Ratio Mathematica Volume 51, 2024

ARMA and/or SETAR Estimation and Out-of-Sample Forecast of the Mean-Reversion Between Brent Crude Oil and Gasoline Prices on the Ghanaian Market

Jonathan Ayiku*
Emmanuel de-Graft Johnson Owusu-Ansah†
Simon Kojo Appiah‡
Emmanuel Harris§

Abstract

This study investigates linear and nonlinear mean reversion in bivariate price connections using VECM and SETAR models. It estimates relevant parameters and compares Self-Exciting Threshold AutoRegressive (SETAR) prediction performance to linear models. Findings reveal that cointegration outperforms the linear AutoRegressive (AR) models, with the 2-SETAR model showing significant asymmetric behavior and continuous price trends, indicating superior predictive capability.

Keywords: Mean-reversion; VECM; TVECM; SETAR; Time Series; ARMA; Ex-pump Ex-refinery

2020 AMS subject classifications: 62M10,91B84 [1]

*Department of Statistics and Actuarial Science, Kwame Nkrumah University of Science and Technology (KNUST), Kumasi, Ghana); nii.ayittey@gmail.com.

†Dept. of Statistics and Actuarial Science, KNUST Kumasi, Ghana); degraftt@gmail.com

‡Dept. of Statistics and Actuarial Science, KNUST, Kumasi, Ghana); skappiah@yahoo.co.uk.

§Dept. of Statistics and Actuarial Science, KNUST, Kumasi, Ghana); emmaharris2002@yahoo.com.

[1]Received on March 14, 2024. Accepted on May 13, 2024. Published on June 1, 2024. DOI:10.23755/rm.v51i0.1581. ISSN: 1592-7415. eISSN: 2282-8214.

1 Introduction

The observed skyrocketing movements of gasoline prices in positive response to significant fluctuation of the international crude oil market on the Ghanaian downstream market remain disheartening compared to slow smooth return to normal or perhaps decrements of same when the upstream declines Bacon [1991]. As a fast growing middle-income economy, Ghana joined the league of oil producers known as the Organization of Petroleum Exporting Countries (OPEC), when it discovered in commercial quantities the black gold in 2007. With the expectation of general public burnt on the benefit such discovery will inure cushioning on prices, rather it seems ill-disposed toward price mechanism in pricing the ex-pump gas oil for the ordinary citizens. In this study, the aim is to investigate the existence of long-run equilibrium or mean-reversion using bivariate analysis of paired prices and test for linear and nonlinear threshold-type mean-reversion of bivariate relationship, lastly, is to estimate the coefficient parameters of (non)linear VECM and threshold parameter value and to compared the forecast performance accuracies of the SETAR to linear models of the mean-reversion process.

Modeling of energy prices transmission mechanism is as-much a daunting issue essentially when the dynamics of crude oil and gasoline nexuses in-printed on the domestic retail markets are scrutinized daily, having been described as awkward and mischievous process. The puzzles over the facade of the implementation of the deregulation of energy market policy amid an automatic adjustment formula for pricing mechanism, and policy of hedging rewards. With such rolled-out programs being held in thrall by authorities notwithstanding the reduction in prices of international crude oil, tantamount to obvious public slanging and a call for the abrupt abrogation of such policy implementations.

In preceding two decades of the 21st century, large empirical research studies such as the works of Borenstein [1997], and Peltzman [2000] have identified that several of such distinctly separated vertical markets (i.e prices for a commodity at different stages of processing) to depict irregularities in price transmission response due to input differentials associated with consumption locally from their characteristic global value of extraction and production. Meyer and von Cramon-Taubadel [2004] and Tappata [2009] referred to this pricing phenomena as asymmetric price transmission (APT), which is often relaxedly referred to as ”rocket and feathers” in crude oil and gasoline nexus.

