Economics Bulletin

Volume 38, Issue 2

Long-run co-movements between oil prices and rig count in the presence of structural breaks

Nicoleta Iliescu
Iona College

Abstract

In this article, we investigate the existence of long-run common trends (co-movements) shared by the WTI oil prices and the rig count variable in the US. To test for cointegration, we employ the Engle-Granger two-step procedure, the Johansen cointegration test, and the Gregory-Hansen procedure (which takes into account the possibility of structural breaks in the data). Both the Engle-Granger procedure and Johansen tests cannot find cointegration. However, the Gregory-Hansen procedure rejects the hypothesis of no cointegration at 1% in the specification which allows for structural breaks in the constant term, slope coefficients, and the trend term. In addition to cointegration, we examine the existence of any type of Granger-causality running between the two variables of interest. The Granger test identifies a bidirectional causality running between the two variables.

Citation: Nicoleta Iliescu, (2018) "Long-run co-movements between oil prices and rig count in the presence of structural breaks", *Economics Bulletin*, Volume 38, Issue 2, pages 1171-1179
Contact: Nicoleta Iliescu - niliescu@iona.edu.
Submitted: April 04, 2018. **Published:** June 21, 2018.

1. Introduction

Oil is the most widely-traded commodity by value in the world and oil markets are closely monitored by both theoreticians in academia and practitioners in industries. Oil prices are also very volatile, with levels plummeting from a high of 133.88 dollars per barrel to 41.12 dollars per barrel in less than half a year, between June 2008 to December 2008 or from 105.79 dollars per barrel to 47.22 dollars per barrel between June 2014 and December 2014, to evoke just two very recent turbulences in oil markets. To no surprise, oil prices and their determinants have received considerable attention in the economic literature. The aim of this paper is to contribute to the understanding of the relationship between oil prices and drilling activity, proxied by the rig count variable.

The existing literature which focuses exclusively on the relationship between oil price and rig count is relatively scant. However, all the studies examining this relationship have consistently reached similar conclusions, in that both variables are connected, but the strength or the direction of this relationship vary with the circumstances under consideration in these studies.

Thus, Khalifa *et al.* (2017), by using a quantile regression and quantile-on-quantile model, investigate the link between changes in oil prices and changes in the rig count in the United States with an emphasis on the importance of lags. Their findings are that the relationship is not linear and the impact is stronger from changes in oil prices to changes in the rig count of up to one quarter lag.

In an earlier paper, Mohn and Osmundsen (2008) set up an Error Correction Model to estimate the short run effects and derive a structural equation to capture the long run co-movements between oil prices and the drilling activity pertaining to the Norwegian Continental Shelf. Their main finding is that the short run effects are weak whereas the long run effects are strong, one possible explanation being the highly regulated oil market in Norway.

Aiming at providing evidence on the link between changes in investments and uncertainty, Kellog (2014) investigates both theoretically and empirically the extent to which oil prices react to changes in the drilling activity in Texas. His findings indicate that the drilling activity is very sensitive to and negatively related to the volatility of the oil prices.

Following the same line of inquiry, Toews and Naumov (2015) estimate a 3-dimensional VAR (Vector Autoregressive Model) relating oil prices, drilling activity, and drilling costs for the top 25 largest oil companies in the world based on production volume. Their results illustrate that the effect of an oil price demand shock increases the number of wells drilled and the average cost of drilling significantly and permanently after a lag of two quarters. The effects of changes in the drilling activity or costs appear to have only transitory and smaller effects on the oil prices.

In another paper, Ringlund *et al.* (2008) analyze the relationship between oilrig activity and oil price in various non-OPEC regions by using a dynamic regression setup. They conclude that the strength of the relationship varies across regions, with oilrig activity in North and Latin America reacting faster and stronger to changes in prices compared to all the other non-OPEC regions considered in the study.

