Permian Basin Petroleum Association, Texas Tech University*. 4.2

Fan, Y. and Xu, J.-H. (2011). What has driven oil prices since 2000? a structural change perspective. *Energy Economics*, **33** (6), 1082–1094.

Fattouh, B. (2007). Wti benchmark temporarily breaks down: is it really a big deal. *Middle East Economic Survey*, **50** (20), 7. 2.1

— (2010). The dynamics of crude oil price differentials. *Energy Economics*, **32** (2), 334–342. 2.1

Fisher-Vanden, K., Jefferson, G. H., Liu, H. and Tao, Q. (2004). What is driving china?s decline in energy intensity? *Resource and Energy economics*, **26** (1), 77–97. 3.2.3

Giles, D. E. and Godwin, R. T. (2012). Testing for multivariate cointegration in the presence of structural breaks: p-values and critical values. *Applied Economics Letters*, **19** (16), 1561–1565. a

Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, pp. 424–438. 3.3.3

Gregory, A. W. and Hansen, B. E. (1996). Residual-based tests for cointegration in models with regime shifts. *Journal of econometrics*, **70** (1), 99–126. 2.2, 2.2.1, 2.2.1, 2.4

Gülen, G., Browning, J., Ikonnikova, S. and Tinker, S. W. (2013). Well economics across ten tiers in low and high btu (british thermal unit) areas, barnett shale, texas. *Energy*, **60**, 302–315. 4.2

Gülen, S. G. (1999). Regionalization in the world crude oil market: further evidence. *The Energy Journal*, pp. 125–139. 2.1

Hadri, K. (2000). Testing for stationarity in heterogeneous panel data. *The Econometrics Journal*, **3** (2), 148–161. 3.2.2

Hammoudeh, S. M., Ewing, B. T. and Thompson, M. A. (2008). Threshold cointegration analysis of crude oil benchmarks. *The Energy Journal*, pp. 79–95. 2.1

Henriques, I. and Sadorsky, P. (2011). The effect of oil price volatility on strategic investment. *Energy Economics*, **33** (1), 79–87. 4.2

Howard, A. W. and Harp Jr, A. B. (2009). Oil and gas company valuations. *Business Valuation Review*, **28** (1), 30–35. 4.2

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of economic dynamics and control*, **12** (2-3), 231–254. 2.2.2, 2.2.2, 2.4

— (1995). *Likelihood-based inference in cointegrated vector autoregressive models*. Oxford University Press on Demand. 2.2.2, 2.2.2, 2.4

— and Juselius, K. (1990). Maximum likelihood estimation and inference on cointegration?with applications to the demand for money. *Oxford Bulletin of Economics and statistics*, **52** (2), 169–210. 3.3.1

—, Mosconi, R. and Nielsen, B. (2000). Cointegration analysis in the presence of structural breaks in the deterministic trend. *The Econometrics Journal*, **3** (2), 216–249. 2.4, a

Jones, G. W. (1991). Urbanization issues in the asian-pacific region. *Asian-Pacific Economic Literature*, **5** (2), 5–33. 3.1

Kahsai, M. S., Nondo, C., Schaeffer, P. V. and Gebremedhin, T. G. (2012). Income level and the energy consumption–gdp nexus: Evidence from sub-saharan africa. *Energy Economics*, **34** (3), 739–746. 3.2.2

Kao, C.-W. and Wan, J.-Y. (2012). Price discount, inventories and the distortion of wti benchmark. *Energy Economics*, **34** (1), 117–124. 2.1

Kellogg, R. (2011). Learning by drilling: Interfirm learning and relationship persistence in the texas oilpatch. *The Quarterly Journal of Economics*, **126** (4), 1961–2004.

Khalifa, A., Caporin, M. and Hammoudeh, S. (2017). The relationship between oil prices and rig counts: The importance of lags. *Energy Economics*, **63**, 213–226. 4.1, 4.2, 4.4

Khan, M. I. (2017). Falling oil prices: Causes, consequences and policy implications. *Journal of Petroleum Science and Engineering*, **149**, 409–427. 4.1

Kilian, L. (2016). The impact of the shale oil revolution on us oil and gasoline prices. *Review of Environmental Economics and Policy*, **10** (2), 185–205.

Kim, D. and Perron, P. (2009). Unit root tests allowing for a break in the trend function at an unknown time under both the null and alternative hypotheses. *Journal of Econometrics*, **148** (1), 1–13. 3.4

Kim, J., Kim, J. and Heo, E. (2013). Evolution of the international crude oil market mechanism. *Geosystem Engineering*, **16** (4), 265–274.

Kraft, J. and Kraft, A. (1978). On the relationship between energy and gnp. *The Journal of Energy and Development*, pp. 401–403. 3.2.2

Langer, L., Huppmann, D. and Holz, F. (2016). Lifting the us crude oil export ban: A numerical partial equilibrium analysis. *Energy Policy*, **97**, 258–266. 4.1

Levin, A. and Lin, C.-F. (1993). Unit root tests in panel data: new results. *University of California at San Diego, Economics Working Paper Series*. 3.2.2

Leybourne, S., Taylor, R. and Kim, T.-H. (2007). Cusum of squares-based tests for a change in persistence. *Journal of Time Series Analysis*, **28** (3), 408–433. 2.1

Liddle, B. and Lung, S. (2010). Age-structure, urbanization, and climate change in developed countries: revisiting stirpat for disaggregated population and consumption-related environmental impacts. *Population and Environment*, **31** (5), 317–343.

Liu, W.-M., Schultz, E. and Swieringa, J. (2015). Price dynamics in global crude oil markets. *Journal of Futures Markets*, **35** (2), 148–162. 2.1

Mishra, V., Sharma, S. and Smyth, R. (2009). Are fluctuations in energy consumption per capita transitory? evidence from a panel of pacific island countries. *Energy Policy*, **37** (6), 2318–2326. 3.2.1

Narayan, P. K. (2005). The saving and investment nexus for china: evidence from cointegration tests. *Applied economics*, **37** (17), 1979–1990. 3.4

— and Smyth, R. (2009). Multivariate granger causality between electricity consumption, exports and gdp: evidence from a panel of middle eastern countries. *Energy Policy*, **37** (1), 229–236. 3.2.3

Newell, R. G. and Raimi, D. (2014). Implications of shale gas development for climate change. *Environ Sci Technol*. 4.2

Noor, S. and Siddiqi, M. (2010). Energy consumption and economic growth in south asian countries: a co-integrated panel analysis. *International Journal of Human and Social Sciences*, **5** (14), 921–926. 3.2.2

Osmundsen, P., Roll, K. H. and Tveterås, R. (2010). Exploration drilling productivity at the norwegian shelf. *Journal of Petroleum Science and Engineering*, **73** (1), 122–128. 4.2

—, — and Tveterås, R. (2012). Drilling speed? the relevance of experience. *Energy Economics*, **34** (3), 786–794. 4.2

Parikh, J. and Shukla, V. (1995). Urbanization, energy use and greenhouse effects in economic development: Results from a cross-national study of developing countries. *Global Environmental Change*, **5** (2), 87–103. 3.2.2

Pedroni, P. (1999). Critical values for cointegration tests in heterogeneous panels with multiple regressors. *Oxford Bulletin of Economics and statistics*, **61** (S1), 653–670. 3.2.2

Perron, P. (1997). Further evidence on breaking trend functions in macroeconomic variables. *Journal of econometrics*, **80** (2), 355–385. 2.2, a

Pesaran, M. H., Shin, Y. and Smith, R. J. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of applied econometrics*, **16** (3), 289–326. 3.3.1, 3.4