The computational approach to unriddle the problem of this pricing phenomena was premiered by Engle [1987] traditional concept of cointegration that captures the existence of a stable long-run equilibrium relationship for any linear combinations of variables which were individually unit root. The concept has received extensive studies by works of Phillips and Ouliaris [1988], Stock and Watson [1988], and Johansen [1988]. The linear vector error correction model (VECM) technique

has proven to be an appropriate framework suitable for the modeling of symmetric mechanism in identifying the mean-reverting ratios that exist between two or more somewhat homogeneous variables. Unfortunately in the works of Huang et al. [2009], Natanelov et al. [2011] and Zhu et al. [2011] empirical investigations which incorporated the above linear technique to model price transmission from crude oil to gasoline unraveled the deficiency to explain for potential structural breaks or policy intervention acting as exogenous dummy variables inherent in the disequilibrium correction due to positive quick unusual peaking of the price and short-lived collapse readjustment just to soar in same direction.
The study of lack of symmetry in the phenomena of long-run equilibrium relationship was premiered by Tong and Lim [1980] and Balke [1997] for deviation of co-movements corrected within a band or percentage parameter. Their works permitted only threshold adjustment in the short-memory part if only the deviation from mean-reversion is above or below threshold specifications. Nonlinearities in cointegrated vector autoregressive models provides more supportive information about short-run price dynamics to external or internal shocks. Peltzman [2000] perceived such asymmetric response to imply a facetious potential gap in the price theory. Notwithstanding this perception, Bacon [1991] study of the U.K gasoline market and that of gal [2002] on Germany, France, UK, Italy, and Spain countries using an updated monthly data from 1985 to 2000 found evidence asymmetries in price adjustment between gasoline and crude oil. Likewise in the U.S market, Karrenbrock et al. [1991], Shin [1992], Borenstein [1997], Balke [1998], Brown and Yucel [2000], bor [2002], and Bachmeier [2003] all found indications of quick upward response in gasoline when crude oil prices are rising than when it fall. This thesis follows closely to model the hypothesis of asymmetric relationship between before-tax and after-tax premium gasoline prices with one another and to the pass-through effect of international crude oil prices.

1.1 The mechanism of petroleum products price build-up.

The technique in determining prices of all ex-pump products sold in the country commence on the 'ex-refinery differential' which computation comprise of import parity of European prices determined by the Oil and LPG Marketing Companies (OMCs or LPGMCs) based on the cost price received from Bulk Oil Distribution Companies(BDCs) for each window. Afterwards, the build-up to 'ex-pump' constitutes the aggregated ex-depot of levies and margins, some special petroleum tax imputed by the authority. Particularly, the authority regulates the prices through negotiation and review of operational margins and charges of the marketers and dealers which incorporates transaction, distribution and administration cost that ensures reasonable profit margins to the computation of the 'ex-pump'. Shortfalls that emerges out of these negotiation process are covered by

the government funding through cross-subsidies of specific products to cushion consumers.
In June 2001, the policy makers introduced the deregulation policy and mechanization of the price build-up computation to the Automatic Petroleum Product Pricing Formula (APPPF). This policy was to equivocate speculation and believes of pass-on of extra undue cost to consumer and haven been remarked as not only a panacea that eradicate the leverage of cross-subsidy intervention for funding shortfalls, but more readily designed at full-cost recovery to alleviate the fiscal burden of products subsidization on the part of government except for imperative difficult times. The proponents of automatic adjustment mechanism also indicated that 'ex-pump' computations as true reflection of trends of global price indicators. Taxes and levies undulate ex-pump products through petroleum-related funds and road-fund which differ from year-in to year-out based on the fiscal policy of governments. The sum of taxes and levies imposed on products constitute an average of 41.91% from 2000 to 2009 and 27.22% from 2010 to 2015 of the indicative maximum price for ex-pump. Negative taxes and margins represent huge subsidies on products such as kerosine and premix which target rural area development and LPG for residents in cities to mitigate the severe suffering on consumers.
Government revenue received from the sector is earmarked for development from it taxes and levies are in the range of 10 to 14 percent. Subsidy intervention by the government inure infrastructural development which money is defrayed to the petroleum providers. For instant, government owed TOR debt is presently criticized over fairness of magnitude interventions, implications on socio-economic indicators and delays in developmental project. The question of 'whether or not these intervention in either cases were realistic', can be traced to the effects of such measures above, and its ability to keep prices low for a long time in the face of hiking international prices.
The strand of the price build-up computation twisted with the policy of price stabilization and recovery levy known as the Petroleum Debt Service Charge (PDSS) which was effective March 2002. Incorporation of such policy to recoup loses as a result of subsidies intervention and services the TOR debts. It was set at 95% of any decline in Ghana cedis value of oil import cost from the average of the levels prevailing world market prices. Thus, taxing petroleum consumers to pay directly for the cost of past consumed subsidies accrued. These two policy frameworks: *APPPF* and *PDSS* are assumed to absorb government burdensome expenditure of crude oil import and transactional cost subsidy and ensure full cost recovery for the BDCs and OMCs. Moreover, they seek to remove public sector dominance and promote liberalized market to ensure adequate supply.