In this article, we investigate the possibility of long-run co-movements shared by the WTI oil prices and the rig count in the US. To test for cointegration, we employ the Engle-Granger two step procedure, the Johansen cointegration test, and the Gregory-Hansen procedure (which takes into account the possibility of structural breaks in the data). Both the Engle-Granger procedure and Johansen tests cannot find cointegration. However, the Gregory Hansen procedure rejects the hypothesis of no cointegration at 1% in the specification which allows structural breaks in the constant term, slope coefficients, and the trend term. In addition to testing for cointegration, we examine the existence of any type of Granger-causality either bi-directional or uni-directional running between the two variables of interest. The Granger-causality test identifies a bidirectional causality running between the two variables. This article's contribution to the existing literature is to reconsider the existence of long-run common trends between the oil price and one of its important determinants, the drilling activity, by assuming a model formulation that includes structural breaks. Previous studies (for instance, Khalifa *et al.* 2017) did not identify any cointegration, but without entertaining the possibility of structural breaks, the power of conventional cointegration tests falls sharply which can lead to false conclusions regarding the common long-run behavior of the two variables.

The remainder of this article is organized as follows. Section 2 presents the data, the methodology, and the results. Section 3 briefly discusses the implications of the results and provides the concluding remarks.

2. Data, methodology, and results

The two variables of interest are oil prices and the rig count. Brent and WTI (West Texas Intermediate) are the two most important international crude oil types which serve as benchmarks for international oil prices. Brent crude, extracted from the North Sea, trades typically at prices slightly higher than the WTI crude which makes up the most of the crude oil extracted from North America. One possible explanation for this price differential is the quality difference between the two types of crude. Brent, with its superior characteristics, is less expensive to refine and thus it commands a higher price. We use data on monthly WTI oil prices expressed in dollars per barrel obtained from the Federal Reserve Economic Data. In order to transform it into constant dollars, we use the producer price index data from the same source. The US monthly rig count data is obtained from Baker-Hughes. It represents the total number of oil rigs in operation (either for exploratory purposes or for extracting oil) in the US during a given month. The time span covered in this analysis runs from July 1987 to December 2017. As Figures 1, 2, 3, and 4 illustrate, the original series do not exhibit an explicit trend and their autocorrelation functions are indicative of non-stationary processes, with very slow decaying rates. The graph and autocorrelation function of the first difference of the original data resemble those of stationary time series (see Figures 1, 2, 5, and 6). However, in order to determine the stationarity properties of the two variables, one must go beyond a visual inspection and employ more formal unit root tests.

