UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Scienze Economiche ed Aziendali "Marco Fanno"

THE LONG-RUN OIL-NATURAL GAS PRICE RELATIONSHIP AND THE SHALE GAS REVOLUTION

MASSIMILIANO CAPORIN
University of Padova

FULVIO FONTINI
University of Padova

July 2015

"MARCO FANNO" WORKING PAPER N.198

The Long-Run Oil-Natural Gas Price Relationship and the Shale Gas Revolution

Massimiliano Caporin* **Fulvio Fontini***

May 2015

Abstract

The gas extraction technological developments of the 2000s have allowed shale gas production, which in the US has become a significant part of the total gas production. Such a significant change might have affected the long-run relationship between oil and natural gas prices postulated by several authors. By using monthly data of oil and gas prices, as well as gas quantities from 1997 to 2013, we test for the presence of a long-run relationship, allowing also for possible breaks. We first show the stationarity of gas quantity data before the production of shale gas and the existence of a break in the trend (and in the the intercept) on the integrated gas price time series, by the time shale gas enters the market. Then, applying a Vector Error Correction Model, we show that shale gas production has affected the relationship across variables. Gas quantities become relevant in the formation of gas prices after the beginning of shale gas production, while impact of oil prices on the gas ones doubles. However, on the basis of the available data, it is not unequivocally possible to assess whether or not a new long-run relationship between oil and gas has been established.

Keywords: Shale Gas, Natural Gas, Crude Oil, Cointegration, Vector Error Correction Models.
JEL Classification: C01, C32, Q40, Q41.

*University of Padova (Italy), Department of Economics and Management

1 Introduction

From as early as 2005, technological developments, such as the extraction of shale gas from shales using hydraulic fracturing and horizontal drilling, began making the supply of natural gas (from now on “gas” for brevity) to the market economically convenient in the United States (US).[1] Shale gas has been traded in the market, in particular at the US Henry Hub, from the beginning of 2007. Since the end of 2008 US gas prices started to decrease, yielding a tremendous competitive advantage for the country’s manufacturing and chemical industries and eliminating the need to import gas. This development is often referred to as the ”shale gas revolution”, highlighting its importance and hinting at a fundamental (and perhaps irreversible) change in the long-run gas price, and possibly in the dynamic evolution of the gas price over time. Indeed, the so-called shale gas revolution should imply a permanent impact on the trend of gas prices that would be determined by the increase in the gas supply due to the introduction of shale gas. Deviations in the price from the trend should, therefore, be seen as temporary and could be determined by several contingent factors impinging on the gas market to be identified case-by-case (Brown and Yucel, 2008). This article aims at evaluating the alleged revolutionary impact of shale gas on the gas price trend. It is well known that oil and gas prices are linked (Hartley et al., 2008). The degree of correlation depends on price formation rules, contractual arrangements, markets structures and liquidity, justified by technical and economical reasons (Villar and Joutz, 2006). Among the first the need to hedge huge investment risks in the cultivation of oil and gas fields (which are often joint sites), while the second refers to the principle of demand substitutability between oil and gas products, as well as the competition for the same scarce inputs. This can give rise to a covariation of oil and gas prices, even in the absence of a contractual indexation of gas contracts to oil since, a) in the final products, oil and gas are often substitutes and b) when developing non-joint fields, companies compete in the same markets for technical and financial resources, so that, for instance, when the oil price spikes, the cost of inputs for gas field development

[1] A similar technological development has occurred in the oil industry, called shale oil

rises, which in turn increases gas prices. This latter effect can be particularly relevant for shale gas (and oil) development, due to its (their) short time to market and the high financial leverage of the relatively small companies that cultivate it (them), particularly in the case of the US. Several scholars have addressed the question of the oil-gas price relationship. Ghouri (2006), for instance, finds that oil and gas prices do have a long-run relationship. Rosthal (2010) specifies that periods of temporary gas over-supply, e.g., the warm winter of 2006-2007, can determine the short-run departures from the long-run equilibrium between oil and gas. It also assesses that causality works from oil to gas, but the opposite is not true. Pindyck (2004), conducting Granger causality tests, finds a similar result. Foss (2007) conjectures that an oil-gas relationship exists that weakens over time, since gas prices seem to be determined by market fundamentals while oil prices are more volatile. Erdos (2012) finds that the oil and gas market relationship starts to weaken from 2002. Romagus (2012) set the end of the long-term price relationship between oil and gas prices at 2008. All of these studies leave open the question if and when the entrance of shale gas into the market has started to determine a structural change in the market behavior, and in particular, whether it has affected gas price formation, so that the long-run relationship between oil and gas prices, if present, has been permanently affected. Indeed, Wakamatzu and Aruga (2013) ask a similar question for the case of the US and Japan. Applying a structural break test, they set the structural change for US gas at 2005. However, they do not investigate further the nature of the long-run relationship and the impact of shale gas on a new long-run relationship, if present. This is the aim of our study. In order to fulfill this, we need to identify the long-run characteristics and properties of the gas prices, oil prices and gas quantities time series. Indeed, it is well known (Pindyck, 1999; Villar and Joutz, 2006) that energy data might be non-stationary. Modeling energy data without taking into account possible non-stationarity can lead to spurious results. A non-stationary variable is integrated of order 1 (call it $I(1)$) if its first difference is stationary, in which case we can say that a unit root exists. When some variables are $I(1)$, it is possible to test the existence of a stationary linear relationship among them, i. e. a cointegration relationship. The latter denotes the existence of a long-run relationship among the non-stationary variables. In the following, we first assess the stationarity property of the gas prices time series, the gas quantities time series and the oil prices time series, paying attention to evaluate whether the entrance of the shale gas into the market has determined a structural break in the series, and, if so, of what nature. Subsequently, we first test for the existence of a cointegration relationship across the variables and then set a Vector Error Correction Model (VECM) to identify the long-run relationships between these variables. The sign and magnitude of the estimated parameters show the role played by shale gas quantities and oil prices in the gas prices time series long-run property. The paper is structured as follows. Section 2 presents the preliminary data analysis to evaluate data stationarity (subsection 2.1) and the methodology used for the cointegration analysis (subsection 2.2). Section 3 describes the results of the cointegration analyses. Final remarks and references follow.