2 Related Works

The influence of symmetric mean-reversion was investigated by overcoming the restriction of the linear model in cointegrated serie Gregory [1996], Balke [1997], These studies revealed the presence of nonlinear effects and possibly deterministic causes with inherent structural fractures that affects changes in policy interventions.

Failure to immediate adjustment after small shocks in equilibrium error is believed to follow a threshold autoregressive (TAR) process, a specified threshold bandwidth with unit roots lying within this range. As an extension of the Tsay's Tsay [1998] univariate approach, Hansen-Seo Hansen and Seo [2002] described an bivariate case of the Threshold of Vector Error Correction Model (from VECM to TVECM) with both unknown cointegrating vector and threshold value. [3] However, if the adjustment behavior in the outer regimes is strong enough, the cointegration relationship is maintained in the longrun. A cointegration model with threshold adjustment was also proposed by Enders [2001]. Their models is restricted to a regime-specific coefficient of the first lag with threshold variable representing lagged equilibrium error in level series for self-exciting threshold autoregression(SETAR), or the first difference series for momentum threshold autoregression(MTAR) (see Enders [1998], Caner and Hansen [2001]).

The study into the regime-switching behavior of the bivariate cointegration or mean-reversion process permits extraction of a single paired price build-up mechanism of the Ghanaian gasoline markets application of threshold. The empirical research studies such as the works of Borenstein [1997], and Peltzman [2000] have identified that several of such distinctly separated vertical markets depict irregularities in price transmission response due to input differentials associated with consumption locally from their characteristic global value of extraction and production. Meyer and von Cramon-Taubadel [2004] and Tappata [2009] referred to this pricing phenomena as asymmetric price transmission (APT), which is often relaxedly referred to as "rocket and feathers" in crude oil and gasoline nexus.

Computationally approach to unriddle the problem of this pricing phenomena was premiered by Engle [1987] traditional concept of cointegration that captures the existence of a stable long-run equilibrium relationship for any linear combinations of variables which were individually unit root. The concept has received extensive studies by works of Phillips and Ouliaris [1988], Stock and Watson [1988], and Johansen [1988].

The study of lack of symmetry in the phenomena of long-run equilibrium relationship was premiered by Tong and Lim [1980] and Balke [1997] for deviation

[3] Seo [2006] discusses bootstrap tests of the null hypothesis of no cointegration in TVECM, while Gonzalo [2006] develop an asymptotic theory for testing the existence of a threshold effect in a VECM.

of co-movements corrected within a band or percentage parameter. Their works permitted only threshold adjustment in the short-memory part if only the deviation from mean-reversion is above or below threshold specifications.

In our empirical studies we have examined thoroughly the bivariate time series consisting of the monthly NIBOR rates of the maturities tomorrow next and 12 months. When modeling this bivariate time series, we find strong evidence for a two-regime TVECM being superior to a linear VECM, and in our out-of-sample forecasting the two-regime SETAR model gives much better prediction of the cointegration relation than a linear AR model. When testing a two-regime SETAR model for the cointegration relation against a three-regime model, the 2-regime model cannot be rejected at any reasonable significance level. In addition, we show how influential a few outliers may be by removing them from the time series and rerunning some of the statistical tests. Also, we have tested all the 66 possible pairs of Norwegian interest rates for cointegration, and we have tested the term spread of each pair for threshold effects, thus testing a linear model against a 2-regime model, as well as testing a 2-regime model against a three-regime model. We find a lot of cointegrated pairs, and we find evidence for a 2-regime model in approximately 50 % of the cases, and evidence for a three-regime model in some cases in this univariate time series analysis.

