Stability in an Era of Instability: Regime Changes in the Market Relationship Between Liquefied Petroleum Gases, Oil and Natural Gas

Knut Anders Kjelas Johansen

June 29, 2017

Abstract

Petroleum products which serves as substitutes should be related by a long-term relative price equilibrium. This paper motivates the use of a Markov Switching methodology to model the relationship between four U.S. petroleum products between 2001 and 2016. This period is known for high volatility due to the financial crisis in 2008, the shale gas revolution and the plummeting of energy prices in 2014. These events could affect the equilibrium price relationships. Our findings indicate that by adjusting for regime shifts, a long-term equilibrium model is well fitted to four different bivariate relationships even under volatile market conditions. Using a vector error correction model, we show that adjustment back to this equilibrium is adjusted gradually but slowly.

Keywords— Markov switching, Cointegration, LPG

1 Introduction

This paper motivates the use of the Markov Switching (MS) model of Hamilton (1989, 1990) in order to explore the direct links between four different petroleum products from the U.S. market. These are Liquefied Petroleum Gases (LPGs) represented by propane and butane from Mont Belvieu Texas (MTBEL), West Texas Intermediate (WTI) crude and natural gas (NG) from Henry Hub. LPGs are byproducts from NG processing and crude refining. One of the major developments after the introduction of U.S. shale gas, is the increase in U.S. LPG production which has attracted little attention in academic circles. This increase in U.S. supply from both tight oil and shale gas formations, has changed the market trade dynamics and turned the U.S. from a net importer of LPG to a net exporter. Being able to model energy prices is therefore important to develop a clear understanding for both suppliers and consumers of the price movement. For consuming countries such as the U.S., approximate 7.45% of total GDP is used for energy in 2016[1]. Adjustments in the price of energy will therefore have a impact on the macroeconomic output. Literature suggests that there are linkage between the fluctuation of energy prices and GDP and stock market (Hamilton, 2003) (Kilian, 2006). EIA estimate that the world consumption of energy will go through a 48% increase between 2012 and 2040. For non-OECD[2] countries it is expected that the demand for energy rises by 71% in the same time period in contrast to OECD countries where energy demand is expected to grow by 18% (EIA, 2016). In 2015, 57% of all consumption derive from crude oil and NG (BP, 2016). For suppliers it is essential to be able to perform risk assessment before production.

[1] 4.7% on crude and 0.82% on natural gas. (Figure 47. Energy end-use expenditures as a share of gross domestic product, 1970-2040, downloaded January 23, 2017, from `www.eia.gov/forecasts/aeo/excel/fig47_data.xls`.)

[2] Organization for Economic Cooperation and Developent

By incorporating the Engle-Granger cointegration methodology ((Granger, 1981) (Engle and Granger, 1987)) in the MS framework, we believe we are able to more accurately model the long-run relationship independent of the price characteristics of energy price series. The MS regression model allows for different model characteristics dependent on which regime the price is currently in. A regime shift could be permanent (structural break) due to permanent shifts in the economy, or temporary due to wars or other interim events. If cointegration relationship is established in a bivariate setup, we continue by estimating a Vector Error Correction Model (VECM) which enables us to interpret the speed at which the dependent variable adjust back to equilibrium after a disruption. The novelty of this method is that it has the potential to distinguish between lack of market integration and market integration under changing market conditions, something a linear model is not able to do. Distinguishing between these two outcomes is important for companies exposed to the markets, or policy makers evaluating the markets.

Ever since the introduction of the cointegration methodology resources has been spend on analyzing the relationship between different energy prices in order to understand the complex interactions between energy markets. Especially the long-term relationship between crude oil and NG has been extensively analyzed (Asche et al., 2006) (Brown and Yücel, 2008) (Neumann, 2008) (Olsen et al., 2015). Similar to these studies Ramberg and Parsons (2010) also find statistical evidence for cointegration, however they highlight two important factors that classical statistical cointegration methods do not adjust for. First is the difference in volatility between crude and NG. According to their calculations, the annual volatility in log NG price between 1991 to 2010 is 69%, compared to 39% for the crude price. A long-term statistical relationship between the two prices will therefore leave a large portion of the NG price unexplained. Second, the detected long-term equilibrium of the two price series by previous scholars seems only stable for certain periods of time. This creates some time pockets of integration, but this depends on the uncertainty of start and end date for the applied dataset.

Less attention has been given to the relationship between LPGs and WTI and NG. Oglend et al. (2015) look at the relationship between LPGs and crude and NG in the U.S. market. Prior to January 2009, they find a cointegration relationship but this weakens after a structural break in 2008/9. This motivates the use of non-linear approaches that are more robust to structural changes in the markets analyzed.

The MS model is a particularly popular tool for applied work. Modeling of business cycles for open economies where the underlying growth depart from the normal upward trend is maybe the most traditional area of use (Kim and Nelson, 1999) (Kim et al., 2005). Other areas of use is modeling of volatility regimes (Hamilton and Lin, 1996) (Hamilton and Susmel, 1994) (Turner et al., 1989), financial crisis (Coe, 2002), and state dependent returns (Perez-Quiros and Timmermann, 2000). When applying the MS methodology on the energy market, most scholars analyze the shifting behavior in the returns of the commodities. Geng et al. (2016) focus on the impact of the U.S. shale revolution. By first modeling the regional gas markets in the U.S. and Europe, then the spreads between these and the regional crudes, they show that the European market is unaffected by the shale revolution while the U.S. prices has increased the decoupling behavior. As an argument they point at oversupply and a U.S. market not able to consume it all or export it due to exporting bottle necks. They first model the individual price movement regimes of WTI, Henry Hub and National Balancing Point (NBP) using the MS model. Second they use the spreads between the same data, and by that look at how the shale has affected the relationship. Galyfianakis et al. (2016) focus on the period between 2005 and 2015, applying a MS model to explain the behavior of five U.S petroleum products. They identify a crisis and a tranquil regime characterized by high and low returns. They also show that the tranquil regime is at least twice as long as the crisis regime, and that the switches is abrupt. These two distinct regimes is also confirmed by Chan et al. (2011).

There are also a significant segment applying MS models on electricity prices ((Bierbrauer et al., 2004), (Weron et al., 2004), (Mount et al., 2006), (Janczura and Weron, 2012)). Since there

are no economically ways of storing electricity and with a variable end-user demand curve due to for example seasonality, electricity data may have even more extreme return characteristics than petroleum prices.

Fewer studies apply the MS methodology when modeling the cointegration relationship between energy prices. Brigida (2014) use a similar method as this paper when modeling the long-term relationship between U.S. crude and NG. By controlling for the shifts in the cointegrated vector, he proves the long-term relationship between crude and NG. Asche et al. (2017) apply a MS-VECM to model European NBP when it is subjected to periodic decoupling from Brent crude. They find that between 1997 and 2014 the market is for the most part integrated, and by that confirm the law of one price (LOP) in the European petroleum market. They also show that without the regime switching model, this is not true. Related to these papers, we also find that extending the linear model to multiple states confirms the presence of market integration between LPGs and oil and natural gas in the US. The regime switching model highlights how the relative values of the energy products shifts over time, and that after controlling for these shifts in relative value the markets adjust to equilibrium.