Poumanyvong, P. and Kaneko, S. (2010). Does urbanization lead to less energy use and lower co 2 emissions? a cross-country analysis. *Ecological Economics*, **70** (2), 434–444. 3.2.1

Raza, S. A., Shahbaz, M. and Nguyen, D. K. (2015). Energy conservation policies, growth and trade performance: Evidence of feedback hypothesis in pakistan. *Energy Policy*, **80**, 1–10. 3.2.3

Reboredo, J. C. (2011). How do crude oil prices co-move?: A copula approach. *Energy Economics*, **33** (5), 948–955. 2.1

Recalde, M. and Ramos-Martin, J. (2012). Going beyond energy intensity to understand the energy metabolism of nations: The case of argentina. *Energy*, **37** (1), 122–132. 3.2

Ringlund, G. B., Rosendahl, K. E. and Skjerpen, T. (2008). Does oilrig activity react to oil price changes? an empirical investigation. *Energy Economics*, **30** (2), 371–396. 4.2

Sadorsky, P. (2011). Trade and energy consumption in the middle east. *Energy Economics*, **33** (5), 739–749. 3.2.3

— (2012). Energy consumption, output and trade in south america. *Energy Economics*, **34** (2), 476–488. 3.1, 3.2.3

Sathaye, J., Tyler, S. and Goldman, N. (1994). Transportation, fuel use and air quality in asian cities. *Energy*, **19** (5), 573–586. 3.1

Shahbaz, M. (2012). Does trade openness affect long run growth? cointegration, causality and forecast error variance decomposition tests for pakistan. *Economic Modelling*, **29** (6), 2325–2339. 3.2.3, 3.4

—, Arouri, M. and Teulon, F. (2014a). Short-and long-run relationships between natural gas consumption and economic growth: Evidence from pakistan. *Economic Modelling*, **41**, 219–226. 3.2.3

—, Loganathan, N., Sbia, R. and Afza, T. (2015). The effect of urbanization, affluence and trade openness on energy consumption: A time series analysis in malaysia. *Renewable and Sustainable Energy Reviews*, **47**, 683–693. 3.4

—, Nasreen, S., Ling, C. H. and Sbia, R. (2014b). Causality between trade openness and energy consumption: What causes what in high, middle and low income countries. *Energy Policy*, **70**, 126–143. 3.1

Smith, J. L. and Lee, T. K. (2017). The price elasticity of us shale oil reserves. *Energy Economics*, **67**, 121–135. 4.2

Snead, M. C. (2005). *The Economics of Deep Drilling in Oklahoma*. Center for Applied Economic Research, Oklahoma State University. 4.2

Solarin, S. A. and Shahbaz, M. (2013). Trivariate causality between economic growth, urbanisation and electricity consumption in angola: Cointegration and causality analysis. *Energy Policy*, **60**, 876–884. 3.1

Tiwari, A. K. and Mutascu, M. (2011). Economic growth and fdi in asia: A panel-data approach. *Economic analysis and policy*, **41** (2), 173–187. 3.4

Toews, G. and Naumov, A. (2015). The relationship between oil price and costs in the oil industry. *Energy Journal*, **36**. 4.2

Wang, Q. (2014). Effects of urbanisation on energy consumption in china. *Energy Policy*, **65**, 332–339. 3.1

Weiner, R. J. (1991). Is the world oil market "one great pool"? *The Energy Journal*, pp. 95–107.

Wilmot, N. A. (2013). Cointegration in the oil market among regional blends. *International Journal of Energy Economics and Policy*, **3** (4), 424. 2.1

York, R. (2007). Demographic trends and energy consumption in european union nations, 1960–2025. *Social science research*, **36** (3), 855–872. 3.2.1

Yu, E. S. and Choi, J.-Y. (1985). The causal relationship between energy and gnp: an international comparison. *The Journal of Energy and Development*, **10** (2), 249–272. 3.2.2

Zhang, C. and Lin, Y. (2012). Panel estimation for urbanization, energy consumption and co 2 emissions: a regional analysis in china. *Energy Policy*, **49**, 488–498. 3.1

Zivot, E. and Andrews, D. W. K. (2002). Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. *Journal of business & economic statistics*, **20** (1), 25–44. 3.3.1

