The Weak Tie Between Natural Gas and Oil Prices

David J. Ramberg and John E. Parsons***

Several recent studies establish that crude oil and natural gas prices are cointegrated. Yet at times in the past, and very powerfully in the last two years, many voices have noted that the two price series appear to have "decoupled". We explore the apparent contradiction between these two views. We find that recognition of the statistical fact of cointegration needs to be tempered with two additional points. First, there is an enormous amount of unexplained volatility in natural gas prices at short horizons. Hence, any simple formulaic relationship between the prices will leave a large portion of the natural gas price unexplained. Second, the cointegrating relationship does not appear to be stable through time. The prices may be tied, but the relationship can shift dramatically over time. Therefore, although the two price series may be cointegrated, the confidence intervals for both short and long time horizons are large.

Keywords: Oil price, Natural gas price, Cointegration

http://dx.doi.org/10.5547/01956574.33.2.2

1. INTRODUCTION

A number of recent academic studies have established that natural gas and crude oil prices are cointegrated.[1] These results have had an impact on ana-

* MIT Engineering Systems Division and the MIT Center for Energy and Environmental Policy Research, MIT, E19-411, 77 Massachusetts Ave., Cambridge, MA 02139 USA. E-mail: dramberg@mit.edu.

** Corresponding author. MIT Sloan School of Management, MIT Center for Energy and Environmental Policy Research, and MIT Joint Program on the Science and Policy of Global Change, E19-411, 77 Massachusetts Ave., Cambridge, MA 02139 USA. E-mail: jparsons@mit.edu.

This research was supported by a grant from BP Technologies Ventures Inc. as a part of the BP-MIT Advanced Conversion Research program. Helpful comments have been given by our colleagues on the MIT Future of Natural Gas study, and by colleagues at BP, Petrobras and Hess, among others.

1. The term cointegration is used to describe a certain relationship between two or more time series, like oil and natural gas prices, that are likely to be non-stationary—for example, because each

lysts in the business and policy community. For example, the recent World Energy Outlook 2009, published by the International Energy Agency, reprinted a table from of one of these studies showing how an increase in the price of crude oil would be mirrored over the subsequent 12 months by a matching increase in the price of natural gas. No sooner had the results of these academic studies achieved widespread acceptance than the world witnessed a remarkable decoupling between these two prices. At the end of December 2008 the price of crude oil stood at $32.35/bbl and the price of natural gas at $5.44/mmBtu, a ratio slightly less than six. From there the price of oil began a recovery while the price of natural gas continued to decline. At the start of September 2009 the price of crude oil stood at nearly $68.02/bbl and the price of natural gas was $1.88/mmBtu, a ratio of more than 36. At the conclusion of 2010, the price of oil reached $91.38/bbl, while the price of natural gas had recovered to only $4.23/mmBtu, yielding a ratio just above 21. So what happened to the strong tie between the prices that these studies documented? Some believe that the recent price movements reflect a permanent rupture of the old tie between the two price series—a decoupling—caused by fundamental changes in the industry. If so, then studies establishing cointegration are already outdated.

This is not the first time the natural gas price has appeared to decouple from the oil price. Throughout the 1980s and early 1990s, the United States experienced a so-called "gas bubble"–an excess supply of deliverable gas—that kept natural gas prices low relative to the then prevailing price of crude oil. The situation reversed itself in the late 1990s and early 2000s so that the price of natural gas was regularly above the level one might have predicted based on the historical relationship. Both times there was industry talk of a decoupling.

Nevertheless, throughout these periods of ups and downs, statistical analysis establishes that the two time series were cointegrated, at least until recently. How is one to rationalize these seemingly contradictory facts? What do we really mean by cointegration if the ratio of prices is shifting so consequentially across decades, sometimes in one direction, and sometimes another?

We attempt to answer these questions by elaborating on exactly what has been documented as cointegration, and putting it into context with the historically changing relationship between the two price series. We also seek to clarify what is meant when industry analysts assert that the two prices have "decoupled." These assertions are often vague and open to alternative interpretations:

- (i) the prices have *temporarily* broken away from the usual relationship to which they will later return, or,

series may grow unboundedly with the growth of the economy or with inflation, or because the variance of each series grows or shrinks with time. It is difficult to properly characterize a relationship between non-stationary time series. Sometimes underlying the non-stationary series is a single process (or combination of processes) causing them to be non-stationary. When this is the case, the relationship between the two can be represented by a fixed functional form such as a linear combination, or a line. Then the series are considered cointegrated, and the function describing their relationship is the cointegrating equation. See, for example, Hendry and Juselius (2000).

- (ii) the prices have *permanently* broken away from the old relationship and moved into a new relationship, or,
- (iii) the two prices no longer maintain a relationship with one another at all.

Which is it? While we cannot guess the intended definition of decoupling by those who declare it has occurred, we do shed light on which of the three possible definitions of decoupling fit the data and describe the relationship between the crude oil and natural gas price series.

In this paper, we address these questions by revisiting the cointegration analyses reported by several researchers over the last twelve years. These include Serletis and Herbert (1999), Villar and Joutz (2006), Brown and Yücel (2008) and Hartley, Medlock and Rosthal (2008). Each of these papers implements a complicated set of statistical analyses of the two data series, plus a number of related conditioning variables, in order to determine if a relationship can be found with any statistical reliability, and, if so, to determine what that relationship is.[2] These analyses involve testing for a cointegrating relationship between the two variables and estimating a vector error correction model (VECM) and a conditional error correction model (conditional ECM). The results of the four papers are broadly consistent with one another, although the details of the modeling and the parameter estimates vary. In this paper we report the results of our own modeling and tests constructed along the lines of Brown and Yücel (2008). We include some more recent data than was available at the time of their analysis. We focus the discussion in the text on an exposition of the results, without walking the reader through the full set of statistical tests performed. However, these are detailed in the working paper version of this paper accessible on-line (Ramberg and Parsons, 2011). Our conclusions are as follows.

Although the two price series appear to be cointegrated, this statistical fact needs to be tempered with two additional points that we think have been insufficiently emphasized in the previous literature.

First, there is an enormous amount of unexplained volatility in natural gas prices. The raw price series for natural gas—without controlling for cointegration and any explanatory variables—is approximately twice as volatile as the raw oil price series. Hence, any simple formulaic relationship between the price of oil and the price of natural gas leaves a large portion of the short-run movements in the price of natural gas unexplained. The more statistically sophisticated approach of constructing a conditional ECM, which includes the

2. Villar and Joutz (2006) and Brown and Yücel (2008) directly model the relationship between natural gas and crude oil prices. Serletis and Herbert (1999) model the relationship between natural gas and fuel oil prices, among other energy prices, but do not include crude oil specifically. Hartley, Medlock and Rosthal (2008) model the relationship between natural gas and crude oil prices, but use the price of fuel oil as an intermediate step. The time windows examined vary across the studies, as does the role of exogenous conditioning variables.

cointegrating relationship and a set of exogenous explanatory variables, and which accounts for the reversion of natural gas prices back to the cointegrating relationship, still leaves a large portion of the volatility in changes in natural gas prices unaccounted for.

Second, the cointegrating relationship does not appear to be stable through time. Natural gas prices may be tied to oil prices, but the relationship can shift dramatically over time. While the previous literature documented that the price of natural gas seemed to be shifting up compared to the price of oil during the period 1989–2005, we show that since early 2006 this trend reversed. The period since the start of 2009 may also reflect a further decoupling of the relationship between the two series, although we may not have enough data to know yet exactly how the relationship has been redefined.