The rest of this paper is structured as follows. Section 2 describe the background for the price characteristics found in the petroleum market in the analyzed time period. In section 3 we present the models applied in this paper. We first do a quick walk through of the two unit root tests Augmented Dickey Fuller (ADF) and Zivot-Andrews test. These are used in order to ensure that all price series are integrated at the same order, and by that is appropriate for cointegration analysis. Second we present the Engle-Granger cointegration methodology and lay out how this is incorporated in the Markov Switching model. Third we outline the Vector Error Correction model for examining market integration. Section 4 applies this methodology on the energy price data form the U.S. energy market. Section 5 concludes.

2 Background

Propane and butane is part of a broader category of petroleum products called Liquefied Petroleum Gases (LPG), not to be confused with Liquefied Natural Gas (LNG) or Compressed Natural Gas (CNG). Globally 60% of all LPG is by-products from crude oil and NG and 40% from refinery production. This range varies widely depending on the region, size of the resource and on the complexity of the refining system (Leija and Gist, 2013). Propane and butane are fairly similar products and traded LPG are often a mix between these two. Because of the similarities between propane and butane, there should be a long term substitution effect among them. This means that apart from short term fluctuations, the prices should always revert back to a long term balance. Before the decoupling in 2008, this was also the case for crude and NG (Brown and Yücel, 2008) (Asche et al., 2017). After the introduction of shale gas, NG has become an increased component in U.S. LPG production. The shale revolution has also made the U.S. into a powerful player and a net LPG exporter, rising from 53 thousand b/d in 2008, to almost 616 thousand b/d in 2015[3]. Earlier most of U.S. exports was sold to Latin America, but the growth of LPG demand among Asian countries increases. The opening of the new enlarged Panama Canal in 2016 reduced the voyage time of U.S. Gulf/Asia export from 45 to 25 days. This allows U.S. LPG to compete for Asian market share with the more traditional suppliers from the Middle East, generating a supply shock in the Asian market. Before 2016 the originally Panama Canal was only able to accommodate 73000 m^3 LPG carriers, but are now available for very large gas carriers (VLGCs) carriyng up to 83000 m^3.

Since the turn of the twenty-first century, energy markets has experienced dramatic changes. Before this the main perception was that the Organization of the Petroleum Exporting Countries (OPEC), with Saudi Arabia as the soul guardian could maintain price stability. By serving as a

[3]U.S. Exports of Propane and Propylene, downloaded January 25, 2017, from `http://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=MPREXUS2&f=M`

swing producer for the global petroleum market, possessing spare capacity, the organization was able to capture a market share above 40% in the late 1990s. This resulted in a stable evolution of the crude price, which from the 1970's to mid 2003 averaged around 20 to 30 $/bbl.

The petroleum market is influenced by both private and governmental interests. The supply side of petroleum is devided into international oil companies (IOCs) and national oil companies (NOCs). For the privately owned IOCs, the goal is to maximize share holder revenue, and by that producing at maximum capacity as long as the selling price is higher than the marginal relevant production price. From a daily global supply of 92 million barrels daily in 2015, 42% came from IOCs. After dominating the petroleum market after World War 2, the 1990s saw a rising influence of NOCs forcing many consolidations among the IOCs. NOCs, operating on behalf of a national government, may in addition to balancing profit oriented goals, also operate with other incentives. These may be energy security for domestic use and assist for the local labor market. This let NOCs hold on to spare capacity with Saudi Arabia serving as the biggest swing state producer.

After 2003 the energy commodity market has been affected by an increased global gross domestic product (GDP) growth averaging 5% a year between 2003 and 2007, with oil consumption growing with 6.5 million bbl daily in the same period. Especially emerging markets such as China powered the growth due to an industrial and urbanization process, increasing the countries total consumption of crude oil by almost 35% (BP, 2016). Together with a weak supply side and low spare production capacity, these events lead to strong turbulence in most energy commodities. For crude oil prices this meant a historical peak north of $140/bbl in July/August 2008 before a collapsing real estate bubble ripped the floor away sending the world into a massive recession and the crude price down to $30/bbl five months later.

In the post financial crisis era the crude prices managed again to stabilize around $100 and $125 until the first half of 2014. However in fear of loosing market share, the willingness of OPEC and especially Saudi Arabia to operate as a swing producer was no longer the case. This made it harder for the petroleum market to balance in case of over supply. In 2014 the supply side would be struck by large significant findings of what is known as unconventional fuels leading to what could be called the start of the U.S. shale revolution. Applying a Bai-Perron structural break test Wakamatsu and Aruga (2013) and Aruga (2016) actually trace this revolution all the way back to 2006. However, the major output from this supply side came first in 2014. With fracking, many small and independent U.S. producers were suddenly able to extract massive amounts of shale petroleum[4] from wells that was previous unavailable for extraction in Texas and North Dakota. Unlike the natural production of conventional wells which rely on the pressure in the reservoir, shale extraction could more or less be turn on and off from one weak to the other. This flexibility created a global supply shock and an entirely new global market equilibrium where many small U.S. producers together could act as swing producers based on the break even price of production. The reaction of Saudi Arabia (OPEC) was to depart from the previous swing producer policy in order to hold on to market share, and instead of swinging down it swinged up. This lead to a massive price drop only comparable with the situation in 2008.

Events such as described above demonstrate that the petroleum market is not controlled either for large price spikes (2008) or price busts (2014). Saudi Arabia is no longer interested in defending a price celling, and the floor is regulated on a day to day production break even prices in low cost extraction areas such as the U.S. shale. This leads to a more volatile unmanaged market driven price development for one of the worlds most strategic commodity (Mcnally, 2017). According to The U.S. Energy Information Administration (EIA) there are huge reserves of technical recoverable shale resources in the world. With an estimated volume of 18.8 trillion cubic meter (t/m^3) the U.S. is only third on the list after China (31.6 t/m^3), Argentina (22.7 t/m^3) and Algeria (20 t/m^3) (EIA, 2013)(Melikoglu, 2014). Due to the success of shale extraction in the U.S., other countries with the same geology preconditions also want to capitalize on this.

[4]After succesively extracting shale gas, the fracking method was also introduced to extraction of shale oil.

But unlike the U.S. with a high degree of privatization and community resistance, the situation in especially European countries are different due to political and environmental concerns (Stevens, 2012).

