Cointegration, Exogeneity and Isolating Long-run Price behaviour

Seyedeh Asieh Tabaghdehi
Department of Economics and Finance, Brunel University, School of Social Science, Uxbridge, Middlesex, UB8 3PH, UK

Abstract: In this research the cointegration analysis is used to detect key features of long--run structure in the gasoline market.
The main purpose of this study is to investigate possible long-run price leadership in the US gasoline market and the characteristics relevant to a competitive market using the VAR (Vector Autoregressive) model. After examining the stationarity and cointegration of the weekly gasoline prices in eight different regions of the US we research long-run price leadership and parallel pricing in the framework of the Cointegrated Vector Auto-Regression.

Keywords: Law of one price, Cointegration, Error Correction Model (ECM), Long-run relationship, Causality, Weak exogeneity, price dispersion, arbitrage.

INTRODUCTION

Gasoline is one of the most price fluctuant products in the world and the current dramatic changes in gasoline prices significantly affect the consumer and business confidence in the market. The gasoline price is significantly influenced by innovation, technological progress, and political instability in the global economy.

The gasoline market has been usually considered as competitive market as it is a homogeneous product with consumers are not particularly concerned over the brand, many suppliers and consumers and price related information commonly available for the consumers. Nevertheless the pump price at the gas station do differ in terms of location, local tax levels and services provided by the outlet.

The main concern of this research is long-run price differentiation in the gasoline market. Price disequilibria in the long-run in the neighbouring regions would affect regional activity and consumer might react radically towards high price differentials by moving job or houses to reduce travel costs, purchasing more fuel efficient vehicles etc., but the persistent price differential suggest discrimination and identifies the possibility of market power and informational inefficiency.

Analysing price properties may be effective in testing for “market definition” when the persistence of the volatility is reduced by this transformation of the data. If volatility is persistent the Johansen test may be converge to the asymptotic distribution when the sample is above 600-1000. (see the simulations in Rahbek et al (2002))

Forni (2004) suggests univariate stationary tests of log price differentials to determine a broad or a narrow market. However the regulation literature is not supporting the stationary test alone as an appropriate detector of the unfair pricing system.

In this paper we examine the Johnsen (1988 & 1996) cointegration in vector autoregressive model (VAR) to examine the existence of the long-run price leadership and its association with stochastic trend and weak exogeneity in the US gasoline market. The main hypothesis is studying on the long-run behaviour or the long-run efficiency of the US gasoline market. The key investigation of this study is why the gasoline price segregates in the different regions of one country at the specific time which it is against the market efficiency and antitrust laws. Price differentiation in different regions of a country identifies that prices are not fully reflecting all available information at any point in time. Any specific patterns of behaviour in the market cannot survive for the long time and it withdraw over time (Fama, 1998), but in US gasoline market the specific pattern of regional price differentiation constantly affecting the market efficiency. Hence we employ the cointegration test to empirically define a market and test for the relative integration.

In section I we introduced the gasoline market and econometric model used in the analysis. In section II we review of essential literature. Section III identified the data for the empirical analysis. Section IV we reviewed the price analysis and cointegration. In section V we carry out the estimation of arbitrage correction in gasoline market. In part VI we estimated weak exogeneity, long-run exclusion, and strictly exogeneity to investigate the potential parallel pricing in the gasoline market. Finally, in Section VII we conclude.

REVIEW OF ESSENTIAL LITERATURE

Since most of the econometricians are concern the cointegration regression and associated analysis are very important in the context of nonstationary variables. Cointegration certifies that there is a linear transformation of integrated variables which is stationary in order to avoid any spurious regression (Engle and Granger, 1981)[1]. If a linear combination of two or more non-stationary variable is stationary then the variable are supposed to be cointegrated complying with an equilibrium relationship in the long-run. Fundamentally the economic variables may float away from equilibrium in short-run but the economic forces will drive the variables to the equilibrium in the long-run.

Cointegration indicates a correlation among variables while spurious regression obtains from non-stationary data specifies there is no actual correlation, there as we considering the false justification could have a dangerous approach in economy respectively. Hence, if the residuals of the regression between two variables have a pattern this signifies that proposed regression is misspecified, but if the residuals are stationary identifying that two variables are cointegrated and there is a long run equilibrium relationship between two variables. Non-stationary time series might be caused by technological progress, economic evolution, or crises, changes in the consumer's preference and behaviour, policy or regime alteration, and organizational or institutional improvement. Regression based on stochastic non-stationary series made by continual cumulating of past effect called 'nonsense regression' and can cause many significant forecasting problems (Hendry and Juselius, 2000).

Cointegration is multivariate analysis and designed to find linear combinations of variables that remove unit roots meaning that the cointegration is an econometrics transformation of unit root data to the stationary format. Furthermore the existence of cointegration signifies the further estimation apart from static cointegration relationship such as dynamic relationship which includes both the equilibrium and the short-run adjustments respectively.

Hendry and Anderson (1977) argued that apart from differencing method there are other ways to accomplish stationary and they specified that often past equilibrium errors are stationary even when the individual time series are non-stationary. In 1987 Error Correction Mechanisms were introduced by Davidson, Hendry, Srba and Yoe and in 1981 Granger introduced the cointegration theory to comprehend the status of ECMs. Furthermore Engle and Granger (1987) proved that ECMs and cointegration were in fact the same things.

Following De Vanya and Walls (1999), Hendry and Juselius (2001), and Forni(2004) proposed that the single cointegration vector suggests a broad market definition. Abroad market implies ***(N-1)*** cointegration vector for ***N*** price series where a narrow market (segmented market) implies less than ***(N-1)*** cointegration vector (Forni, 2004). However with ***(N-1)*** cointegrating vector we might find long-run causality (cointegration exogeneity) or identify that one of the prices is weak exogenous (WE) for all cointegration vectors, or some of the prices ***(n2)*** are WE

[1] The null hypothesis of the Engle and Granger cointegration test is two variables do not cointegrate and the rejection of the null hypothesis is beneficial to avoid the spurious regression.

for the first ***(n1)*** vectors. Buccarossi (2006) and Forni (2004) suggest that in bivariate analysis competitive behaviour is consistent with parallel pricing which this can be confirm by testing the stationarity in the log price proportion or differences. There is an opportunity of parallel pricing only when ***N-1*** series react to a single price series which it is a weak exogenous factor for the vector of cointegration relationships, see Johansen (1992).

However Hunter and Burke (2007) determine that the stationary test on the price differential of the multivariate series is not an appropriate methodology for analysing the competitive behaviour in the market where it is only applicable in bivariate case.

ARBITRAGE CORRECTION IN GASOLINE MARKET

It is clear if the gasoline prices of different regions in the US could be identical then the associated market will be in equilibrium, otherwise there would be an arbitrage opportunity across all regions. This trading mechanism will be inclined to equalize the prices in long term by raising prices in the low-price regions and lowering the prices in the high-price regions. In empirical modeling cointegration analysis is multivariate time series study to estimate the equilibrium relationship and the Error Correction Model (ECM) is an appropriate method to investigate the incentive of the change in the variables determining equilibrium, see Patterson (2000). We investigative the long-run equilibrium in the US gasoline market by using the error correction model expressed as an “Arbitrage Correction Model”. The hypothesis of this argument is to permit the possibility that the regional gasoline prices deviate from the equilibrium value which the arbitrage opportunity is tend to correct disequilibrium over time.

The cointegration analysis only considering the long-run relationship of series without allowing for the short-run dynamics but an appropriate time-series modelling should describe both short-run dynamics and long-run equilibrium simultaneously. The Granger representation theorem demonstrates an important relation between the existence of cointegration and an error correction specification (Granger, 1997). The Error Correction Model (ECM) is basis on the dynamic relationship that combines both long-run and short-run behaviour of the series by identifying the potential equilibrium and the short-run adjustment towards the equilibrium. However according to Engle and Granger (1987) the ECM and cointegration are the same but with different names.

The error correction model is a dynamical coherent methodology that causes thrashing some restriction from the model which we named it as “arbitrage correction”. According to Kremers, Ericsson, and Dolado (1992) ECM is good model and can provide good cointegration. The ECM procedures is an ordinary starting point for modelling which it binds the cointegration relationship in long-run and it is robust to many information of trivial process such as specific lag lengths and dynamic involved whereas weak exogeneity is mostly valid empirically (Ericsson and Mackinnon, 2002).

However the ECM might not accurate the suitable long run relationship due to the structural change by estimating inconsistent cointegration vector and consequently leads us to the poor prediction (Clements and Hendry, 1995). For investigating the short-run dynamics of the relationship in the gasoline price of different regions in US we employed a Vector Error Correction Model (VECM) specification to explain the dynamics of the series (US gasoline price) considering that ECM clarifies short-run and long-run price reaction to the external shocks. Bachmeier and Griffin (2006) signified that the prices of the crude oil in the different geographical parts of the world are cointegrated. De Vany and Walls (1999) by using ECM approach identified a significant cointegration relation between eleven regions of the US between electricity prices.