Therefore, although the two price series have been cointegrated, the confidence intervals for both short and long time horizons are large. This paper explores the nature of this apparent contradiction in an attempt to better characterize the relationship between crude oil and natural gas prices.

2. THE NATURAL GAS AND OIL PRICE RELATIONSHIP

What is the structure of the relationship of the natural gas price with the oil price? It seems natural to imagine that the price of oil and the price of natural gas would tend to rise or fall in tandem. They are both energy carriers, with one barrel (bbl) of crude oil having approximately the same energy content as six million Btu (mmBtu) of natural gas.[3] This rough logic would argue that the price of a barrel of crude oil should equal six times the price of an mmBtu of natural gas. If the price of natural gas rises by $1/mmBtu, then the price of crude oil should rise by $6/bbl.

Economists would quibble with the presumption that the ratio of prices ought to be determined exactly by the energy content equivalence. For example, Adelman and Watkins (1997) and Smith (2004) warn against valuing reserves in terms of "barrel of oil equivalent" or "gas equivalent". The two fuels have different costs of production, transportation, processing and storage, and they serve different portfolios of end uses with only a modest overlap. The two fuels also have different environmental costs. One should expect these factors to enter into the determination of any relationship between the prices of the two commodities, and the equilibrium relationship is unlikely to match the energy content equivalence ratio. Perhaps for this reason, the industry press contains a variety of other rules-of-thumb, including the simple 10-to-1 ratio, as well as more sophisticated, burner tip parity rules. One burner tip parity rule is based on competition between natural gas and residual fuel oil, while the other is based on competition between natural gas and distillate fuel oil. Both account for the transportation cost differential from the wellhead to power plants and industrial users. Both then translate

3. To be precise, one barrel of West Texas Intermediate crude oil contains 5.825 mmBtu.

the relationship back to the price of crude oil based on the typical ratio between the price of the fuel oil and the price of crude.[4] What is the formula that best describes the relationship, if any?

In fact, nothing like an energy content equivalence nor any other simply defined relationship has been persistently observed. Figure 1 shows the real spot price series in 2010 dollars for West Texas Intermediate (WTI) crude oil and Henry Hub (HH) natural gas from 1991–2010 plotted together on the same graph.[5,6] The scale for the price of natural gas is shown on the left-hand-side, while the scale for the price of crude oil is on the right-hand-side. Looking only as far back as the 1990s, the ratio of the price of oil ($/bbl) to the price of natural gas ($/mmBtu) has sometimes been as low as 2.5-to-1, and other times as high as 36-to-1. Natural gas prices sometimes spike dramatically, without there being any noticeable change in crude oil prices.

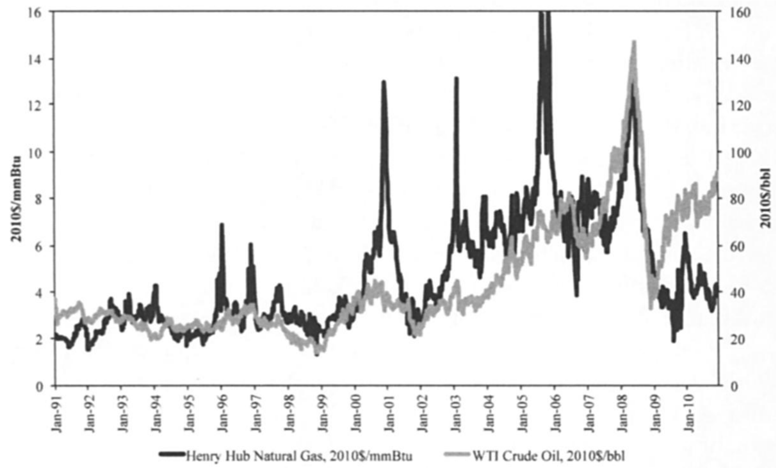
Figure 2 shows each of the four pricing rules-of-thumb mentioned earlier. The horizontal axis is the price of oil and the vertical axis is the price of natural gas. The line for each rule gives the predicted price of natural gas as a function of the given price of crude oil. Figure 2 also shows the scatterplot of the actual combinations of crude oil and natural gas prices in our data series. Each point in the scatterplot represents a different week's pair of prices, with the week's crude oil price determining the point's location along the horizontal axis, and the week's natural gas price determining the point's location along the vertical axis. It is clear that when the oil price has been above $80/bbl all four of the rules have overestimated the natural gas price, although the 10-to-1 rule is clearly the best of the lot. In order to examine the low oil price range more clearly, Figure 3 reproduces the rule-of-thumb graphs and the scatterplot, but focused only on the lower portion of the range of oil prices, i.e., those below $30/bbl. In this range, the actual prices are arrayed widely around the residual fuel oil burner-tip-rule

4. In Brown and Yücel (2008), the relationship generated by competition with residual fuel oil at the burner tip is given as $P_{HH,t} = -0.25 + (85\%/6.287)P_{WTI,t}$, where $P_{HH,t}$ is the price of natural gas at the Henry Hub, and $P_{WTI,t}$ is the price of West Texas Intermediate crude oil at Cushing Oklahoma. The relationship generated by competition with distillate fuel oil at the burner tip is given as $P_{HH,t} = -0.80 + (120\%/5.825)P_{WTI,t}$.

5. The starting point for our data is dictated by the history of the natural gas market in the US. The Natural Gas Policy Act of 1978 gradually led to the removal of price controls on the interstate sale of natural gas in the United States. As of January 1, 1985, ceilings were removed on the sale of new gas. This was followed by the 1987 repeal of sections of the Power Plant and Industrial Fuel Use Act that restricted the use of natural gas by industrial users and electric utilities and the Natural Gas Wellhead Decontrol Act of 1989 which completed the decontrol of US natural gas prices. In addition, the Federal Energy Regulatory Commission pursued a policy of encouraging open access to natural gas pipelines, especially through Order 636. Market depth grew quickly. By April 1990, the New York Mercantile Exchange initiated trading in a natural gas futures contract.

6. Both series are weekly day-ahead prices of commodities as sampled by Bloomberg. The natural gas prices are volume-weighted averages in $/mmBtu for delivery at Henry Hub in Louisiana. The crude oil prices are the arithmetic averages in $/bbl for West Texas Intermediate (WTI) crude oil traded at Cushing, Oklahoma. All prices were subsequently converted into real 2010 dollars.

Figure 1: Natural Gas and Crude Oil Spot Prices, 1991–2010 (real 2010 dollars)



Both series are weekly day-ahead prices of commodities as sampled by Bloomberg. The natural gas prices are volume-weighted averages in $/mmBtu for delivery at Henry Hub in Louisiana. The crude oil prices are the arithmetic averages in $/bbl for West Texas Intermediate (WTI) crude oil traded at Cushing, Oklahoma. All prices were subsequently converted in to real 2010 dollars.

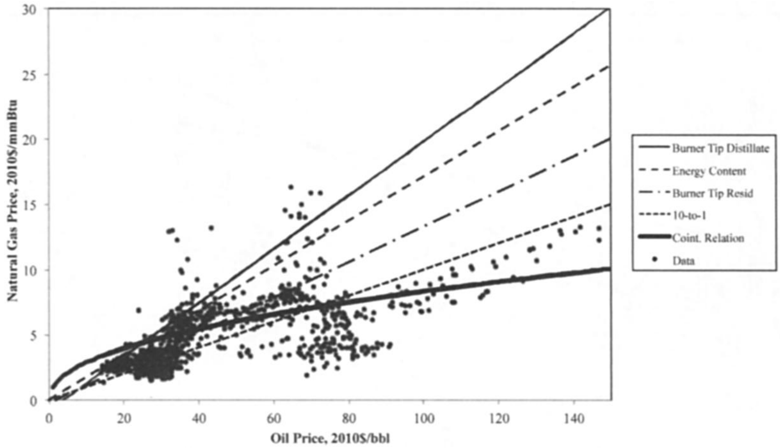
and the 10-to-1 rule, and only occasionally in the neighborhood of the energy-content-equivalence rule or the distillate fuel oil burner-tip-rule.