Market conditions for a commodity which together with classical supply and demand is so influenced by geopolitical tensions as described above, gives energy price series unusual characteristics. First of all this leads to high volatility. Regnier (2007) show that petroleum products are more volatile than 95% of other products sold domestically in the U.S. This is due to frequently large shocks geopolitical events has on the commodities, shown as erratic behavior on a frequent basis in the log-returns price series. This outlying observations tends also to occur in clusters leading to asymmetrical data (high excess kurtosis), where often followed by large negative returns showed by negative skewness. All together this leads to non-Gaussian behavior (Balke and Fomby, 1994) (Franses and Van Dijk, 2000). Another hallmark of petroleum prices is mean reversion toward marginal cost which should change over a longer time span. The intuition is that the petroleum products are sold in competitive markets with supply and demand, and price shocks should therefore only be temporary. Pindyck (1999) examine the long-run behavior of different U.S. non renewable energy commodities. With a dataset containing up to 127 years of data, he finds evidence of slow mean reversion to a stochastic trend line. Due to this characteristics, modeling energy prices is a challenging task. It may be hard to estimate a parsimonious linear model with constant coefficients, yet still much analysis use models which assume Gaussian innovations. The classical example is the standard Box-Jenkins methodology which focus entirely on the two first moments and because of this miss the important effect of large shocks to the price series. A solution is to model certain time periods of the price series where this characteristics are dampen. Examples are the steady growth period between 2003 and 2007, or the stabilized regime between 2011 and 2014. A second solution, which is used in this paper, is to consider non-linear models simply because linear models would not be able to cope with these characteristics.

These large changes in energy markets casts doubt about the ability of linear models to properly model the relationship between energy markets.

3 Research Methods and Models

In this section, we outline the Markov Switching approach used in this paper to investigate energy price relationships.

3.1 Unit root

As is conventional, the first part of the analysis consists of establishing the order of integration of the data. The applied cointegration methodology of Engle and Granger demand stationary data, or else the we will end up with high R^2 and highly significant t-statistics also known as spurious regression results (Granger and Newbold, 1974). Testing for unit root we apply the standard Augmented Dickey-Fuller (ADF) test denoted:

$$\Delta y_t = a_0 + \gamma y_{t-1} + a_2 t + \sum_{i=2}^{p} \beta_i \Delta y_{t-i+1} + \varepsilon_t$$

where a_0 is a drift term, $\gamma = a_1 - 1$ and a_2 adds a linear trend (Dickey and Fuller, 1979) (Said and Dickey, 1984). $\sum_{i=2}^{p} \beta_i \Delta y_{t-i+1}$ is added to capture autoregressive components that if else would have ended up in the residuals. Akaike information criteria (AIC) is applied to choose lag length on the differenced terms (Akaike, 1981). The null hypothesis is then to test if $\gamma = 0$ with the alternative that it is less than zero.

Since this paper is build around the possibility of regime shifts in the data, we also apply the Zivot-Andrews (ZA) test for structural breaks (Zivot and Andrews, 1992). This because

the standard ADF test is biased towards the non-rejection of a unit root in the pressence of structural changes. Perron (1989) modified the standard ADF test by adding a dummy variable making the break date fixed. However, knowing this date a priori is a very difficult task.

The ZA test is a continuation of the work of Perron by letting the shift be decided endogenously instead of exogenously, meaning that the date of the shift is estimated rather than fixed. This is done by finding the point where the evidence of unit root is least favorable, meaning where the t-statistic from the ADF-test is at a minimum (Byrne et al., 2006). The ZA test is therefore a standard ADF test with a dummy variable to account for structural changes in the regression model. The order of the test is similar to the standard ADF-test decided by AIC. The least restricted ZA-test allow for the possibility of change in both the intercept and trend and is denoted:

$$y_t = \mu + \beta t + \alpha y_{t-1} + \theta \mathrm{DU}_t + \gamma \mathrm{DT}_t + \sum_{i=1}^{k} \delta_i \Delta y_{t-i} + \varepsilon_t$$

$$\mathrm{DU}_t = \begin{cases} 1 \text{ if } t > \lambda T \\ 0 \text{ otherwise} \end{cases} \qquad \mathrm{DT}_t = \begin{cases} t - \lambda T \text{ if } t > \lambda T \\ 0 \text{ otherwise} \end{cases}$$

where DU_t and DT_t are dummy variables for shifts in the mean and trend at time T_B. The α parameter is the estimated break date equal to T_B/T. This is the least restricted model and by equaling θ or γ equal to 0, one allows for break only in the trend or intercept. The null hypothesis is a unit root process with no structural breaks versus the alternative of a trend stationary process with a break.

3.2 Markov switching model

The primary hypothesis behind this paper is that the price of the derived biproduct from one energy carrier and the initial energy carrier should have a long term equilibrium. Empirically, this means that non-stationary prices are cointegrated. The concept of cointegration was introduced in 1981 by Granger (1981) and Engle and Granger (1987), henceforth referred to as EG. Here the formal definition of cointegration is presented as: The components of the vector x_t are said to be cointegrated of order d, b, denoted $x_t \sim \mathrm{CI}(d, b)$, if (a) all components of x_t are $I(d)$ and (2) there exists a vector $\alpha(\neq 0)$ so that $y_t = \alpha' x_t \sim I(d-b)$, $b > 0$. The vector α is called the cointegrating vector. Following this definition the bivariate relationship between two vectors y_t and x_t is cointegrated if ε_t from $y_t = \alpha x_t + \varepsilon_t$ is stationary. If this holds, they would not drift to much apart from each other but revert back to a long term equilibrium. The standard EG model is a pure long-run linear equilibrium model and by that does not allow for non-normality in the included data. Based on the previous mentioned data characteristics we therefore introduce a MS framework to allow for structural change in the dependent data (Brigida, 2014).

The MS framework was introduced by Hamilton (1989, 1990) as an extension of the previous work of Quandt (1958), Goldfeld and Quandt (1973), and Cosslett and Lee (1985) to an autoregression. By this approach, the world is divided into m states in which the endogenous variable is allowed to switch between states according to a state indicator variable, S_t, for $i = 1, ..., m$. The state is an unobservable variable, and we therefore need to make inference about past regimes. The regime at any point in time is unobserved, but the process that establish a new regime is assumed known. As indicated by the name, the state transition in MS models assume a Markov property for S_t:

$$P(S_t = i | S_{t-1} = j, S_{t-2} = q, ...) = P(S_t = i | S_{t-1} = j) = p_{ij}$$

where p_{ij} is the transition probabilities and $\sum_N^{i=1} p_{ij} = 1$ conditional on a value of j. With constant transition probabilities the regimes evolve exogenously over time.

Incorporating the MS model in the EG method we estimate a long-term equilibrium model:

$$y_t = \alpha_{0,S_t} + \phi_{1,S_t} x_t + \varepsilon_t$$

where y_t is the endogenous variable and x_t is assumed to be a exogenous explanatory variable of log energy price series, and $\varepsilon_t \sim N(0, \sigma^2_{S_t})$. In our analysis we incorporate the LPGs as endogenous variables, and WTI or NG as exogenous variables. We assume a two-regime model, although more regimes are possible. This means that $N = 2$ so that $S_t = 1$ or $S_t = 2$. The transition probabilities is constrained to lie in [0,1] and form the transition matrix:

$$\mathbf{P} = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} = \begin{pmatrix} p_{11} & 1 - p_{11} \\ 1 - p_{22} & p_{22} \end{pmatrix}$$

This translates to:

$$P[S_t = 1 | S_{t-1} = 1] = p_{11}$$

$$P[S_t = 2 | S_{t-1} = 1] = 1 - p_{11}$$

$$P[S_t = 2 | S_{t-1} = 2] = p_{22}$$

$$P[S_t = 1 | S_{t-1} = 2] = 1 - p_{22}$$

The parameters of the model is collected in the vector $\theta = (\alpha_1, \phi_{1,1}, \sigma_1, \alpha_2, \phi_{1,2}, \sigma_2, p_{11}, p_{22})$. The estimation of the parameters is carried out with Maximum-likelihood estimation (MLE) presented by Hamilton (1989), but a Bayes (Gelfand and Smith, 1990) and Semi-parametric method (Campbell, 2002) is also possible. Hamilton (1989) developed a iterative algorithm constructing the likelihood function:

$$L(\theta) = \sum_{T}^{t=1} \log f(y_t | \Omega_{t-1}; \theta)$$

where Ω_{t-1} is observations available up to time $t-1$. The conditional density function is then computed by the Hamilton Filter. From this we are able to obtain the filtered and smoothed probabilities (Piger, 2009).