The result of the Augmented Dickey Fuller (ADF) test and Error Correction Model (ECM) with **q** lags in the estimation is compared in Table 1. The results of the ADF test imply that the price proportion of most series is stationary as the critical values are significant at 5% level. Significance results indicate that series move in proportion to each other in the long-run but with no indication of them following each other. For testing the importance of the cointegration we analyse all equations in the system by testing for stationarity on the error correction term. In most cases the error correction model is significant but in one case the ECM is insignificant which it can arise due to lack of cointegration or WE[2], or the cointegration relation cannot be identified by the ECM, consequently these tests cannot be relied on to reject cointegration. Thus the prices move in proportion both in long-run and short-run confirming that the market is efficient in both terms.

Table 1- Summary of ADF tests, ECM test of regional price proportion. (With intercept and no trend)

Log price differential (q)[3]	ADF (q)/ OLS t-statistic	CM (q)/ OLS t-statistic	VECM (q)/ OLS t-statistic
P_{NE-MW} (25)	**-3.81 ***	-14.48 **| P_{MW}	1.27 | P_{MW}
P_{MW-CA} (25)	**-4.93 ***	8.70 **| P_{CA}	2.07 | P_{CA}
P_{MW-EC} (25)	**-4.72 ***	-10.15 **| P_{EC}	0.81 | P_{EC}
P_{LA-GC} (23)	-2.22	-5.63 **| P_{GC}	2.73 | P_{GC}
P_{RM-WC} (16)	**-5.81 ***	-6.62**| P_{WC}	1.01 | P_{WC}
P_{MW-GC} (20)	**-3.36***	-8.46 **| P_{GC}	1.19 | P_{GC}
P_{GC-RM} (16)	**-5.21***	-1.22| P_{RM}	3.96 | P_{RM}
P_{GC-WC} (20)	**-3.78****	-2.65 ** | P_{WC}	1.65 | P_{WC}
P_{MW-RM} (24)	**-4.43***	-3.76 **| P_{RM}	5.81 | P_{RM}

Note: Critical value at 1% is -3.44, at 5% is -2.87 computed in Oxmetrics Professional (Doornik and Hendry, 2009). * Significant at the 95% confidence level and ** significant at the 99% confidence level

For the further investigation we followed Boswijk (1992), Hunter and Simpson (1995), and Bauwens and Hunter (2000) and applied restrictions on α, β, and α as well as β to study the causal and exogeneity structure of data and identify the potential weak exogeneity and parallel pricing in the gasoline market.

[2] The single equations may suggest more WE variables than can arise when the test is applied at the level of the system.
[3] q is the lag order of each series which had been selected by using same process as the previous study via inspection of the correlogram.

EXOGENEITY AND CAUSALITY ANALYSIS OF GASOLINE REGIONAL MARKET

Modeling VAR in US Gasoline Market- Test of WE and Parallel Pricing

According to Granger (1969) cointegration implies causality, which the cointegration and causality mechanism are following the same hypothesis with different empirical approach for testing. Granger causality test focus on whether one time series is useful in forecasting another and identifies that if one variable provides statistically significant information in order to improve the estimation of the other variable. Testing for causality is usually useful for defining market boundaries (Horowitz 1981; Ravallion 1986; Slade 1986; Gordon, Hobbs, and Kerr 1993).

The most general test for cointegration is the multivariate test based on the vector autoregressive model (Johansen, 1988). By fallowing the VAR analysis we verify the interrelationships between variables and modelling multivariate ECM by identifying r cointegration combination with set of k variables. The adjustment parameters in the VAR system clarify the potential causality and weak exogeneity in the market and provide information on the price leadership and the dominion of one region in term of supply.

By analysing single equations from the VAR, econometrically and theoretically the framework is less restrictive as the stationary tests impose a common factor's restriction. Accordingly it binds the short-run respond to be same as the long-run and by implication causes arbitrage to be imposed on short-run parameters. Hence by estimating VAR as a system and relating to an error correction model, we illuminate market segmentation and potential arbitrage opportunity. In this part of the study we discuss exogeneity and identification, exogeneity implies restriction on the long-run parameters of the model.

Following the VAR model below:

$$p_t = \sum_{i=1}^{k-1} \mathit{\Pi}_i \, p_{t-i} + \Pi_k \, p_{t-k} + \mu + e_t$$

Where Π is an n×n matrix. The above VAR system can be rewritten in error correction form as:

$$\Delta\, p_t = \sum_{i=1}^{k-1} \mathit{\Gamma}_i \, \Delta p_{t-i} + \mathit{\Gamma}_k \, p_{t-k} + \mu + e_t$$

Considering $\boldsymbol{\Delta P_t}$ and $\boldsymbol{\Delta P_{t-i}}$ are stationary where error term assumed to be stationary, therefore for having the above equation applicable $\mathbf{p_t}$ must contains cointegration vectors or $\boldsymbol{\Gamma_k = 0_{n\times n}}$ with **r** rank which **r** determines the number of the linear combination of $\mathbf{p_t}$ which are stationary respectively. If r=N this indicates that N linear combination of series are stationary. If the **r=0** then there is no linear combinations that are stationary and in this study the rejection of the null of r=0 confirms that the relationship between regional oil prices is not spurious.

When series are cointegrated there is a restricted long-run parameter matrix as:

$$\Pi = \alpha\beta'$$

Where α and β are n×r matrix and this could be identified by setting α=I or β′=I, then we either find the β specifying the long-run relationship, or we identify α which it shows the adjustment for each variable in the short-run.

By normalising the proposed model the β has a unit values, then in I(0) case each equation has a single identified error correction or cointegration term. The 8 × 8 matrices express the long-run relationship in gasoline price of eight different geographical areas in US.

$$\Pi=\alpha\begin{bmatrix}\alpha 11 & 0 & \dots & 0\\ 0 & \alpha 22 & \cdots & 0\\ \vdots & 0 & \ddots & \vdots\\ 0 & 0 & \cdots & \alpha nn\end{bmatrix}\begin{bmatrix}\beta_1^{'}\\ \beta_2^{'}\\ \vdots\\ \beta_n^{'}\end{bmatrix} \text{ where } \beta_{ij}=[\beta_{1j} \quad \cdots \quad j \quad \cdots \quad \beta_{nj}]$$

If we have non-stationary series this is the approach to identification considered by Boswijk(1992) and consistent with single equation model considered above.

The existence of cointegration in a VAR system specifies that common stochastic trends are integrated in the system, where in ***N***-dimensional I(1) conitegrated VAR system with ***r*** cointegration relations there are ***N-r*** common stochastic trends.

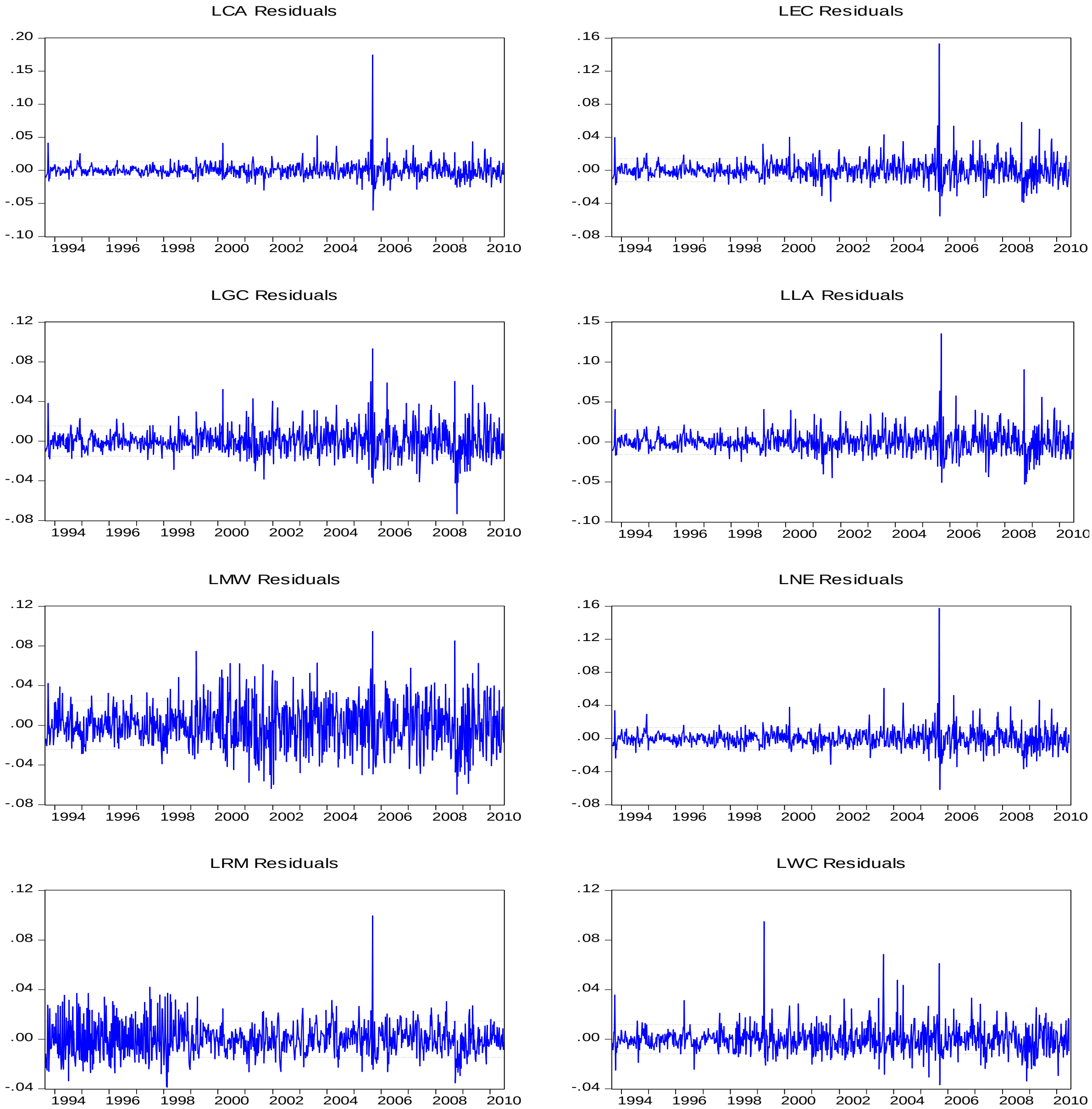
Table 2 indicates some descriptive statistics for the log gasoline price, differences, and for the residuals of the VAR model. Noticeably the gasoline price changes (price differences seems to be stationary around a constant mean of approximately zero. Using Jarque-Bera test the Normality is tested under the null of data is from normal distribution and it is strongly rejected in most of the price changes series and VAR residuals. The significant rejection of the null hypothesis could be explained as excess kurtosis or skewness.