The term spread is modeled by a SETAR(3) model. When testing one regime against two regimes, the null hypothesis of one regime, i.e., a linear model, is strongly rejected, but when testing 2 regimes against 3 regimes, the null hypothesis of 2 regimes cannot be rejected at any reasonable significance level. Therefore, we choose a two-regime SETAR(3) model for the term spread. In the out-of-sample forecasting of the term spread, the SETAR(3) model with two regimes gives much better prediction of the term spread than the linear models.

In lieu of this, the paper set out to empirically model the asymmetric price build-up mechanism of the Ghanaian premium gasoline in nexuses with of Brent crude oil by investigating existence of long-run equilibrium or mean-reversion using bivariate analysis of paired prices; and to compared the forecast performance accuracies of the SETAR to linear models of the mean-reversion process to ARMA models.

The objective of this study is to empirically investigate asymmetric price build-up within the Ghanaian premium gasoline market. The question of whether the pricing mechanism of ex-refinery and ex-pump in nexuses with Brent crude oil mean-reverts in asymmetric manner. The study employs both Johansen and Juselius [1990] and Hansen and Seo [2002] techniques to test for the existence of long-run equilibrium and presence of threshold-type nonlinear effect in the disequilibrium correction term. The further estimation of coefficient parameters of the particular vector error correction model addresses the issue of percentage of adjustment speed and time elapsed to reoccur in mean-reversion. The techniques are lucid on the insidious spontaneous adjustment responds or contemporaneity, and the extent of lag-delayed market prices respond or spontaneity to both external and internal shock.

The rest of the study in this paper is structured as follows: Section 2 discusses briefly the results of some empirical investigation that employed techniques in gasoline market. Section 3 outlines the statistical models and misspecification test methods employed to unravel the linear and nonlinear dynamics of the multivariate time series analysis. Section four contains the results and interpretation of the empirical analysis of investigation into both symmetric or asymmetric price build-up mechanism engulfed in the Ghanaian petroleum market.

3 Methodology

Following that each level pairs of price build-up transmission obeys the Engle-Granger [Engle, 1987] and Johansen [Johansen, 1988] traditional concept of cointegration: $w_t = X_t\beta \sim I(0)$ for each $x_{kt} \sim I(1)$ $k = 1, 2$-dimensional or bivariate time series, and also exhibit presence nonlinear effects given the Hansen-Seo [Hansen and Seo, 2002] two step approach for 2R-TVECM(p) compact equation:

$$\begin{aligned} \triangle X_t = \quad & \alpha^{(1)}\beta' X_{t-1}.I_{\omega_{t-d}(\beta)\leq\gamma} & (1) \\ & + \sum_{i=1}^{p-1} \Gamma^{(1)} \triangle X_{t-i}.I_{\omega_{t-d}(\beta)\leq\gamma} + \epsilon_t^{(1)} & (2) \\ + \quad & \alpha^{(2)}\beta' X_{t-1}.I_{\omega_{t-d}(\beta)>\gamma} & (3) \\ + \quad & \sum_{i=1}^{p-1} \Gamma^{(2)} \triangle X_{t-i}.I_{\omega_{t-d}(\beta)>\gamma} + \epsilon_t^{(2)} & (4) \end{aligned}$$

where α-coefficient of the loading matrix, β-cointegrating vectors of the coefficient matrix of the long-run equilibrium impact $\Pi = \alpha\beta'$, and $\Gamma^{(m)}$-accumulating $p-1$ lagged coefficient matrix of the short-run impact of the level representation for an $m = 2(or3)$-regime switches defined for $I_{\{.\}}$ is a dummy of the heavy-side indicator function:

$$I_{\omega_{t-d}(\beta)} = \begin{cases} 1, & \texttt{for} \quad \omega_{t-d}(\beta) \leq \gamma^{(1)} \\ 0, & \texttt{for} \quad \gamma^{(2)} < \omega_{t-d}(\beta) \end{cases} \tag{5}$$

and ϵ_t is a vector error such that $\epsilon_t \sim IN(0, \Sigma)$. This discrete linear models representation is similar to Tsay's Tsay [1998] generalization of Self-Exciting AutoregRessive (m-SETAR(p)), m-SETAR(p) model due to discontinuity or structural breaks in policy implementation.