If y_t and x_t are $I(1)$ and the stated-weighted/conditional residuals from the MS model is stationary we conclude that y_t and x_t is cointegrated, CI$(1, 1)$. The conditional residuals from the linear regression in levels can then be defined as the equilibrium errors between y_t and x_t. We use the conditional residuals to implement a vector error correction model (VECM), which is an appropriate model when the endogenous variable is expected to exhibit short run deviations in response to changes in the exogenous variable (Durr, 1993). The VECM display changes in the dependent variable as a function of lagged changes in the dependent and independent variable togheter with an error correction term (ECT), which in our model is the equilibrium errors from the linear regression. The VECM is denoted:

$$\Delta y_t = \alpha_1 + \alpha_y \hat{\varepsilon}_{t-1} + \sum_{i=1} \alpha_{11}(i) \Delta y_{t-i} + \sum_{i=1} \alpha_{12}(i) \Delta x_{t-i} + \epsilon_{yt}$$

where α_y is the ECT which indicate the speed of adjustment from the long-run cointegration relationship. Following Asche et al. (2017), x_t is a exogenous variable which is either lagged values of crude or NG. Because of the previous stated cointegration relationship, this term should be negative, since a potential deviation from the long-run equilibrium should correct gradually back. In our model the α_y is the weekly adjustment after a shock to the system has driven the LPG price out of the equilibrium with either the crude or NG price.

4 Empirical Anlysis and discussion

4.1 Time series properties

Our dataset consist of four prices series from the U.S. energy market running from June 2001 to April 2016. These are propane and butane from MTBEL, WTI crude and NG from Henry Hub. Originally the dataset was sampled at daily frequency, however we have converted into weekly basis with the same energy equivalent ($/MMbtu). This correspond to a total of 776 observations. Figure 2 and 3 show all four price series in log-levels.

The first moments in log-returns are displayed in the first part of Table 1 for all four price series. With a mean close to zero and a approximate standard deviation of 0.03, all series behave similar. As expected all series, especially propane, show high excess kurtosis. Only WTI resembles anything close to a normal distrubution (Ex.Kurt=3). This indicate that all series are leptokurtic in that large observations occur often. Aside from NG all series also have negative skewness implying that the left tail of the distribution is fatter than the right tail. As mentioned previously, in financial data large negative returns tend to occur more often than large positive ones. Together, this leads to high Jarque-Bera tests which rejects the null hypothesis of normal distribution.

Table 1: Descriptive statistics and unit root tests

					ADF		ZA	
	Mean[(a)]	Std.	Ex.Kurt	Skew.	Levels (lag)	First diff.	Teststat. (lag)	Breakmonth
Propane	-0.048	0.027	17.26	-1.90	-0.18 (12)	-7.57***	-3.51 (1)	02.2014
Butane	-0.001	0.026	4.46	-0.71	-0.18 (8)	-8.65***	-3.02 (1)	01.2014
WTI	0.166	0.025	3.43	-0.47	-0.03 (8)	-7.76***	-3.57 (1)	07.2014
NG	-0.173	0.031	8.47	0.18	-0.19 (11)	-8.81***	-3.82 (5)	07.2008

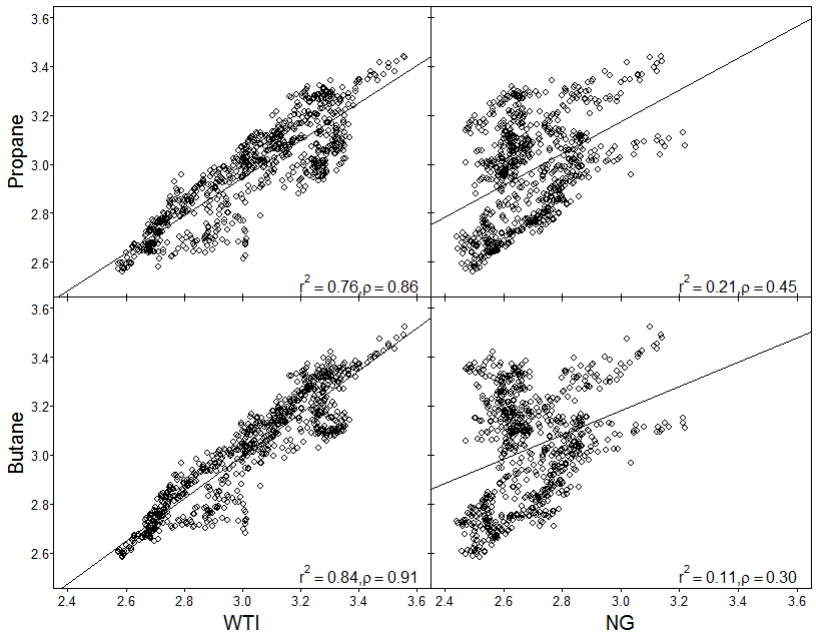
(a): Multiplied with 1000. ***, **, *, indicates statistical significanse at 1%, 5% and 10% level.

In order to ensure that both series are integrated at the same order, two tests for unit root are performed on the level and first difference on all four series. With lag order chosen on the basis of AIC performance, the ADF test of the series in levels including a constant show that we cannot reject the null of a unit root. This is confirmed with the first difference series indicating stationarity. Since the time period analyzed include some major events we also include the ZA test. By estimating the ZA only allowing for a break in the intercept we find support for the the ADF test by failing to reject the null hypothesis of unit root for all series.

4.2 Markov switching model

Figure 1 displays the scatter plot between the four different price series where each point representing an ordered par between the price series. Propane and butane series is displayed on the vertical axis and WTI and Henry Hub NG on the horizontal axis. Based on the previous discussion about the LPG market, we treat the propane series as endogenous variables and WTI and NG as exogenous variables. The plot clearly indicate a positive relationship between the two LPG prices and WTI. This means that low values of the LPGs are associated with low values in WTI and so on. The plot also include the least-squares line which confirms this strong positive relation. Based on visual inspection the LPGs and WTI should therefore be cointegrated, and the least-square line should in fact be the long-run relationship between the price series. This is not the case for the relationship between the LPGs and NG. There is still a positive gradient, but the ordered pair are much more random indicating weak/no long-run relationship between the variables. This is also confirmed by a low coefficient of determination for both propane and butane.