Table 2- Descriptive statistics

1993 - 2010	$\bar{x}$	$S\bar{x}$	Skew	Kurt	Jarq.Bera	Min	Max
P_{CA}	5.123	0.374	0.56	2.10	77.27	4.58	6.04
P_{EC}	5.095	0.384	0.57	2.10	78.50	4.54	6.03
P_{GC}	5.063	0.375	0.64	2.21	84.41	4.52	6.00
P_{LA}	5.072	0.393	0.58	2.10	81.11	4.51	6.02
P_{MW}	5.076	0.383	0.58	2.14	77.32	4.49	6.02
P_{NE}	5.127	0.366	0.56	2.16	74.02	4.60	6.04
P_{RM}	5.117	0.361	0.65	2.24	84.15	4.60	6.03
P_{WC}	5.194	0.370	0.52	2.03	75.51	4.70	6.11
ΔP_{CA}	0.001068	0.018	2.14	33.25	35004.83	-0.08	0.23
ΔP_{EC}	0.001078	0.019	1.17	20.94	12274.99	-0.09	0.21
ΔP_{GC}	0.001008	0.020	0.12	10.43	2070.90	-0.14	0.14
ΔP_{LA}	0.001074	0.021	0.60	13.96	4554.93	-0.11	0.19
ΔP_{MW}	0.001042	0.028	0.01	5.42	219.65	-0.12	0.15
ΔP_{NE}	0.001073	0.018	1.59	25.64	19599.10	-0.09	0.21
ΔP_{RM}	0.001046	0.021	0.11	7.35	712.90	-0.10	0.14
ΔP_{WC}	0.0011	0.020	0.61	9.82	1799.24	-0.11	0.14
$\hat{\varepsilon}_{PCA}$	2.02E-18	0.012	4.00	59.50	119404.7	-0.06	0.17
$\hat{\varepsilon}_{PEC}$	1.34E-18	0.013	2.43	29.62	26846.12	-0.056	0.15
$\hat{\varepsilon}_{PGC}$	-5.46E-19	0.014	0.75	7.83	930.77	-0.07	0.09
$\hat{\varepsilon}_{PLA}$	-1.41E-18	0.014	1.50	15.29	5872.92	-0.05	0.14
$\hat{\varepsilon}_{PMW}$	-7.33E-19	0.022	0.23	3.56	19.37	-0.07	0.09
$\hat{\varepsilon}_{PNE}$	-4.53E-19	0.012	2.87	39.29	49500.32	-0.06	0.16
$\hat{\varepsilon}_{PRM}$	2.00E-18	0.013	0.74	6.56	546.11	-0.04	0.10
$\hat{\varepsilon}_{PWC}$	-8.11E-19	0.011	1.53	13.46	4353.39	-0.04	0.09

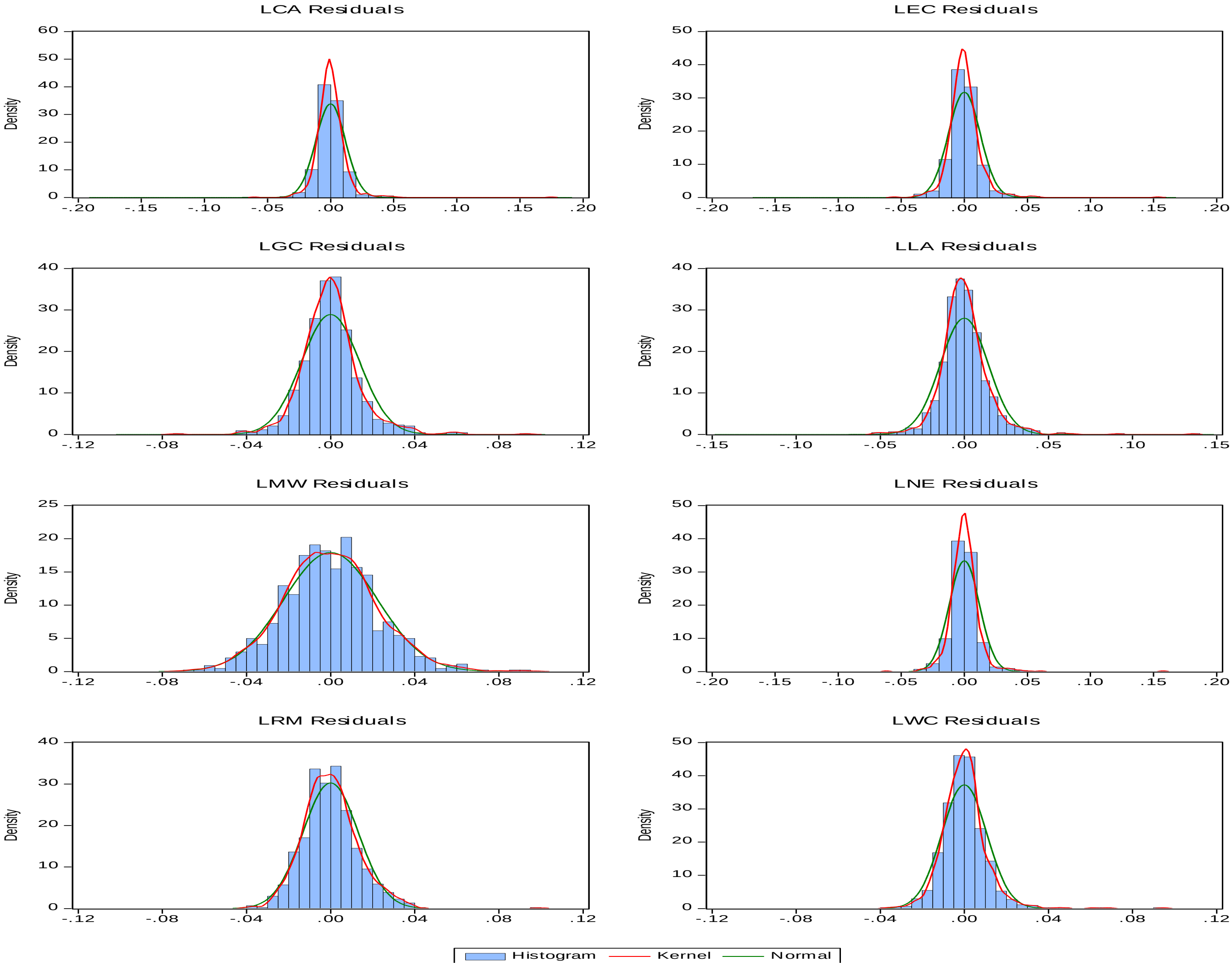
Considering the plot 1, there are some outlier observations in the residuals and the largest outlier is at end of 2003 and beginning of 2004 which this needs to be effectively explained in the specification of the VAR model. These observed outliers could be caused by extreme demand and supply shocks in regions.