Figure 4 provides a time-series representation of the performance of each of the rules-of-thumb, graphing the prediction errors through time, i.e. the actual log natural gas price minus the predicted log natural gas price. These graphs call attention to a key problem that will undermine any simple relationship between the price of natural gas and the price of oil: there is much more volatility in the natural gas price than can be accounted for by movements in the oil price. This fact is also evident in Figure 1. The annualized volatility of the log natural gas price series is 72%, while the annualized volatility of the log crude oil price series is 39%, so natural gas was a little less than twice as volatile as crude oil.[7] Much of the volatility in the natural gas price series appears to take the form of temporary spikes in the price. These spikes have a relatively short duration.

Although no simple relationship with the oil price can account for all of the variation in the natural gas price, nevertheless, the eye can spot some rough

7. Volatility is annualized using this formula: *Ann vol NG* = Standard deviation($\log P_{HH,t} - \log P_{HHt-1}$)$*\sqrt{52}$. Assuming that the time series has some element of mean reversion, then the standard deviation of the *annual* price changes is expected to be less than the *annualized* standard deviation of the *weekly* price changes. If the time series is a pure geometric Brownian motion, then annual and the annualized standard deviations are expected to be the same.

Figure 2: Pricing Rules-of-Thumb Versus Observed Prices, 1991–2010

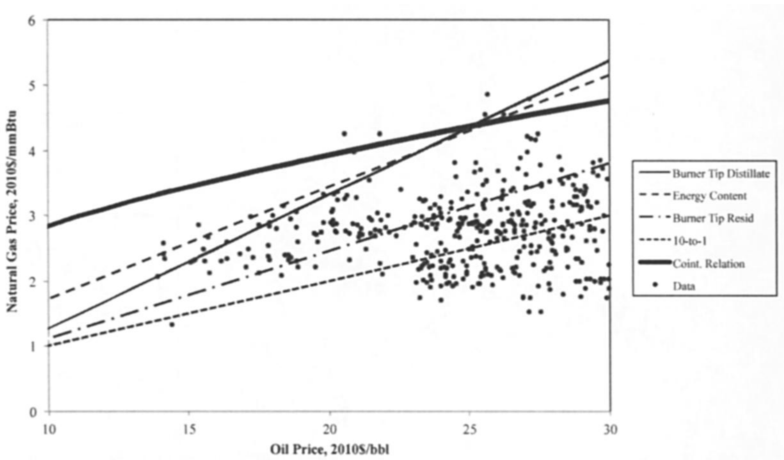


The figure charts natural gas prices as a function of oil prices. The four straight lines show the four pricing rules-ofthumb. The top line (using the ordering of the lines at the right of the figure) is the burner-tip parity rule based on natural gas competing with distillate fuel oil, the second line is the energy-content equivalence rule, the third line is the burner-tip parity rule based on natural gas competing with residual fuel oil, and the fourth line is the 10-to-1 rule. The dark black, slightly curved line is the estimated cointegrating equation from the VECM. The scatterplot of data points are the actual price combinations observed over the 1991–2010 period. All observed prices are quoted in real terms in 2010 dollars.

relationship between the two price series. The price spike of 2008 is the most dramatically clear example of this, as the two price series seem to move almost in lock step. The more lasting price run-up from 2003 through 2007 also clearly reflects some tie between the two price series. Even in the time period before 2002 this rough relationship seems to show up, though with less clarity. So, is the price of natural gas tied to the price of oil, or not?

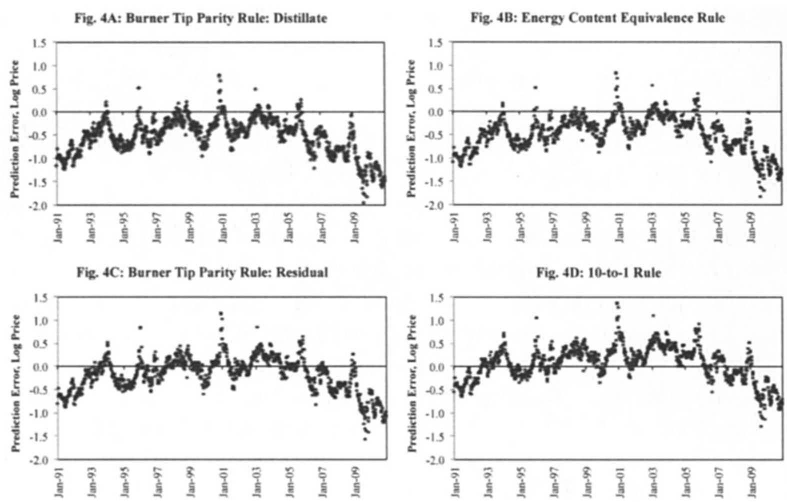
Part of the problem is that a number of other variables have some short-run influence either on the price of oil or on the price of natural gas. Fluctuations in one or more of these variables can lead to the price of either natural gas or oil temporarily diverging from its long-run level. These short-run fluctuations mask whatever long-run relationship may exist, making the relationship a complicated one to properly identify. The simplest of these other variables is the seasonal fluctuation in the price of natural gas in the United States. The price of crude oil is not seasonal, so the ratio of the prices must vary through the calendar year. Other variables, too, temporarily shift the supply and demand for natural gas relative to oil. While the WTI crude oil price is for delivery in Cushing, Oklahoma and refers to a specific type of oil produced in that region, it remains a benchmark price for crudes traded globally and it fluctuates primarily with factors affecting

Figure 3: Pricing Rules-of-Thumb Versus Observed Prices (Low Oil Price Range)



The figure shows the same data as Figure 2, except that it focuses in on the low range of oil prices so as to make visible the different observed prices and the comparison to the different rules-of-thumb.

Figure 4: Prediction Errors for Four Rules of Thumb



Prediction error calculated by subtracting the predicted log natural gas price from the actual log natural gas price.

global demand and supply. In contrast, the price of natural gas for delivery into the Henry Hub, Louisiana is impacted much more strongly by fluctuations in supply and demand specific to the North American marketplace. These include weather events such as unexpectedly severe winter storms that cause the price of natural gas to spike, or surprisingly mild winter weather that causes the price to fall. These also include temporary interruptions to supply caused by hurricanes that shut-in production, and similar events. While the natural gas price in North America is also linked to the fluctuations in supply and demand elsewhere in the globe, prices in different regions of the world can move markedly apart from one another at times.

Identifying the underlying tie between the two prices—if any—requires filtering out the effect of these various factors. This is the challenging task to which we now turn.

3. COINTEGRATION ANALYSIS

Our analysis breaks the natural gas and oil price relationship into four components. First, there is the underlying or fundamental tie between the natural gas price and the oil price. This is called the cointegrating equation. When we say this is the fundamental tie, we mean that this is the relationship that is generally reestablished after periods in which the two prices move away from one another. Second, there is the error correction mechanism. Whenever the natural gas price has been pulled away from the fundamental tie, the price will predictably drift back towards the fundamental tie. The model estimates the rate at which this drift back occurs. Third, there are a few identifiable and recurrent exogenous factors—such as seasonality, episodic heat waves and cold waves and intermittent supply interruptions from hurricanes—that cause the natural gas price to deviate from this fundamental tie in predictable ways. The statistical analysis attempts to identify and filter out these three identifiable components. The fourth component is the residual volatility or price movement not accounted for by the first three components. These are the unexplained shocks remaining after the three identifiable components of the movements in the natural gas and oil prices have been filtered out. This residual volatility reflects the myriad temporary disruptions to the supply and demand for natural gas or oil, which, much like the identifiable and recurrent exogenous factors, pull the two prices away from the fundamental tie.