Figure 1: WTI, Brent monthly Prices and their Spread

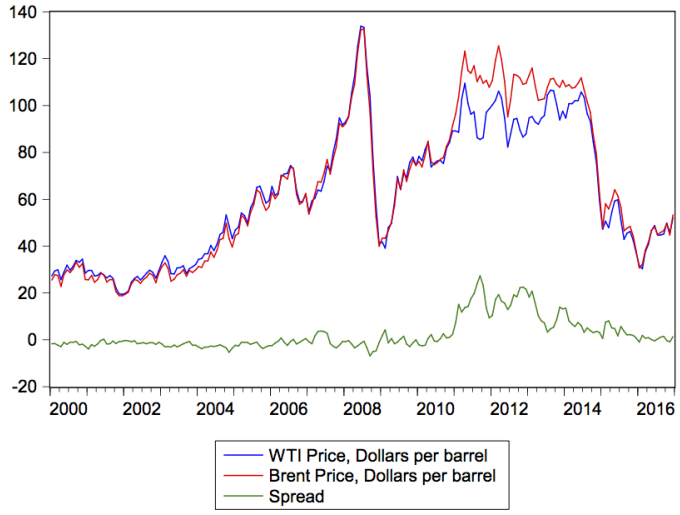

Figure 2: US Crude oil and Shale oil Production

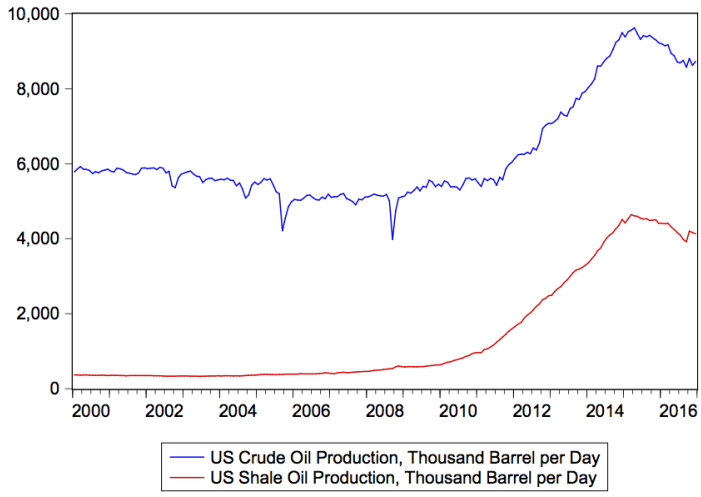

Figure 3: US Export of Crude oil

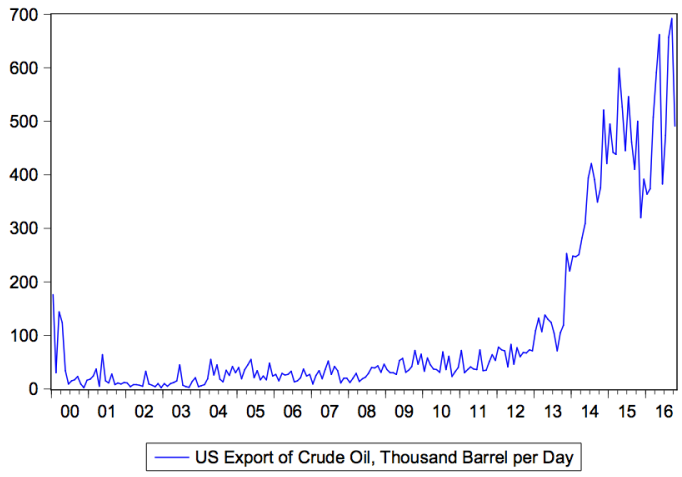

Figure 4: Impulse Response Function - Full-Sample

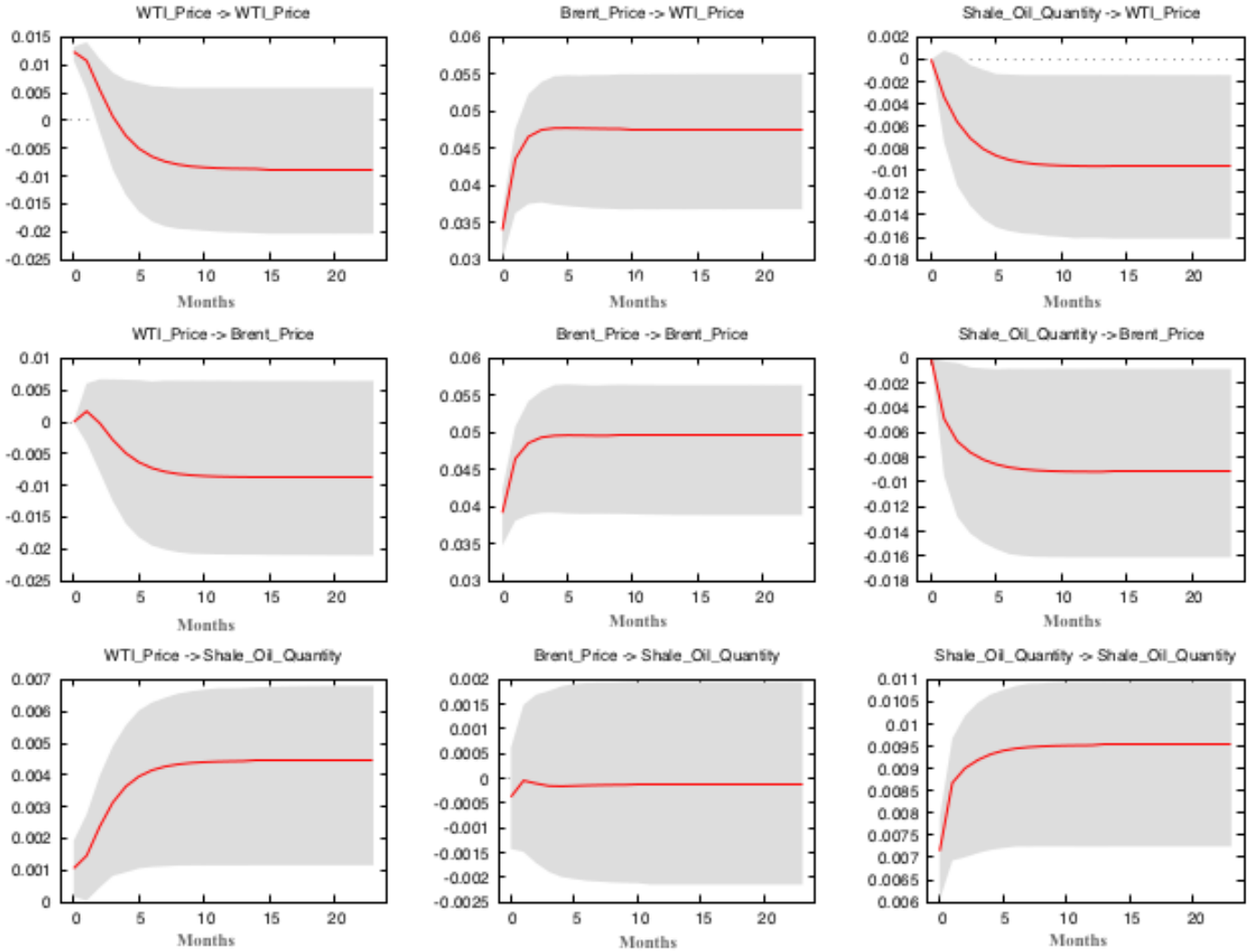

Figure 5: Impulse Response Function - First Sub-Sample

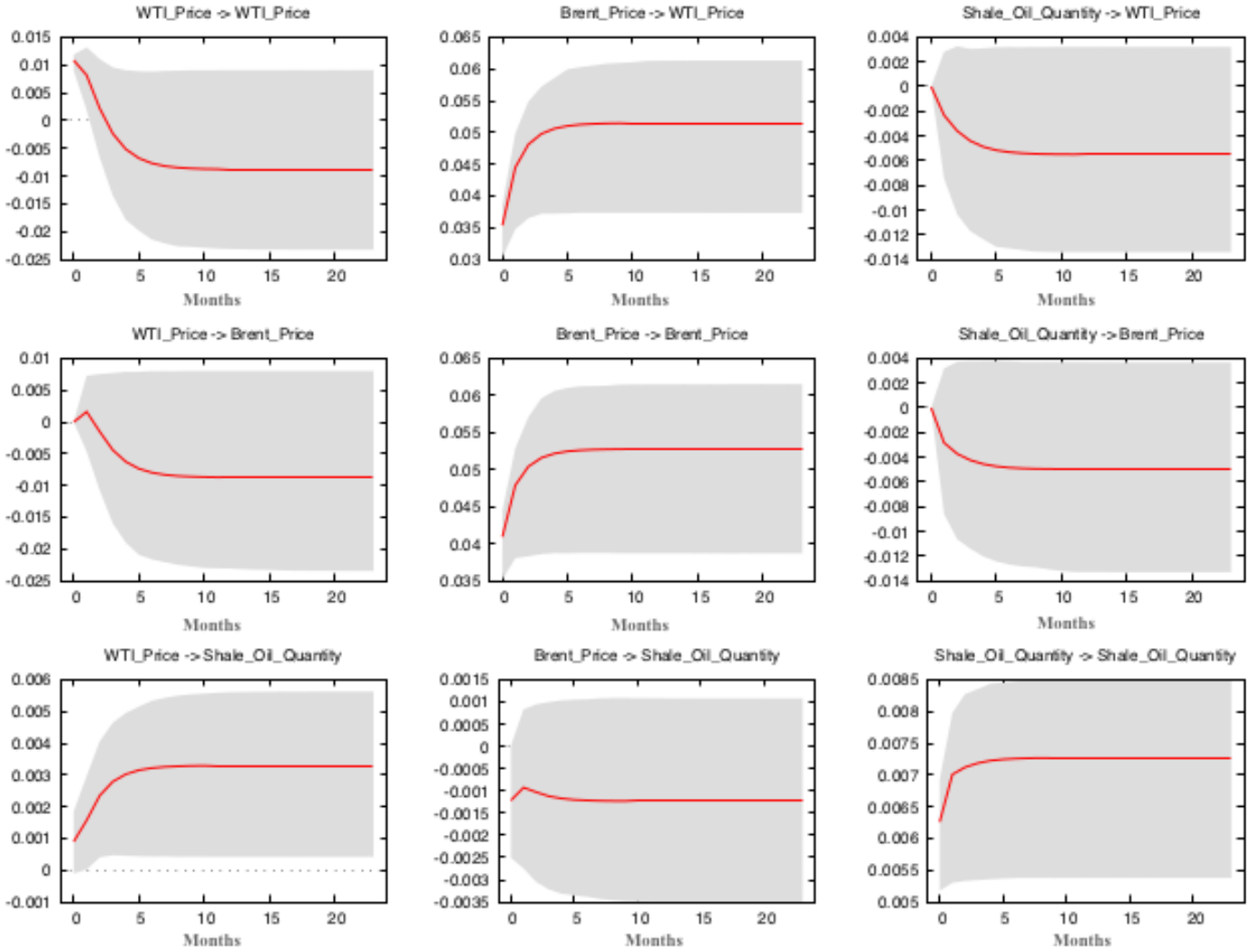

Figure 6: Impulse Response Function - Second Sub-Sample

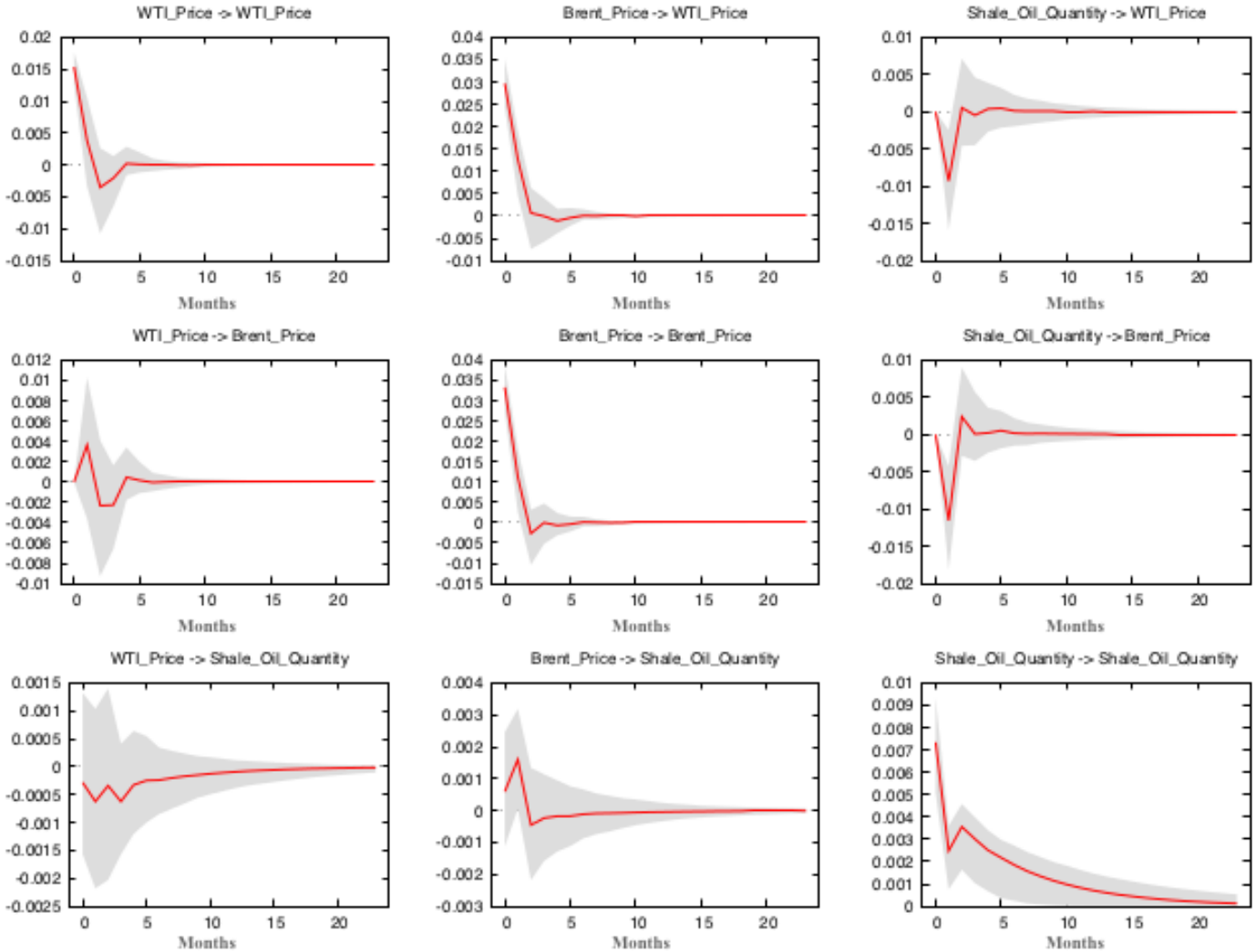

Figure 7: Trends of the variables from 1960 to 2015

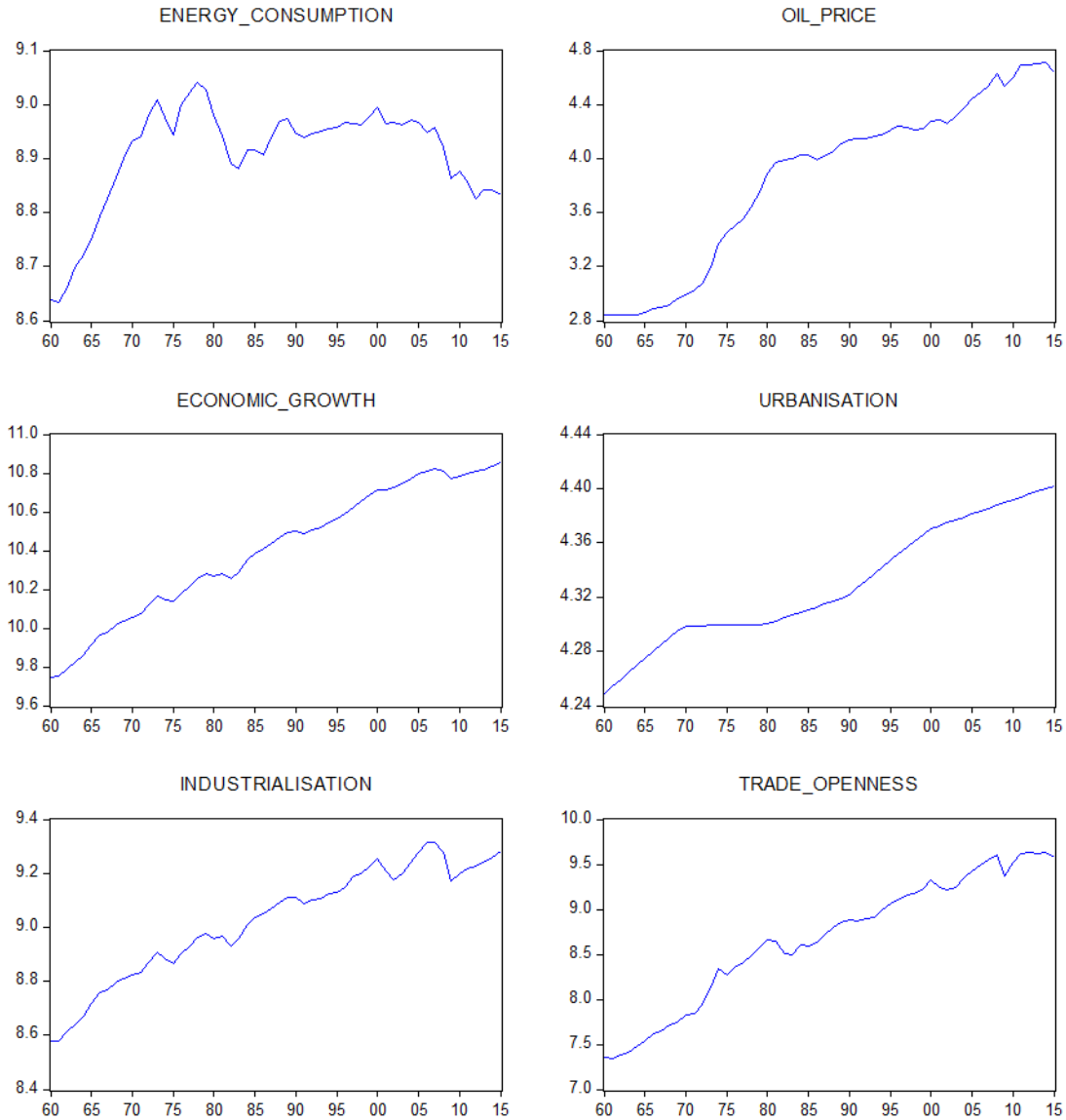

Figure 8: $CUSUM$ and $CUSUM_{SQ}$ test

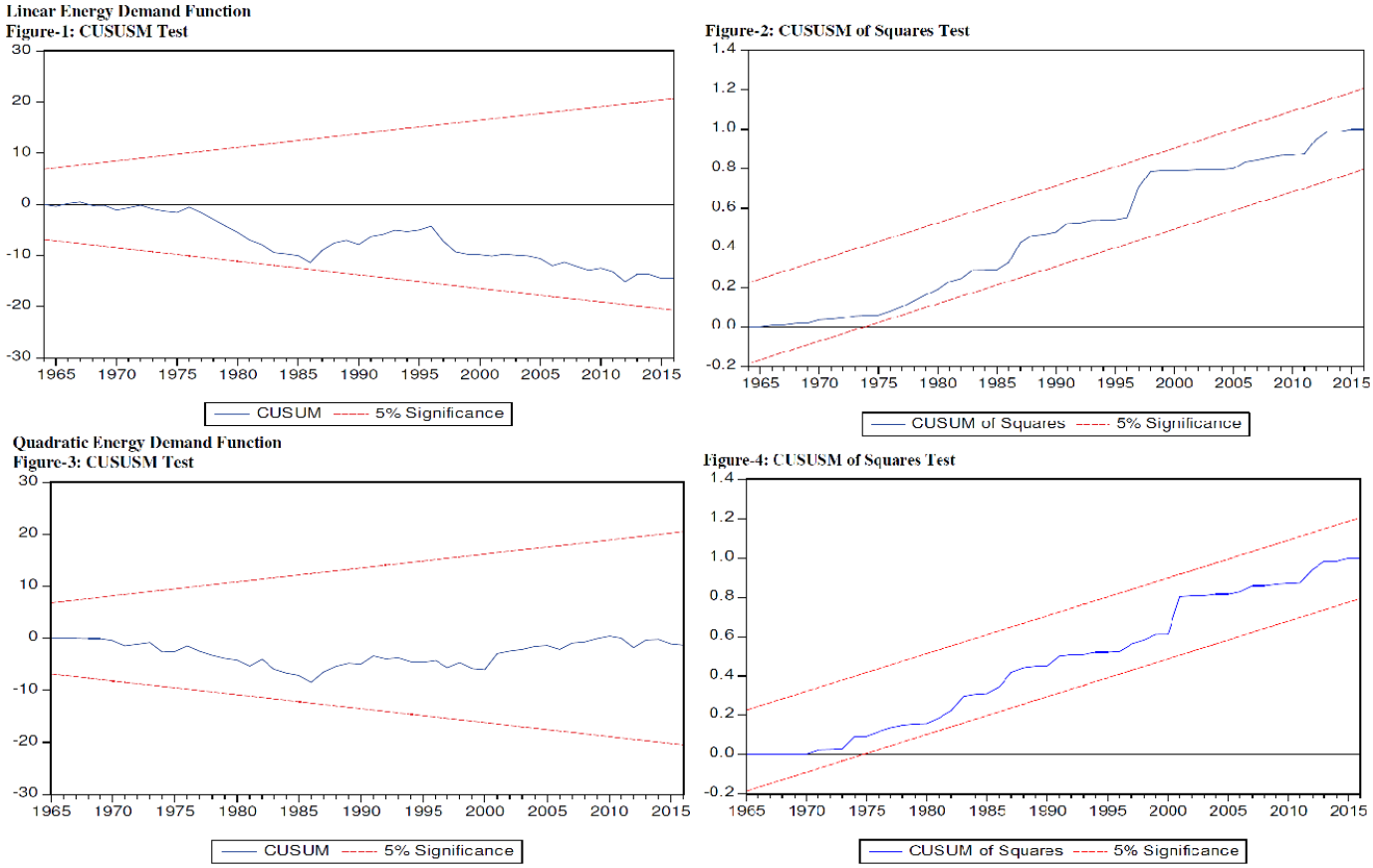

Figure 9: Levels and the changes of the variables

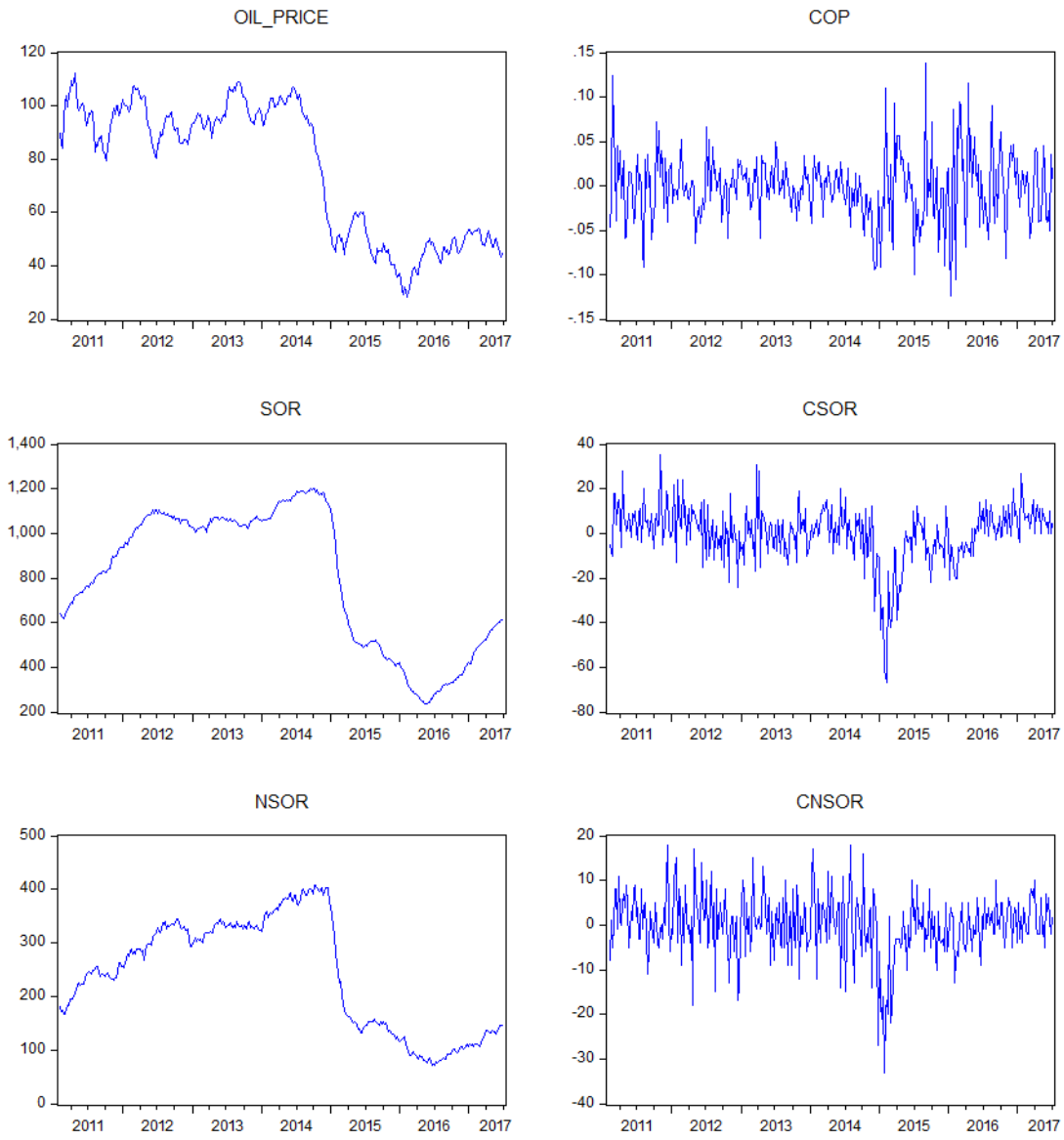

Table 1: Descriptive analyses of monthly growth rates

Table 1: Descriptive analyses of monthly growth rates

	WTI Price	Brent Price	WTI Quantity	Shale Oil Quantity	WTI without Shale Oil	WTI Price	Brent Price	WTI Quantity	Shale Oil Quantity	WTI without Shale Oil
			Full-sample							
Mean	0.003	0.004	0.002	0.012	-0.001					
Median	0.014	0.018	0.002	0.009	-0.002					
Maximum	0.214	0.198	0.175	0.081	0.189					
Minimum	-0.332	-0.311	-0.230	-0.026	-0.264					
Std. Dev.	0.088	0.092	0.032	0.020	0.035					
Skewness	-0.825	-0.892	-2.478	0.618	-2.428					
Kurtosis	4.463	4.181	29.568	3.155	28.727					
			2000M01 - 2011M01					2011M02 - 2017M11		
Mean	0.009	0.010	-0.001	0.007	-0.001	-0.008	-0.008	0.007	0.021	0.000
Median	0.022	0.030	0.000	0.004	-0.001	0.004	-0.003	0.007	0.023	-0.002
Maximum	0.204	0.198	0.175	0.073	0.189	0.214	0.196	0.056	0.081	0.060
Minimum	-0.332	-0.311	-0.230	-0.022	-0.264	-0.245	-0.266	-0.027	-0.029	-0.052