Figure 1: Scatter plot of log petroleum prices



However visual inspection is not enough and we start to implement the EG theory by estimating the long run equilibrium relationship and testing for pairwise cointegration. The results of all four regressions are displayed in Table 2. Except for the regression between butane and WTI with 10% significance level, we are not able to reject for unit root in any of the residuals. We are therefore not able to prove a $CI(1,1)$ relationship between the LPGs and either oil or natural gas.

To motivate the choice of methodology in this paper, we also conduct a recursive version of the F-test proposed by Chow (1960) on the long-run equilibrium models. The Chow test is testing for non linear behavior in the dependent variable in a linear regression at time t^*. This is implemented by first estimating two equations:

$$y_1 = \alpha_1 x_1 + \varepsilon_1 \qquad \text{for } t = 1, ..., t^*$$

$$y_2 = \alpha_2 x_2 + \varepsilon_2 \qquad \text{for } t = t^* + 1, ..., T$$

from which we get the unrestricted residual sum of squares (RSS_U) by adding the residual sum of squares from both models. The restricted residual sum of squares (RSS_R) is obtained by estimeating one simple regression over all the data, and by that assuming equal parameters for the whole sample. The Chow test for a structural break is then denoted:

$$\frac{k^{-1}(RSS_R - RSS_U)}{(T-k)^{-1}RSS_U} \sim F_{k,T-k}$$

where k is the number of regressors. With a 5% critical value of 3.01, we are not able to reject the null hypothesis of no structural break for any of the price series. It is therefore safe to say that the relationship between the price series is not stable over time and thereby the application of non-linear models are valid.

Table 2: EG and MS cointegration models

	Propane - WTI			**Butane - WTI**		
	EG Model	MS Model		EG Model	MS Model	
α_0	0.64***	0.54***	0.87***	0.38***	0.66***	0.36***
α_1	0.77***	0.82***	0.66***	0.87***	0.74***	0.89***
σ	0.10	0.05	0.06	0.09	0.06	0.04
$\overline{Chow}$	384.79			303.17		
AIC	-1371.34	-2360.86		-1580.82	-2474.51	
$P(s_t = 1 \mid s_{t-1} = 1)$		0.9913			0.9853	
$E[S_t = 1]$		115.2			68.0	
$P(s_t = 2 \mid s_{t-1} = 2)$		0.9867			0.9919	
$E[S_t = 2]$		75.4			123.8	
ADF (lag)	-2.69 (8)	-7.13*** (4)		-3.11*(5)	-6.12***(5)	
	Propane - NG			**Butane - NG**		
	EG Model	MS Model		EG Model	MS Model	
α_0	1.22***	1.73***	0.68***	1.69***	0.79***	2.58***
α_1	0.65***	0.52***	0.79***	0.50***	0.76***	0.23***
σ	0.18	0.10	0.06	0.21	0.08	0.11
$\overline{Chow}$	401.02			432.73		
AIC	-452.17	-2360.86		-234.54	-1467.89	
$P(s_t = 1 \mid s_{t-1} = 1)$		0.9924			0.9922	
$E[S_t = 1]$		132.2			127.9	
$P(s_t = 2 \mid s_{t-1} = 2)$		0.9921			0.9923	
$E[S_t = 2]$		126.0			130.4	
ADF (lag)	-1.96 (8)	-4.39*** (12)		-1.87(8)	-4.86***(1)	

***, **, *, indicates statistical significanse at 1%, 5% and 10% level. Critical values for the ADF tests are from MacKinnon (1991)

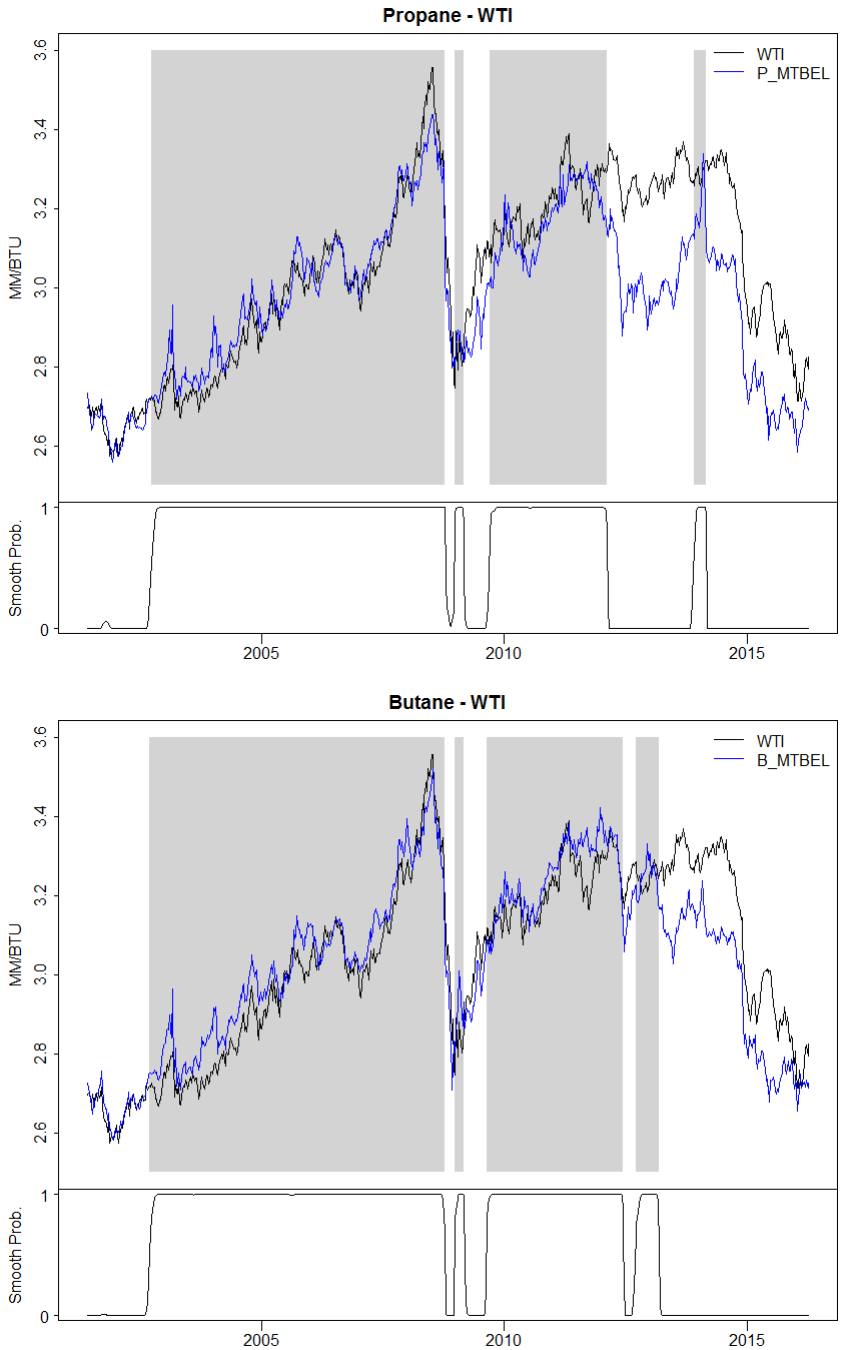
Given the results from the simple long-run equilibrium one state model we now turn to the MS model of the same relationships (Table 2). The regression coefficients of all four regressions are statistically significant. For all four regressions the AIC improves compared to the one state EG model indicating a better fit when allowing for regime shifts. The residuals from all four models are also stationary, indicating that the series are cointegrated of order (1, 1) once the shifts in the cointegrated vector are controlled for.

