The Switching Relationship between Natural Gas and Crude Oil Prices

Matthew Brigida*

Abstract:

In this analysis we more accurately capture the cointegrating relationship between natural gas and crude oil prices by endogenously incorporating shifts in the cointegrating vector into the estimation of the cointegrating equation. Specifically, we allow the cointegrating equation to switch between *m* states, according to a first-order Markov process. First, we find evidence that regime-switching exists in the relative pricing relationship, and that two is the optimal number of states. Once we control for shifts in the cointegrating vector, we find natural gas and crude oil prices are cointegrated, and an error correction model (ECM) of their long-term equilibrium relationship is properly specified. This finding broadens the ECM model of their relationship to longer and more varied sample periods. Also, in a direct comparison of the two and one state cointegrating equations, we found evidence of the potential superiority of the two-state equation, in that it may be robust to shifts in the cointegrating vector which are missed by standard tests for a unit root. Further, our analysis finds evidence that natural gas and crude oil prices did not permanently 'decouple' in the early 2000s, but rather experienced a temporary shift in regimes. We find that forecasts of the relative pricing of natural gas and crude oil should be conditioned on state probability.

[This version: December, 2013]

*Corresponding Author: Associate Professor of Finance, Clarion University of Pennsylvania. Mailing address:

Department of Finance, 840 Wood St., Clarion, PA 16214; email: mbrigida@clarion.edu; phone: 814.393.2603

The Switching Relationship between Natural Gas and Crude Oil Prices

1. Introduction

Many recent studies on the long-term relationship between natural gas and crude oil prices have found that the series are generally cointegrated (Serletis and Herbert (1999); Villar and Joutz (2006); Bachmeier and Griffin (2006); Brown and Yucel (2008); Hartley, Medlock and Rosenthal (2008); Ramberg and Parsons (2012))[1]. However, noted in many of these analyses is the likely presence of structural breaks in the relationship.

This is prompted by the observation that over the last 30 years there has been wide variation in the ratio of crude oil to natural gas prices. This ratio was above 10 for much of the 1985—1995 'gas bubble' period, and then below 10 until 2005. Since 2009 the ratio has spiked above 30. This marked change in the ratio since 2009 has renewed speculation of 'decoupling' between the price series. In addition, Ramberg and Parsons (2012) found evidence for structural breaks in 2006 and 2009 using the Gregory and Hansen (1996) test for a single structural break in a cointegrating relationship[2].

Our present analysis builds on the prior literature by endogenously incorporating shifts in the cointegrating vector into the estimation of the cointegrating equation. That is, we model the structural breaks in the relative pricing relationship as switches between cointegrating regimes, and these switches are endogenously determined according to a first-order Markov process. This

---

[1] For more on the methodology employed in these cointegrating analyses see: Hendry and Juselius (2000) and (2001); Engle and Granger (1987).

[2] They used the version of the test based on the augmented Dickey-Fuller test statistic. While the test is for a single structural break, Ramberg and Parsons (2012) used the test over the period 1997—2010 and found a break in February 2009. They then used the test again over the interval 1997 through February 2009 and found a break in March 2006. They did not choose to repeat this procedure over the sample period 1997 through March 2006.

approach will afford a probability law over the entire data generating process which takes into account distinct changes in the cointegrating vector.

Once such regime changes are controlled for in this manner, one can model the long-term equilibrium relationship between natural gas and crude oil over wider and more varied sample periods. This affords a better measure of present energy market integration, possibly better forecasts of relative prices, and a more thorough understanding of how technological changes affect the natural gas and crude oil pricing relationship.

That the regime-switching is endogenous is an important point. In our model, changes in regime are determined solely by the underlying data generating process. This obviates biases an econometrician may have in determining whether regime changes exist, and the timing of such changes.

Some events which may induce regime switches in the relationship between natural gas and oil are technological changes and legislation, among others. For example, Hartley et al (2008) found evidence that the marked increase in the use of the combined-cycle combustion turbines for electricity generation in the late 1990s made natural gas electricity generation more cost effective, thereby substantially increasing demand for natural gas and thereby increasing prices. More recently, an increase in the supply of shale gas because of the introduction of hydrofracking has driven North American natural gas prices lower.

Including such structural breaks into any model of the relationship is important because, as Villar and Joutz (2006) note, structural changes in the cointegrating equation can cause forecast failure. The present analysis incorporates such structural changes into the cointegrating equation, and thereby the error correction model (ECM).

In this analysis we further offer an answer to the idea that the relative pricing relationship between natural gas and oil permanently 'decoupled'. Using the regime-switching model, and data though 2012, we will show that the parameters governing the relationship between natural gas and crude oil did indeed change from 2000-2009[3], however the parameters have since reverted to their pre-2000 values. That is, the 'decoupling' was a temporary shift in regimes.

Throughout this paper we use a standard error correction model (ECM) analysis similar to prior literature, but with our regime-switching cointegrating equation. We first estimate an ECM over our entire sample and review the results. We then estimate an ECM over a subinterval for which we can also estimate a control (the standard nonswitching cointegrating equation), and compare our results to this control.

Lastly, understanding the relative pricing of natural gas and oil is important for both corporate managers and policymakers. Models of the pricing relationship are necessary to estimate cash flows in the long-term capital budgeting plans of both energy producers and consumers. For instance dynamics of the relationship may dictate whether an energy producer should drill wells to target natural gas or oil. Alternatively, the relative pricing may determine the type of fuel to use when building a power plant. For policymakers the relative pricing may affect decisions from permitting energy transportation infrastructure to setting royalty payments.

The paper is as follows: section 2 presents the Markov-switching cointegration equation, the determination of the number of states, and results; section 3 describes the ECM and results; section 4 concludes.

[3] The relationship switched to a second state in August 2000, and stayed in this state until April 2009, with one interruption from September 2001 to October 2002 which coincides with the Enron collapse and its attendant effect on natural gas markets.

2. Markov-Switching Cointegrating Equation

The cointegrating equation with first-order, M-state, endogenous Markov-switching parameters may be written as:

$$P_{HH} = \beta_{0,S_t} + \beta_{1,S_t} P_{WTI} + e_t, \quad e_t \sim N(0, \sigma_{S_t}^2) \tag{1}$$

$$P(S_t = j | S_{t-1} = i) = p_{ij}, \qquad \forall i,j \in 1,2, \dots M, \text{ and } \sum_{j=1}^{M} p_{ij} = 1 \tag{2}$$

$$\beta_{0,S_t} = \beta_{0,1} S_{1t} + \beta_{0,2} S_{2t} + \cdots + \beta_{0,M} S_{Mt} \tag{3}$$

$$\beta_{1,S_t} = \beta_{1,1} S_{1t} + \beta_{1,2} S_{2t} + \cdots + \beta_{1,M} S_{Mt} \tag{4}$$

$$\sigma_{0,S_t} = \sigma_{0,1} S_{1t} + \sigma_{0,2} S_{2t} + \cdots + \sigma_{0,M} S_{Mt} \tag{5}$$

where for $m \in 1,2, \dots, M$, if $S_t = m$, then $S_{mt} = 1$, and $S_{mt} = 0$ otherwise. P_{HH} and P_{WTI} refer to the log of natural gas and crude oil prices respectively. $\beta_{0,St}$, $\beta_{1,St}$, and σ_{St} are parameters to be estimated for each state S_t, and p_{ij} is the transition probability from sate i to state j.