Std. Dev.	0.091	0.097	0.037	0.018	0.041	0.087	0.087	0.017	0.025	0.022
Skewness	-1.078	-1.053	-2.280	0.929	-2.460	-0.318	-0.550	0.470	-0.004	0.339
Kurtosis	5.040	4.217	23.900	4.205	25.010	3.368	4.118	3.212	2.381	2.927

Table 2: Perron unit root test

Table 2: Perron unit root test

	Break in Trend & Intercept				Break in Intercept				Break in Trend			
	Break dates	Lags	Test statistics	C.V. at 5%	Break dates	Lags	Test statistics	C.V. at 5%	Break dates	Lags	Test statistics	C.V. at 5%
WTI Price	2011M08	1	-3.87	-5.59	2014M06	1	-4.16	-5.23	2012M07	1	-3.77	-4.83
Brent Price	2012M06	1	-3.83	-5.59	2014M06	1	-4.07	-5.23	2012M08	1	-3.70	-4.83
Shale oil Q.	2005M03	4	-2.34	-5.59	2011M02	4	-4.25	-5.23	2005M08	4	-2.07	-4.83

[1] The table reports the critical values of the Perron (1997) unit root test in the presence of a structural break in a series intercept and/or linear trend. We consider three different cases: break in both trend and intercept; break in intercept only; break in trend only. The table also includes the optimal break date for each series and the optimal number of lags used for the computation of the test statistic. The null hypothesis of the test is the presence of a unit root.

Table 3: Gregory - Hansen cointegration test results

Table 3: Gregory - Hansen cointegration test results

	ADF^* Test statistics	Break dates	Z_t^* Test statistics	Break dates	Z_a^* Test statistics	Break dates	ADF^*, Z_t^* C.V. at 5%	Z_a^* C.V. at 5%
WTI and Brent								
C	-5.66	2011M02	-5.76	2010M10	-58.69	2010M10	-4.61	-40.48
C/T	-6.04	2010M10	-6.11	2010M10	-65.17	2010M10	-4.99	-47.96
C/S	-6.15	2010M10	-6.25	2011M02	-67.56	2010M10	-4.95	-47.04
WTI, Brent, Shale oil								