2 Data and Methodology

2.1 Data analysis

The empirical analysis is based on the US natural gas market data. We use monthly observations of the following variables: Natural Gas Real Spot Prices (USD/MMBtu) at Henry Hub; Natural Gas Quantities (Tcf) including Shale Gas, i.e. US natural gas gross withdrawals; Crude Oil Real Spot Prices (dollars per barrel) at WTI – Cushing, Oklahoma. The observation period ranges from January 1997 to December 2013, for a total of 204 observations per variable.[2] We also construct the series of natural gas quantities without shale gas subtracting in each period the quantity of shale gas from the overall natural gas quantities, i.e. US natural gas gross withdrawals less US natural gas gross withdrawals from shale. Note that the gas quantity presents a mild seasonal pattern

[2]Data for 2014 are not included, since shale gas production data for 2014 had not yet been released by the time this article was completed. Shale gas data are released with more than a one-year delay.

Table 1: Descriptive analyses of monthly growth rates

	Gas Price	Oil Price	Gas Quantity	Gas Price	Oil Price	Gas Quantity
			Full sample			
Mean	0.003	0.009	0.001			
Median	0.000	0.019	0.002			
Maximum	0.477	0.207	0.064			
Minimum	-0.470	-0.342	-0.093			
St. Dev.	0.137	0.087	0.017			
Skewness	-0.095	-0.818	-0.916			
Exc. Kurtosis	4.322	4.898	10.208			
		1997-2006			2007-2013	
Mean	0.008	0.010	0.000	-0.004	0.007	0.003
Median	0.006	0.021	0.002	-0.022	0.017	0.003
Maximum	0.477	0.204	0.039	0.376	0.207	0.064
Minimum	-0.470	-0.191	-0.078	-0.299	-0.342	-0.093
St. Dev.	0.151	0.081	0.014	0.114	0.096	0.021
Skewness	-0.328	-0.214	-1.371	0.549	-1.298	-0.778
Exc. Kurtosis	4.190	2.629	10.366	3.784	6.193	8.647

which emerges by looking at the correlograms. This deterministic component has been eliminated by means of a simple multiplicative approach. All of the results are provided for the seasonally adjusted total gas production. In the following, as needed, we split the full data sample into sub-samples to take into account the introduction of shale gas into the market.

Figures 1 and 2 display the time evolution of oil and gas prices, which show high differences in the analysed sample. In particular, we point out that while the gas price seems to be characterized by a decreasing trend in the second part of the sample, in the same period (up to 2013), the oil price sensibly increases, partially recovering the big drop of 2007-2008.[3] This is confirmed by Table 1, where the gas price growth rates are, on average, negative in the second subsample, while the oil price growth rates are positive. A second notable finding is the evident change in gas production: the gas quantity increases

[3]Notice that the data for 2014 are not included in the analyses, although oil prices have shown a high decrease. We excluded the oil data for 2014 from the analyses because of the lack of shale gas production data for the same period.

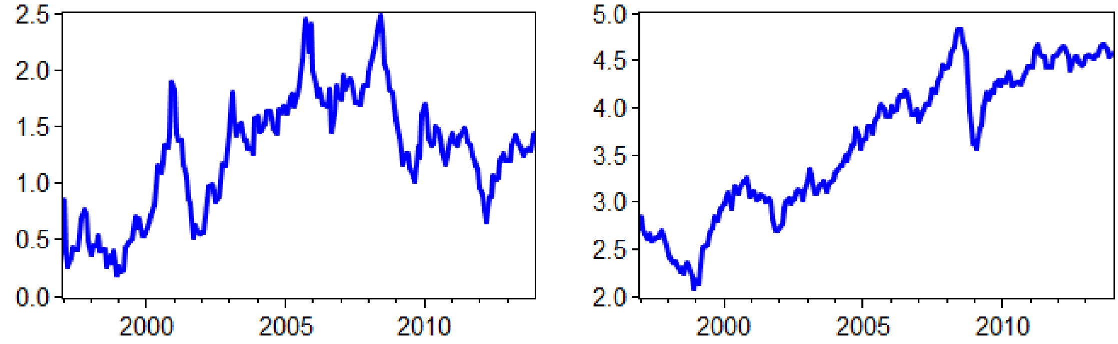

Figure 1: Natural Gas (left) and Oil (right) monthly real log-prices from 1997 to 2012.