Construction of the likelihood function for the above Markov Switching cointegrating equation was done using the Hamilton filter (see Hamilton (1994) or Kim and Nelson (1999)). Minimization of the negative log-likelihood was done using the *optim* function in the R programming language. The minimization was unconstrained.

The residuals of the *m*-state model are weighted by *filtered* state probability, which is the probability that the relationship is in state S_t given information only through time $t-1$. This means probabilities in the residual are not biased by using information through time T, as would be the case if we used smoothed state probabilities[4]. The time *t* weighted residual in the *m*-state case is:

$$e_t = e_{S_t=1}P(S_t = 1|\varphi_{t-1}) + \cdots + e_{S_t=m}P(S_t = m|\varphi_{t-1}) \qquad (6)$$

where φ_{t-1} denotes the information available at time *t-1*, and $e_{S_t=m}$ denotes the time *t* residual of the state *m* model.

The data used to estimate the model are monthly and weekly logged prices for rolling front month NYMEX full-size natural gas and oil futures. The crude oil contract is for west Texas intermediate deliverable in Cushing Oklahoma, and the natural gas contract is for delivery at the Henry Hub in Louisiana. The data are available from the Energy Information Agency of the U.S. Department of Energy.

2.1 Determining the number of states

The prior literature has generally alluded to two regimes in the natural gas and crude oil relationship: one regime where crude oil prices are relatively high compared to natural gas (1985—1995 gas bubble and post-2009), and another regime where natural gas prices are relatively high (the interval from 1995—2005). Ramberg and Parsons (2012) also estimate two cointegrating equations over the interval 1997—2010.

However, to determine the appropriate number of states we estimated the cointegrating equation allowing for the number of states to range from one to three. Note, a one-state equation

[4] Smoothed probabilities are $P(S_t = m|\varphi_T)$, which are the probabilities that the model is in state *m* at time *t* given information through time *T*.

is the standard, non-switching, cointegrating equation. We compared the results of each model based on the behavior of the residuals, the estimated state probabilities, and Akaike's Information Criterion (AIC)[5]. Ultimately, we concluded that two states best described the process.

Comparing the one and two state models, we cannot reject a unit root in the residuals of the one-state cointegrating equation using either monthly or weekly prices. This is evidence that the logs of natural gas and oil prices are not cointegrated in the one-state model. Importantly, this means an ECM of the relationship between crude oil and natural gas is internally inconsistent[6] with respect to the residuals in the one-state cointegrating equation.

Alternatively, using the two-state model we reject the null of a unit root in the residuals of the cointegrating equation at the 1% level for both the Augmented Dickey-Fuller and the Phillips-Perron tests, and for both weekly and monthly prices. This is evidence that the two series are cointegrated in the two-state model. That is, the series are cointegrated once we control for shifts in the cointegrating vector. An ECM is therefore appropriate in this case.

[INSERT TABLES 1 and 2 HERE]

Using t-tests on the parameters estimated in the two-state model, we find the β_{0S_t}, β_{1S_t} and σ_{S_t} coefficients are all significantly different at the 1% level depending on the state. That is, $\beta_{01}$is significantly different from β_{02}, etc. Lastly, the AIC is lower for the two-state (-0.2086 weekly; -0.0032 monthly) than the one-state (0.9902 weekly; 0.9919 monthly) model.

---

[5] Given unresolved issues in implementing likelihood ratio tests (such as in Hansen (1992)) to determine the number of states, Psaradakis and Spagnolo (2003) employ a Monte Carlo analysis to test the performance of methods based on complexity-penalized likelihood criteria in Markov Switching autoregressive models. They found that the AIC was generally successful in choosing the number of states, so long as (1) the sample size, and (2) the parameter changes, are both sufficiently large.

[6] Because the error correction term, and hence the right hand side, has a unit root.

Turning to the two-state versus three-state model, residuals from both models reject a unit root at the 5% level of significance. However, note that in the two-state model the dates of the regime switches are consistent when using price series sampled and the monthly and weekly frequency. Conversely, in the three-state model the regimes switch at differing times when using monthly and weekly series. The two-state model also exhibits distinct regime switching in that at each time point the model is confident of either being in state 1 or 2. However, the three-state model using monthly series is usually unclear about which state the process is in at each time point. Figures 1, 2, 3, 4 contain filtered state probabilities for the two and three state models. The AIC is also lower for the two-state (-0.2086 weekly; -0.0032 monthly) rather than the three-state (0.8620 weekly; 0.7385 monthly) model.

Lastly, note that the state probabilities in the two-state model are stable in that once they are in a given state they remain in that state for some time[7]. For the two-state model, using monthly prices, the expected duration is 6.94 years and 4.66 years for state 1 and 2 respectively. Using weekly prices, the expected duration is 3.10 years and 2.13 years for state 1 and 2 respectively[8]. In fact, once the model switched to state 2 in August of 2000, it remained in that state until April 2009, except for one interval from September 2001 to October 2002 which coincided with the Enron collapse and its effect on natural gas markets. In sum, the two-state model exhibits stable states with distinct regime switching. We conclude, by every measure, the

[7] using the estimated transition probabilities, we may calculate the expected duration of a given state *m* as:

$$E(D_m) = \sum_{j=1}^{\infty} j(P[D_m = j]) = \frac{1}{1-p_{mm}}$$

where $E(D_m)$ is the expected duration of state *m*, and p_{mm} is the probability of being in state *m* at time *t* given the process was in state *m* at time *t-1*.
[8] We expect the expected state duration to be shorter for weekly prices given the increased number of opportunities to transition in a fixed time period.

two-state model is more appropriate for modeling the cointegrating equation of natural gas and oil prices.

[INSERT FIGURES 1, 2, 3, AND 4 HERE]

3. Error Correction Model with Regime Dependent Residuals

In this section we use the state-weighted residuals from the switching cointegration regression to estimate an ECM of natural gas and crude oil prices over the entire sample period (June 1997 – September 2012). We report results from this ECM, however standard practice is to not compare them to a similar ECM using the nonswitching residuals, because as we saw in the earlier section we cannot reject a unit root in the nonswitching residuals[9].