Figure 1: WTI Oil Price level and 1st difference data

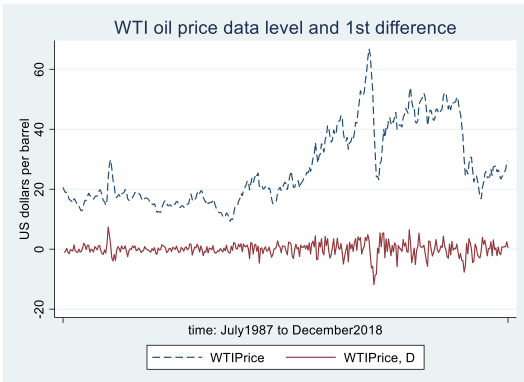


Figure 2: Rig count level and 1st difference data

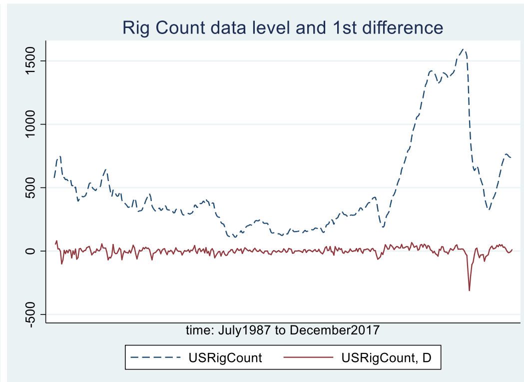


Figure 3: Autocorrelation WTI oil price

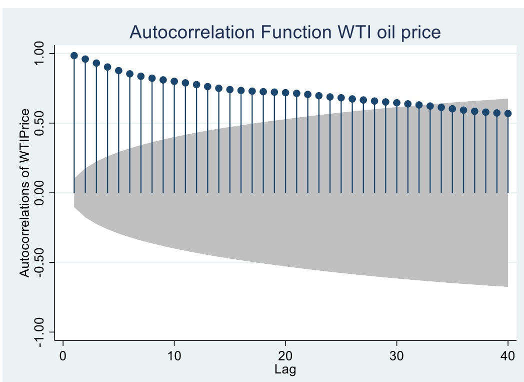


Figure 4: Autocorrelation Rig count

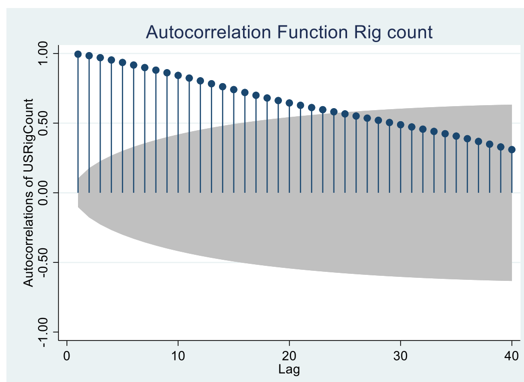


Figure 5: Autocorrelation differenced WTI price

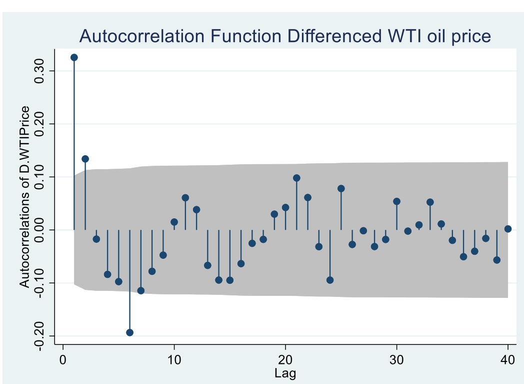
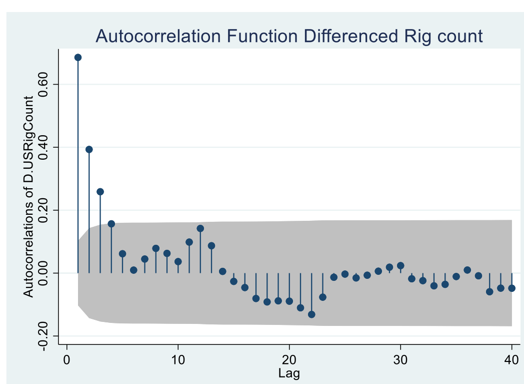


Figure 6: Autocorrelation differenced Rig count



Because of the low power of the unit root tests in detecting non-stationarity, especially in the presence of structural breaks, we conduct a series of unit root tests: the Augmented-Dickey-Fuller (ADF) test, the Phillips-Perron (PP) test, and the Zivot-Andrews (ZA) test.

One of the most widely-used methods for testing the unit root in a time series, the ADF test, is conducted on the following equation:

$$y_t = \alpha + \rho y_{t-1} + \delta t + \sum_{i=1}^{k} \beta_i \Delta y_{t-i} + \varepsilon_t \quad (1)$$