To identify these four components we implement a complicated set of statistical analyses that are described fully in the working paper version of this paper, Ramberg and Parsons (2011), accessible on-line. Here we focus on just the main result, which is the estimation of this Vector Error Correction Model (VECM):

$$P_{HH,t} = \gamma + \beta P_{WTI,t} + \mu_t, \tag{1}$$

$$\Delta P_{HH,t} = a_{HH} + \alpha_{HH}\mu_{t-1} + \sum_{i=1}^{9} b_{HH,i}\Delta P_{WTI,t-i} + \sum_{i=1}^{9} c_{HH,i}\Delta P_{HH,t-i} + \sum_{j=1}^{6} d_{HH,j}X_{j,t} + \varepsilon_{HH,t}, \quad (2)$$

$$\Delta P_{WTI,t} = a_{WTI} + \alpha_{WTI}\mu_{t-1} + \sum_{i=1}^{9} b_{WTI,i}\Delta P_{WTI,t-i} + \sum_{i=1}^{9} c_{WTI,i}\Delta P_{HH,t-i} + \sum_{j=1}^{6} d_{WTI,j}X_{j,t} + \varepsilon_{WTI,t}, \quad (3)$$

Equation (1) is the cointegrating equation which captures the first component, the hypothesized fundamental tie between Henry Hub natural gas and WTI crude oil prices. $P_{HH,t}$ is the log natural gas price in week t, $P_{WTI,t}$ is the log crude oil price in week t, γ is a constant to be estimated, β is a parameter to be estimated, and μ_t is the cointegrating error term in week t. $\Delta P_{HH,t}$ is the change in the log natural gas price from week $t-1$ to week t, μ_{t-1} is the lagged cointegrating error term from equation (1), $\Delta P_{WTI,t}$ is the change in the log crude oil price, X_j is the matrix of six exogenous variables representing additional drivers of the Henry Hub natural gas price, the variously subscripted parameters a, α, b, c, and d are to be estimated. Finally, $\varepsilon_{HH,t}$ and $\varepsilon_{WTI,t}$ are the error terms. After estimating the VECM, we examine the coefficients for the error-correction term for both the change in natural gas prices and the change in crude oil prices. The coefficient on the error correction term for changes in the crude oil price when the natural gas and oil prices are not at the long-run equilibrium is positive and not statistically significant. This runs counter to the theory and purpose of an error correction mechanism, and provides statistical evidence that oil prices are not affected by natural gas prices. In order to treat the oil prices as (at least weakly) exogenous, we then estimate the matching conditional Error Correction Model (conditional ECM),

$$\Delta P_{HH,t} = a_{ECM} + \alpha_{ECM}\mu_{t-1} + b_{ECM}\Delta P_{WTI,t} + \sum_{i=1}^{9} c_{ECM,i}\Delta P_{HH,t-i} + \sum_{j=1}^{6} d_{ECM,j}X_{j,t} + \varepsilon_{ECM,t}. \quad (4)$$

Equation (4) uses as an input, μ_{t-1}, the previously estimated error term from equation (1) together with the contemporaneous change in the price of oil, the lagged changes in the price of natural gas, and the matrix of six exogenous variables. These capture the second and third components of the relationship between the natural gas and oil prices, and the error term captures the fourth component, the residual volatility. In the conditional ECM, the effect of natural gas prices on oil prices is omitted.

The results of the estimation are shown in Tables 1 and 2. Following Brown and Yücel (2008), the exogenous variables used are the number of heating degree days (HDD), the number of cooling degree days (CDD), the deviation from the normal number of heating degree days (HDDDEV), the deviation from the normal number of cooling degree days (CDDDEV), the deviation of the amount of natural gas in storage from its average (STORDIFF), and the amount of natural gas production shut-in, e.g. due to storms (SHUTIN).[8] Because certain of the exogenous variables are only available starting on June 13, 1997, when we report results for the cointegration analysis, the results are based on estimation over the June 13, 1997 to December 31, 2010 period of time, and do not include the period 1991 up to June 13, 1997.

We now turn to discussing the results in more detail, focusing one at a time on the separate components, beginning with the first.

The Fundamental Tie Between the Natural Gas Price and the Oil Price

The estimated cointegrating equation is:

$$P_{HH} = -0.0333 + (0.468 \times P_{WTI}). \tag{5}$$

This relationship is graphed in Figures 2 and 3. The cointegrating relationship is linear in the logged prices. Converted back into dollars, the log-linear relationship is a slightly concave curve. As can be seen in the Figures, when the oil price is \$10/bbl, the cointegrating relationship predicts a natural gas price of \$2.84/mmBtu. At \$60/bbl, the predicted natural gas price is \$6.57/mmBtu. Were the oil price to reach \$150/bbl, the cointegrating relationship predicts a corresponding natural gas price of \$10.09/mmBtu.

One can see in Figures 2 and 3 that the cointegrating equation attempts to fit the data better than the various rules-of-thumb by crafting a compromise out of slightly overestimating natural gas prices when the oil price is low and slightly underestimating them when the oil price is high. Nevertheless, it is impossible to escape the problem of the great volatility in natural gas prices. Figure 5 provides a time-series representation of the performance of the cointegrating relationship—equation (5)—at predicting the natural gas price from January 1991 through December 2010, i.e. both before and during the sample period used in the estimation. The centered mean absolute error for the cointegrating relationship is 0.394, which is approximately the same as for the burner tip distillate rule-of-thumb discussed above.[9] This repeats the earlier observation that the natural gas

8. Normal Heating Degree Days or Cooling Degree Days reflect the average value for each week from 1971–2000.

9. Note that the VECM was estimated over the period June 13, 1997–December 31, 2010, so the 0.394 mean absolute error incorporates errors both in- and out-of-sample. Focusing just on the 1997–2010 data used for the estimation, the centered mean absolute error for the cointegrating relationship is 0.341. Over this shorter window, the centered mean absolute error for the rules-of-thumb ranged

Table 1: Parameter Estimates for the Vector Error Correction Model. Full Period, June 13, 1997–December 31, 2010

Variable	Coefficient	Estimate	p−value
	γ	−0.0333	
P_{WTIt}	β_{WTI}	0.4680	0.001**