We therefore also find the longest, and most recent, time period where the residuals from both the one and two state cointegrating equations admit rejection of the null hypothesis of a unit root at a 5% significance level. This is the subinterval from October 2004 to September 2012. We then estimate ECMs, using both one and two state cointegrating equation residuals, over this subinterval of the full sample period and compare the results.

We estimate a matching conditional ECM, wherein crude oil is treated as exogenous. Our ECM is:

$$\Delta P_{HHt} = \beta_0 + \beta_1 e_{t-1} + \beta_2 \Delta P_{WTIt} + \sum_{i=1}^{p} \rho_i \Delta P_{HHt-i} + \sum_{n=1}^{N} \mu_n X_{nt} + \xi_t \qquad (8)$$

where X is a matrix of exogenous control variables. These variables are: cooling degree days (CDD); heating degree days (HDD); deviations from the average number of CDDs and HDDs;

[9] We do however include in an appendix a comparison of the full sample ECM results despite a unit root in the nonswitching residuals. The results confirm that switching residuals perform better.

the natural gas storage differential from the 5-year average; Baker Hughes' North American rotary rig count.

All of the control variables are standard except for the rig count. We included this variable because there is anecdotal evidence[10] that, because natural gas prices have dropped below marginal production costs, rig count has become more sensitive to natural gas prices. As the natural gas price increases toward marginal costs, more rigs are brought online, and conversely when prices decline rigs are idled. We therefore expect that changes in the rig count to be positively related to contemporaneous changes in the price of natural gas.

3.1 Data

Natural gas and oil prices are, as earlier mentioned, rolling front-month logged futures prices from the NYMEX and available from the EIA. The natural gas storage data are also from the EIA, and are included in the ECM as monthly deviations from that month's prior 5-year average (*STORAGE DIFF*). CDD and HDD are population-weighted national averages and are available from the U.S. National Weather Service's Climate Prediction Center. They are included in the ECM in levels (*CDD* and *HDD*), and as 'monthly deviations from the norm' as defined by the Climate Prediction Center (*CDDDEV* and *HDDDEV*). Rig count data are available from Baker Hughes, and the count is first-differenced (*RIG COUNT DIFF*). Augmented Dickey-Fuller and Phillips-Perron tests reject a unit root in all exogenous variables.

[10] From a recent article in Bloomberg news, "Gas producers in North America including Chesapeake Energy Corp. (CHK) are killing their commodity's biggest rally in 10 months by opening more wells, putting the U.S. on track to have record gas supplies this year". http://www.bloomberg.com/news/2012-11-14/gas-prices-doomed-to-stay-low-as-producers-pump-faster.html

3.2 Full sample ECM results: Switching cointegrating equation[11]

The coefficient of the cointegrating term is about -0.15 for the two-state model. This implies a 90% adjustment to equilibrium takes approximately 14 months. The coefficient of the cointegrating term is negative and significant, and is robust to the number of lagged changes in natural gas included in the ECM.

Consistent with theory and prior literature the coefficient of the lagged change in logged crude oil (WTI) is positive and significant across all ECM specifications, except for one where fuel oil is included. This reflects that, when all else is held constant, an increase in crude oil prices tends to increase natural gas prices in subsequent periods.

Regarding the rig count explanatory variable, the change in rig count is positively and significantly related to changes in natural gas prices in the full ECM (12 natural gas price lags). The coefficient of 0.0002 implies an increase in the rig count of 100 rigs coincides with a 2.02% increase in natural gas prices. The average monthly variation in the rig count is 94. The exogenous factors in the ECM (heating and cooling degree day variables, and natural gas storage) are all of the appropriate sign and most are significant. Lastly, the full-sample ECMs explain about 27% to 38% of the variation in the logged differences in natural gas prices.

We conclude that the residuals from the two-state cointegrating equation may be used to model the long-term relationship between natural gas and crude oil in the ECM framework. The results of using the two-state residuals are consistent with the underlying economic theory and prior literature. Conversely, over our sample, an ECM with a standard one-state cointegrating equation would otherwise be rendered not useable because of a unit root in the residuals induced

[11] Following are results of the ECM on monthly prices. Estimates of the ECM on weekly data are similar and available on request.

by regime-switching. However, over the same sample, a two-state Markov-switching cointegrating equation successfully controls for shifts in the cointegrating vector and affords a properly specified ECM of the relationship between natural gas and crude oil prices. This broadens the applicability of the ECM to longer, and more varied, sample periods.

For robustness, we include results of an ECM with contemporaneous and once lagged changes in logged residual fuel oil[12] in columns 3 and 4 of table 3. The results are robust to the inclusion of fuel oil.

[INSERT TABLE 3 HERE]

3.3 Subsample results: Switching versus nonswitching cointegrating equation

We estimated ECMs using residuals from the both two and one state cointegrating equations over the months from October 2004 to September 2012. This is the longest, and most recent, subinterval over which can reject a unit root at 5% in the residuals from both the two and one state cointegrating equations. These ECMs enable us to directly compare the performance of the two cointegrating equations.

Notably, the error correction term is negative and significant in the ECMs which use the two-state residuals. However, using the one-state residuals we find the error correction terms are insignificant, and in one ECM the term is positive[13]. We can conclude, over this sample, the two-state residuals are consistent with the ECM form, though the one-state residuals are not.

The crude oil term is positive and significant in the ECMs using both the two and one state residuals. Also, the exogenous terms are all of the appropriate sign. Lastly, the ECMs with

---

[12] 'U.S. Residual Fuel Oil Retail Sales by Refiners (Dollars per Gallon)' available from the EIA.
[13] Note, a positive error correction term is inconsistent with an ECM, because it implies deviations from equilibrium are compounded, not reduced, over time.

two-state residuals explain slightly more of the variation in natural gas than the ECMs with one-state residuals. The ECM with two-state residuals and 12 months of lagged changes in logged natural gas explains 43.7% of the logged change in natural gas.

[INSERT TABLE 4 HERE]

A possible explanation of the poor performance of the ECMs using the one-state cointegrating equation residuals may be seen in figure 5 below. Figure 5 shows the state probabilities for the two-state cointegrating equation over the October 2004 to September 2012 subsample. This figure shows clear regime switching, despite the one-state cointegrating equations having stationary residuals at the 5% level of significance (-3.9324 augmented Dickey-Fuller test statistic with a p-value of 0.0156)[14]. That is, the two-state cointegrating equation compensated for the shift in the cointegrating vector, which led to a properly specified ECM. Conversely, this shift was missed using standard tests for a unit root in the one-state residuals, which led to an improperly specified ECM.