Seo [2006] suggests that the cointegration relation of a TVECM is could be modeled with univariate time series function such as SETAR model for either $m = 2(or3)$-regimes, otherwise expressed as linear model. The sections below describes the expression for possible (non)linear univariate time series functions of the cointegrating relations, $w_t = X_t\beta$.

3.1 Test and estimation for linear or ARIMA models of the cointegrating process

From the Box and Jenkins [1994] procedure for fitting an autoregressive moving average denoted as ARMA(p, q) models of the random spread process is given as:

$$(1 - \sum_{j=1}^{p} \phi_j L^j)\omega_t = \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \epsilon_t \tag{6}$$

which follows that the residual series, $\epsilon_t \sim i.i.dN(0, \sigma^2)$ of uncorrelated white noise such that $\phi_j \neq \theta_j$ and $\phi_j, \theta_j \neq 0$ satisfying the stationary and linear stochastic difference property above. ARMA process has both causal and invertible properties. Thus, a restriction of the lag $q = 0$ gives an ARMA($p, 0$) or AR(p) model expressed in the form:

$$\omega_t - \sum_{j=1}^{p} \phi_i L^j \omega_t = \epsilon_t \tag{7}$$

such that $\phi_i \neq 0$, $\phi_i < 1$ and $\phi_0 < 1$ is causal process. The Markov process is AR(1) of the form $\omega_t = \phi_1 \omega_{t-1} + \epsilon_t$. Likewise, if $p = 0$ leads to a restricted Moving Average process, ARMA($0, q$) or MA(q) model in the form:

$$\omega_t = \sum_{j=0}^{q} \theta_j L^j \epsilon_t \tag{8}$$

such that $\theta_i \neq 0$, $\theta_i < 1$, and $\theta_0 = 1$ is invertible process. The MA(∞) representation has unique solution

$$\omega_t = \frac{\sum_{j=0}^{q} \theta_j L^j}{1 - \sum_{j=1}^{p} \phi_j L^j} \epsilon_{t-j} \tag{9}$$

such that $\mu = 0$ and

$$(1 - \sum_{j=1}^{p} \phi_j L^j)^{-1} = 1 + \sum_{j=1}^{p} \phi_j L^j < \infty$$

if the characteristic root of the polynomial $1 - \sum_{j=1}^{p} \phi_j(z_j) \neq 0$ lie outside the unit circle, $|z| \leq 1$. The restrictions also apply.

However, for $\omega_t \sim I(1)$ non-stationary or unit root process may take several interesting deterministic components, D_t in form:

$$\omega_t = \sum_{j=0}^{\infty} \phi^j (\epsilon_{t-j} + D_{t-j}) \tag{10}$$

such that $\sum_{j=0}^{\infty} \phi^j \leq \infty$ with $E[\omega_t] = 0$ and $Var[\omega_t] = \dfrac{\delta_\epsilon^2}{1-\phi}$. Thus, ϕ_ϵ does not decrease exponentially with ACF and PACF of ω_t persistently varying over time.

The deterministic components of integrated processes may be identified as D_t=zero for purely random walk process, $\triangle\omega_t = \epsilon_t$, or with constant drift $\triangle\omega_t = \delta + \epsilon_t$. D_t could also stem from either a linear trend-stationary $D_t = \delta + t$ or quadratic trend process respectively, $\delta + t^2$. An innovation dummy with an indicator function

$$|D_t| = \begin{cases} 1 & \text{for } D_t \geq 0 \\ 0 & \text{for } D_t < 0 \end{cases} \tag{11}$$

occurring due to some policy intervention causing break instability with an event phenomenon or perhaps seasonal component.