Plot 1- Graph of residuals from a VAR (20) of US regional gasoline price and 99% confidence bands



In plot 2, the normal density of 8 VAR residuals is reported. The kernel density should not deviate excessively from the normal density; whereas residuals need to be homoscadstic with constant variance over time. However having longer tails in kernel density comparing with normal density indicates the existence of the outliers and confirms the non-normality, finding from Jarque-Bera test in Table **2**.

Plot 2- Normal Density of 8 VAR residuals of US regional gasoline price



Hence with 8 price series and 4 cointegration rank the corresponding unrestricted matrices is as below:

$$\Delta\begin{bmatrix} p_{1t} \\ \vdots \\ p_{8t} \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{14} \\ \vdots & \ddots & \vdots \\ \alpha_{81} & \cdots & \alpha_{84} \end{bmatrix} \begin{bmatrix} 1 & \dots & \beta_{14} & \beta_{81} \\ \vdots & \ddots & \vdots & \vdots \\ \beta_{41} & \dots & 1 & \beta_{48} \end{bmatrix} \begin{bmatrix} p_{1t-1} \\ \vdots \\ p_{8t-1} \end{bmatrix} - \begin{bmatrix} \gamma_{11} & \cdots & \gamma_{18} \\ \vdots & \ddots & \vdots \\ \gamma_{81} & \cdots & \gamma_{88} \end{bmatrix} \Delta x_{t-i} + \begin{bmatrix} \epsilon_{1t} \\ \vdots \\ \epsilon_{8t} \end{bmatrix}$$

According to Forni (2004) and De Vanya and Walls (1999) for N number of I(1) series there is at most one (N-r) long-run relationship as:

$$P_{At} - P_{Bt} \sim I(0)$$

For studying the gasoline market structure and identifying the number of long-run relationship with Johansen statistics test, it is required to impose some restrictions in the above VAR model. Following Forni (2004), De Vanya and Walls (1999), and Hendry and Juselius (2001) we analysed the competitiveness of gasoline market and abroad market definition by investigating a potential single trend among the series. Following, Johansen (1992), Hunter and Simpson (1995), Bauwens and Hunter (2000), Burke and Hunter (2012) weak exogeneity in the long-run had been identified by imposing restriction in $\alpha_{ij} = 0$, for i=1, ..., 4 and j=1, ..., 8. For weak exogeneity and long-run exclusion there are **r** restrictions on **α** and **β** for each variable excluded while for strict exogeneity, there are **2r** restrictions on **α** and **β** for each variable exclude.

Moreover the Long-run Exclusion can be test by imposing restrictions in $\beta_{ji}=0$, for j=1, …, 8 and i=1, …, 4 . Followed by testing for the Normalisation of the data by imposing below restriction:

$$\beta_{ii}=1, \text{ for } i=1, \ldots, 4$$

$$\beta_{ij}=0, \text{ for } \begin{cases} i = 1, \ldots, 4 \\ j = 1, \ldots, 8 \\ i \neq j \end{cases}$$

parallel pricing can be tested by imposing the restriction on $\boldsymbol{\beta'}_{1\times4}= [\mathbf{-1} \quad \ldots \quad \mathbf{-1}]$ to confirm a long-run correspondence between the price series.

In Table 4, tests of cointegration are derived from the VAR model and the result of imposed restrictions in α or β or both α and β at the same time are presented accordingly. The sample includes 901 observation and the results relate to Weak Exogeneity tests, Long-run Exclusion test, Strict Exogeneity and Parallel Pricing with 20 lags in the estimation. The first block of results in Table 4 relate to weak exogeneity test with r=4 and from p-values it can be determined that log price of Golf Coast (GC), Lower Atlantic (LA), East Coast (EC) and Mid West (MW) are potential WE for β, where essentially must be only N-r=1 weak exogenious variable. Order by WE test suggests considerable WE in the Golf Coast and Lower Atlantic gasoline prices identifying that the gasoline price of GC and LA determines the other regions prices, consistence with Burke and Hunter (2012). However more significantly Golf Coast price changes will effect directly to the other region's prices and everything is conditioned on the GC correspondingly. Following the Juselius (1995) in the next section of the table long-run exclusion is tested and the results are strongly significant in most regions indicating the rank condition and the likely robustness of proposition of cointegrating vectors.

As an alternative analysis we order the system using the long-run exclusion (LE)[4] disinclined to normalise on a variable that may be LE(Boswijk, 1996). Here requiring LE to interact with all the other variables and this is the least likely to cointegrate with all the variables.
As a next alternative analysis, in the next section of the Table 4 we order the system based on the imposition of Stricted Exogeneity (SE) by combining LE and WE simultaneously to reorder the system respectively. The result of the SE test provides a rational for reorder a system from less strictly exogeneity to most SE and position LA last as a least SE series.

In the last part of the Table 3 followed by testing of normalization and WE conditioned on LA log prices as most WE variable and then testing for N+WE conditioned on GC log price as most WE variable. The same values of χ^2 from the N+WE test and the WE test identifies that the normalisation is innocuous. This Hence, the result gives rise to the same conclusion as the KPSS stationary tests, suggesting that we could get Long-run Equilibrium Price Targeting (LEPT) when we cannot find parallel pricing.
To this end the regional gasoline price behaviour may not be consistence with the fully functioned gasoline market across the US. There may be geographical or structural reasons for this to occur. For investigating the market structure in more depth we study the US company gasoline price and analyse the WE of the company data in the following section of our study. A penalty associated with analysing the price series, is the volatility in the data which it is addressed in the next model.

---

[4] Long-run Exclusion test is an estimation to exclude the potential redundant variable from the cointegration space.

Table 3- Test of cointegration, WE, LE, SE and Parallel Pricing of US Gasoline Price 1993-2010

Hypothesis		Null: r≤4	Statistics [p-value]
(WE)| r=4	P_{CA}	α_{1i}= 0, for i=1, …, 4	$\chi^2(4)$ = 15.266 [0.0042] **
	P_{EC}	α_{2i}= 0, for i=1, …, 4	$\chi^2(4)$ = 7.0753 [0.1320]
	P_{GC}	α_{3i}= 0, for i=1, …, 4	$\chi^2(4)$ =1.7056 [0.7897]
	P_{LA}	α_{4i}= 0, for i=1, …, 4	$\chi^2(4)$ =3.5660 [0.4679]
	P_{MW}	α_{5i}= 0, for i=1, …, 4	$\chi^2(4)$ =8.7465 [0.0678]
	P_{NE}	α_{6i}= 0, for i=1, …, 4	$\chi^2(4)$ =10.803 [0.0289] *
	P_{RM}	α_{7i}= 0, for i=1, …, 4	$\chi^2(4)$ =32.663 [0.0000] **
	P_{WC}	α_{8i}= 0, for i=1, …, 4	$\chi^2(4)$ = 19.215 [0.0007] **
(LE)| r=4	P_{CA}	β_{j1}= 0, for j=1, …, 4	$\chi^2(4)$ = 18.758 [0.0009] **
	P_{EC}	β_{j2}= 0, for j=1, …, 4	$\chi^2(4)$ = 10.883 [0.0279] *
	P_{GC}	β_{j3}= 0, for j=1, …, 4	$\chi^2(4)$ = 15.663 [0.0035] **
	P_{LA}	β_{j4}= 0, for j=1, …, 4	$\chi^2(4)$ = 8.9190 [0.0632]
	P_{MW}	β_{j5}= 0, for j=1, …, 4	$\chi^2(4)$ = 24.574 [0.0001] **
	P_{NE}	β_{j6}= 0, for j=1, …, 4	$\chi^2(4)$ = 0.89249 [0.9256]
	P_{RM}	β_{j7}= 0, for j=1, …, 4	$\chi^2(4)$ = 36.858 [0.0000] **
	P_{WC}	β_{j8}= 0, for j=1, …, 4	$\chi^2(4)$ = 26.188 [0.0000] **
Normalization (N) + (WE) P_{GC} | r=4		β_{ii}=1, for i=1, …, 4 $\beta_{ij}= 0$, for $\begin{cases} i = 1, \dots, 4 \\ j = 1, \dots, 4 \\ i \neq j \end{cases}$ α_{3i}= 0, for i=1, …, 4	$\chi^2(4)$ = 1.7056[0.7897]
Normalization (N) + (WE) P_{LA} | r=4		β_{ii}=1, for i=1, …, 4 $\beta_{ij}= 0$, for $\begin{cases} i = 1, \dots, 4 \\ j = 1, \dots, 4 \\ i \neq j \end{cases}$ α_{4i}= 0, for i=1, …, 4	$\chi^2(4)$ = 3.5698[0.4673]
SE = (LE) + (WE)| r=4	P_{CA}	α_{1i}= 0, for i=1, …, 4 β_{j1}= 0, for j=1, …, 4	$\chi^2(8)$ = 53.103 [0.0000] **
	P_{EC}	α_{2i}= 0, for i=1, …, 4 β_{j2}= 0, for j=1, …, 4	$\chi^2(8)$ = 35.607 [0.0012] **
	P_{GC}	α_{3i}= 0, for i=1, …, 4 β_{j3}= 0, for j=1, …, 4	$\chi^2(8)$ =44.477 [0.0000] **
	P_{LA}	α_{4i}= 0, for i=1, …, 4 β_{j4}= 0, for j=1, …, 4	$\chi^2(8)$ =30.611 [0.0063] **
	P_{MW}	α_{5i}= 0, for i=1, …, 4 β_{j5}= 0, for j=1, …, 4	$\chi^2(8)$ =49.294 [0.0000] **
	P_{NE}	α_{6i}= 0, for i=1, …, 4 β_{j6}= 0, for j=1, …, 4	$\chi^2(8)$ =33.533 [0.0024] **
	P_{RM}	α_{7i}= 0, for i=1, …, 4 β_{j7}= 0, for j=1, …, 4	$\chi^2(8)$ =70.029 [0.0000] **
	P_{WC}	α_{8i}= 0, for i=1, …, 4 β_{j8}= 0, for j=1, …, 4	$\chi^2(8)$ = 62.893 [0.0000] **