[INSERT FIGURE 5 HERE]

[14] This prompts a questions about why the ADF test rejected the null of a unit root despite the structural break in the cointegrating relationship. Gregory, Nason, and Watt (1996), in a Monte Carlo experiment, investigated the rejection frequency of the ADF test in the presence of such structural breaks. In an experiment with a cointegrating relationship of length 100 with low serial correlation in the error term they found the rejection frequency was: 96% if there was no structural break; 68% if a structural break occurred at time 0.25 (where time ranges from 0 to 1), 39% if the break was at time 0.5; 31% if the break was at time 0.75. While the marked drop in rejection frequency is evidence that the ADF test correctly indicates against the constant parameter cointegrating relationship, there is still substantial probability that the test will reject the null in the presence of a structural break.

4. Conclusions

In this analysis we have found evidence for a regime-switching relationship between natural gas and oil prices. This evidence is, firstly, that the cointegrating equation is well-suited to a regime-switching model with two states. The cointegrating equation's residuals have a stationary distribution when using the switching model, and contain a unit root when the relationship is constrained to one state. This is evidence that a two-state cointegrating equation is more broadly applicable to longer, and more varied, sample periods than the one state cointegrating equation. The filtered probabilities of the two-state equation also exhibit stable states with distinct regime-switching.

In an ECM over the full sample, which would not be properly specified with one-sate residuals, the two-state cointegrating equation provided an estimate of the long-term relationship between crude oil and natural gas which is consistent with theory and prior literature. This is evidence that the two-state cointegrating equation successfully controls for parameter shifts in the cointegrating vector over long sample periods.

Additional evidence for the superiority of the regime-switching model is found by a direct comparison of the results of an ECM of natural gas and crude oil prices with two-state residuals, to the results of the same ECM using one-state residuals. Specifically, we found evidence that the two-state cointegrating equation controls for shifts in the parameters of the cointegrating vector which may be missed in standard tests for a unit root in the one-state residuals.

This is evidence that if you specifically choose a time interval where the residuals from the one-state cointegrating equation reject a unit root (the two-state residuals also rejecting a unit root) then the ECM model implied by the two-state cointegrating equation is *possibly better* than the ECM implied by the one-state cointegrating equation. These results imply, at a minimum, that tests for changes in regimes should accompany any analysis of the long-term relationship between natural gas and crude oil prices.

These results have practical implications beyond motivating the use of Markov-switching cointegrating equations. Firstly, this analysis shows that there is a stronger, and longer lasting, relationship (reversion to a long-term equilibrium) between natural gas and crude oil prices once one controls for endogenous regime switching. This implies these energy markets are more integrated that one would otherwise estimate.

The results also imply that natural gas and crude oil prices did not permanently decouple in the early 2000s, but rather exhibited a temporary shift in August of 2000 to a regime wherein natural gas prices performed relatively better than crude oil prices. This regime lasted, with one interruption coinciding with the Enron collapse, until approximately May of 2009, after which the relationship has reverted to its original state with oil price increases outpacing natural gas prices.

In sum, the two-state cointegrating equation allows for a more thorough and accurate understanding of the long-term equilibrium relationship of natural gas and crude oil prices. Moreover this understanding spans broader and more varied sample periods. This is evidence that models of the relative pricing of natural gas and crude oil should be conditioned on state probability.

References

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

Brown, S.P.A., and M.K. Yucel, (2008) "What drives natural gas prices?" *The Energy Journal* 29(2): 45-60.

Engle, R. F., and C.W.J. Granger (1987). "Cointegration and Error Correction: Representation, Estimation, and Testing." *Econometrica 50: 987-1007.*

Gregory, A.W., and B.E. Hansen (1996). "Residual-based Tests for Cointegration in Models with Regime Shifts." *Journal of Econometrics* 70: 99-126.

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

Kim, C.J., and C.R. Nelson, (1999) *State-Space Models with Regime Switching*, The MIT Press.

Hamilton, J.D. (1994). *Time Series Analysis.* Princeton: Princeton University Press.

Hansen, B.E. (1992). "The likelihood ratio test under nonstandard conditions: Testing the markov switching model of GNP." *Journal of Applied Econometrics* 7, S61—S62.

Hartley, P.R., K.B. Medlock III, and J.E. Rosthal (2008). "The relationship of natural gas to oil prices." *The Energy Journal* 29(3): 47-65.

Hendry, D.F. and K. Juselius (2000). "Explaining Cointegration Analysis: Part I." *The Energy Journal,* 21(1): 1-42.

Hendry, D.F., and K. Juselius. (2001). "Explaining Cointegration Analysis: Part II." *The Energy Journal,* 22(1): 75-120.

Psaradakis, Z., and N. Spagnolo. (2003). "On the determination of the number of regimes in Markov-switching autoregressive models." *Journal of Time Series Analysis* 24, 237-252.

Ramberg, D.J., and J.E. Parsons (2012). "The weak tie between natural gas and oil prices" *The Energy Journal* 33(2): 13-35.

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

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

Tables

Table 1. Results of Estimating Cointegrating Equation Regressions (equations 1-5) with Monthly Prices

There are 225 monthly observations for natural gas and oil prices. The one-state model has 222 degrees-of-freedom, and the two-state model has 217. The alternative hypothesis in the augmented Dickey-Fuller and Phillips-Perron tests is to conclude the residual series is stationary. Residuals in the two-state model are weighted by the filtered state probability. p-values are below the coefficient in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels respectively.

	One-State Model	Two-State Model	
State	1	1	2
β_0	-0.6910 (5e-6)***	-0.0287 (0.7518)	0.3351 (0.2068)
β_1	0.5608 (<1e-10)***	0.2889 (<1e-10)***	0.3949 (1.4e-8)***
σ^2	0.1537 (<1e-10)***	0.2017 (<1e-10)***	0.2634 (<1e-10)***
$P(S_t = 1 \mid S_{t-1} = 1)$		0.9880	
$P(S_t = 2 \mid S_{t-1} = 2)$		0.9821	
-ln(Likelihood)	108.59	-8.36	
AIC	0.9919	-0.0032	
Augmented Dickey-Fuller Test	-1.9168 (0.6109)	-4.4297 (0.01)***	
Phillips-Perron Test	-10.062 (0.5434)	-46.2419 (<0.01)***	

Table 2. Results of Estimating Cointegrating Equation Regressions (equations 1-5) with Weekly Prices

There are 978 weekly observations for natural gas and oil prices. The one-state model has 975 degrees-of-freedom, and the two-state model has 970. The alternative hypothesis in the augmented Dickey-Fuller and Phillips-Perron tests is to conclude the residual series is stationary. Residuals in the two-state model are weighted by the filtered state probability. p-values are below the coefficient in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels respectively.