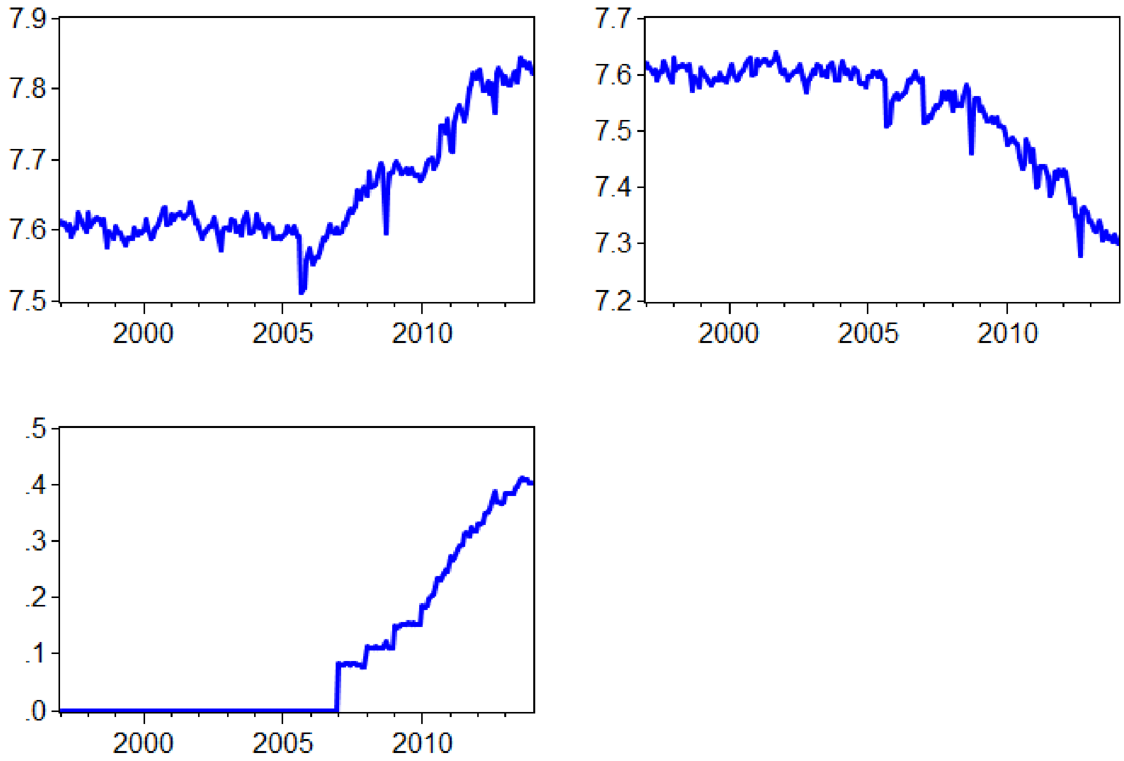

Figure 2: Natural gas production: seasonally adjusted log-level including shale gas production (top left); seasonally adjusted log-level excluding shale gas production (top right); relative weight of shale gas production (bottom).

from 2007 onward with an upward sloping trend. This is due to the increase in shale gas production that reaches about 40% of the total quantity of gas produced at the end of our sample.

Looking at the series behavior, we note that, in the full sample, both the prices as well as the quantity log-levels might be integrated. Such a potential finding would be, on the one hand, a pre-requisite for the possible existence of a long-run equilibrium relationship between the variables; on the other hand, it would suggest that the modeling of the series should be considered on their first difference; that is, on their growth rates. In the graphic evaluation of the data, we pointed out the existence of an evident change in the trend for the gas quantity. This indicates the need to verify the null hypothesis of non-stationarity, i.e. integration, by means of tests which are valid also in the presence of a break in the trend and/or in the intercept of a given random variable. In our case, the break date is known, and is located, roughly, at the end of 2006 since shale gas production (as monitored by the shale gas production data we consider) starts at the beginning of 2007, while it is null before that date. However, the exact identification of the break might not be precise. The market could anticipate, even partially, the effect of shale gas production on both prices and quantities, due to the announcement effect, or might postpone its impact due to the rolling over of the investments and their time-to-market.

Table 2 reports the results of the Perron (1997) unit root test that allows for a break at an unknown location either on the intercept, on the trend, or on both the trend and the intercept. The test also accounts for an additional dynamic in the series, represented by the dependence of prices or quantities on their past values. We report in Table 2 the lag structure identified by the test procedure. Moreover, Table 2 also indicates the estimated break date. We find empirical evidence supporting the existence of a unit root for all the variables under study. The break dates oscillate between 2005 and 2008. However, by looking at the plots in Figure 3, adopted within the test procedure to identify the break date in the specific case of gas prices, we observe that the test statistic behavior has a minimum in a range from 2005 and 2008. Notably, the end of 2006 is just in

Table 2: Perron unit root tests

	Break in Trend and Intercept			Break in Intercept			Break in Trend		
	Break date	Lags	T-stat	Break date	Lags	T-stat	Break date	Lags	T-stat
Oil Real Price	2008-9	2	-4.82	2008-7	2	-4.93	2007-11	2	-4.38
Natural Gas Real Price	2005-5	9	-4.43	2008-6	9	-3.61	2006-12	9	-4.21
Gas Quantity Total	2005-6	4	-4.82	2010-7	4	-2.29	2006-9	4	-3.41
Gas Quantity Without Shale	2008-1	11	-3.03	2009-11	11	-1.86	2007-12	11	-2.71

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, columns 5 and 8. 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. For the break in both intercept and trend the 1% critical value is -6.32 and the 5% critical value equals -5.59. In the case of break only in the intercept, the critical values are equal to -5.92 and -5.23 at the 1% and 5% confidence levels, respectively. Finally, in the case of a break only in the trend, the two critical values are, -5.45 at the 1% level and -4.83 at the 5% level.

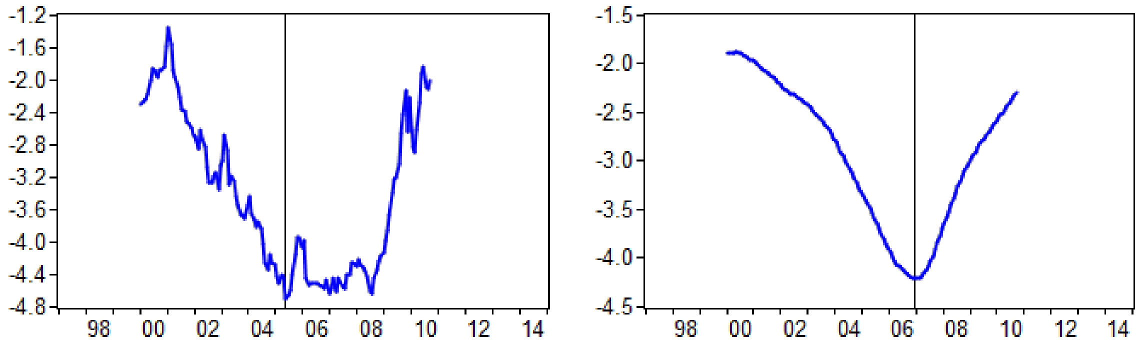

Figure 3: Perron (1997) unit root test for alternative break dates in the case of natural gas real log price with a break in trend and intercept (left) and in the trend only (right).