Variable	Coefficient	Estimate	p−value	Variable	Coefficient	Estimate	p−value
	a_{HH}	−0.0018	0.902		a_{WTI}	−0.0076	0.369
μ_{t-1}	α_{HH}	−0.0468	0.001**	μ_{t-1}	α_{WTI}	0.0109	0.203
ΔP_{WTIt-1}	b_{HH1}	0.0843	0.199	ΔP_{WTIt-1}	b_{WTI1}	−0.1124	0.004**
ΔP_{WTIt-2}	b_{HH2}	−0.0146	0.824	ΔP_{WTIt-2}	b_{WTI2}	−0.0959	0.014*
ΔP_{WTIt-3}	b_{HH3}	0.0647	0.327	ΔP_{WTIt-3}	b_{WTI3}	0.0837	0.032*
ΔP_{WTIt-4}	b_{HH4}	−0.0726	0.270	ΔP_{WTIt-4}	b_{WTI4}	−0.0322	0.409
ΔP_{WTIt-5}	b_{HH5}	−0.0320	0.627	ΔP_{WTIt-5}	b_{WTI5}	0.0207	0.596
ΔP_{WTIt-6}	b_{HH6}	0.0660	0.317	ΔP_{WTIt-6}	b_{WTI6}	−0.0497	0.202
ΔP_{WTIt-7}	b_{HH7}	0.1793	0.007**	ΔP_{WTIt-7}	b_{WTI7}	−0.0711	0.068+
ΔP_{WTIt-8}	b_{HH8}	0.0744	0.260	ΔP_{WTIt-8}	b_{WTI8}	0.1331	0.001**
ΔP_{WTIt-9}	b_{HH9}	0.0833	0.206	ΔP_{WTIt-9}	b_{WTI9}	0.0969	0.013*
ΔP_{HHt-1}	c_{HH1}	−0.1057	0.008**	ΔP_{HHt-1}	c_{WTI1}	0.0547	0.020*
ΔP_{HHt-2}	c_{HH2}	−0.0548	0.163	ΔP_{HHt-2}	c_{WTI2}	0.0076	0.743
ΔP_{HHt-3}	c_{HH3}	−0.1498	0.000**	ΔP_{HHt-3}	c_{WTI3}	−0.0143	0.530
ΔP_{HHt-4}	c_{HH4}	−0.0414	0.283	ΔP_{HHt-4}	c_{WTI4}	0.0142	0.533
ΔP_{HHt-5}	c_{HH5}	−0.1005	0.008**	ΔP_{HHt-5}	c_{WTI5}	0.0162	0.473
ΔP_{HHt-6}	c_{HH6}	0.0498	0.193	ΔP_{HHt-6}	c_{WTI6}	0.0208	0.359
ΔP_{HHt-7}	c_{HH7}	−0.0631	0.096+	ΔP_{HHt-7}	c_{WTI7}	−0.0045	0.842
ΔP_{HHt-8}	c_{HH8}	0.0184	0.628	ΔP_{HHt-8}	c_{WTI8}	−0.0153	0.494
ΔP_{HHt-9}	c_{HH9}	−0.0832	0.028*	ΔP_{HHt-9}	c_{WTI9}	−0.0309	0.167
HDD_t	d_{HH1}	8.58E-05	0.311	HDD_t	d_{WTI1}	5.46E-05	0.276
HDDDev_t	d_{HH2}	1.03E-03	0.000**	HDDDev_t	d_{WTI2}	−7.57E-05	0.528
CDD_t	d_{HH3}	−4.68E-04	0.072+	CDD_t	d_{WTI3}	1.43E-04	0.352
CDDDev_t	d_{HH4}	3.35E-03	0.000**	CDDDev_t	d_{WTI4}	3.92E-04	0.339
StorDiff_t	d_{HH5}	−1.94E-05	0.252	StorDiff_t	d_{WTI5}	1.88E-05	0.060+

$Shutin_t$	d_{HH6}	4.55E-06	0.290	$Shutin_t$	d_{WTI6}	−6.25E-06	0.014*
Variables		Chi^2 Stat.	p−value	Variables		Chi^2 Stat.	p−value
Lagged HH		45.67	0.000**	Lagged HH		10.38	0.321
Lagged WTI		13.57	0.138	Lagged WTI		49.84	0.000**
Lagged HH + WTI		59.68	0.000**	Lagged HH + WTI		56.29	0.000**
Exogenous Variables		50.77	0.000**	Exogenous Variables		12.43	0.053+
Exog + Lagged HH		86.03	0.000**	Exog + Lagged HH		24.23	0.061+
Exog + Lagged WTI		61.09	0.000**	Exog + Lagged WTI		62.07	0.000**
Exogenous + Lagged		98.91	0.000**	Exogenous + Lagged		69.94	0.000**

Equation	Parameters	RMSE	R^2	Chi^2 Stat.	p−value
ΔP_{HH}	26	0.0958	0.1458	114.7092	0.0000
ΔP_{WTI}	26	0.0567	0.0952	70.6978	0.0000

+ = 0.1, * = 0.05, ** = 0.01 significance levels. Number of Observations: 698

Table 2: Parameter Estimates for the Conditional Error Correction Model. Full Period, June 13, 1997–December 31, 2010

Variable	Coefficient	Estimate	p–value
	a_{ECM}	0.0044	0.750
μ_{t-1}	α_{ECM}	−0.0497	0.000**
ΔP_{WTIt}	b_{ECM1}	0.2458	0.000**
ΔP_{HHt-1}	c_{ECM1}	−0.1040	0.007**
ΔP_{HHt-2}	c_{ECM2}	−0.0463	0.216
ΔP_{HHt-3}	c_{ECM3}	−0.1424	0.000**
ΔP_{HHt-4}	c_{ECM4}	−0.0472	0.200
ΔP_{HHt-5}	c_{ECM5}	−0.1090	0.003**
ΔP_{HHt-6}	c_{ECM6}	0.0563	0.125
ΔP_{HHt-7}	c_{ECM7}	−0.0393	0.279
ΔP_{HHt-8}	c_{ECM8}	0.0376	0.299
ΔP_{HHt-9}	c_{ECM9}	−0.0686	0.059+
HDD_t	d_{ECM1}	3.91E-05	0.630
$HDDDev_t$	d_{ECM2}	1.06E-03	0.000**
CDD_t	d_{ECM3}	−5.29E-04	0.037*
$CDDDev_t$	d_{ECM4}	3.18E-03	0.000**
$StorDiff_t$	d_{ECM5}	−1.94E-05	0.239
$Shutin_t$	d_{ECM6}	4.40E-06	0.289

Variables	Chi² Stat.	p–value
Lagged HH	48.32	0.000**
Exogenous Variables	50.35	0.000**
Exog + WTI	65.37	0.000**
Exog + Lagged HH	89.70	0.000**
Exog + WTI + Lagged	104.74	0.000**

Equation	Parameters	RMSE	R^2	Chi² Stat.	p–value
ΔP_{HH}	18	0.0951	0.1479	121.1898	0.0000

+ = 0.1, * = 0.05, ** = 0.01 significance levels. Number of Observations: 698

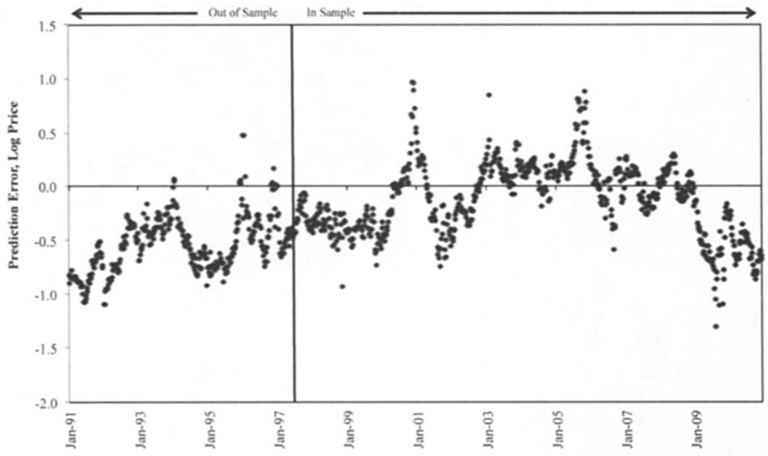
price series is just too volatile to be accounted for by any simple tie to the oil price including this cointegrating equation. Only by somehow accounting for this additional volatility could we reduce this error. The other components of the VECM and the conditional ECM attempt to provide this accounting, and we now turn to examine how successfully they do so.