where ε_i is independently and identically distributed error with mean zero and constant variance, y_t is either the WTI oil price or the rig count, y_{t-i} is its lagged value at lag i, α is the drift term, and t is the time trend. The optimal lag length (k) was determined based on the Akaike Information Criterion (AIC). We also conducted the ADF test by using the optimal lag length indicated by the Schwartz Bayes Information Criterion (SBIC) as we did for all the other tests in this paper. Although in most situations AIC and SBIC produce different optimal lag lengths, all stationarity, cointegration, and Granger-causality tests yielded similar results when run based on the two different number of lags. For brevity, we present only the results obtained using the lags given by the AIC, with the other results available upon request. Table I presents the results on 3 different specifications in equation (1): one specification assumes no drift ($\alpha = 0$ and $\delta = 0$), one allows a drift but no trend ($\delta = 0$) and one allows both a drift and trend. In all three specifications, the null hypothesis of a unit root is $\rho = 1$ against the alternative that the times series is stationary. We also run the ADF test on the same specifications and hypotheses on the first differenced data. Table I indicates that at 1% we fail to reject the null hypothesis for the level series and we reject the null for the first differenced data. It is clear that both WTI oil price and rig count are integrated of order 1, *I(1)* processes, i.e., non-stationary at their level but stationary at their first differences.

Table I: Unit Root ADF test results

		Oil Price	Rig Count	1st Difference Oil Price	1st Difference Rig Count
		7 lags (AIC)	12 lags (AIC)	6 lags (AIC)	3 lags (AIC)
Specification ADF Test Statistic (1% Critical Value)	Without Drift (-3.451) With Drift (-2.337) With Trend (-3.986)	-1.797 -1.797 -2.216	-2.235 -2.235 -2.809	-8.409 -8.409 -8.407	-7.124 -7.124 -7.158

For the PP test we start with the same equation as in the ADF test by leaving out the terms that include the lagged values of y_t. The null and the alternative hypotheses in the PP method are similar to the ADF test. The results presented in Table II (for two specifications, without drift and with trend) are consistent with those obtained in the ADF test.

Table II: Unit Root PP test results

		Oil Price	Rig Count	1st Difference Oil Price	1st Difference Rig Count
Specification PP Test Statistic (1% Critical Value)	Without Drift (-3.451) With Trend (-3.985)	-2.130 -2.822	-1.457 -1.862	-13.502 -13.483	-8.267 -8.275

One serious limitation of both ADF and PP tests is that they do not allow for structural breaks in the original time series. However, both the WTI oil price and the rig count test positively for structural breaks and thus the validity of the results obtained in the ADF and PP tests might be compromised. To this end, a third unit root test, the ZA test, is employed to account for structural breaks. The ZA test allows for and endogenously determines the break date (t_{break}) and its test statistics is based on the following equation:

$$y_t = \alpha + \theta \mathrm{D}U_t + \delta t + \gamma DT_t + \rho y_{t-1} + \sum_{i=1}^{k} \beta_i \Delta y_{t-i} + \varepsilon_t \quad (2)$$

Where $DU_t = I(t > t_{break})$ and $DT_t = (t - t_{break})I(t > t_{break})$ and I is the indicator function. Based on equation (2), three specifications are tested: one with break in the level ($\gamma = 0$), one with break in the trend ($\theta = 0$), and one with breaks in both level and trend. The null hypothesis is $\rho = 1$ against the alternative that the times series is stationary. Consistent with the previous outcomes, the results of the ZA test indicate that the original time series is non-stationary across all specifications. Table III summarizes the ZA test results.

Table III: Unit Root ZA test results

		Oil Price	Rig Count	1st Difference Oil Price	1st Difference Rig Count
		1 lag (AIC)	3 lags (AIC)	0 lags (AIC)	2 lags (AIC)
Specification	With break in drift (-5.34)	-4.276	-4.157	-13.846	-7.671
ZA Test Statistic	With break in trend (-4.93)	-4.172	-2.700	-13.591	-7.411
(1% Critical Value)	With break in drift and trend (-5.57)	-4.777	-4.103	-13.846	-7.688

Because both series are *I(1)*, there is a possibility that they are cointegrated, i.e., there exists a linear combination of them which is a stationary process, a property that can be established via cointegration tests. The immediate implication of cointegration is that it allows for common long-run trends in both variables and regressing one variable on the other would not yield spurious results.