the middle of this period. This confirms the location of the break and presents three different interpretations: the possible anticipation in the price behaviors of shale gas production, whose impact starts to emerge before January 2007; the need for a period of adjustment after the shock before the market reverts back to a long-term behavior; and the need of an adjustment period after the beginning of shale gas production before the market reaches a new long-run equilibrium relationship. The distinction between the last two cases is of crucial interest and corresponds to the absence/existence of the shale gas revolution. Reverting back to a previous equilibrium (after a relevant but transitory shock) will evidence a temporary effect of the shale gas revolution, while a new equilibrium relationship denotes a structural change and a permanent impact of the shale gas revolution. The fact that the adjustment is not immediate can easily be motivated by the need for some time before the market perceives shale gas extraction as a mature technology with a stable (as opposed to a rapidly changing) production.

To shed further light on the possible change in the series behavior before/after the shale gas revolution, we run unit root tests on selected sub-samples, reported in Table 3.

We first consider the two subsamples we identify on the basis of the break date, namely, from 1997 to 2006 and from 2007 to 2013. Then, we change the possible break date of the full sample by first considering the possible market anticipation of the shale gas impact, and second, its possible delayed (or transitory) impact. We stress the latter

Table 3: Augmented Dickey-Fuller and Philips-Perron unit root tests.

	ADF test			PP test		
Range	Gas Price	Gas Quantity	Oil Price	Gas Price	Gas Quantity	Oil Price
1997-2003	0.2683	**0.0002**	0.4190	0.1771	**0.0002**	0.3790
1997-2004	0.1608	**0.0001**	0.3145	0.1108	**0.0001**	0.2656
1997-2005	0.2125	**0.0008**	0.2774	0.1029	**0.0008**	0.2489
1997-2006	0.0701	**0.0008**	0.1703	0.6457	**0.0012**	0.1497
2005-2007	0.1546	0.5792	0.9859	0.1375	0.6861	0.5384
2005-2008	0.1974	0.0664	0.0821	0.1318	0.0931	0.9311
2006-2008	0.1048	**0.0039**	0.1581	0.0845	**0.0039**	0.9657
2007-2013	0.4253	**0.0018**	**0.0114**	0.5492	**0.0022**	0.2300
2008-2013	0.3115	**0.0048**	**0.0210**	0.7052	**0.0048**	0.3053
2009-2013	0.4499	0.0871	0.2115	0.4499	0.0993	0.2311
2010-2013	0.5241	0.1249	0.2282	0.7704	0.1489	0.2580

P-values of the ADF unit root test. The testing equations includes the optimal number of lags as identified by means of the Schwarz information criteria and trend and intercept if they were statistically significant at the 5% confidence level. Boldface values indicate samples and series for which there are evidences of stationarity at the 5% confidence level.

is also motivated by the evident jumps in the quantity at the end of 2007 and 2008 (see Figure 2). Before commenting on the results, we point out that we adopt the asymptotic critical values of the test statistics. Consequently, the true p-values might differ from those reported, being in general larger since short sample critical values are normally characterized by thicker tails.

We have some relevant findings. First of all, the gas quantity is stationary in the first sub-sample, irrespective of the ending year (from 2003 to 2006) and for both test statistics. Then, we have evidence of non-stationarity in the 2005-2007 and 2005-2008 ranges, but this might be due to the jump in the quantity and, thus, should not be considered much as a true evidence of non-stationarity. If we move to the most recent years, we note that the gas quantity becomes non-stationary once shale gas production reaches a significant fraction of total gas production. The most relevant evidence of non-stationarity is provided from 2009. The stationarity of the gas quantity has a relevant consequence on the existence of long-run relationships. In fact, we cannot postulate the possibility of cointegration between gas price, oil price and gas quantity from 1997 to 2006 as the quantity was, in that range, stationary. The gas price is always non-stationary, independent on the period considered. The result holds for both unit root tests. In contrast, the oil price shows some mild evidence of stationarity in the second part of the sample for the ADF test. This evidence reduces in the case of the Philips-Perron test. We thus conclude that both the gas and the oil price are integrated over the various samples we consider. Moreover, we also have some evidence of a transitory instability in the test outcomes when the data for 2007 and 2008 are included in the computation of the test statistics.

Summarizing the findings of the preliminary data analysis, we first see that there is evidence of non-stationarity over the full sample and for all of the three variables under study. However, the Perron unit root tests in the presence of unknown breaks confirm the existence of a break that coincides with the beginning of shale gas production, or better, with the beginning of the impact of shale gas on the time series. When splitting the data into different subsamples, to take into account different possible starting dates of shale gas impacting the market, the gas quantity emerges as non-stationary only for the most recent years, i.e. focusing on the second sub-sample, and in particular, when considering data from 2009 onward. This contrasts with the full-sample analyses. We interpret it as possible evidence of an instability due to the start of shale gas production that might not have been fully recovered even in the last years. In contrast, we see that the gas quantity is stationary in the range 1997-2006. Moreover, both oil and gas log-prices are stationary irrespective of the samples considered, even if for oil we do have some mild evidence of instability in the test outcome when the sample period starts just after the entrance into the market of shale gas production in 2007.

2.2 Methodology

We are interested in evaluating the existence and the strength of long-run relationships across the three considered variables. We thus need to evaluate the possible existence of cointegration between the prices of oil and gas, or among prices and total gas quantity.