The two α_1 parameters for each MS model indicate the price elasticity between the dependent variables (LPGs) and the independent variables (crude and NG). Fore one dollar increase in WTI, propane and butane will, averaging over both states, have an increase of 0.74% and 0.81%. For one dollar increase in NG, the propane and butane price will increase on average by 0.65% and 0.49%.

The probabilities in the transition matrix indicate that the holding (sojurn) time after entering a regime is large. This is shown by high probabilities in the diagonal elements with all four models having probabilities above 0.98, and are therefore nearly absorbing/recurrent states. This also indicate that high (low) regimes tends to be followed by high (low) regimes. The switching process between states is following Bernoulli trials with a geometric distribution. Expected mean holding time after entering a regime is calculated by $S_i = \frac{1}{1-p_{ii}}$. These are included in Table 2 and show that the average time spent in a regime for all models are 112 weeks.

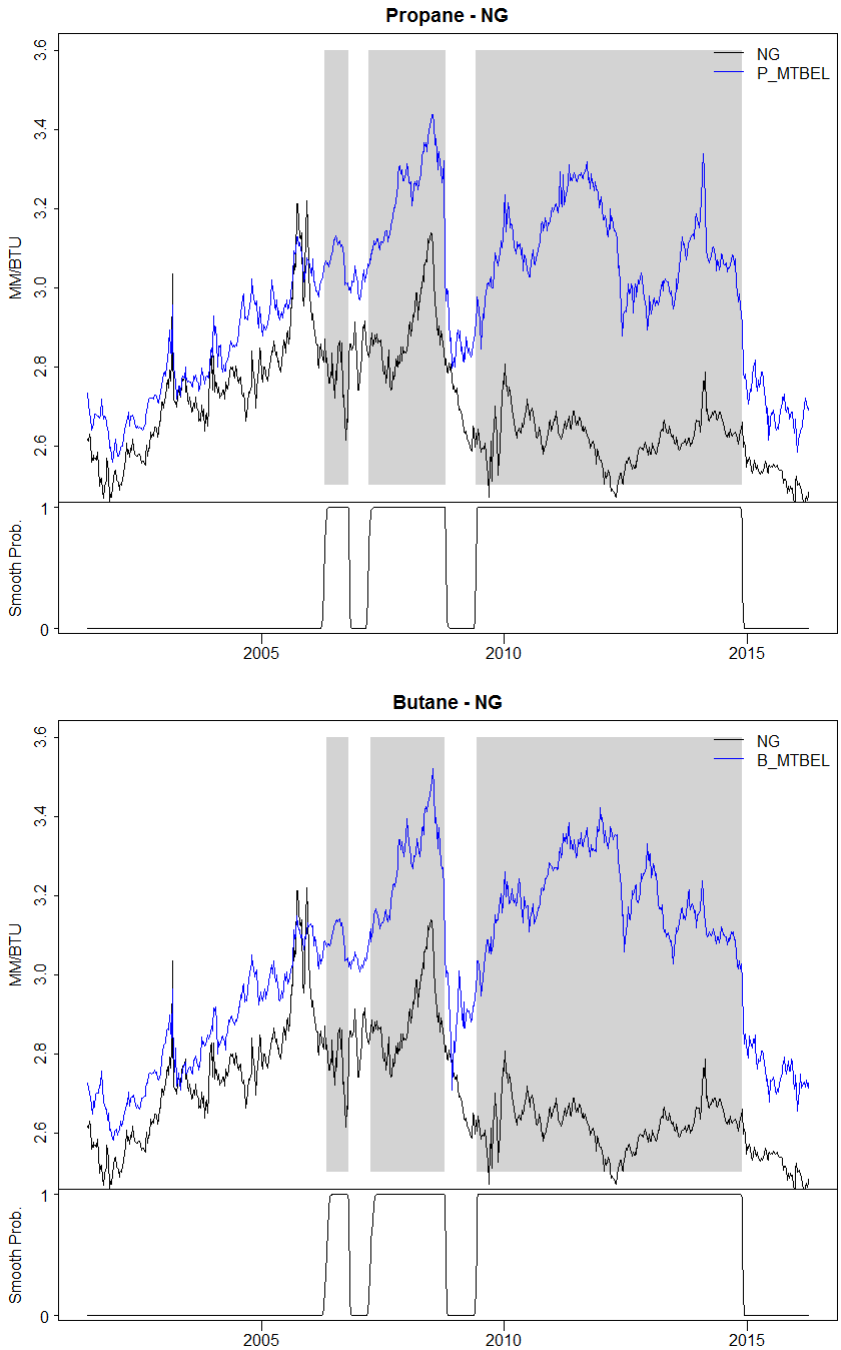
Figure 2 and 3 show the log-level price series with the detected regimes indicated by the

Figure 2: Regimes and Smooth Probabilities LPG - WTI



shaded areas (p>0.5) in the first panel. The break dates are fairly similar conditioned on the same exogenous variable. For the two LPG - WTI models we find four distinct regimes between 2002 - 2008, 2008 - 2009, 2009 - 2012 and 2013. It is worth noticing that the butane series always switch before the propane series with a time period from a couple of weeks up to a month. For the LPG - NG models we find three regimes between 2006, 2007 - 2008 and 2009 - 2014. Based on visual inspection it may seem that periods with high price evolution similarity is subjected to one state, and periods where prices decouple is another state. In the lower panel of Figure 2 and 3 we have included the smoothed probabilities of the regimes which is the probability that the unobserved MC is in a particular regime, conditioned on all previous information. This allows us to see when the different regimes starts to emerge. The smoothed probabilities are