	One-State Model	Two-State Model	
State	1	1	2
β_0	-0.6867 (<1e-10)***	-0.1350 (0.0050)***	0.2806 (0.0043)***
β_1	0.5591 (<1e-10)***	0.3214 (<1e-10)***	0.4074 (<1e-10)***
σ^2	0.1566 (<1e-10)***	0.2116 (<1e-10)***	0.2025 (<1e-10)***
$P(S_t = 1 \mid S_{t-1} = 1)$		0.9938	
$P(S_t = 2 \mid S_{t-1} = 2)$		0.9910	
-ln(Likelihood)	481.2024	-110.1142	
AIC	0.9902	-0.2086	
Augmented Dickey-Fuller Test	-2.3475 (0.4312)	-5.2500 (0.0100)***	
Phillips-Perron Test	-14.7862 (0.2849)	-75.7543 (<0.01)***	

Table 3. Full-Sample, Two-State (Switching), Error-Correction Model Results: Below are estimates of the error correction model on monthly data from June 1997 to September 2012. There are 184 observations. The response variable is the change in the log Henry Hub natural gas price from time *t-1* to time *t*. *, **, *** denote significance at the 10%, 5%, and 1% levels respectively.

	Does not include fuel oil		**Includes fuel oil**	
Intercept	0.1117 (0.0028)***	0.0850 (0.0103)**	0.1202 (0.0016)***	0.1018 (0.0033)***
e_{t-1}	-0.1295 (0.0182)**	-0.1527 (0.0054)***	-0.1363 (0.0146)**	-0.1568 (0.0049)***
ΔP_{WTIt}	0.3457 (0.0004)***	0.3603 (0.0003)***	0.2625 (0.0499)**	0.2241 (0.1063)
ΔP_{RFOt}			0.1921 (0.2704)	0.3059 (0.0863)*
ΔP_{RFOt-1}			-0.1799 (0.1898)	-0.2636 (0.0595)*
ΔP_{HHt-1}	0.1135 (0.0820)*	0.1883 (0.0119)**	0.1389 (0.0867)*	0.1877 (0.0163)**
ΔP_{HHt-2}	-0.0550 (0.4773)	-0.0007 (0.9925)	-0.0147 (0.8594)	0.0390 (0.6217)
ΔP_{HHt-3}	-0.0832 (0.2845)	-0.0134 (0.8562)	-0.0751 (0.3376)	-0.0061 (0.9340)
ΔP_{HHt-4}	0.0540 (0.4498)	0.0908 (0.2070)	0.0775 (0.2918)	0.1184 (0.1088)
ΔP_{HHt-5}	-0.1064 (0.1330)		-0.1024 (0.1492)	
ΔP_{HHt-6}	0.0064 (0.9274)		0.0065 (0.9277)	
ΔP_{HHt-7}	-0.0201 (0.7677)		-0.0161 (0.8141)	
ΔP_{HHt-8}	-0.1159 (0.0851)*		-0.1184 (0.0809)*	
ΔP_{HHt-9}	-0.2603 (0.0002)***		-0.2662 (0.0002)***	
ΔP_{HHt-10}	-0.0113 (0.8746)		0.0003 (0.9967)	
ΔP_{HHt-11}	-0.0578 (0.4070)		-0.0493 (0.4818)	
ΔP_{HHt-12}	-0.1261 (0.0721)*		-0.1247 (0.0764)*	
CDD_t	-0.0006 (0.0001)***	-0.0005 (0.0002)***	-0.0006 (0.0001)***	-0.0006 (0.0000)***
$CDDDEV_t$	0.0025 (0.0000)***	0.0022 (0.0000)***	0.0025 (0.0000)***	0.0022 (0.0000)***
HDD_t	-0.0001 (0.0998)*	-6.9e-05 (0.1621)	-0.0001 (0.0586)*	-9.7e-5 (0.0612)*
$HDDDEV_t$	0.0008 (0.0000)***	0.0008 (0.0000)***	0.0008 (0.0000)***	0.0008 (0.0000)***
STORAGE $DIFF_t$	-0.0001 (0.0233)**	-2.2e-05 (0.5597)	-0.0001 (0.0475)**	-1.4e-5 (0.7012)
RIG COUNT $DIFF_t$	0.0002 (0.0446)**	6.2e-05 (0.4234)	0.0002 (0.0490)**	6.9e-5 (0.3742)

F-statistic	6.316***	6.622***	5.851***	6.095***
adj. R^2	0.3848	0.2748	0.3871	0.2872
Q-stat (12 lags)	6.3823	17.6161	6.2582	20.6658*
Breusch-Pagan	24.1399	11.9726	26.5444	16.2316

Table 4. Subsample Error-Correction Model Results: Switching versus Nonswitching Below are estimates of the error correction model on monthly data from October 2004 to September 2012. There are 96 observations. Over this subinterval both the one and two state cointegrating equations have stationary residuals, and so each ECM is internally consistent. The response variable is the change in the log Henry Hub natural gas price from time *t-1* to time *t*. *, **, *** denote significance at the 10%, 5%, and 1% levels respectively.

Variable	**Switching**		**No switching**	
Intercept	0.0941 (0.0746)*	0.1123 (0.0235)**	0.0869 (0.1098)	0.1270 (0.0163)**
e_{t-1}	-0.0669 (0.0943)*	-0.0929 (0.0187)**	0.0501 (0.2391)	-0.0375 (0.3060)
ΔP_{WTIt}	0.2620 (0.0359)**	0.3085 (0.0184)**	0.2223 (0.0728)*	0.2738 (0.0408)**
ΔP_{HHt-1}	-0.0468 (0.6622)	0.0954 (0.3329)	-0.0929 (0.4051)	0.0964 (0.3433)
ΔP_{HHt-2}	-0.0025 (0.9809)	0.1034 (0.2919)	-0.0532 (0.6201)	0.0893 (0.3761)
ΔP_{HHt-3}	-0.0443 (0.6826)	0.0251 (0.8056)	-0.1015 (0.3563)	-0.0007 (0.9949)
ΔP_{HHt-4}	-0.0077 (0.9348)	0.0415 (0.6721)	-0.0533 (0.5765)	0.0147 (0.8841)
ΔP_{HHt-5}	-0.1154 (0.2397)		-0.1855 (0.0661)*	
ΔP_{HHt-6}	0.0512 (0.5801)		0.0004 (0.9962)	
ΔP_{HHt-7}	-0.0987 (0.2966)		-0.1679 (0.0962)*	
ΔP_{HHt-8}	0.0043 (0.9628)		-0.0599 (0.5371)	
ΔP_{HHt-9}	-0.2440 (0.0093)***		-0.2801 (0.0039)***	
ΔP_{HHt-10}	-0.2544 (0.0128)**		-0.2962 (0.0059)***	
ΔP_{HHt-11}	-0.0434 (0.6577)		-0.0558 (0.5769)	
ΔP_{HHt-12}	-0.1075 (0.2629)		-0.1400 (0.1577)	
CDD_t	-0.0008 (0.0006)***	-0.0009 (0.0000)***	-0.0009 (0.0002)***	-0.0009 (0.0000)***
$CDDDEV_t$	0.0034 (0.0000)***	0.0031 (0.0000)***	0.0040 (0.0000)***	0.0031 (0.0001)***
HDD_t	-0.0001 (0.1324)	-0.0001 (0.0560)*	-0.0001 (0.0776)*	-0.0002 (0.0257)**
$HDDDEV_t$	0.0007 (0.0073)***	0.0010 (0.0002)***	0.0008 (0.0023)***	0.0011 (0.0004)***
STORAGE $DIFF_t$	-0.0001 (0.2887)	-6.5e-6 (0.9245)	-2.7e-05 (0.7247)	-9.7e-7 (0.9908)
RIG COUNT $DIFF_t$	0.0001 (0.2167)	0.0001 (0.2232)	0.0001 (0.2975)	0.0001 (0.3134)
F-statistic	4.1830***	4.6880***	4.0170***	4.0310***