For this purpose, we refer to the tests of Johansen (1988, 1996). The presence of a single cointegrating relationship allows for taking into account the VECM described as:

$$\Delta X_t = \alpha\beta' X_{t-1} + \sum_{j=1}^{p} \Phi_j \Delta X_{t-j} + \delta D_t + \varepsilon_t \tag{1}$$

where X_t is the vector of the modeled variables (in logs), $\beta' X_{t-1}$ is the disequilibrium error, and β contains the cointegration coefficients; that is, the coefficients of the long-run relationship between the variables. The vector α contains the adjustment coefficients to past disequilibrium, while the summation monitors the short-run dynamic of the series growth rates. Finally, D_t contains a set of deterministic variables, namely a constant and a linear trend. For the present paper, we have verified that the most appropriate specification of the model is that satisfying a general restriction on the coefficients δ leading to a model with the possible presence of a trend in the cointegrating equation, thus excluding the possible presence of a quadratic trend in the variables. Therefore, we impose $\delta = \alpha\gamma$, leading to a disequilibrium equal to $\beta' X_{t-1} + \gamma D_t$. The changes in the series behaviors shown in the previous section indicate the need to account for a possible break in the long-run analyses. This topic has been discussed in the econometrics literature by several authors, including Hansen (1992), Gregory and Hansen (1996a,b), Gregory, et al. (1996), Hansen and Johansen (1999), Johansen et al. (2000), Hansen (2003) and Carrion-i-Silvestre and Sanson (2006). When dealing with cointegration across variables, structural breaks might have an impact on the various elements characterizing the VECM model: changes in the number of cointegrating relationships (instead of the assumed single relationship), changes in the long-run coefficients β, changes in the deterministic behaviors coefficients γ, changes in the adjustment coefficients α, changes in the short-term dynamic (either the lag order p or the coefficient matrices Φ_j). This will be analysed in the next section.

3 The Gas-Oil Long-Run Relationship: Empirical Analysis

We start the empirical analysis by evaluating the possible presence of cointegration in the full sample. We first test for the existence of a long-run relationship between gas and oil prices. Then, we test for cointegration in the presence of a third variable, total gas quantity. Given the evidence shown in the previous section which denotes the existence of a break in the deterministic components, we test for the presence of cointegration allowing for a change in the trend slope, following the approach of Johansen et al. (2000). We thus test for full sample cointegration across the gas and oil prices by introducing a break in the form of a step dummy assuming value 1 from January 2007, and interacting the dummy with the cointegration equation intercept and trend (the latter being included in the model only from 2007). Results are reported in the first column of Table 4, where we provide the Johansen (1988, 1996) test statistics. Notice that, as shown in Johansen et al. (2000), the critical values of the Trace test are influenced by the presence of the break in the deterministic components. The Trace test 1% critical value for the presence of a single cointegration relationship and a break date located after about 60% of the total sample equals 23.3 (see Johansen et al., 2000, and Giles and Godwin, 2012). We thus have evidence for the presence of a full-sample cointegration when accounting for a break in the linear trend, since the test statistics equal approximately 27. Estimating the associated VECM model, we recover the long-run relationship, which is characterized by an oil price coefficient almost equal to 1 and highly significant. Moreover, the equation adjustment coefficient for gas is negative, close to -0.17 and significant, while the oil price equation adjustment coefficient is not statistically significant. This suggests that the oil price does not react to disequilibriums with respect to the long-run equation, while, on the contrary, gas prices do react.

We also test for the presence of full-sample cointegration across the three considered variables. Results are reported in the second column of Table 4. As in the previous case,

Table 4: Full sample cointegration estimation.

Included Variables	Gas Price Oil Price	Gas Price Oil Price Gas Quantity
Lags	2	4
Rank	1	1
Trace	26.98*	40.82*
	0.0006	0.0018
Max. Eig.	23.52	24.73
	0.0013	0.0149
Deterministic	IC	IC
Exogenous	D07, D07xT	D07, D07xT
Cointegration equation: $Gas_t - \mu - \beta_1 Oil_t - \beta_2 GasQ_t = \varepsilon_t$		
Gas Price	1	1
Gas Quantity		1.776
		(0.709)
Oil Price	-1.004	-0.968
	(-9.703)	(-9.907)
Adjustment coefficients		
Gas Price	-0.169	-0.217
	(-3.960)	(-4.375)
Gas Quantity		-0.003
		(-0.449)
Oil Price	0.049	0.015
	(1.717)	(0.464)

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 linear trend in the cointegration equation (TC), intercept in the cointegration equation (IC), step dummy from January 2007 (D07) and interaction between the step dummy and a linear trend. The table then reports estimated coefficients and the associated T-statistics (in parentheses). Note that the cointegration relationship, reported in the upper part of the table, gives the coefficients leading to the computation of the cointegration error. P-values reported in the table for the cointegration test statistics are derived under the assumption of no break in the intercept and trend. A star denotes rejections of the null hypotheses under the appropriate critical values (see Johansen et al., 2000, and Giles and Godwin, 2012).

we introduce a step dummy capturing the structural break and its interaction with the linear trend. The critical value of the Trace test corrected for the break in the trend is now equal to 42.8 at the 1% confidence level and 36.9 at the 5% level. Empirical evidence thus suggests the presence of a single cointegration relationship but only with 5% confidence given that the test statistic equals 40.8. We can further interpret such a result by looking at the estimated cointegration equation coefficients. In the VECM model estimation output, we see that the oil price coefficient remains almost unchanged and highly significant compared to the previous specification of the model, similar to the oil and gas price equation adjustment. The latter, in particular, rises to -0.217 suggesting that gas prices react more quickly to disequilibrium. Note that the gas quantity coefficient, even if it has the expected sign, is non-significant. This is possibly due to the misspecification of the full model, where the gas quantity can have a role sensibly differing before/after the break. Alternatively, taking into consideration the stationarity of the gas quantity up to 2006, we might have a more relevant change in the relationship across variables over time: the existence of a cointegration relationship might be supported for the second part of the sample, while in the first part such a relationship could not exist. A deeper evaluation of such an issue is deserved, and this can be made by looking at sub-sample estimates. Note that, by construction, moving to shorter samples, we will exclude a priori the possible presence of a cointegration equation across the three variables up to 2006 given the stationarity of the gas quantity.