The Error Correction Mechanism and the Rate of Recovery

Many factors may pull the price of natural gas away from the fundamental relationship. The model then allows for an error correction mechanism by

from 0.398 for the 10-to-1 and the energy content equivalence rules to 0.411 for the residual burner tip parity rule, to 0.428 for the distillate burner tip parity rule. So, not surprisingly, the cointegrating equation does fit the data in-sample better than any of the rules-of-thumb.

Figure 5: Prediction Errors for the Stand-alone Cointegrating Relationship



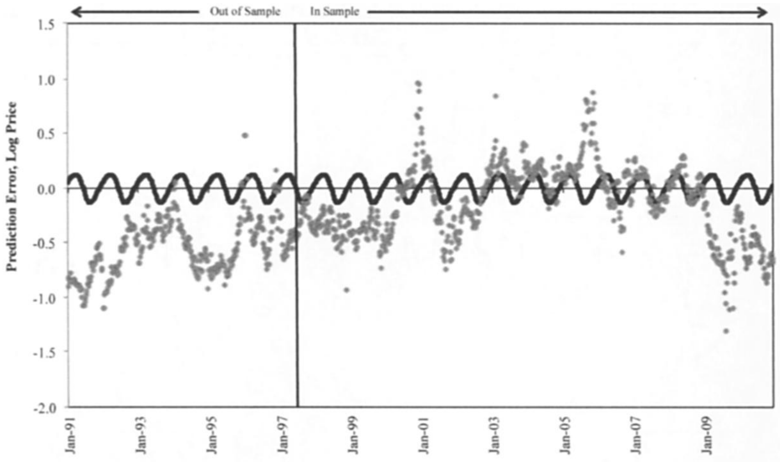
Prediction error calculated by subtracting the predicted log natural gas price from the actual log natural gas price using equation (5).

which the natural gas price is pulled back to the fundamental relationship. This reversion to the fundamental relationship is a predictable part of the price movement captured in the estimated error correction mechanism. For example, when the crude oil price rises 20%, from \$50/bbl to \$60/bbl, and all other variables are held constant, then, according to the cointegrating relationship, the price of natural gas should rise approximately 9%, from \$6.04/mmBtu to \$6.56/mmBtu. This occurs gradually, however, with the half-life of the rise being nearly 22 weeks. Alternatively, if the price of natural gas price spikes up by 166%, from \$6.04/mmBtu to \$10/mmBtu, while the crude oil price is steady at \$50/bbl, then the natural gas price is expected to eventually fall back to \$6.04/mmBtu. The half-life for the return of the natural gas price to its cointegrating relationship is nearly eight weeks.

Exogenous Factors

The first two of our exogenous variables, HDD and CDD, capture the well known seasonality of natural gas prices in the U.S. To generate the estimated seasonal fluctuations, we simulate the path of natural gas prices through the average annual cycle of HDD and CDD in our dataset using our estimated conditional ECM and holding the crude oil price, the other four exogenous factors and the error terms all fixed. The resulting natural gas price settles into a cycle around an average point that occurs in the first week of July and again in the third week of December. The price peaks at about 113% in April. The trough is at 87% in

Figure 6: Seasonality Relative to Prediction Errors



September. The total amplitude of the seasonal variation in the natural gas price is 26 percentage points. At a base price of \$7/mmBtu (the July and December prices), this is a range of \$1.82/mmBtu. Figure 6 shows this seasonal variation overlayed on the observed prediction errors for the estimated cointegrating relationship. This allows one to see how much larger is the actual range of variation than can be accounted for by the predictable seasonal component. For example, the standard deviation of the logged error series for the cointegrating relationship from 1991–2010 is 0.394. The seasonality coefficient, however, only ranges as high or low as +/– 0.131, or about a third of a standard deviation. Using two standard deviations as a benchmark for capturing the vast majority of the range in gas volatility, the seasonal component could not account for any more than 16.6% of natural gas volatility.

The next two of our exogenous variables, HDDDEV and CDDDEV, capture the impact of unseasonably cold or warm weather on demand and therefore price. These variables, too, only account for a modest amount of the volatility in the change in the natural gas price. To illustrate this, Figure 7 compares the actual changes in the price of natural gas prices around a typical cold spell—the two weeks of March 18–25, 2005—against the portion of the price change attributable to the cold spell. A typical spell last two weeks, with the first week exhibiting an HDD level 20 degree days above normal and the second week exhibiting an HDD level 12 degree days above normal. The portion attributable to the cold spell is calculated using the estimated conditional ECM in equation (4) using the actual deviation in HDD, and holding the crude oil price fixed, setting the initial natural gas price so that the error term in equation (1) is zero, setting the other exogenous factors in equation (4) to zero, setting the error terms in equation (4) to zero, and, simulating how the natural gas price evolves in

Figure 7: Predicted Impact of Unseasonal Cold Snap Versus Actual Change in Price

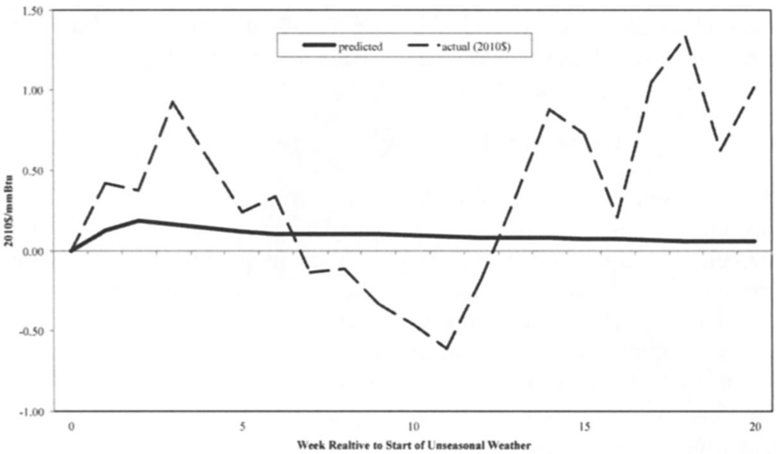


Figure shows actual cumulative price change around a typical unseasonal cold snap. Week 0 is March 11, 2005 and weeks 1 and 2, March 18 and March 25, 2005 are the weeks of the cold snap. A typical spell lasts two weeks with the first week exhibiting an HDD level 20 degree days above the average and the second week exhibiting an HDD level 12 degree days above the average. We chose this set of dates because it matched a typical cold snap. The figure also shows the predicted cumulative price impact of the cold spell. This is calculated using the estimated conditional ECM in equation (4) using the actual deviation in HDD, and holding the crude oil price fixed, setting the initial natural gas price so that the error term in equation (1) is zero, setting the other exogenous factors in equation (4) to zero, setting the error terms in equation (4) to zero, and simulating how the natural gas price evolves in response to the shock to the exogenous variable of HDD deviations.

response to the shock to the exogenous variable HDDDEV. In our example, the price on week zero, March 11, 2005, was $7.51/mmBtu. On March 18, week one of the cold spell, the price had increased by $0.42/mmBtu. The cold spell is predicted to have increased the price by $0.13/mmBtu, so that absent the cold spell the price would have increased by $0.29/mmBtu. By March 25, week two of the cold spell, the price fell slightly, yielding a cumulative increase of $0.37/mmBtu. The estimated model attributes a cumulative increase of $0.19 to the cold spell, accounting for about half of the actual cumulative increase. From there on out, the cumulative increase attributable to the cold spell gradually dissipates. The cumulative change in the actual price swings far below and far above zero. Note that while, for comparison, we attempted to identify a point in our actual dataset where the variables other than HDDDEV were fairly stable, the real-life changes in the Henry Hub price were nonetheless affected by minor changes in these variables. Even after accounting for the movements of additional variables, however, a large portion of the movements in the actual natural gas price series remains unaccounted for.