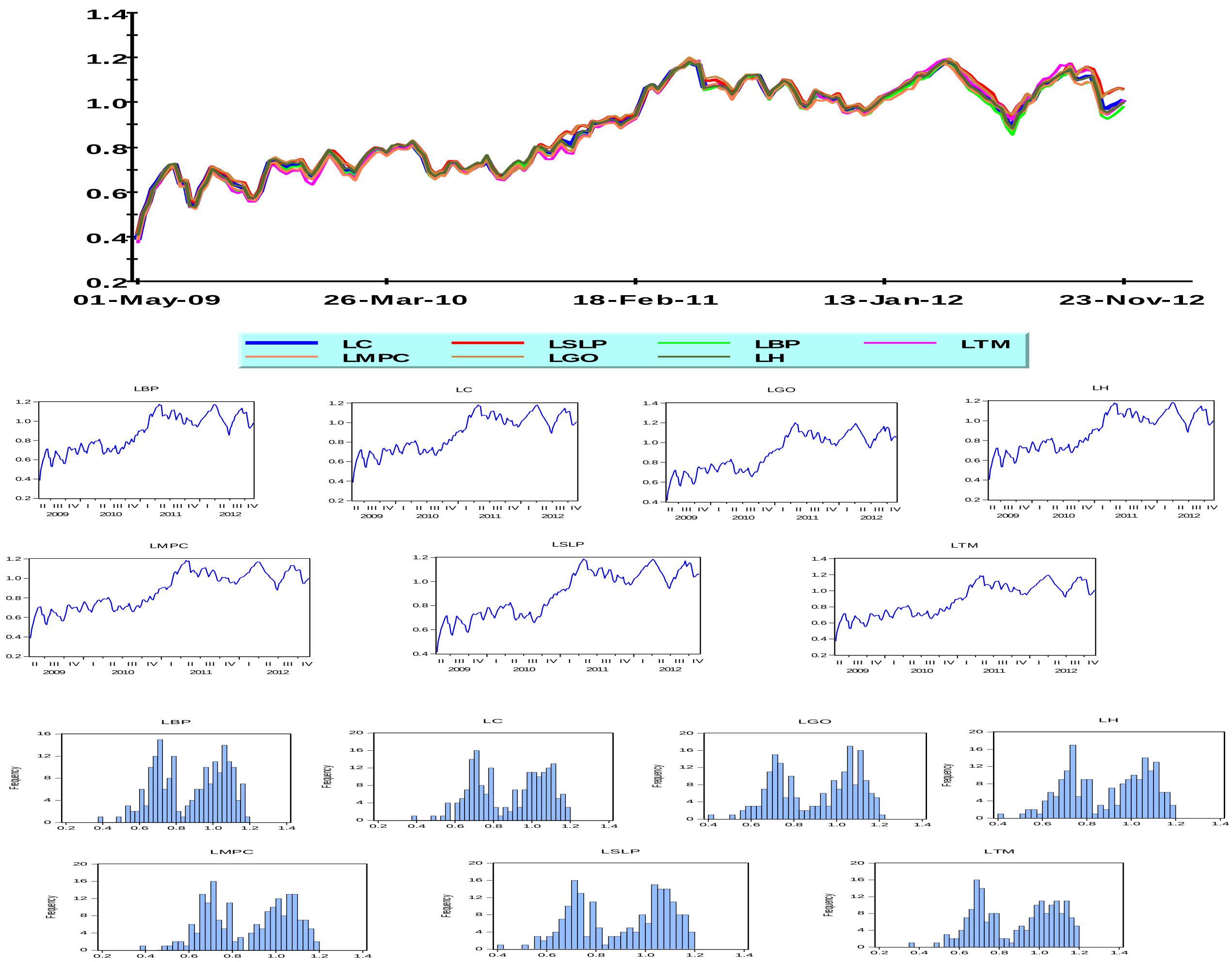
Note: Weak Exogeneity (WE), Long-run Exclusion (LE), and Strict Exogeneity (SE). * significant at the 5% level and ** significant at the 1% level.

COINTEGRATION AND EXOGENEITY ANALYSIS OF GASOLINE MARKET USING COMPANY DATA

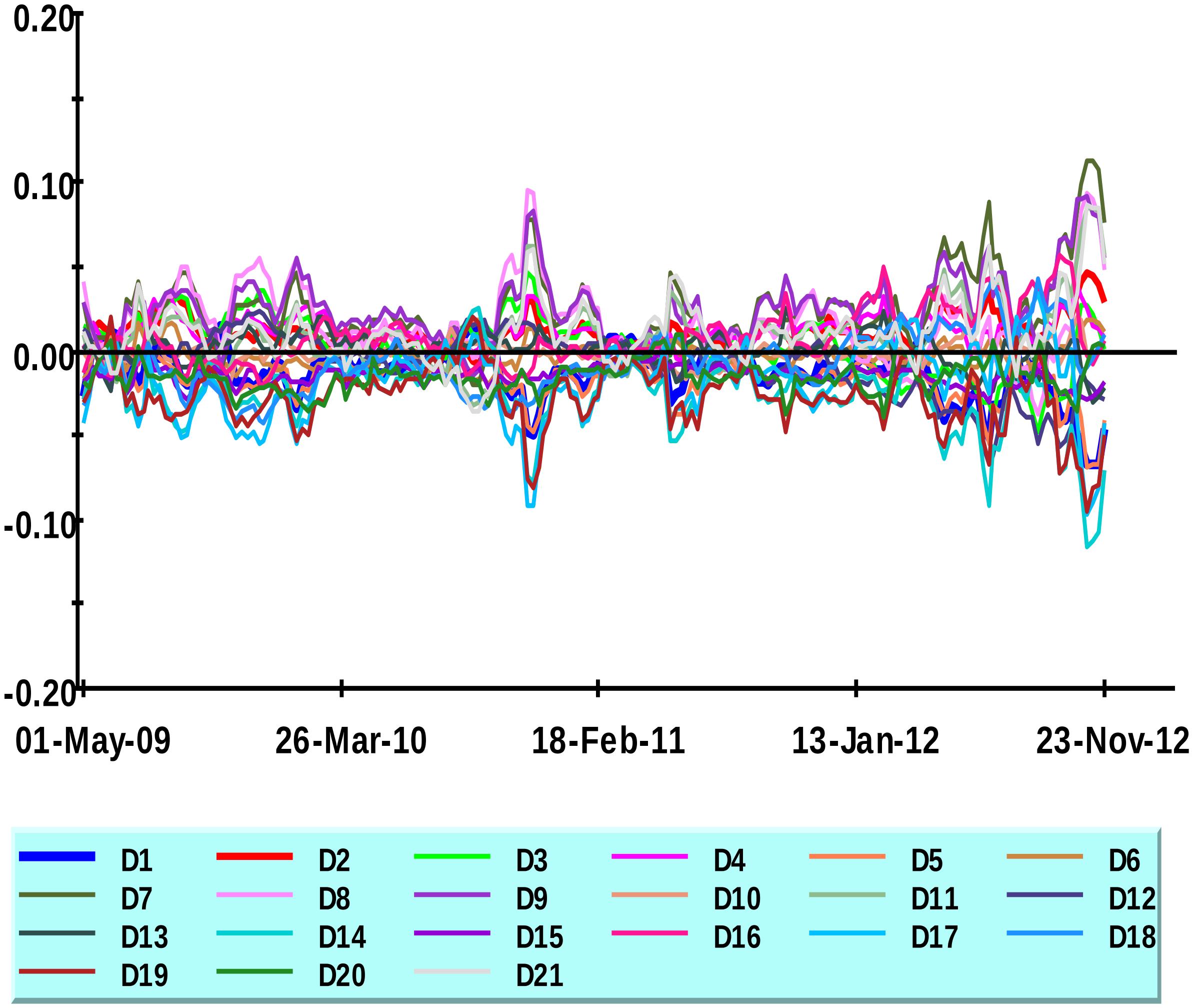
Following Hendry and Juselius (2001) and Hunter and Burke (2012) we consider cointegration relationship and the dynamic model using company data across the US to identify the market leader behavior respectively. The main concern of this part of our study is to focus on exogeneity by applying the cointegration and WE test to explicate the gasoline market structure and price behavior in the market using the company data.