We investigate our conjecture by first testing for the existence of cointegration relationships in the first sub-sample from January 1997 up to December 2006. Then, we test for cointegration in the second sub-sample from 2007 to 2013. In this second case, we allow for cointegration between gas and oil prices, and also between gas, oil prices and gas quantity. Moreover, taking into account the abrupt increase of shale gas production in 2007 and 2008, we contrast the results of this latter analysis with those obtained by focusing on shorter samples. Such an evaluation allows for highlighting possible changes in the relationship across variables. However, it could still be affected by the instability due to the start of shale gas production and by the uncertainty in test outcomes due to the use of shorter samples. We consider testing for cointegration between gas, oil prices and gas quantity for several shorter samples. In particular, we select samples with a size ranging from 3 to 5 years in order to exclude either the first years of shale gas production data or the last years in which shale gas production begins to stabilize. The choice of the samples, even if it is different from the one adopted before (see Table 3), is nonetheless coherent with the findings reported there, which suggest the existence of changes in the stationarity properties at the beginning of shale gas production and in the last years of the sample. Table 5 contains all of the relevant results.

We stress that for all tests the specification of the deterministic terms as well as the number of lags has been kept fixed. For the case of gas and oil log-prices only, we tested for the presence of cointegration allowing for the presence of a liner trend in the cointegrating equation and intercepts in both the cointegrating equation and the VAR. On the contrary, in order to test the existence of a long-run relationship between gas price, gas quantity and oil price, we maintain the two lags, introduce the intercepts in the cointegrating equation and in the VAR and exclude the trend in the cointegration equation. We motivate the difference in the test structure to the presence of a trend in the gas quantity data over the last part of the sample. Introducing a trend in the cointegrating relationship would have not improved the model fit, and would have also led to nonsignificant coefficients (for the trend). Finally, the results for the gas and oil price case that include the trend are not affected much by the exclusion of the trend variable, apart from a few cases specifically mentioned in the table.

Looking at Table 5, we note that the presence of a long-run equilibrium, i.e. cointegration, is detected by the Trance and Maximum Eigenvalue tests of Johansen (1988, 1996) if we focus on the results referring to the first sample (1997-2006) and the second sample (2007-2013). In the latter, we have the same finding including or excluding the gas quantity. We also stress that the presence of cointegration is clearer when the quantity is included in the long-run equilibrium relationship. However, when moving to shorter

Table 5: Trace and maximum eigenvalue cointegration tests on selected sub-samples.

Variables	Sample	Lags	Determ.	Rank	Trace	Max Eig.
Gas and Oil Price	1997-2006	4	IC, IV	1	**0.0297**	**0.0188**
Gas and Oil Price	2007-2013	2	IC, IV, TC	1	**0.0349**	0.0511
Gas and Oil Price	2007-2009	2	IC, IV, TC	1	**0.0443**	**0.0422**
Gas and Oil Price	2008-2010	2	IC, IV, TC	1	**0.0271**	**0.0324**
Gas and Oil Price	2009-2011	2	IC, IV, TC	1	0.0932[a]	0.2922
Gas and Oil Price	2010-2012	2	IC, IV, TC	1	0.8745	0.9192
Gas and Oil Price	2011-2013	2	IC, IV, TC	1	0.4415	0.3269
Gas and Oil Price	2007-2010	2	IC, IV, TC	1	**0.0081**	**0.0128**
Gas and Oil Price	2008-2011	2	IC, IV, TC	1	**0.0109**	**0.0143**
Gas and Oil Price	2009-2012	2	IC, IV, TC	1	0.1792	0.2990
Gas and Oil Price	2010-2013	2	IC, IV, TC	1	0.6354	0.7754
Gas and Oil Price	2007-2011	2	IC, IV, TC	1	**0.0035**	**0.0076**
Gas and Oil Price	2008-2012	2	IC, IV, TC	1	**0.0199**	0.0670
Gas and Oil Price	2009-2013	2	IC, IV, TC	1	0.1817[a]	0.2506
Gas and Oil P. + Gas quant.	2007-2013	2	IC, IV	1	**0.0097**	**0.0075**
Gas and Oil P. + Gas quant.	2007-2009	2	IC, IV	1	**0.0073**	**0.0054**
Gas and Oil P. + Gas quant.	2008-2010	2	IC, IV	1	0.0860	**0.0240**
Gas and Oil P. + Gas quant.	2009-2011	2	IC, IV	1	0.0872	0.2385
Gas and Oil P. + Gas quant.	2010-2012	2	IC, IV	1	0.4217	0.4611
Gas and Oil P. + Gas quant.	2011-2013	2	IC, IV	1	0.3460	0.2441
Gas and Oil P. + Gas quant.	2007-2010	2	IC, IV	1	**0.0267**	**0.0099**
Gas and Oil P. + Gas quant.	2008-2011	2	IC, IV	1	0.0823	**0.0269**
Gas and Oil P. + Gas quant.	2009-2012	2	IC, IV	1	0.0618	0.2423
Gas and Oil P. + Gas quant.	2010-2013	2	IC, IV	1	0.3256	0.3722
Gas and Oil P. + Gas quant.	2007-2011	2	IC, IV	1	**0.0264**	**0.0099**
Gas and Oil P. + Gas quant.	2008-2012	2	IC, IV	1	**0.0204**	**0.0137**
Gas and Oil P. + Gas quant.	2009-2013	2	IC, IV	1	0.0720	0.1905