There are two other exogenous variables—the level of natural gas storage and the level of shut-in production in the Gulf of Mexico due to hurricanes. These, too, account for a modest amount of the volatility of natural gas price changes, although for economy of space we provide detail on the impact of these two individual variables in the working paper version of this paper, Ramberg and Parsons (2011), accessible on-line.

The Residual or Unexplained Volatility

The implementation of the VECM and conditional ECM modeling techniques improves the fit of the predicted natural gas price over the rules-of-thumb. Nevertheless, a large amount of the volatility in natural gas price changes could not be explained by the combination of the cointegrating relationship with the crude oil price, the error correction mechanism, and the identified exogenous variables. The portion of the volatility in the natural gas price explained by the conditional ECM is approximately 15%. That means the fraction of variance of changes in natural gas prices unexplained by our model is nearly 85%. Therefore, although the two series are cointegrated, this statistical fact should not be taken to mean that the two series are tightly coupled. Over short horizons there is significant unexplained volatility in changes in the natural gas price. The two prices regularly decouple, sometimes significantly, although this decoupling is not long lasting.

4. A CHANGING RELATIONSHIP OVER TIME?

One possible explanation for the weak explanatory power of the model is that we are trying to identify a single relationship across a long window of time, when in fact the relationship has evolved over this period. As we noted earlier, the natural gas and oil prices are not likely to be equated simply on the basis of energy equivalence because of the different costs of production, transportation, processing and storage, and because of the different end use markets they serve. These different underlying technical and economic factors make the cointegrating relationship diverge from a strict energy equivalence. But these factors are themselves shifting over time, sometimes gradually and sometimes swiftly. Villar and Joutz (2006) examined the 1989–2005 period and found that the cointegrating relationship between logged oil and gas prices shifted up by nearly half of a percent per month, with the price of natural gas relative to crude oil having increased. Hartley, Medlock and Rosthal (2008) examined a substantially overlapping period, 1990–2006, which exhibited a similar increase in the price of natural gas relative to the price of crude oil. Hartley, Medlock and Rosthal go a step further to specifically attribute this to the increased demand for natural gas arising from the installation of advanced combined cycle gas turbine (CCGT) power plants with significantly improved heat rates. Figure 8 shows the dramatic shift up in the cointegrating relationship documented by Villar and Joutz, contrasting the estimated relationship at the start of their data set, in 1989, with the estimated relationship at the end, in 2005.

Figure 8: Villar-Joutz Shifting Cointegrating Relationship, 1989 and 2005

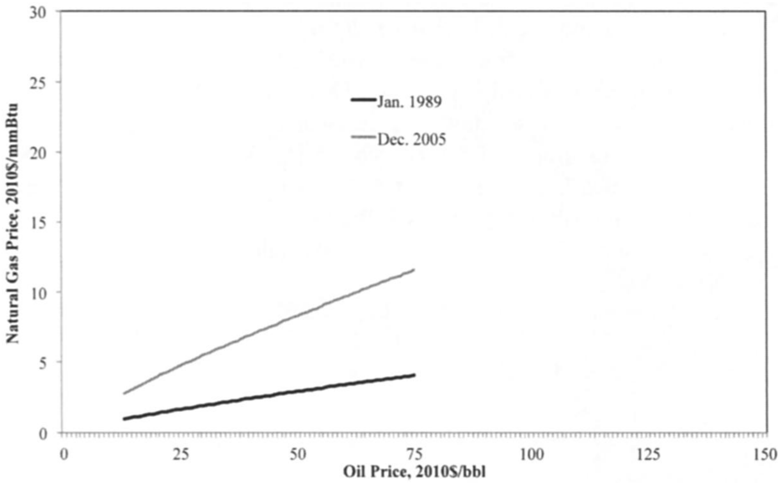


Figure graphs the estimated cointegrating relationship from Villar and Joutz (2006) at the start of their period, in January 1989, and at the end of their period, in December 2005. We only graph the relationship in the range of oil prices relevant to that window of time. However, we preserve the same scale for the overall graph as in Figure 2 in order to keep them comparable.

Has the cointegrating relationship shifted once again, but this time in the opposite direction? A major technological innovation in recent years has been improvements in horizontal drilling and hydraulic fracturing making possible the low cost exploitation of natural gas in shales. Production from shales has dramatically increased in the U.S. in recent years, and is almost certainly the cause of the most recent drop in the price of natural gas relative to oil. Simultaneously, the price of oil has reached a higher level than before, and oil use is more and more dominated by the transport sector. Each of these developments shapes the competition between the two energy carriers and therefore the equilibrium relationship between them. Is this shift statistically identifiable in our data, taking into account the error correction mechanism and exogenous factors?

To address this question, we examined our data as follows. First, as reported in the previous section, we fit a single cointegrating relationship over our full dataset, June 13, 1997 to December 31, 2010.[10]

10. We ran the Johansen test for cointegration over our full window of time, from June 13, 1997 through December 31, 2010, with lag length of ten weeks. The test selects a rank of one over a rank of two at the 1% significance level, but also selects a rank of zero (not cointegrated) over a rank of one (cointegrated) at the 5% level. The Schwartz-Bayesian Information Criterion and the Hannan-Quinn Information criteria disagree on the rank, with the former selecting a rank of zero and the latter selecting a rank of one. If we select a rank of one (cointegrated) and fit our VECM over the full window of time and then evaluate the errors for a unit root using the Augmented Dickey Fuller tests or the Phillips-Perron tests, we reject the null hypothesis of a unit root in the errors, which is evidence that the identified relationship is a cointegrating relationship.

Second, we considered the possibility of cointegration, but allowing for breakpoints in the structure of the relationship. We identify two breakpoints, one at February 6, 2009 and one at March 10, 2006.[11]

Looking at the data from June 13, 1997 to February 6, 2009, the evidence for a cointegrating relationship is strong. This is true whether we fit a single relationship across the full window of time, June 13, 1997–February 6, 2009, or we fit two separate relationships for the two sub-segments, June 13, 1997–March 10, 2006 and March 17, 2006–February 6, 2009. However, the evidence clearly argues that the relationship shifted across the two sub-segments.[12] The two cointegrating relationships we estimate for our two windows of time are:

$$\log P_{HH} = -1.2007 + (0.7261 \times \log P_{WTI}) \tag{6}$$

for the June 13, 1997–March 10, 2006 period, and

$$\log P_{HH} = 0.1969 + (0.4621 \times \log P_{WTI}) \tag{7}$$

for the March 17, 2006–February 6, 2009 period. Figure 9 graphs these two relationships. The cointegrating equation has shifted downward in the latter period, predicting a lower price of natural gas given the price of crude oil. This shift is in exactly the opposite direction from the shift documented by previous authors for the earlier era, 1989–2005.

Finally, looking at the short window of time from February 13, 2009 through December 31, 2010, we cannot say much since neither series displays sufficient evidence of non-stationarity for the tests of cointegration to be meaningful. The on-line working paper version of this paper, Ramberg and Parsons (2011), contains a description of the full set of tests and investigations performed across all segments and combinations.[13]

11. We employ the Gregory and Hansen (1996) tests for cointegration with regime shift. The null hypothesis is no cointegration across the full time period, and the alternative hypothesis is cointegration, where the cointegrating vector is allowed to change at a single unknown time during the sample period. The alternative includes the possibility of no change or break. Their tests do not require ex ante information on the timing of a break, nor a presumption about whether or not there is a break. All three tests accept the alternative of cointegration, including the possibility of a breakpoint. Although there is minor disagreement among the tests about the exact dating of a likely breakpoint, we chose February 6, 2009 based on the ADF test. We then repeated the Gregory and Hansen (1996) test on the shorter window of time from June 13, 1997 through February 6, 2009 and identified the earlier breakpoint at March 10, 2006.