For robustness check, we run 3 different cointegration tests, the Johansen procedure (henceforth JP), the Engle-Granger (EG) procedure, and finally, the Gregory-Hansen (GH) procedure.

The JP is testing for the maximum number of stationary linear relationships between variables which are *I(1)*. In a 2-variable model, the number of cointegrating vectors is at most one. The procedure starts with the following VAR:

$$\Delta Y_t = \boldsymbol{\Gamma} Y_{t-1} + \Sigma \boldsymbol{\Gamma}^i \Delta Y_{t-k} + E_t \tag{3}$$

Where ΔY_t is a *2x1* vector of the first differences of the WTI oil price and rig count variables, $\boldsymbol{\Gamma}$ and $\boldsymbol{\Gamma}^i$ are *2x2* matrices of coefficients, k is the number of lags in the VAR system determined by the AIC, E_t is a vector of normally distributed errors.

Matrix $\boldsymbol{\Gamma}$ can be re-written as a product of two vectors such that $\Gamma = \alpha\beta^T$ where $\beta^T = (\beta_1, \beta_2)$. If there exists a linear combination such that $\beta_1 WTIOilPrice + \beta_2 RigCount$ is an *I(0)* process, then WTI oil price and rig count are cointegrated variables. The stationarity of the process is determined by the rank (*r*) of the $\boldsymbol{\Gamma}$ matrix.

The estimation of the parameters of α and β vectors is done by maximum likelihood and so are two log-likelihood tests (trace statistics and the maximum eigenvalue statistics). The null hypothesis of the trace statistic is that there are at most *r* cointegrating vectors versus the alternative that there is one cointegrating vector. As the results in Table IV indicate, we fail to reject the null of no cointegration.

Table IV: Cointegration test-Johansen procedure

Lags in the VAR system	H_0 hypothesis (no cointegration)	H_1 hypothesis (number of cointegrating equations)	Trace statistics	5% Critical Value	1% Critical Value
4 (AIC)	r = 0	r > 0	12.1188	15.41	20.04
	r less than 1	r = 1	3.9887	3.76	6.65

The second cointegration test we employ is the 2-step EG procedure. Because the choice of dependent and independent variables in the EG procedure matters, the following 2 equations are separately estimated at the 1st step:

$$WTIOilPrice_t = a_0 + a_1 t + a_2 RigCount_t + u_{1t} \tag{4}$$

$$RigCount_t = b_0 + b_1 t + b_2 WTIOilPrice_t + u_{2t} \tag{5}$$

Two specifications are tested, one without trend ($a_1 = 0$ and $b_1 = 0$) and one with trend.

At the 2nd step of the procedure, the OLS residuals, $\hat{u}_{it}$, are tested for stationarity by conducting an ADF test in the equation below:

$$\Delta\hat{u}_{it} = \rho_{i0}\hat{u}_{it-1} + \Sigma\rho_{ij}\Delta\hat{u}_{it-j} + \vartheta_{it} \tag{6}$$

With *i = 1,2.* The number of lags of the residuals (*j*) in equation (6) was determined by the AIC.

If $\hat{u}_{it}$ are stationary, then WTI oil price and rig count are cointegrated. Therefore, the null hypothesis is that the residuals $\hat{u}_{it}$ are non-stationary (no cointegration) and the alternative hypothesis is that the residuals $\hat{u}_{it}$ are stationary (variables are cointegrated).

Table V summarizes the results from the 1st and 2nd steps of the EG procedure. In line with the results obtained in the JP, a cointegrating relationship cannot be established between WTI oil price and rig count.