adj. R^2	0.4370	0.3505	0.4239	0.3073
Q-stat (12 lags)	4.4443	8.0882	4.0019	8.499
Breusch-Pagan	18.2154	7.9087	23.2796	5.3393

Figures

Figure 1. Monthly Data: Both the natural gas and crude oil price series below are in logs and less the mean of the logged series. Below the price series is the probability that the relative pricing relationship is in state 2.

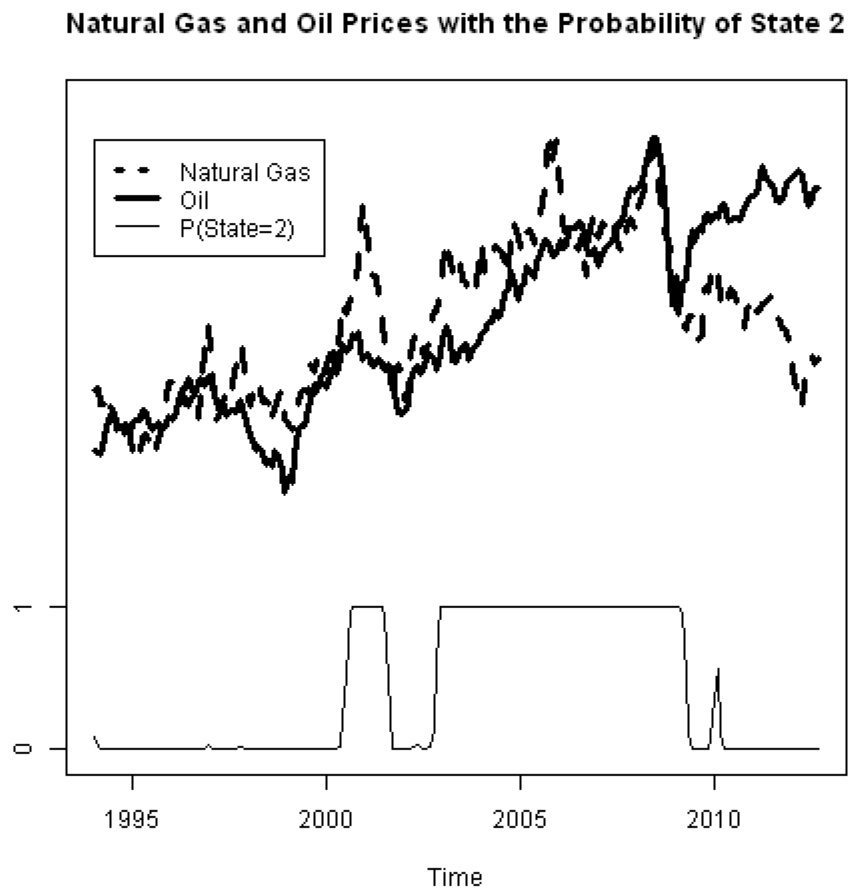


Figure 2. Weekly Data: Both the natural gas and crude oil price series below are in logs and less the mean of the logged series. Below the price series is the probability that the relative pricing relationship is in state 2.

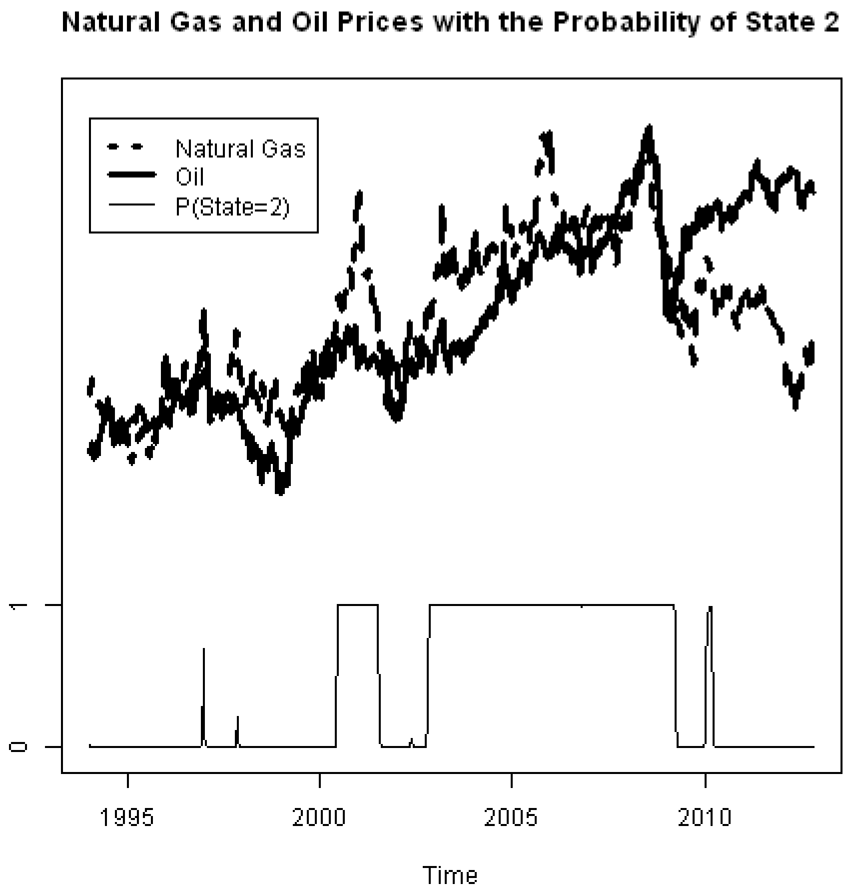


Figure 3. 3 states, monthly series: Below is the probability that the relative pricing relationship between natural gas and crude oil is in each state in the 3-state cointegrating equation.