The data is weekly gasoline company prices in the US. The data are for seven major gasoline producing companies in the US for the period of May 2009 - November 2012 (187 observations). Here seven companies are selected, Citgo, Sunoco, BP, Transmontaigne, Marathone. Gulf Oil, and Hess Corporation. Plot 5shows the log company prices, Citgo (C), Sunoco Logistic Partners (SLP), BP, Transmontaigne (TM), Marathon Petroleum Corporation (MPC). Gulf Oil (GO), and Hess Corporation (HC). Plots below show the US log companies price differentials.

Plot 3- The log gasoline price of Citgo (C), Sunoco Logistic Partners (SLP), BP, Transmontaigne (TM), Marathon Petroleum Corporation (MPC). Gulf Oil (GO), and Hess Corporation (HC)

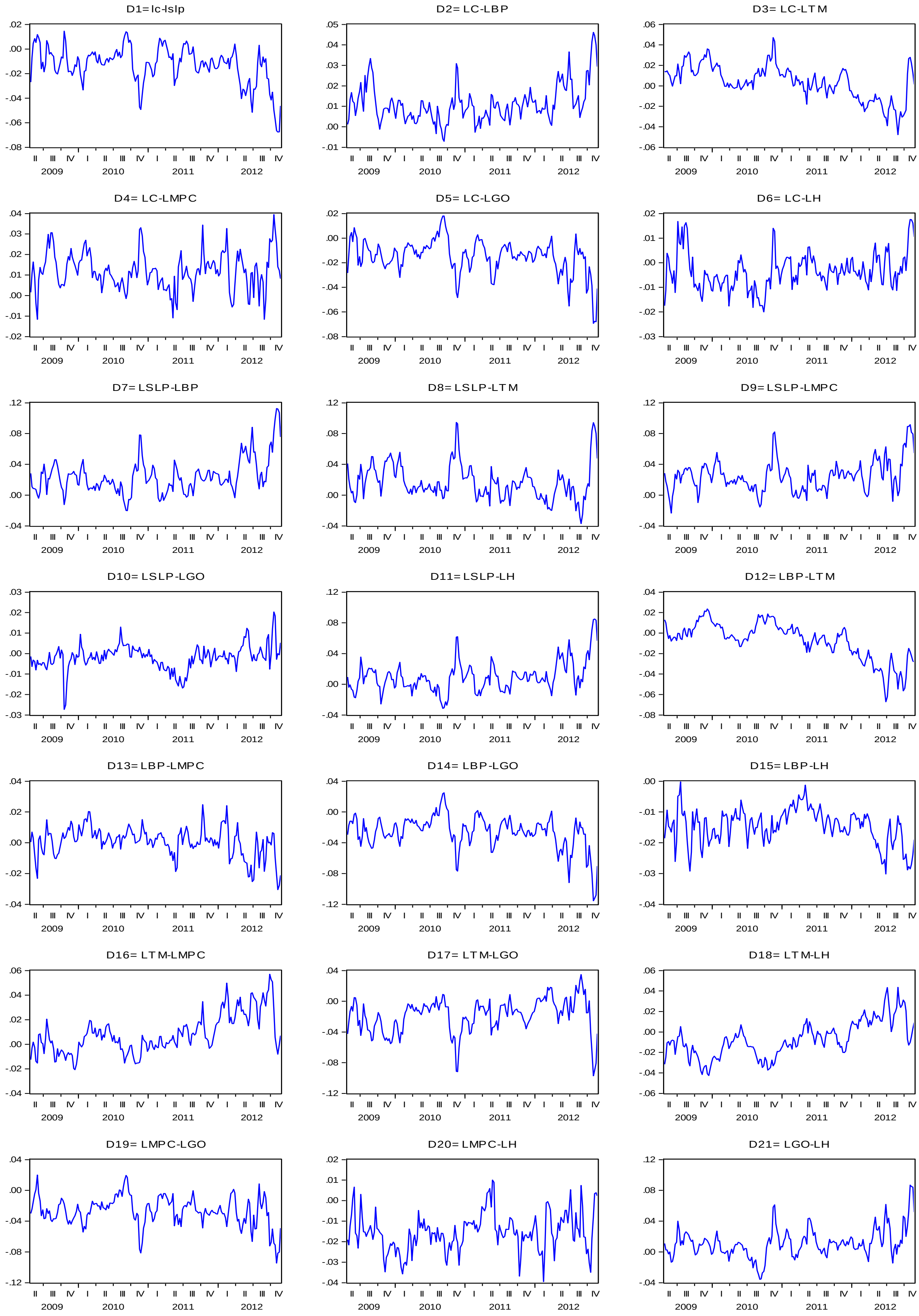


Plot 4- The log of gasoline price differential within seven different companies[5]

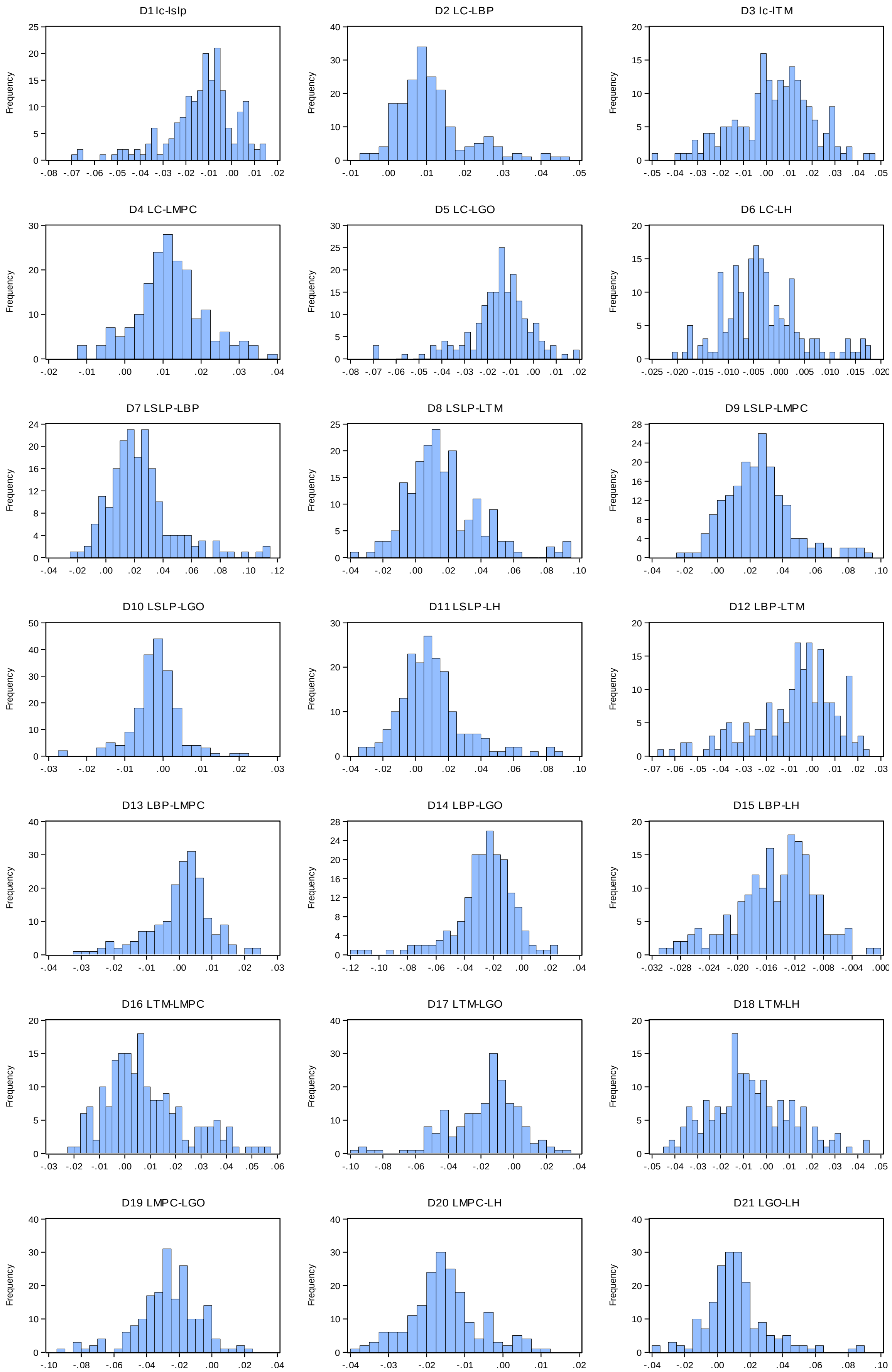


[5] C and SLP, C and BP, C and TM, C and MPC, C and GO, C and H, SLP and BP, SLP and TM, SLP and MPC, SLP and GO, SLP and H, BP and TM, BP and MPC, BP and GO, BP and H, TM and MPC, TM and GO, TM and H, MPC and GO, MPC and H, GO and H

Plot 5- The log of gasoline price differential within seven different companies



Plot 6- Frequancy distribution of the gasoline log price differential within seven different companies



The test of stationary on company data is obtained from single equation using sample of T=187 weekly observations and the result of ADF tests and ECM with q lags are presented in table below.