C	-6.53	2013M04	-6.61	2013M04	-73.20	2013M04	-4.92	-46.98
C/T	-6.79	2013M04	-6.92	2013M04	-79.06	2013M04	-5.29	-53.92
C/S	-7.26	2011M03	-7.30	2011M02	-86.81	2011M02	-5.50	-58.33

[1] Note that C.V. stand for critical value. C = Level shift, C/T = Level shift with trend, C/S = Regime shift.

Table 4: Full-sample cointegration estimation

Table 4: Full-sample cointegration estimation

Included Variables	WTI Price Brent Price	WTI Price Brent Price Shale Oil Quantity
Lags	2	2
No. of cointegration	1	1
Trace	45.76***	69.66***
P-value	0.0042	0.0003
Critical value at 1%	41.98	64.57
Critical value at 5%	36.06	57.43
Max. Eig.	39.01***	44.92**
P-value	0.0002	0.0246
Critical value at 1%	31.26	49.57
Critical value at 5%	26.10	43.22
Deterministic	IC, IV	IC, IV
Exogenous	D11, D11*T	D11, D11*T
Cointegration equation: $WTI_t - \mu - \beta_1 Brent_t - \beta_2 ShaleQ_t = \epsilon_t$		
WTI Price	1	1
Shale Oil Quantity		0.058*
		(0.023)
Brent Price	-0.968***	-0.983***
	(0.019)	(0.019)
Adjustment coefficients:		
WTI Price	-0.34***	-0.514***
	(0.129)	(0.181)
Shale Oil Quantity		0.044
		(0.032)
Brent Price	-0.040	-0.108
	(0.136)	(0.187)

[1] The first panel includes the cointegration test on the full-sample, and indicates the structure of the VECM model in terms of lags and deterministic and exogenous components, including intercept in the cointegration equation (IC) and test VAR (IV), step dummy from February 2011 (D11) and interaction between the step dummy and a linear trend. Star denotes rejections of the null hypotheses under the appropriate critical values (see Johansen *et al.* (2000) and Giles and Godwin (2012)). The table also reports estimated coefficients, the standard errors (in parentheses). *** denotes the significant level at 1%, ** denotes significant level at 5% and *denotes significant level at 10%.