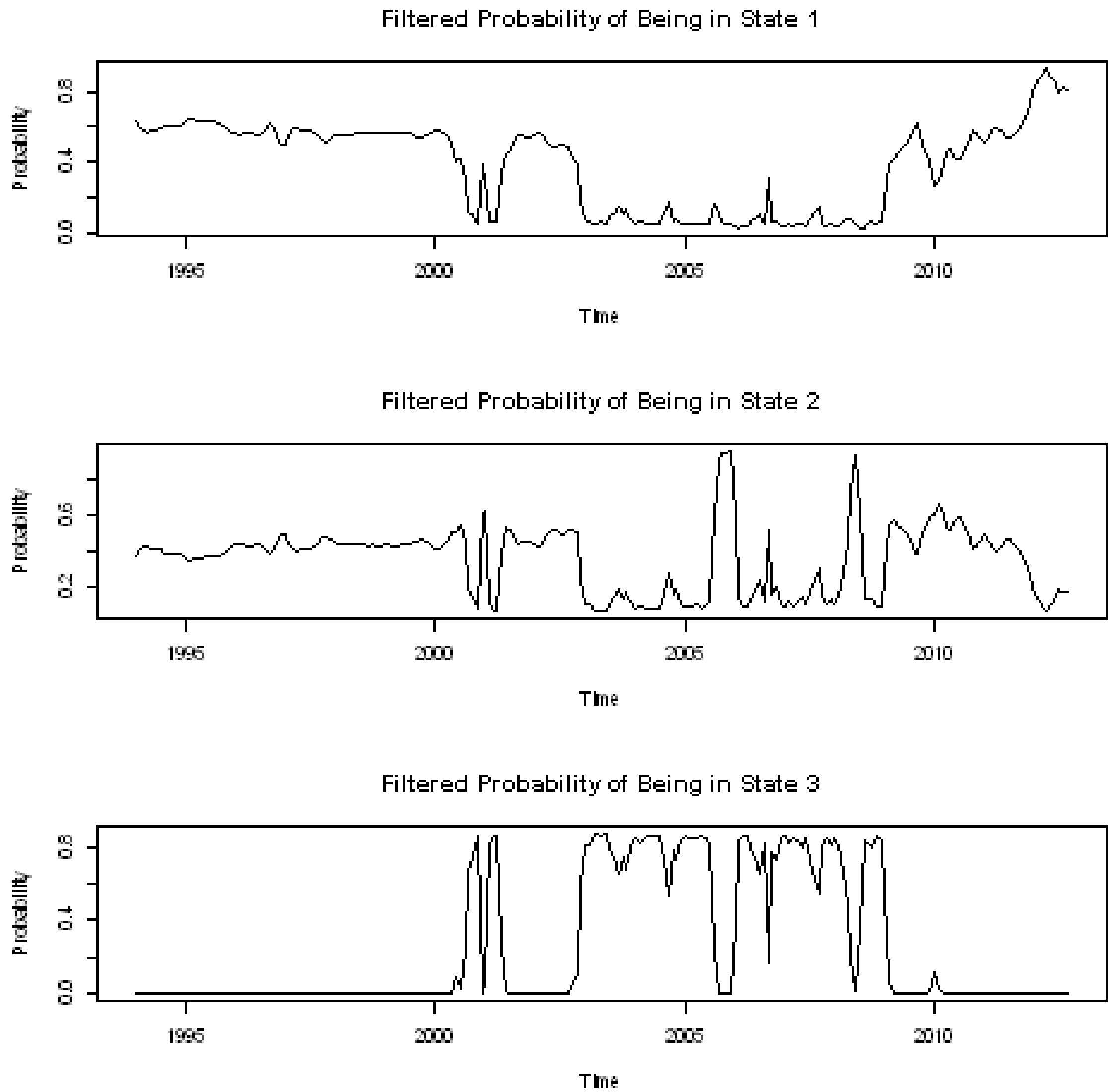


Figure 4. 3 state, weekly series: Below is the probability that the relative pricing relationship between natural gas and crude oil is in each state in the 3-state cointegrating equation.