The ADF test result shows out of 21company price proportions in 19 price proportion the critical values are significant which this implies the price proportions are stationary. Stationarity of the company price proportions indicates that in the long-run the series move in proportion to each other and follow the same stochastic trend suggesting that there is arbitrage opportunity in the long-run. Furthermore we applied ECM to identify that whether prices move in proportion in the both long-run and short-run as the Long and the short-run performance fixes in efficient behavior of the market. The result of this restriction is presented in table 4 indicating that many of q lagged price inflation terms are significant and this implies that mostly market is efficient in the short-run.

However in 4[6] price proportions there is discrepancy between the result of ADF test and ECM test. For further investigation we applied VECM test which it is an autoregressive model related to the error correction model. From the last column of table 4, the test statistics for the all price proportions are insignificant and this implies that the differences from ADF test result and ECM test in 4 price proportion series mentioned above relates to ΔP_{TM}, ΔP_{BP}, and ΔP_{C}. In general this insignificant result indicates that either there is no cointegration relation or it is unfeasible to identify the cointegration relation respectively.

Table 4- Summary of ADF tests, ECM test of company price proportion. (With intercept and no trend)

Log price differential	ADF (q)/ OLS t-statistic	:M (q)/ OLS t-statistic	VECM (q)/ OLS t-statistic
$P_{C \text{ and } SLP}$	-3.68** (2)	-5.14 ** |P_{SLP}	0.13 |P_{SLP}
$P_{C \text{ and } BP}$	-3.43* (2)	5.63 ** | P_{BP}	1.26 | P_{BP}
$P_{C \text{ and } TM}$	-2.90* (2)	-3.93 ** | P_{TM}	1.11 | P_{TM}
$P_{C \text{ and } MPC}$	-4.83** (2)	-6.22 ** | P_{MPC}	1.05 | P_{MPC}
$P_{C \text{ and } GO}$	-4.24** (2)	-5.03 ** | P_{GO}	0.06 | P_{GO}
$P_{C \text{ and } H}$	-3.79** (2)	-5.72 ** | P_{H}	1.29 | P_{H}
$P_{SLP \text{ and } BP}$	-3.51** (2)	-5.19 ** | P_{BP}	0.37 | P_{BP}
$P_{SLP \text{ and } TM}$	-3.86** (2)	-5.30 ** | P_{TM}	1.52 | P_{TM}
$P_{SLP \text{ and } MPC}$	-3.91** (2)	-5.56 ** | P_{MPC}	1.35 | P_{MPC}
$P_{SLP \text{ and } GO}$	-4.34** (2)	-6.29 ** | P_{GO}	0.22 | P_{GO}
$P_{SLP \text{ and } H}$	-3.66** (2)	-5.23 ** | P_{H}	1.53 | P_{H}
$P_{BP \text{ and } TM}$	**-2.08 (2)**	-2.24 * | P_{TM}	0.07 | P_{TM}
$P_{BP \text{ and } MPC}$	-4.29** (2)	-5.00 ** | P_{MPC}	0.23 | P_{MPC}
$P_{BP \text{ and } GO}$	-3.88** (2)	-4.35 ** | P_{GO}	0.06 | P_{GO}
$P_{BP \text{ and } H}$	-4.22** (2)	-5.31 ** | P_{H}	0.39 | P_{H}
$P_{TM \text{ and } MPC}$	-2.98* (2)	-3.72 ** | P_{MPC}	-0.08 | P_{MPC}
$P_{TM \text{ and } GO}$	-3.88** (2)	-4.06 ** | P_{GO}	-0.26 | P_{GO}
$P_{TM \text{ and } H}$	**-2.47 (2)**	**-2.79 | P_{H}**	0.17 | P_{H}
$P_{MPC \text{ and } GO}$	-4.34** (2)	-4.77 ** | P_{GO}	0.20 | P_{GO}
$P_{MPC \text{ and } H}$	-4.74** (2)	-5.23 ** | P_{H}	0.93 | P_{H}
$P_{GO \text{ and } H}$	-4.00** (2)	-4.75 ** | P_{H}	1.63 | P_{H}

Note: ADF Critical value at 1% is -3.44, at 5% is -2.87 computed in Oxmetrics Professional (Doornik and Hendry, 2009).

* Significant at the 95% confidence level and ** significant at the 99% confidence level

[6] $P_{C \text{ and } BP}$, $P_{C \text{ and } TM}$, $P_{BP \text{ and } TM}$, $P_{TM \text{ and } MPC}$

For an efficient estimation we conditioned the single equation to one or more exogenous variable. As a first step using companyies weekly data we defined the cointegration rank and present the result in the table 5. The Johansen trace test result indicates that there are 4 cointegrating equations at the 5% level.

***Table5: Unrestricted Cointegration Rank Test – Trace* of US Gasoline company Price 2009-2012**

H_0 : rank ≤	Trace test	P-value
rank ≤0	210.44	[0.000] **
rank ≤1	147.27	[0.000] **
rank ≤2	97.290	[0.000] **
rank ≤3	58.469	[0.003] **
rank ≤4	30.732	[0.039] *
rank ≤5	13.612	[0.094]
rank ≤6	3.1218	[0.077]

Note: * significant at the 5% level and ** significant at the 1% level.

We followed the related system as an unrestricted VAR in error correction form and applies weak exogeneity (WE) test, long-run exclusion (LE), and strict exogeniety (SE) test using r=4. The P-value of the WE test in the table 6 indicates that the log price for all major companies in US might be WE for β where for the Gulf Oil company and Sunoco Logistic Partners WE is accepted at highly significance level. In the second part of the table 6 the insignificant *p-values* shows that none of the prices could be long-run excluded.

Table 6- Test of WE, LE and SE of US Gasoline company Price 2009-2012

Hypothesis		Null: r≤4	Statistics [p-value]
(WE)| r=4	P_{BP}	α_{1i}= 0, for i=1, …, 4	$\chi^2(4)$ = 3.1709 [0.5296]
	P_C	α_{2i}= 0, for i=1, …, 4	$\chi^2(4)$ = 2.3903 [0.6644]
	P_{GO}	α_{3i}= 0, for i=1, …, 4	$\chi^2(4)$ =0.81513[0.9364]
	P_H	α_{4i}= 0, for i=1, …, 4	$\chi^2(4)$ =4.0462 [0.3998]
	P_{MPC}	α_{5i}= 0, for i=1, …, 4	$\chi^2(4)$ =2.1942 [0.7001]
	P_{SLP}	α_{6i}= 0, for i=1, …, 4	$\chi^2(4)$ =0.63963 [0.9586]
	P_{TM}	α_{7i}= 0, for i=1, …, 4	$\chi^2(4)$ =1.9057 [0.7531]
(LE)| r=4	P_{BP}	α_{1i}= 0, for i=1, …, 4	$\chi^2(4)$ = 40.356 [0.0000] **
	P_C	α_{2i}= 0, for i=1, …, 4	$\chi^2(4)$ = 21.873 [0.0002] **
	P_{GO}	α_{3i}= 0, for i=1, …, 4	$\chi^2(4)$ =23.744 [0.0001] **
	P_H	α_{4i}= 0, for i=1, …, 4	$\chi^2(4)$ =36.885 [0.0000] **
	P_{MPC}	α_{5i}= 0, for i=1, …, 4	$\chi^2(4)$ = 27.844 [0.0000] **
	P_{SLP}	α_{6i}= 0, for i=1, …, 4	$\chi^2(4)$ =19.293 [0.0007] **
	P_{TM}	α_{7i}= 0, for i=1, …, 4	$\chi^2(4)$ = 13.696 [0.0083] **
SE = (LE) + (WE)| r=4	P_{BP}	α_{1i}= 0, for i=1, …, 4 β_{j1}= 0, for j=1, …, 4	$\chi^2(8)$ = 44.336 [0.0000] **
	P_C	α_{2i}= 0, for i=1, …, 4 β_{j2}= 0, for j=1, …, 4	$\chi^2(8)$ = 24.454 [0.0019] **
	P_{GO}	α_{3i}= 0, for i=1, …, 4 β_{j3}= 0, for j=1, …, 4	$\chi^2(8)$ =24.308 [0.0020] **
	P_H	α_{4i}= 0, for i=1, …, 4 β_{j4}= 0, for j=1, …, 4	$\chi^2(8)$ =39.472 [0.0000] **
	P_{MPC}	α_{5i}= 0, for i=1, …, 4 β_{j5}= 0, for j=1, …, 4	$\chi^2(8)$ =31.088 [0.0001] **
	P_{SLP}	α_{6i}= 0, for i=1, …, 4 β_{j6}= 0, for j=1, …, 4	$\chi^2(8)$ =19.758 [0.0113] *
	P_{TM}	α_{7i}= 0, for i=1, …, 4 β_{j7}= 0, for j=1, …, 4	$\chi^2(8)$ =16.585 [0.0347] *

Note: Weak Exogeneity (WE), Long-run Exclusion (LE), and Strict Exogeniety (SE). * significant at the 5% level and ** significant at the 1% level.