Table 5: Selected sub-samples cointegration estimation

Table 5: Selected sub-samples cointegration estimation

Included Variables	WTI Price Brent Price	WTI Price Brent Price	WTI Price Brent Price Shale Oil Quantity	WTI Price Brent Price Shale Oil Quantity
Sample	2000M01 - 2011M01	2011M02 - 2017M11	2000M01 - 2011M01	2011M02 - 2017M11
Lags	2	1	2	1
No. of cointegration	1	0	1	0
Trace	23.77***	12.80	42.77***	29.63
P-value	0.0023	0.1222	0.0062	0.1675
Critical value at 1%	19.94	19.94	41.08	41.08
Critical value at 5%	15.49	15.49	35.01	35.01
Max. Eig.	21.98***	11.92	31.06***	18.74
P-value	0.0025	0.1137	0.0054	0.2264
Critical value at 1%	18.52	18.52	29.26	29.26
Critical value at 5%	14.26	14.26	24.25	24.25
Deterministic	IC, IV	IC, IV	IC, IV, TC	IC, IV, TC
Cointegration equation:				
	$WTI_t - \mu - \beta_1 Brent_t - \beta_2 ShaleQ_t = \epsilon_t$			
WTI Price	1		1	
Shale Oil Quantity			0.100***	
			(0.037)	
Brent Price	-0.969***		-0.945***	
	(0.012)		(0.020)	
Adjustment coefficients:				
WTI Price	-0.962***		-0.721**	
	(0.289)		(0.287)	
Shale Oil Quantity			0.111**	
			(0.051)	
Brent Price	-0.698		-0.369	
	(0.308)		(0.322)	

[1] The panel includes the cointegration test on the selected sub-samples, and indicates the structure of the VECM model in terms of lags and deterministic components, including linear trend in the cointegration equation (TC), intercept in the cointegration equation (IC) and test VAR (IV). The table then reports estimated coefficients, the standard errors (in parentheses). *** denotes the significant level at 1%, ** denotes significant level at 5% and *denotes significant level at 10%.