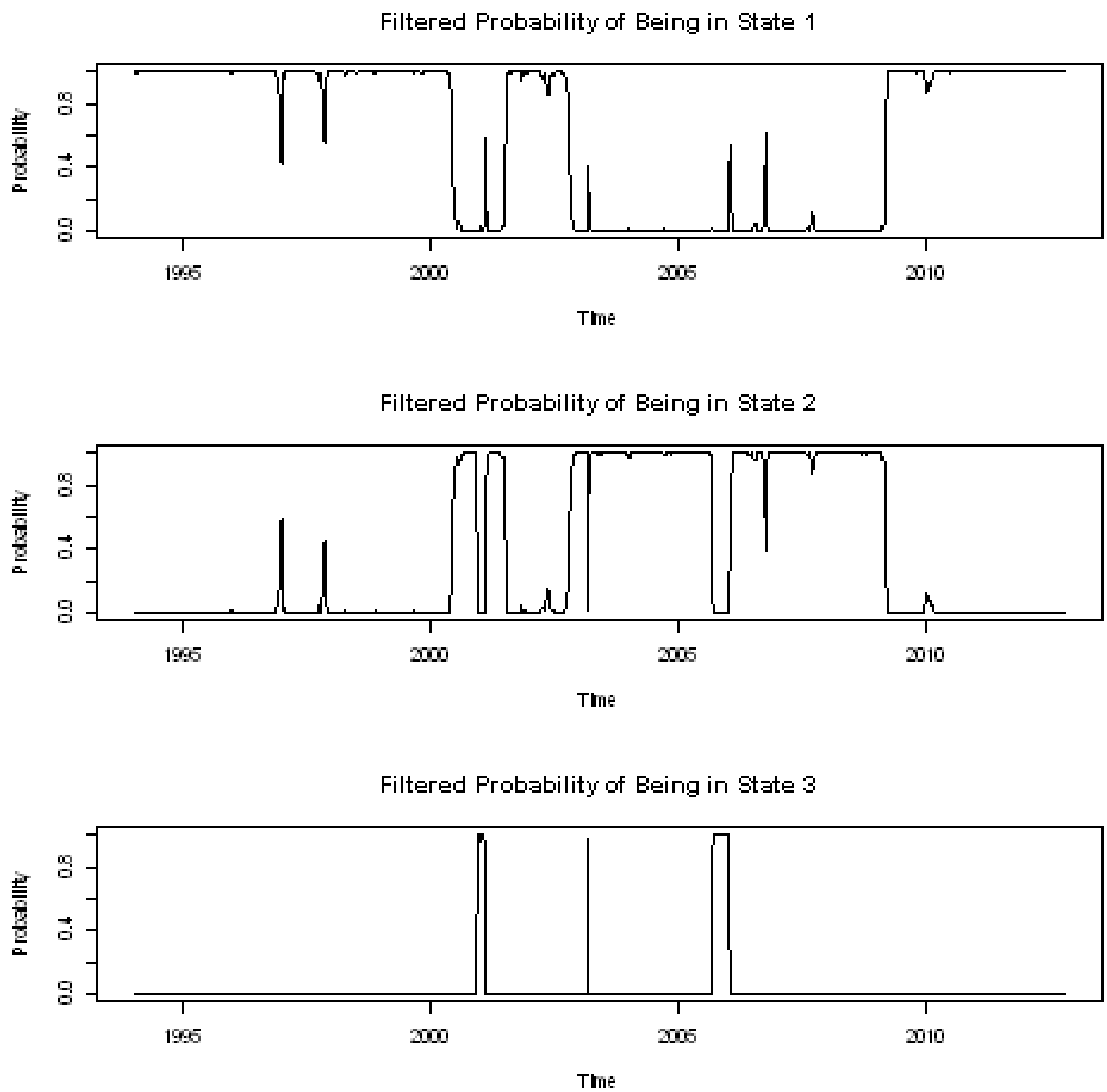
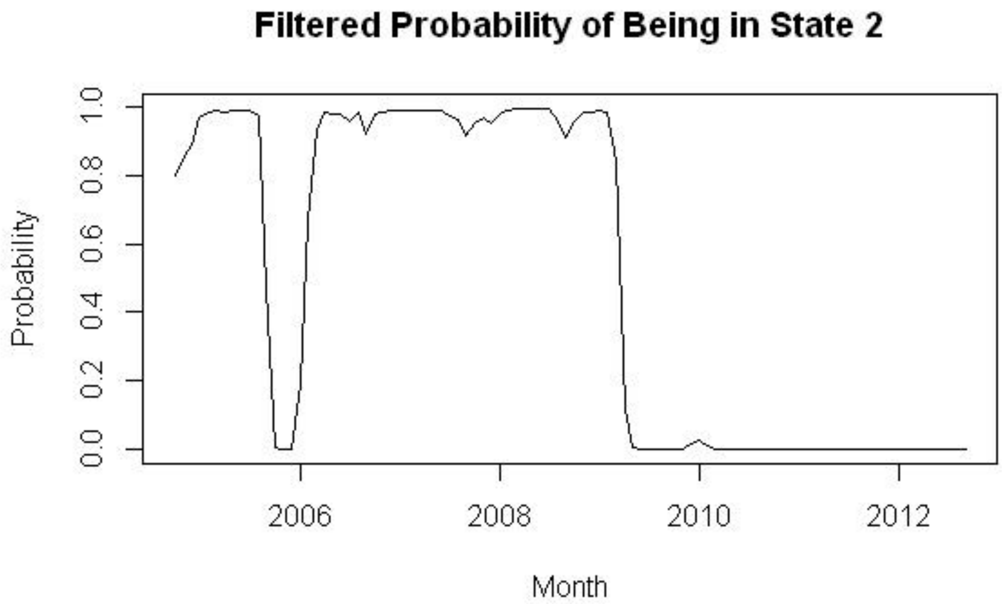


Figure 5. Filtered probability of state 2 in a 2-state cointegrating equation estimated on monthly data over the interval October 2004 to September 2012.



Appendix:

Error-Correction Model Results over the Full-sample: Switching vs. Non-switching Below are estimates of the error correction model on monthly data from June 1997 to September 2012. There are 184 observations. The response variable is the change in the log Henry Hub natural gas price from time *t-1* to time *t*. *, **, *** denote significance at the 10%, 5%, and 1% levels respectively.

Variable	Switching Residuals		Non-Switching Residuals	
Intercept	0.1117 (0.0028)***	0.0850 (0.0103)**	0.1212 (0.0014)***	0.0944 (0.0051)***
e_{t-1}	-0.1295 (0.0182)**	-0.1527 (0.0054)***	-0.0359 (0.1670)	-0.0664 (0.0120)**
ΔP_{WTIt}	0.3457 (0.0004)***	0.3603 (0.0003)***	0.3710 (0.0002)***	0.3994 (7.2e-5)***
ΔP_{HHt-1}	0.1135 (0.0820)*	0.1883 (0.0119)**	0.0752 (0.2954)	0.1208 (0.0798)*
ΔP_{HHt-2}	-0.0550 (0.4773)	-0.0007 (0.9925)	-0.1038 (0.1645)	-0.0582 (0.4100)
ΔP_{HHt-3}	-0.0832 (0.2845)	-0.0134 (0.8562)	-0.1252 (0.0995)*	-0.0611 (0.3959)
ΔP_{HHt-4}	0.0540 (0.4498)	0.0908 (0.2070)	0.0176 (0.8012)	0.0475 (0.5000)
ΔP_{HHt-5}	-0.1064 (0.1330)		-0.1318 (0.0618)*	
ΔP_{HHt-6}	0.0064 (0.9274)		-0.0174 (0.8054)	
ΔP_{HHt-7}	-0.0201 (0.7677)		-0.0299 (0.6646)	
ΔP_{HHt-8}	-0.1159 (0.0851)*		-0.1188 (0.0837)*	
ΔP_{HHt-9}	-0.2603 (0.0002)***		-0.2525 (0.0004)***	
ΔP_{HHt-10}	-0.0113 (0.8746)		0.0035 (0.9613)	
ΔP_{HHt-11}	-0.0578 (0.4070)		-0.0515 (0.4656)***	
ΔP_{HHt-12}	-0.1261 (0.0721)*		-0.1187 (0.0960)*	
CDD_t	-0.0006 (0.0001)***	-0.0005 (0.0002)***	-0.0006 (0.0002)***	-0.0005 (0.0004)***
$CDDDEV_t$	0.0025 (5e-7)***	0.0022 (2e-05)***	0.0024 (3.6e-6)***	0.0020 (0.0001)***
HDD_t	-0.0001 (0.0998)*	-6.9e-05 (0.1621)	-0.0001 (0.0780)*	-7.1e-5 (0.1528)
$HDDDEV_t$	0.0008 (3e-5)***	0.0008 (5e-06)***	0.0008 (4.7e-5)***	0.0008 (1.4e-5)***
STORAGE $DIFF_t$	-0.0001 (0.0233)**	-2.2e-05 (0.5597)	-0.0001 (0.004)***	-7.8e-5 (0.0722)*
RIG COUNT $DIFF_t$	0.0002 (0.0446)**	6.2e-05 (0.4234)	0.0001 (0.0765)*	4.3e-5 (0.5792)
F-statistic	6.316 (5e-12)***	6.622 (1e-09)***	5.981 (2.3e-11)***	6.450 (2.3e-9)***
adj. R^2	0.3848	0.2748	0.3695	0.2687