CONCLUSION

For non-stationary variables cointegration test is an appropriate method for the perfect market investigation. The test for the market integration is employed by analysing the US gasoline market. The empirical finding indicates that gasoline price of different regions are cointegrated and this rejects the distinct market definition. Considering cointegration relation indicates that error correction test is more powerful than the Dickey-Fuller test (Dolado, Ericsson and Kremers, 1992).
The causality test indicates the same result and it is not following the hypothesis of one identical and global market for gasoline. To this end the observed market behaviour in long-run could be due to the geographical condition of state and refineries capacity. Considering the empirical results we are suggesting a change in the regulation of the gasoline market by using an appropriate strategy such as tax break in the smaller firms in order to achieve a single and efficient market.

Considerably the availability and accessibility of the consumer's market information could significantly affect the price relationship in the gasoline market. There is a need of an authoritative body for utilizing required political decision in order to harmonize the gasoline price a cross of one country.

VIII. REFERENCES

Adams III, A.F., 1997, Search costs and price dispersion in a localized, homogeneous product market: Some empirical evidence, Review of Industrial Organization 12, 801-808.

Alan T. Sorensen, Equilibrium Price Dispersion in Retail Markets for Prescription Drugs, *Journal of Political Economy,* 2000, vol. 108, no. 4

Asplund,M. and R. Friberg: "The Law of One Price in Scandinavian Duty-Free Stores", American Economic Review, vol. 91, September 2001, pp. 1072-1083.

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.

Barron, J. M., Taylor, B. A., Umbeck, J. R. "Number of Sellers, Average Prices, and Price Dispersion", International Journal of Industrial Organization, 2004, Vol. 22 (8-9), 1041-1066.

Borenstin S., Bushnell J., Knittel C. R., Wolfarm C., (2008), " Inefficiencies and Market Power in Financial Arbitrage: A study of California's Electricity Markets", The Journal of Industrial Economics. Vol. 56 (2), 347-378.

Boswjik, H.P., (1996), "Cointegration, identification, and exogeneity: inference in structural error correction models", Journal of Business and Economic Statistics 14, 153-160.

Burke,S.P., and Hunter, J., (2005), "Modelling Non-stationary Economic Time Series: A Multivariate Approach", Palgrave, Basingstoke.

Carlson, J.A. and R.P. McAfee, 1982, Discrete equilibrium price dispersion: Extensions and technical details, Institute Paper no. 812, Purdue University.

De Vanya, A.S. and Walls, D. (1999). 'Cointegration analysis of spot electricity prices: insights on transmission efficiency in the western US', Energy Economics, vol 21, pp.435-448.

Dreher, A., Krieger, T., 2008. Do prices for petroleum products converge in a unified Europe with non-harmonized tax rates. Energy Journal 29, 41–68.

Ericsson, N.R., and MacKinnon, J.G., (2002), " Distribution of Error Correction Tests for Cointegration", Econometrics Journal 5, 285-318.

Fix, A. and Buika, K.: "Geographic market definition in the US airline industry". October 2010. ,JOB MARKET PAPER.

Fran Asche, Daniel V. Gordon, Rognvaldur Hannesson, (2004), "Tests For Market Integration and the Law of One Price: The Market For Whitefish in France" *Marine Resource Economics*, Volume 19, pp. 195–210.

Gerd KUPPER. Bert WILLEMS., (2011), "Arbitrage in energy markets: competing in the incumbent's shadow," Center for Economic Studies - Discussion papers

Giulietti, M. and M. Waterson, 1997, Multiproduct firms' pricing behaviour in the Italian grocery trade, Review of Industrial Organization 12, 817-832.

Goldberg, Pinelopi K. & Verboven, Frank, 2005. "Market integration and convergence to the Law of One Price: evidence from the European car market" Journal of International Economics Elsevier, vol. 65(1), pages 49-73

Gordon, D.V., J.E. Hobbs, and W.A. Kerr. 1993. A Test for Price Integration in the EC Lamb Market. *Journal of Agricultural Economics* 44:126–34.

Guimãraes, P., 1996, Search intensity in oligopoly, Journal of Industrial Economics 44, 415-426.

Hendry, D., Juselius, K., 2000. Explaining cointegration analysis: part I. Energy Journal 21, 1–42.

Hendry, D., Juselius, K., 2001. Explaining cointegration analysis: part II. Energy Journal 22, 75–120.

Hosken, D and C.T. Taylor (2004), 'Discussion of: Using stationarity tests in antitrust market definition', American Law and Economics Review, 6, 465-475.

Horowitz, I. 1981. Market Definition in Antitrust Analysis: A Regression-Based Approach. *Southern Economic Journal* 48:1–16.

Kremers, J.J.M., Ericsson, N.R. and Dolado, J.J., (1992), "The Power of Cointegration Tests", Oxford Bulletin of Economics and Statistics 54, 349-67.

Küpper,G., and Willems, B., (2010), " Arbitrage in Energy Markets: Price Discrimination under Congestion" *The Energy Journal* 31, 41-66.

Marvel, H.P., 1976, The economics of information and retail gasoline price behavior: An empirical analysis, Journal of Political Economy 84, 1033-1060.

Miller, I. J., and Ratti, R. A., (2009), "Crude Oil and Stock Markets: Stability, instability, and bubbles", Energy Economics 31, 559-568.

Mohammadi, H., (2009), "Electricity Prices and Fuel Costs: Long-run relations and short-run dynamics", Energy Economics 31, 503-509.

Monica Giulietti, Catherine Waddams Price, and Michael Waterson, CONSUMER CHOICE AND COMPETITION POLICY:A STUDY OF UK ENERGY MARKETS, 2005, The Economic Journal, 115 (October), 949–968.

Pinkse, J., Slade, M.E., Brett, C. (2002). Spatial Price Competition: A Semiparametric Approach, Econometrica, 70, 3, 1111-1153.

Png, I.P.L. and D. Reitman, 1994, Service time competition, RAND Journal of Economics 25, 619-634.

Ravallion, M. 1986. Testing Market Integration. *American Journal of Agricultural Economics* 68:102–9.

Rosa M. Abrantes-Metz, Luke M. Froeb, John F. Geweke, Christopher T. Taylor, 2005, A Variance Screen for Collusion, working papers

Rosenthal, R.W., 1980, A model in which an increase in the number of sellers leads to a higher price, Econometrica 48, 1575-1579.

Szymon Wlazlowski , Monica Giulietti , Jane Binner , Costas Milas, “Price dynamics in European petroleum markets”, Energy Economics 31 (2009) 99–108.

Sche_man,David and Pablo Spiller, (1987), “Geographic Market Definition under the US Department of Justice Merger Guidelines," Journal of Law and Economics, 30 (1), 123-147.

Schmidt, S., 2001, Market structure and market outcomes in deregulated rail freight markets, International Journal of Industrial Organization 19, 99-131.

Serletis, A., Herbert, J., 1999. The message in North American energy prices. Energy Economics 21, 471–483.

Slade, M.E. 1986. Exogeneity Test of Market Boundaries Applied to Petroleum Products. *Journal of Industrial Economics* 34:291–304.

Slade, M.E. (1992). Vancouver’s Gasoline-Price Wars: An Empirical Exercise in Uncovering Supergame Strategies, The Review of Economic Studies, 59, 2, 257-276.

Sorensen, A.T., 2000, Equilibrium price dispersion in retail markets for prescription drugs, Journal of Political Economy 108, 833-850.

Stahl, D.O., 1989, Oligopolistic pricing with sequential consumer search, American Economic Review 79, 700-712.

Stigler, G.J. 1969. “*The Theory of Price”*. London: Macmillan.

Stigler, G.J., and R.A. Sherwin. 1985. The Extent of a Market. *Journal of Law and Economics* 28:555–85.

Stiglitz, J.E., 1987, Competition and the number of firms in a market: Are duopolies more competitive than atomistic markets, Journal of Political Economy 95, 1041-1061.

Theodore Panagiotidis , Emilie Rutledge, “Oil and gas markets in the UK: Evidence from a cointegrating approach” , Energy Economics 29 (2007) 329–347

Van Hoomissen, Theresa. “Price Dispersion and Inflation: Evidence from Israel.” *J.P.E.* 96 (1988): 1303–14.

Varian, H.R., 1980, “A model of sales”, American Economic Review 70, 651-659.