12. Having chosen a specific break point, it is appropriate to apply the Chow test to determine stability of the estimated intercept and slope coefficient in the cointegrating relationship. This establishes that the values are not constant across the two periods.

13. Since tests of stationarity are asymptotic, it is not surprising that we found disagreement among the test results over short time periods. This is because the tests are less accurate over shorter periods, and may not be able to identify components of a non-stationary series that move very slowly in such a short window of time.

Figure 9: Shifting Cointegrating Relationship, 1997–2009

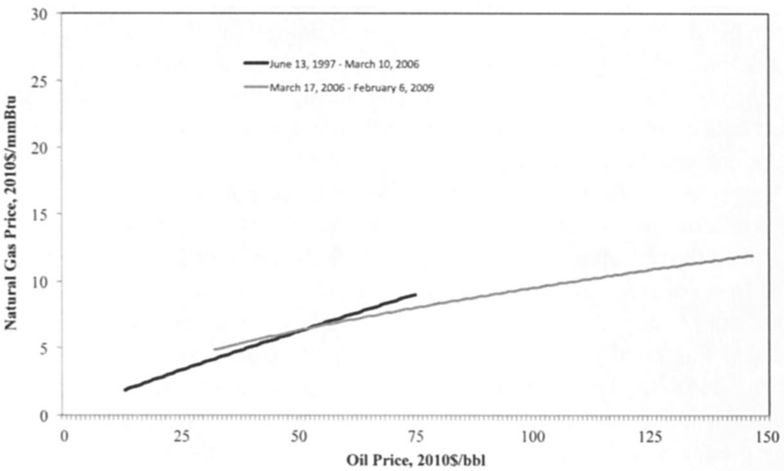


Figure graphs our estimated cointegrating relationship on two sub-periods of our sample suggested by the Gregory Hansen (1996) test. Each relationship is estimated as a part of the full Vector Error Correction Model. A Chow test confirms that the intercept and slope coefficients in the two sub-periods are different from one another.

Of course, the conditional ECMs based on each of the segmented cointegrating relationships also account for a greater portion of the volatility in natural gas price changes than the model covering the period as a whole. The June 13, 1997–March 10, 2006 model accounts for nearly 21% of natural gas price change volatility through the crude oil price and the included conditioning variables. The March 17, 2006–February 6, 2009 model accounts for 26% of the price change volatility in natural gas. Nevertheless, there remains a large amount of unexplained volatility in changes in the natural gas price even in each of these separately estimated time windows.

These results support the hypothesis that whatever relationship might characterize the prices of natural gas and oil, that relationship is not stable over long periods of time. Earlier researchers documented a statistically reliable relationship through a window of years when the price of gas shifted upward relative to the price of oil, and we have documented a statistically reliable relationship during subsequent years when the price of gas was lower relative to the price of oil. Today's tie between the price of natural gas and the price of oil may not be very predictive of tomorrow's tie.

5. CONCLUSION

A number of recent academic studies have established that natural gas and crude oil prices are cointegrated. However, recent years have witnessed a price of natural gas that seems decoupled from the price of oil, reaching new lows relative to the price of oil. In this paper we have confronted the apparent contradiction between these two facts by examining more closely what is and is not established by the cointegration tests. While we are able to reconfirm the presence of a statistically significant relationship between the two price series, our results emphasize two other points that are important to any discussion about a relationship or a decoupling.

First, there is an enormous amount of unexplained volatility in natural gas price changes. The raw price series for natural gas is approximately twice as volatile as the raw oil price series. Applying a VECM and estimating a conditional ECM to account for the predictable error correction and for exogenous variables which temporarily disturb the relationship still leaves an enormous amount of volatility in natural gas price changes unaccounted for. Our model of the 1997–2010 period only accounts for about 15% of the volatility in natural gas price changes, leaving 85% unaccounted for. Splitting the sample up into shorter periods produces only a modest improvement in the fit, in-sample. There is no escaping the significant size of the short-run swings in the natural gas price that cannot be accounted for. At short horizons, the cointegrating relationship is statistically identified, but not very reliable for predicting the natural gas price with any precision.

Second, the cointegrating relationship itself has changed over time, shifting upward in one era and downward again in a later era. These shifts are likely due to shifts in the underlying technological and economic forces determining an equilibrium relationship between the two prices. In other words, variables relating to other technological or economic forces were omitted from our model, and these may be responsible for movements in the natural gas price that the model could not account for. Therefore, the historical cointegrating relationship may not be a very reliable predictor of the future natural gas price, at least not at longer horizons over which shifts in the underlying forces are unpredictable.

This analysis can inform the repeated discussions about how the natural gas price has "decoupled" from the oil price. First, our documentation of the unaccounted for volatility points out that there are likely to be many occasions when the prices temporarily break away from the usual relationship to which they will later return. These decouplings can be severe, but they are also not very long lasting—less than one season typically—and the old relationship is reestablished. Second, our documentation that the cointegrating relationship has shifted over time, first in one direction and then in another, points out that prices can decouple from one relationship only to recouple in a new relationship. Third, there is not yet any evidence that the relationship between the two price series has been severed completely. Indeed, it is hard to imagine that natural gas and oil prices could decouple completely and permanently. For example, while conversion of gas to liquids may seem expensive now, the technological possibility of conversion does place a cap on the degree to which oil prices can rise relative to natural gas prices. Other technological and economic constraints act similarly to prevent a complete decoupling. However, the freedom of motion is large.

REFERENCES

Adelman, M.A., and G.C. Watkins (1997). "The value of United States oil and gas reserves." *Advances in the Economics of Energy and Resources,* JAI Press Inc., 10: 131–184.

Brown, Stephen P.A. and Mine K. Yücel (2008). "What drives natural gas prices?" *The Energy Journal* 29(2): 45–60. http://dx.doi.org/10.5547/ISSN0195-6574-EJ-Vol29-No2-3.

Hartley, Peter R., Kenneth B. Medlock III, and Jennifer E. Rosthal (2008). "The relationship of natural gas to oil prices." *The Energy Journal* 29(3): 47–65. http://dx.doi.org/10.5547/ISSN0195-6574-EJ-Vol29-No3-3.

Hendry, David F., and Katarina Juselius (2000). "Explaining cointegration analysis: Part 1" *The Energy Journal* 21(1): 1–42, and "Part 2" *The Energy Journal* 22(1): 75–120.

IEA (International Energy Agency) (2009). *World Energy Outlook.* OECD. Paris, France.

Ramberg, David J. and John E. Parsons (2011). "The Weak Tie Between Natural Gas and Oil Prices." *CEEPR Working Paper* 10-017, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA. http://web.mit.edu/ceepr/www/publications/workingpapers/2010-017.pdf.

Serletis, Apostolos and John Herbert (1999). "The message in North American energy prices." *Energy Economics* 21: 471–483. http://dx.doi.org/10.1016/S0140-9883(99)00015-8.

Smith, J.L. (2004) "Petroleum property valuation." *Encyclopedia of Energy*, Elsevier, 4: 815–816.

Villar, Jose, A., and Frederick L. Joutz (2006). "The relationship between crude oil and natural gas prices." Energy Information Administration, Office of Oil and Gas.