Table V: Cointegration test-Eagle Granger (EG) procedure

Lags- 2nd step regression	Trend- 1st step regression	Estimated model-1st step regression	Test statistic value (1% critical value)- 1st step regression	Cointegrated Variables
12 (AIC)	No	WTI = 18.259 + 0.018RigCount	-1.906 (-3.927)	No
12 (AIC)	Yes	WTI = 9.806 + 0.069t + 0.010RigCount	-1.926 (-4.370)	No
7 (AIC)	No	RigCount = 81.826 + 14.914WTI	-1.753 (-3.927)	No
7 (AIC)	Yes	RigCount = 66.919 + 0.478t + 12.247WTI	-1.884 (-4.370)	No

Similar to the EG test, the third cointegration test conducted in this study, the GH procedure, is also a residual-based cointegration test but additionally, the GH test allows for the presence of a structural break in the series. As Gregory and Hansen (1996) argue, conventional ADF tests fail to detect cointegration in the most instances where time series exhibit structural breaks. It is also the case of our time series, where both the JP and EG tests have failed to show evidence of cointegration.

In this article, we will estimate three specifications of the GH procedure, one which assumes a break in the constant term (labeled level in equation (7)), another one which assumes a break in the constant term and slope coefficient (labeled regime in equation (8)), and finally, one which assumes a break in the constant term, slope, and trend (labeled regime, trend in equation (9)). The optimal lag length in the ADF stationarity test of the residuals was obtained via AIC. The null and the alternative are the same as in the EG procedure described above. The 1st step regression equations for the three specifications are:

$$y_t = c_0 + c_1 D_t + c_2 x_t + u_t \text{ (level)} \quad (7)$$

$$y_t = c_0 + c_1 D_t + c_2 x_t + c_3 x_t D_t + u_t \text{ (regime)} \quad (8)$$

$$y_t = c_0 + c_1 D_t + c_2 x_t + c_3 x_t D_t + c_4 t + c_5 t D_t + u_t \text{ (regime, trend)} \quad (9)$$

Where $y_t / x_t =$ WTI oil price/rig count or rig count/WTI oil price, $D_t = I(t < t_{break})$, and $I =$ indicator function.

The 2nd step of the GH procedure is conducted in the same manner as in the EG procedure. Table VI illustrates that the GH procedure also fails to reject the null of no cointegration at 1% in two specifications, namely those labeled level and regime. However, in the third specification (labeled regime, trend) the null of no cointegration is rejected at 1%, thus the time series appear to be cointegrated. The implication of this result is that there are long-run co-movements between the two variables but they are detectable only in the model where the cointegrating vector is allowed to be time-variant at the constant term, slope coefficient, and trend level at an unknown point in time.

Table VI: Cointegration test-Gregory-Hansen procedure (for time series with structural breaks)

Dependent variable - independent variable (1st step GH test) /lags in the ADF test (2nd step GH test)	Break specification	Break date	ADF Test statistic (1% critical value)	Cointegrated variables
WTI-Rig Count/ 2 lags (AIC)	level	March 2004	-4.57 (-5.13)	No
WTI-Rig Count/ 2 lags (AIC)	regime	March 2004	-4.62 (-5.47)	No
WTI-Rig Count/ 2 lags (AIC)	**regime, trend**	**August 2005**	**-6.64 (-6.02)**	**Yes**
Rig Count-WTI/ 2 lags (AIC)	level	December 2010	-3.40 (-5.13)	No
Rig Count-WTI/ 2 lags (AIC)	regime	December 2010	-5.40 (-5.47)	No
Rig Count-WTI/ 2 lags (AIC)	**regime, trend**	**December 2010**	**-6.88 (-6.02)**	**Yes**

Although the cointegration tests could not establish cointegration across all tests and all specifications with one exception, running a Granger-causality test is a valid exercise and it would determine the direction of causality (if any) between the two variables. The test is conducted on the following VAR:

$$WTIOilPrice_t = \Sigma\lambda_{11,p} WTIOilPrice_{t-p} + \Sigma\lambda_{12,p} RigCount_{t-p} + e_1 \quad (10)$$

$$RigCount_t = \Sigma\lambda_{21,p} WTIOilPrice_{t-p} + \Sigma\lambda_{22,p} RigCount_{t-p} + e_2 \quad (11)$$