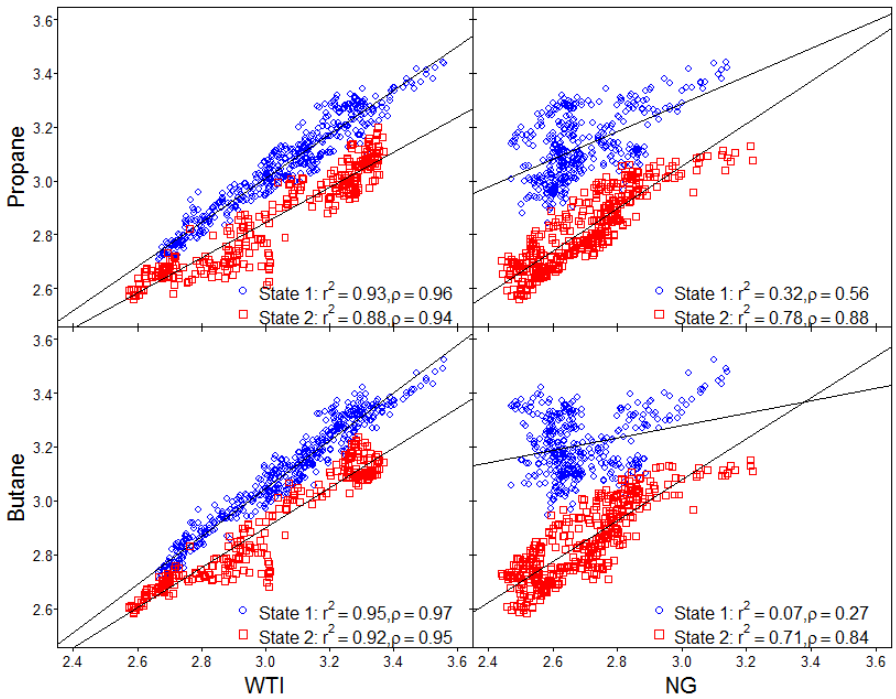
Figure 3: Regimes and Smooth Probabilities LPG - NG



characterized with abrupt regime switches, confirming the findings of Galyfianakis et al. (2016).

Finally we include a new scatter-plot of the log-price in Figure 4 displaying the relationships between LPGs on the y-axis and crude and NG on the explanatory x-axis. Unlike the first scatter-plot in Figure 1, we have now divided the points into which state the MS model has put them in with all smooth probabilities above 0.5 serving as state one. We also include two least-squares lines, one for each state. After adjusting for the regime shifts, all four plots still have a positive gradient. The linear association is also evident except for state 1 in the two LPG - NG relationships. This is obviously a extreme state since both the R^2 and correlation is quite low compared with state one in the propane - WTI relationships. However compared to Figure 1, the least-square lines in all four relationships seems to be much better fitted when

Figure 4: Scatter plot of log petroleum prices with MS regimes



adjusting the long-term equilibrium model to regime shifts. Low R^2 in state 1 in the LPG - NG relationships is consistent with the findings of Oglend et al. (2015) which detects lack of cointegration in the same time period.

4.3 Vector Error Correction Model

While the cointegration analysis gives the long-term relationship, the VECM measure the short run dynamics. Even though we did not find a cointegration relationship in the linear EG model, we start by estimating a VECM of the linear relationship for comparison reasons. The results are displayed in Table 3 and confirm that none of the ECT in the four regressions is significantly different from zero.

Table 3: Linear-VECM models

	Propane - WTI	**Butane - WTI**	**Propane - NG**	**Butane - NG**
α_1	0.00	0.00	0.00	0.00
$\hat{\varepsilon}_{t-1}$	-0.01	-0.01	-0.01	-0.01
Δy_{t-1}	-0.05	-0.08*	-0.00	-0.02
Δx_{t-1}	0.13***	0.13***	0.02	0.02
F-stat.	3.31**	2.68**	1.26	1.03
AIC	-3383.75	-3447.73	-3377.64	-3442.79
LogLik.	1696.88	1728.87	1693.82	1726.40
LB(20)	35.69 (0.69)	24.09 (0.24)	37.53 (0.01)	24.65 (0.22)

***, **, *, indicates statistical significanse at 1%, 5% and 10% level.

Having accounted for the regime shifts in the long-term equilibrium model with the MS-model, we now model the dynamic relationship with applying a VECM on the conditional

residuals from the MS model over the whole sample period. The results are displayed in Table 4. In order to achieve a parsimonious model with respect to autocorrelation, it is necessary to apply different lag length specifications for each model. The ECT is significant and negative for all four models. This indicates that all four LPG series adjust to short term fluctuations in order to restore the long term equilibrium. This confirm the law of one price (LOP) between WTI and NG and the LPG series. However, the ECT is quite low for all four models. With an average ECT of 5% for all four models, the adjustment period from a shock occurs and back to the long-term equilibrium is quite long. We are therefore able to say that there are a significant amount of market friction between WTI and NG and the LPGs. By adjusting 5% every week it will take 44 weeks for a 90% adjustment. The adjustment process after a shock is adjusted exponentially at the rate of the ECT. The half life of this adjustment period is calculated by $ln(0.5)/ECT$ (Bachmeier and Griffin, 2006). The shortest half-life period is between propane and WTI which needs 8.87 weeks to adjust a shock by 50%. With 25.62 weeks, the longest half-life period is between butane and NG.

Table 4: MS-VECM models

	Propane - WTI	**Butane - WTI**	**Propane - NG**	**Butane - NG**
α_1	0.00	0.00	0.00	0.00
$\hat{\varepsilon}_{t-1}$	-0.08***	-0.05**	-0.04***	-0.03***
Δy_{t-1}	-0.07*	-0.10 **	-0.01	-0.03
Δy_{t-2}	-0.02		-0.03	0.01
Δy_{t-3}	0.08**		0.07**	0.08**
Δx_{t-1}	0.14***	0.15***	0.02	0.02
Δx_{t-2}			0.01	
Δx_{t-3}			0.02	
F-stat.	5.58***	4.46***	2.29**	2.19**
AIC	-3387.55	-3453.01	-3371.95	-3436.29
LogLik.	1700.77	1731.51	1694.98	1725.14
LB(20)	31.68 (0.05)	25.83 (0.17)	31.70 (0.05)	20.91 (0.40)
Half-life	8.87	14.24	18.95	25.62

***, **, *, indicates statistical significanse at 1%, 5% and 10% level.

5 Concluding Remarks

Energy commodity prices have been the subject of research for many years. This paper make an effort to incorporate the cointegration theory of Engle and Granger ((Granger, 1981) (Engle and Granger, 1987)) into the Markov Switching methodology of Hamilton (1989, 1990). By doing this we have shown that it is possible to estimate bivariate market integrations despite changing market conditions, which is impossible with a standard linear model.

Our results indicate that between 2001 and 2016 there is a long-term relationship between all four prices. A price shock to the market is therefore gradually adjusted back to equilibrium confirming the law of one price. However, this adjustment is quite low indicating a long adjustment period for all price series.

References

Akaike, H. (1981). Likelihood of a model and information criteria. *Journal of econometrics*, 16(1):3–14.

Aruga, K. (2016). The us shale gas revolution and its effect on international gas markets. *Journal of Unconventional Oil and Gas Resources*, 14:1–5.

Asche, F., Oglend, A., and Osmundsen, P. (2017). Modeling uk natural gas prices when gas prices periodically decouple from the oil price. *Forthcoming Energy Journal*.

Asche, F., Osmundsen, P., and Sandsmark, M. (2006). The uk market for natural gas, oil and electricity: are the prices decoupled? *The Energy Journal*, pages 27–40.

Bachmeier, L. J. and Griffin, J. M. (2006). Testing for market integration crude oil, coal, and natural gas. *The Energy Journal*, pages 55–71.

Balke, N. S. and Fomby, T. B. (1994). Large shocks, small shocks, and economic fluctuations: Outliers in macroeconomic time series. *Journal of Applied Econometrics*, 9(2):181–200.