The first column reports the variables among which we test for the existence of a single cointegration relationship. The second column includes the sample we consider, while the third column indicates the number of lags we have included in the testing equation. Note that the number of lags has been chosen combining different information criteria and taking also a balance between the sample size and the significance of the coefficients in the various lags (using lag-exclusion tests). The fourth column specifies the deterministic components included in the testing equation: linear trend in the cointegration equation (TC), intercept in the cointegration equation (IC) and intercept in the VAR (IV). The fifth column indicates the number of cointegration relationships detected among the variables in the first column. The last two columns report the p-values for the Trace and Maximum Eigenvalue test statistics described in Johansen (1988, 1996). [a] denotes cases where, by excluding the trend (marginally significant), the test leads to the presence of one cointegration relationship at the 5% confidence level. Rows are grouped for samples of equal length: half sample; 3-, 4- and 5-year samples. Boldface values identify cases where we do have evidences of cointegration at the 5% confidence level.

samples, we have a high difference between the first part of the second sub-sample (up to 2011) and the various sub-sets that include the years 2012 and 2013. For the first part of the second sub-sample, cointegration is present in almost all cases, while when including the years 2012 and 2013, there is empirical evidence suggesting that the oil and gas prices (with and without the gas quantity) are not linked by a long-run equilibrium relationship. See the results in Table 5 for the ranges 2009-2012, 2009-2013 and 2010-2013.

There are alternative explanations for these findings. It is possible to conjecture that a decoupling between the oil and gas prices exists. Such an interpretation comes from the latest years, while for the first years the relevant increase of shale gas production might have led to a period of instability before full decoupling was reached. This can explain the results for the samples including the years 2007-2009. Alternatively, we can suppose that the long-run equilibrium is still present and may differ from the former one. However, the 2007-2009 period, which is characterized by the shale gas production boom, could affect the power of statistical tests, while the last part of the sample might not be long enough to identify the new long-run relationship between the variables of interest.

In order to assess the role played by shale gas, we replicate the analysis run for the full sample period by looking at the sub-samples 1997-2006 and 2007-2013 and estimating for the first sub-sample the VECM model described in section 2.2. for the gas and oil prices. For the second period, we estimate the VECM model for the gas, oil prices and gas quantity. Note that the gas quantity is excluded from the first sub-sample as it is stationary. For the second sub-sample, despite the evidence shown in the preliminary data analysis, we allow for the inclusion of the gas quantity starting from the beginning of shale gas production in 2007, since the cointegration tests run above (see Table 5) show that there is no long-run relationship for the period 2009-2013. Table 6 presents selected results.

The sub-sample evidence confirms that the relationship between oil and gas prices has been affected by shale gas production. In fact, the first sub-sample coefficients (first column of Table 6) are close to those of the full sample analyses (see Table 4), with a long-

Table 6: Selected coefficients for the VECM model.

Included Variables	Gas Price Oil Price	Gas Price Oil Price	Gas Price Oil Price Gas Quantity
Sample	1997-2006	2007-2013	2007-2013
Cointegration equation: $Gas_t - \mu - \beta_1 Oil_t - \beta_2 GasQ_t = \varepsilon_t$			
Gas Price	1	1	1
Gas Quantity			8.475
			(7.283)
Oil Price	-0.958	-2.773	-2.272
	(-10.551)	(-5.493)	(-6.915)
Adjustment coefficients			
Gas Price	-0.263	-0.009	-0.027
	(-4.043)	(-0.377)	(-3.266)
Gas Quantity			-0.001
			(-0.115)
Oil Price	-0.004	0.078	0.119
	(-0.092)	(4.350)	(4.958)

Note that in the cointegration equation Gas_t is the natural gas real log-price, Oil_t is the oil real log-price, $GasQ_t$ is the natural gas total quantity (in logs), and ε_t is a stationary error term. The natural gas quantity is not included in the first and second columns. We report in the table only the long-run and adjustment coefficients and the corresponding T-statistics. Note that for oil we expect a negative coefficient and for the quantity a positive coefficient given the form adopted for the cointegration equation.

run relationship between prices with a unit coefficient and the adjustment of gas prices only to disequilibrium. In contrast, the second sub-sample estimates (second and third columns of Table 6) show a different pattern: the long-run coefficient linking gas prices to oil prices is larger than 2 and highly significant, showing that the impact of oil price on the formation of the gas price has more than doubled after the entrance of shale gas into the market; the gas price quantity coefficient is statistically significant and positive, i.e. a rise in gas quantity has a negative impact on the long-run gas price; the gas price adjustment coefficient to disequilibrium is negative, and still significant; when including gas quantity into the equation, oil prices adjust to disequilibrium, compared to both the first-sample and the full-sample analysis; the gas quantity adjustment coefficient is nonsignificant (as it was for the full-sample analysis). These findings differ from the full sample analysis, which seems to be influenced by the first part of the sample. In particular, the high and significant role played by gas quantity after the entrance of the shale gas shows that shale gas plays an important role in the long-run gas price formation. This effect was not present in the full-sample analysis due to the absence of shale gas in the first part of the sample period. Recall, however, that the cointegration test results of Table 5, and in particular, those associated with the last part of the sample and excluding 2007 and 2008 data, show evidence suggesting the absence of a long-run equilibrium relationship across the variables of interest. Those mixed results, i.e. a lack of cointegration when focusing on the most recent years, and a relevant change in the long-run coefficients when focusing from 2007 onward, seem to suggest that the instability created by the start of shale production has not yet been fully recovered by the market. This challenges a full evaluation on the real effect of the shale gas revolution and does not allow for an unequivocal assessment of whether a new long-term equilibrium has been reached by the market. Further analyses on longer samples (including data beyond 2013) might help shed light on this issue.