Table 6: Descriptive Statistics and Correlation Matrix

Table 6: Descriptive Statistics and Correlation Matrix

Variables	$\ln E_t$	$\ln P_t$	$\ln Y_t$	$\ln I_t$	$\ln U_t$	$\ln O_t$
Mean	8.8396	3.7081	10.2629	8.9211	4.3137	8.4390
Median	8.9210	3.9930	10.2889	8.9766	4.3066	8.6079
Maximum	9.0405	4.710	10.8641	9.3123	4.4041	9.6381
Minimum	8.3144	2.6705	9.1632	8.0634	4.1934	6.7466
Std. Dev.	0.1862	0.7095	0.4705	0.3259	0.0578	0.8936
Skewness	-1.3808	-0.2006	-0.5930	-0.9261	-0.2193	-0.3411
Kurtosis	3.7866	1.4752	2.3589	2.9673	2.1957	1.8049
Jarque-Bera	2.3018	2.9399	3.0742	2.5809	2.3429	4.2868
Probability	0.3109	0.2910	0.3030	0.3209	0.3099	0.1011
Sum	592.2580	248.4474	687.6165	597.7153	289.0238	565.4142
Sum Sq. Dev.	2.2889	33.2282	14.6160	7.0118	0.2207	52.7115
$\ln E_t$	1.0000					
$\ln P_t$	-0.4536	1.0000				
$\ln Y_t$	0.3997	0.5505	1.0000			
$\ln I_t$	0.6087	0.4179	0.3916	1.0000		
$\ln U_t$	-0.2640	0.3307	0.3835	0.3609	1.0000	
$\ln O_t$	0.4787	0.2821	0.5070	0.2672	0.3723	1.0000

Table 7: Unit Root Analysis

Table 7: Unit Root Analysis

Variables	ADF at Level		ADF at First Difference	
	T-Statistic	P-Value	T-Statistic	P-Value
$\ln E_t$	-1.1850 (1)	0.6863	-6.0729(2)*	0.0000
$\ln P_t$	-1.4598 (2)	0.8332	-4.9564(1)*	0.0008
$\ln Y_t$	-2.7261 (3)	0.2300	-4.8554(2)*	0.0011
$\ln I_t$	-2.2955 (2)	0.1616	-5.2805(3)*	0.0003
$\ln U_t$	-1.2170 (1)	0.8984	-6.5031(2)*	0.0000
$\ln O_t$	-2.7277 (2)	0.2292	-4.0298(1)**	0.0123
Variables	**ADF at Level with Break**		**ADF at First Difference with Break**	
	T-Statistic	P-Value	T-Statistic	P-Value
$\ln E_t$	-2.6976 (1)	1963	-6.8286(2)*	1973
$\ln P_t$	-2.8371 (2)	1973	-5.7514(3)*	1971
$\ln Y_t$	-2.8351 (1)	2007	-6.4389(2)*	1969
$\ln I_t$	-4.1767 (3)	2008	-6.7467(1)*	2009
$\ln U_t$	-4.4004 (2)	1969	-6.0277(3)*	1969
$\ln O_t$	-3.6532 (1)	1972	-8.2910(2)*	1974

[1] Note: The asterisks * and ** represent significant at the 1%, 5%, and 10% levels, respectively. The optimal lag length used are shown in parentheses.

Table 8: The Results of ARDL Cointegration Test

Table 8: The Results of ARDL Cointegration Test

Bounds Testing Approach to Cointegration				Diagnostic Tests					
Estimated Models	Lag Length	Break Year	F-Statistic	χ^2_{NORMAL}	χ^2_{ARCH}	χ^2_{RESET}	χ^2_{SERIAL}	CUSUM	CSUSUMsq
$E_t = f(P_t, Y_t, U_t, I_t, O_t)$	2,2,2,2,2,2	1963	8.957*	0.7160	1.8285	2.7011	0.9076	Stable	Stable
$P_t = f(E_t, Y_t, U_t, I_t, O_t)$	2,2,2,1,2,2	1973	10.880*	0.6017	2.3101	0.4009	1.1006	Stable	Stable
$Y_t = f(E_t, P_t, U_t, I_t, O_t)$	2,2,2,2,2,1	2007	7.886**	0.1563	1.6080	1.1836	2.1032	Stable	Stable
$U_t = f(E_t, P_t, Y_t, I_t, O_t)$	2,2,1,2,1,2	2008	2.405	2.1506	2.0751	0.3261	0.1352	Unstable	Stable
$I_t = f(E_t, P_t, Y_t, U_t, O_t)$	2,2,2,1,2,2	1969	6.997**	1.3300	4.0207	2.1431	0.3143	Stable	Stable
$O_t = f(E_t, P_t, Y_t, U_t, I_t)$	2,1,2,1,2,2	1972	8.901*	1.2901	2.0107	2.1341	0.3104	Stable	Stable
$E_t = f(P_t, Y_t, Y_t^2, U_t, I_t, O_t)$	2,2,2,2,2,2	1963	8.256*	0.7065	1.8080	2.3013	0.9171	Stable	Stable
$P_t = f(E_t, Y_t, Y_t^2, U_t, I_t, O_t)$	2,2,2,1,2,2	1973	9.808*	0.6177	2.3701	0.4039	1.1000	Stable	Stable
$Y_t, Y_t^2 = f(E_t, P_t, U_t, I_t, O_t)$	2,2,2,2,2,1	2007	7.806**	0.1541	1.7060	1.1338	2.1671	Stable	Stable
$U_t = f(E_t, P_t, Y_t, Y_t^2, I_t, O_t)$	2,2,1,2,1,2	2008	2.024	2.1545	2.7010	0.3621	0.1532	Unstable	Unstable
$I_t = f(E_t, P_t, Y_t, Y_t^2, U_t, O_t)$	2,2,2,1,2,2	1969	6.957**	1.3450	4.2007	2.1313	0.3411	Stable	Stable
$O_t = f(E_t, P_t, Y_t, Y_t^2, U_t, I_t)$	2,1,2,1,2,2	1972	8.876*	1.2607	2.155	2.1355	0.3017	Stable	Stable

Significance Level	Critical Values (T=49) Lowe bounds $I(0)$	Upper bounds $I(1)$
1 % Level	6.873	8.163
5 % Level	5.110	6.190
10 % Level	4.330	5.243

[1] Note: The optimal lag length is determined by AIC. Critical values are collected from Narayan (2005). The asterisks * and ** denotes significant at the 1% and 5%, respectively.