Bierbrauer, M., Trück, S., and Weron, R. (2004). Modeling electricity prices with regime switching models. In *International Conference on Computational Science*, pages 859–867. Springer.

BP (2016). Bp statistical review of world energy.

Brigida, M. (2014). The switching relationship between natural gas and crude oil prices. *Energy Economics*, 43:48–55.

Brown, S. P. and Yücel, M. K. (2008). What drives natural gas prices? *The Energy Journal*, pages 45–60.

Byrne, J. P., Perman, R., et al. (2006). Unit roots and structural breaks: a survey of the literature. *Paper provided by Business School-Economics, University of Glasgow in its series Working Papers with*, (2006_10).

Campbell, S. D. (2002). Specification testing and semiparametric estimation of regime switching models: An examination of the us short term interest rate. Technical report, Working Paper, Brown University, Department of Economics.

Chan, K. F., Treepongkaruna, S., Brooks, R., and Gray, S. (2011). Asset market linkages: Evidence from financial, commodity and real estate assets. *Journal of Banking & Finance*, 35(6):1415–1426.

Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica: Journal of the Econometric Society*, pages 591–605.

Coe, P. J. (2002). Financial crisis and the great depression: a regime-switching approach. *Journal of Money, Credit, and Banking*, 34(1):76–93.

Cosslett, S. R. and Lee, L.-F. (1985). Serial correlation in latent discrete variable models. *Journal of Econometrics*, 27(1):79–97.

Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, 74(366a):427–431.

Durr, R. H. (1993). What moves policy sentiment? *American Political Science Review*, pages 158–170.

EIA (2013). Thechnically recoverable shale oil and shale gas resources: An assesment of 137 shale formations in 41 countries outside the United States. `http://www.eia.gov/analysis/studies/worldshalegas/pdf/overview.pdf`. Accessed: 2017-01-23.

EIA (2016). International energy outlook 2016. *US Energy Information Administration, Washington, DC.*

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

Franses, P. H. and Van Dijk, D. (2000). *Non-linear time series models in empirical finance.* Cambridge University Press.

Galyfianakis, G., Drimbetas, E., Sariannidis, N., et al. (2016). Modeling energy prices with a markov-switching dynamic regression model: 2005-2015. *Bulletin of Applied Economics*, 3(1):11–28.

Gelfand, A. E. and Smith, A. F. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American statistical association*, 85(410):398–409.

Geng, J.-B., Ji, Q., and Fan, Y. (2016). The impact of the north american shale gas revolution on regional natural gas markets: Evidence from the regime-switching model. *Energy Policy*, 96:167–178.

Goldfeld, S. M. and Quandt, R. E. (1973). A markov model for switching regressions. *Journal of econometrics*, 1(1):3–15.

Granger, C. W. (1981). Some properties of time series data and their use in econometric model specification. *Journal of econometrics*, 16(1):121–130.

Granger, C. W. and Newbold, P. (1974). Spurious regressions in econometrics. *Journal of econometrics*, 2(2):111–120.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica: Journal of the Econometric Society*, pages 357–384.

Hamilton, J. D. (1990). Analysis of time series subject to changes in regime. *Journal of econometrics*, 45(1):39–70.

Hamilton, J. D. (2003). What is an oil shock? *Journal of econometrics*, 113(2):363–398.

Hamilton, J. D. and Lin, G. (1996). Stock market volatility and the business cycle. *Journal of Applied Econometrics*, 11(5):573–593.

Hamilton, J. D. and Susmel, R. (1994). Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics*, 64(1):307–333.

Janczura, J. and Weron, R. (2012). Efficient estimation of markov regime-switching models: An application to electricity spot prices. *AStA Advances in Statistical Analysis*, 96(3):385–407.

Kilian, L. (2006). Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market.

Kim, C.-J., Morley, J., and Piger, J. (2005). Nonlinearity and the permanent effects of recessions. *Journal of Applied Econometrics*, 20(2):291–309.

Kim, C.-J. and Nelson, C. R. (1999). Has the us economy become more stable? a bayesian approach based on a markov-switching model of the business cycle. *Review of Economics and Statistics*, 81(4):608–616.

Leija, F. and Gist, R. L. (2013). Shale gas development altering lpg demand, trade. *Oil & gas journal*, 111(6):92–97.

MacKinnon, J. G. et al. (1991). *Critical values for cointegration tests.* Department of Economics, University of California.

Mcnally, R. (2017). *Crude Volatility.* Columbia University Press.

Melikoglu, M. (2014). Shale gas: analysis of its role in the global energy market. *Renewable and Sustainable Energy Reviews*, 37:460–468.

Mount, T. D., Ning, Y., and Cai, X. (2006). Predicting price spikes in electricity markets using a regime-switching model with time-varying parameters. *Energy Economics*, 28(1):62–80.

Neumann, A. (2008). Linking natural gas markets-is lng doing its job?

Oglend, A., Lindbäck, M. E., and Osmundsen, P. (2015). Shale gas boom affecting the relationship between lpg and oil prices. *Energy Journal*, 36(4):265–286.

Olsen, K. K., Mjelde, J. W., and Bessler, D. A. (2015). Price formulation and the law of one price in internationally linked markets: an examination of the natural gas markets in the usa and canada. *The Annals of Regional Science*, 54(1):117–142.

Perez-Quiros, G. and Timmermann, A. (2000). Firm size and cyclical variations in stock returns. *The Journal of Finance*, 55(3):1229–1262.

Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis. *Econometrica: Journal of the Econometric Society*, pages 1361–1401.

Piger, J. (2009). Econometrics: models of regime changes. In *Complex Systems in Finance and Econometrics*, pages 190–202. Springer.

Pindyck, R. S. (1999). The long-run evolution of energy prices. *The Energy Journal*, pages 1–27.

Quandt, R. E. (1958). The estimation of the parameters of a linear regression system obeying two separate regimes. *Journal of the american statistical association*, 53(284):873–880.

Ramberg, D. J. and Parsons, J. E. (2010). The weak tie between natural gas and oil prices. *Center for Energy and Environmental Policy Research (CEEPR) No*, pages 10–017.

Regnier, E. (2007). Oil and energy price volatility. *Energy Economics*, 29(3):405–427.

Said, S. E. and Dickey, D. A. (1984). Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, 71(3):599–607.

Stevens, P. (2012). The shale gas revolution: Developments and changes. *Chatham House*, pages 2–3.

Turner, C. M., Startz, R., and Nelson, C. R. (1989). A markov model of heteroskedasticity, risk, and learning in the stock market. *Journal of Financial Economics*, 25(1):3–22.

Wakamatsu, H. and Aruga, K. (2013). The impact of the shale gas revolution on the us and japanese natural gas markets. *Energy Policy*, 62:1002–1009.

Weron, R., Bierbrauer, M., and Trück, S. (2004). Modeling electricity prices: jump diffusion and regime switching. *Physica A: Statistical Mechanics and its Applications*, 336(1):39–48.

Zivot, E. and Andrews, D. W. (1992). Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. *Journal of Business & Economic Statistics.*