4 Conclusions

The paper analyses the long-run relationship between oil and gas prices in the US Henry Hub market and the impact that shale gas has had on it. A necessary but yet not sufficient condition for time series to exhibit a long-run relationship is their non-stationarity (the presence of a unit root). Considering the time series of gas prices at Henry Hub, oil (WTI) and gas quantities ranging from 1997 to 2013, we show that effectively they are non-stationary. However, applying the Perron unit root test in the presence of a structural break, we show that the entrance of shale gas has determined a break from 2007, even though its precise date identification can be influenced by market anticipation or delays (and transition). We replicate the unit root test for several sub-samples, before and after shale gas rise, taking into account the possible market anticipation and/or its possible delayed (or transitory) impact, and show that the gas quantities time series has unstable stationarity properties, with changes occurring at the start of shale gas production and at the end of the sample, when its production starts stabilizing. This shows that a possible long-run impact of gas quantities on prices, if present, arises only after the shale gas production has started. We then test the long-run behavior applying a VECM. We consider first the full data sample (with and without gas quantity) and show that oil and gas prices exhibit a positive long-run one-to-one relationship. This is in line with what one could expect when taking into account that oil and gas are substitutes as inputs and compete in the same market for the factors of production. We then replicate the analysis by splitting the data into sub-samples, before and after the start of shale gas. Not all years have been included in the analysis, given the evidence of stationarity for the gas quantity up to 2006. However, the data of recent years show instability in the stationarity of the quantity; thus the results should be taken with caution. We confirm evidence of a negative impact of gas quantities on gas prices in the second sub-period, i.e. after the beginning of Shale Gas) the role referred to as the shale gas revolution. However, the importance of oil prices in determining gas prices rises, too. The impact of oil prices on gas prices more than doubles after the entrance of shale gas. Several factors can explain such results. We conjecture that an important role could have been played by shale oil production. Oil prices can be affected by the rise of shale oil production (not analysed here), which relies on a similar technology to shale gas production. It is well possible that the same factors that affect the shale gas industry also influence the shale oil industry, which would reinforce their relationship. The rise in the shale gas and shale oil quantity might therefore explain our results. However, a caveat needs to be placed here. Our analysis shows that the oil-gas long-run relationship ceases to hold from 2009 onward. It is not possible to assess whether this depends on an insufficient time series length or, on the contrary, denotes a true end to the oil-gas price coupling. The problem of the length of the time series could also affect the result of the adjustment coefficient. Further analysis on longer time series are needed to evaluate this point.

Acknowledgement. The authors wish to thank Gianluca Galetti for his help in data collection. Clearly, they are the sole responsible for what is written in the article. Fulvio Fontini acknowledges the MIUR-PRIN Research grant n. 2010S2LHSE_005.

References.

Brown, S. P. A. and Yücel, M. K (2008), What Drives Natural Gas Prices?, The Energy Journal, 29, 45-60.

Carrion-i-Silvestre, J.L., and Sanso, A., 2006, Testing the null of cointegration with structural breaks, Oxford Bulletin of Economics and Statistics, 68, 623-646.

Erdos, P. (2012) Natural gas prices on three Continents, Energies, 5, 4040-4056.

Foss M. M., (2007), United States natural gas prices to 2015, Oxford Institute for Energy Studies.

Ghouri, S. S. (2006), Forecasting natural gas prices using cointegration technique, OPEC Review, 30, 249-69

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

Gregory, A.W., and Hansen, B.E., 1996a, Residual-based tests for cointegration in models with regime shifts, Journal of Econometrics, 70, 99-126.

Gregory, A.W., and Hansen, B.E., 1996b, Residual-based tests for cointegration in models with regime and trend shifts, Oxford Bulletin of Economics and Statistics, 58, 555-560.

Gregory, A.W., Nason, J.M., and Watt, D.G., 1996, Testing for structural breaks in cointegrated relationships, Journal of Econometrics, 71, 321-341.

Hansen, B.E., 1992, Tests for parameter instability in regressions wit I(1) processes, Journal of Business and Economic Statistics, 10, 321-335.

Hansen, H., and Johansen, S., 1999, Some tests for parameter constancy in cointegrated VAR models, Econometrics Journal, 2, 306-333.

Hansen, P.R., 2003, Structural changes in the cointegrated vector autoregressive model, Journal of Econometrics, 114, 261-295.

Hartley, P. H., Medlock III, K. B. and Rosthal, J. E. (2008), The relationship of natural gas to oil prices, The Energy Journal, 29, 47-65.

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

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

Johansen, S., R. Mosconi and Nielsen B. (2000), Cointegration analysis in the presence of structural breaks in the deterministic trend, Econometrics Journal, 3, 216-249.

Perron, P., 1997, Further Evidence on Breaking Trend Functions in Macroeconomic Variables, Journal of Econometrics , 80 (2), pp.355-385.

Pindyck R. S., (1999), The Long-run evolution of energy prices, The Energy Journal, 20 (2), 1-27.

Pindyck R.S., (2004), Volatility in natural gas and oil markets, Journal of Energy and Development., 30 (2004), pp. 1–19.

Romagus, G. (2012), Convergence and divergence of crude oil and natural gas prices, Ph.D. Dissertation, University of Houston. ProQuest, UMI Dissertations Publishing, Ann Arbor.

Rosthal J. E., (2010), The relationship between crude oil and natural gas prices and its effect on demand, Ph.D. Dissertation, Rice University, ProQuest, UMI Dissertations Publishing, Ann Arbor.

Villar J. A., Joutz F. L., (2006), The relationship between crude oil and natural gas prices. Energy Information Administration, Office of Oil and Gas, Washington DC.

Wakamatzu, H. and Aruga, K. (2013) The impact of the shale gas revolution on the U.S. and Japanese natural gas markets, Energy Policy 62, 1002–1009.