Table 9: Long Run Analysis

Table 9: Long Run Analysis

Dependent Variable = ln E_t				
Variables	Coefficient	T-Statistic	Coefficient	T-Statistic
Constant	23.2536*	2.7658	10.2345*	4.3578
$\ln P_t$	-0.5435*	-7.9567	-0.4001*	-7.1714
$\ln Y_t$	0.5437**	2.2347	-6.984*	-4.6075
$\ln Y_t^2$			0.2997*	4.5225
$\ln U_t$	-0.6263*	-7.9756	51.0192*	4.4156
$\ln U_t^2$			-5.8611*	-4.3831
$\ln I_t$	0.5502**	2.4569	-5.5078*	-4.0809
$\ln I_t^2$			0.2949*	3.9884
$\ln O_t$	0.4908*	5.6724	5.1369*	8.2155
$\ln O_t^2$			-0.2882*	-7.9895
D_t	0.1283*	5.5158	0.0179*	4.1768
R^2	0.9429		0.9801	
Adj-R^2	0.9382		0.9774	
F-Statistic	19.8504*		35.8856*	
Durbin Watson	2.0445		1.8298	
Stability Test				
Test	F-Statistic	Prob. Value	F-Statistic	Prob. Value
χ^2_{Normal}	3.6346	0.1624	3.0341	0.1720
χ^2_{Serial}	1.1371	0.2430	2.2371	0.2030
χ^2_{ARCH}	2.0413	0.3531	2.4414	0.3351
χ^2_{Hetero}	1.8179	0.9356	2.8179	0.8316
χ^2_{Remsay}	1.2806	0.2571	2.2806	0.2070
CUSUM	Stable		Stable	
CUSUM	Stable		Stable	

[1] Note: * and ** show significance at 1%, 5%, respectively.

Table 10: Short Run Analysis

Table 10: Short Run Analysis

Dependent Variable = $\Delta \ln E_t$				
Variables	Coefficient	T-Statistic	Coefficient	T-Statistic
Constant	-0.0197***	-1.7447	-0.0046	-0.3927
$\Delta \ln P_t$	-0.1509**	-1.9818	-0.1997**	-2.5213
$\Delta \ln Y_t$	0.8740*	3.8667	0.8063*	2.5819
$\Delta \ln Y_t^2$			-2.2825	-0.4922
$\Delta \ln U_t$	-0.5090	0.3280	-1.1075**	-2.2640
$\Delta \ln U_t^2$			18.3573**	2.3282
$\Delta \ln I_t$	-0.1100	-0.6306	-0.111	-0.5790
$\Delta \ln I_t^2$			1.0387	0.3658
$\Delta \ln O_t$	0.1514*	2.7543	0.1750*	2.6356
$\Delta \ln O_t^2$			-0.1511	-0.4080
D_t	0.0044	0.6632	0.0057	0.7807
ECM_{t-1}	-0.0748*	-2.5678	-0.2488**	-2.2623
R^2	0.7218		0.7725	
Adj-R^2	0.6877		0.7261	
F-Statistic	21.1346*		16.7601*	
Durbin Watson	1.5187		1.6051	
Stability Test				
Test	F-Statistic	Prob. Value	F-Statistic	Prob. Value
χ^2_{Normal}	1.2842	0.5357	0.8851	0.6423
χ^2_{Serial}	2.1896	0.1216	2.4431	0.1098
χ^2_{ARCH}	1.5876	0.2124	2.1343	0.3531
χ^2_{Hetero}	0.5458	0.9555	1.3434	0.9111
χ^2_{Remsay}	0.4875	0.6277	1.1680	0.2476
CUSUM	Stable		Stable	
CUSUM	Stable		Stable	

[1] Note: *, ** and *** show significance at 1%, 5% and 10%, respectively.

Table 11. VECM Granger Causality Analysis

Table 11. VECM Granger Causality Analysis

Dependent variable	Short-run						Long-run			
	$\sum \Delta \ln\ E_{t-1}$	$\sum \Delta \ln\ P_{t-1}$	$\sum \Delta \ln\ Y_{t-1}$	$\sum \Delta \ln\ U_{t-1}$	$\sum \Delta \ln\ I_{t-1}$	$\sum \Delta \ln\ O_{t-1}$	Break Year	ECM_{t-1}	$CUSUM$	$CUSUM_{SQ}$
$\Delta \ln\ E_t$		0.5688	4.6630**	0.2257	0.2077	2.5093***	1963	-0.1207*	Stable	Stable
		[0.5697]	[0.0138]	[0.7987]	[0.8131]	[0.0913]		[-2.7474]		
$\Delta \ln\ P_t$	1.2327		1.9945	5.1226*	2.0711	4.2482**	1973	-0.1609*	Stable	Stable
	[0.3000]		[0.1466]	[0.0095]	[0.1365]	[0.0154]		[-2.5407]		
$\Delta \ln\ Y_t$	3.3890**	1.7366		0.0764	4.5019**	0.4046	2007	-0.2261**	Stable	Stable
	[0.0174]	[0.1864]		[0.9265]	[0.0140]	[0.6693]		[-2.6027]		
$\Delta \ln\ U_t$	2.1217	1.1958	0.1215		0.7666	1.4001	2008	-0.0871	Stable	Unstable
	[0.1279]	[0.3108]	[0.8858]		[0.4698]	[0.2559]		[-1.2492]		
$\Delta \ln\ I_t$	0.1819	1.3009	5.4447**	0.1177		1.1343	1969	-0.3801*	Stable	Stable
	[0.8342]	[0.2812]	[0.0122]	[0.8895]		[0.3296]		[-3.9307]		
$\Delta \ln\ O_t$	0.8196	1.7776	4.1678**	0.0768	1.5241		1972	-0.2440**	Stable	Stable
	[0.4463]	[0.1794]	[0.0165]	[0.9261]	[0.2499]			[-2.2959]		
	$\sum \Delta \ln\ E_{t-1}$	$\sum \Delta \ln\ P_{t-1}$	$\sum \Delta \ln\ Y_{t-1}, \sum \Delta \ln\ Y_{t-1}^2$	$\sum \Delta \ln\ U_{t-1}$	$\sum \Delta \ln\ I_{t-1}$	$\sum \Delta \ln\ O_{t-1}$	Break Year	ECM_{t-1}	$CUSUM$	$CUSUM_{SQ}$
$\Delta \ln\ E_t$		0.6587	4.9879**	0.3227	0.2976	3.9360**	1963	-0.1170**	Stable	Stable
		[0.5491]	[0.0128]	[0.7781]	[0.8001]	[0.0154]		[-2.6476]		
$\Delta \ln\ P_t$	1.1328		1.884	5.2612*	1.9712	4.4480**	1973	-0.1591*	Stable	Stable
	[0.3102]		[0.1487]	[0.0087]	[0.1378]	[0.0144]		[-3.5498]		
$\Delta \ln\ Y_t, \Delta \ln Y_t^2$	4.0987**	1.3376		0.0674	4.5910**	0.5045	2007	0.1921**	Stable	Stable
	[0.0154]	[0.2063]		[0.9625]	[0.0137]	[0.6687]		[-2.7027]		
$\Delta \ln\ U_t$	2.0207	1.295	0.1512		0.7767	1.4041	2008	-0.078	Unstable	Unstable
	[0.1309]	[0.2008]	[0.8781]		[0.4688]	[0.249]		[-1.3490]		
$\Delta \ln\ I_t$	0.108	1.31	5.0047**	0.107		1.103	1969	-0.3615*	Stable	Stable
	[0.8041]	[0.2809]	[0.0145]	[0.8705]		[0.3301]		[-4.0711]		
$\Delta \ln\ O_t$	0.8196	1.7776	4.1678**	0.0768	1.5241		1972	-0.2440**	Stable	Stable
	[0.4463]	[0.1794]	[0.0165]	[0.9261]	[0.2499]			[-2.2959]		

[1] Note: *, ** and *** show significance at 1%, 5% and 10%, respectively.