Where p is the number of lags included in each equation determined by AIC. If RigCount Granger-causes WTIOilPrice then $\lambda_{12,p}$ are jointly significantly different from zero in equation (10). If WTIOilPrice Granger-causes RigCount then $\lambda_{21,p}$ are jointly significantly different from zero in equation (11). These are tested by performing an F-test of the null hypothesis that $\lambda_{12,p} = 0$ (from equation 10) and $\lambda_{21,p} = 0$ (from equation 11), respectively. The null in both tests would indicate that RigCount DOES NOT Granger-cause WTIOilPrice (in equation 10) and WTIOilPrice DOES NOT Granger-cause RigCount (in equation 11).

The results of the Granger-causality test are presented in Table VII. It is clear that the causality is bidirectional, from the oil price to rig count and vice versa.

Table VII: Granger-causality test results

Lags	Null hypothesis (H_0)	p-value	Conclusions (***1% significance level; **5% significance level)
4 (AIC)	Rig Count ≠> WTI Oil Price	0.045	**Reject H_0 ➔ Rig Count Granger-causes WTI Oil Price
4 (AIC)	WTI Oil Price ≠> Rig Count	0.000	***Reject H_0 ➔ WTI Oil Price Granger-causes Rig Count

3. Conclusions and implications

In this study, we investigated the cointegration properties of two variables, WTI oil prices and the rig count in the USA. We employed a set of three cointegration tests and for each test ran various specifications. Both EG procedure and JP failed to reject the null of no cointegration at any significant level and across all specifications considered. Because these two tests perform poorly in general (i.e., cannot find cointegration where in fact it exists) in the presence of structural breaks in the data, we also conducted the GH procedure which allows for structural breaks in either the constant term and/or slope coefficients and/or time trend. Our results indicate that the GH procedure detects cointegration between the two variables in the specification that allows for breaks in the constant term, slope coefficients, and trend. Cointegration signifies co-movements in the long-run and although the path of the two variables can be quite divergent in the short-run, the cointegration will force each variable to ultimately correct itself towards their long-run equilibrium relationship. The policy implication of cointegration is that each variable could assist in predicting and explaining the long-run behavior of the other one.

The results of the Granger-causality test illustrate the fact that there is a bi-directional causality running between the two variables. The implication of the bi-directional Granger-causality is that lagged values of each variable add predictive power in the estimation of the other variable, a feature that can be exploited in forecasting.

References

Engle, R. F. and C.W. J. Granger (1987) "CoIntegration and Error Correction: Representation, Estimation, and Testing" *Econometrica* **55(2)**, 251-276.

Gregory, A.W. and B.E. Hansen (1996) "Residual-based tests for cointegration in models with regime shift" *Journal of Econometrics* **70(1)**, 99-126.

Johansen, S. (1996) "Likelihood-based Inference in Cointegrated Vector Autoregressive Models" Oxford University Press: Oxford.

Kellogg, R. (2014) "The effect of uncertainty on investment: evidence from Texas oil drilling" *American Economic Review* **104(6)**, 1698–1734.

Khalifa, A., M. Caporin and S. Hammoudeh (2017) "The relationship between oil prices and rig counts: The importance of lags" *Energy Economics* **63(1)**, 213–226.

Mohn, K. and P. Osmundsen (2008) "Exploration economics in a regulated petroleum province: the case of the Norwegian Continental Shelf" *Energy Economics* **30(2)**, 303–320.

Ringlund, B., E. Rosendahl and T. Skjerpen (2008) "Does oil rig activity react to oil price changes? An empirical investigation" *Energy Economics* **30(2)**, 371–396.

Toews, G. and A. Naumov (2015) "The Relationship Between Oil Price and Costs in the Oil and Gas Industry" *Oxford Centre for the Analysis of Resource Rich Economies* 152, University of Oxford, U.K.

Zivot, E. and D. W. K. Andrews (1992) "Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis" *Journal of Business & Economic Statistics* **10(3)**, 251-270.