The effect of D_t's in ω_t is that $\rho \geq 1$ with roots $z \geq 1$ trapped within the complex unit circle without unique stationary solution for certain unstable D_t cumulating behavior or trend embedded in ω_t. The unit root process expressed in form: $\phi^*_{p-1}(L)(1-L)^d\omega_t = \phi^*_{p-1}(L)\triangle^d\omega_t = \theta(L)\epsilon_t$ denoted as ARMA($p-d, q$) or ARIMA(p, d, q) of stationary covariance process with $\phi^*_{p-1}(L)$ roots avoiding the unit circle. The process of differencing, $\triangle^d\omega_t \sim I(d)$ with no deterministic term is said to be *integrated process* of order d stationary with $\epsilon_t \sim i.i.d, N(0, \sigma^2)$ expressed as:

$$\triangle^d\omega_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j} \tag{12}$$

such that $\psi_j(z) = \sum_{j=0}^{\infty} \psi_j z^j \leq \infty$ with $\psi_0 = 1$ convergent at $|z| \leq 1 + \delta$ for $\delta > 0$. [4] Also, detrending or removing deterministic trend to achieve some filtered series of fluctuations around zero-horizontal line. By definition ω_t is trend-stationary if $\omega_t = a_j t + z_t$ where a_j are constants for $j = 0, 1, \ldots$; while z_t is a stationary series such that $\omega_t = z_t$ is level-stationary if $a_j = 0$.

[4]Practically, after the second differencing, $I(d = 2)$, it is assumed that $\triangle^{d=1,2}\omega_t$ is stationary with exponentially decaying tail.

3.2 Estimation of the Parameters of Reduced form VECM

From the work of Johansen [1995] developed an estimate parameters of reduced form VECM using the Gaussian maximum likelihood (ML) procedure. The procedure is a maximum likelihood frame defined on the VECM given by

Removing the unrestricted deterministic variables, Δ_t and lagged differences and possibly additional non-modeled variables through a priori regression. For available T and p sample size and presample values respectively, we denote $\triangle X = [\triangle x_1, \ldots, \triangle x_T]$, $X_{-1} = [x_0, \ldots, x_{T-1}]$, $\varepsilon = [\epsilon_t, \ldots, \epsilon_T]$, and $Y = [X_0, \ldots, X_{T-1}]$ with $Y_{t-1} = [\triangle x'_{t-1}, \ldots, \triangle x'_{t-p+1}]'$ and the parameters, $\mathit{\Gamma} = [\Gamma_1, \ldots, \Gamma_{p-1}, \Psi_0]$; then compact matrix of VECM becomes

$$\triangle X = \alpha\beta' X_{-1} + \mathit{\Gamma} Y + \varepsilon \tag{13}$$

Assume estimates for matrix coefficient $\alpha\beta'$ of long-run part are known, we concentrating on the short-run dynamic adjustment effect using ordinary least square (OLS) estimator for each equation of the system $\hat{\mathit{\Gamma}} = (\triangle X - \alpha\beta' X_{-1})Y'(YY')^{-1}$. By substituting $\hat{\mathit{\Gamma}}$ in the above

$$\triangle X = \alpha\beta' X_{-1} + (\triangle X - \alpha\beta' X_{-1})Y'(YY')^{-1}Y + \varepsilon \tag{14}$$

$$\triangle X(I - Y'(YY')^{-1}Y) = \alpha\beta' X_{-1}(I - Y'(YY')^{-1}Y) + \hat{\varepsilon} \tag{15}$$

$$\triangle XM = \alpha\beta' X_{-}M + \hat{\varepsilon}; \qquad \texttt{for} \quad M = I - Y'(YY')^{-1}Y \tag{16}$$

Now $\hat{\alpha}$ and $\hat{\beta}$ can be obtained from Johansen [1995] reduced rank regression procedure based on the residual, $\hat{\varepsilon}$ above by defining a product of second moment matrix $S_{ij} = T^{-1} \triangle X_i M \triangle X_j'$ for $i, j = 0, 1$ such that:

$$S_{00} = T^{-1} \triangle XM \triangle X', \quad S_{01} = T^{-1} \triangle XMX_{-}', \quad S_{11} = T^{-1} X_{-}MX_{-}' \tag{17}$$

for solving a generalized eigenproblem of the matrix $\lambda S_{11}V = S_{01}' S_{00}^{-1} S_{01}V$ with S_{11} and $S_{01}' S_{00}^{-1} S_{01}$ assumed to be symmetric and positive definite. First assume β is known, α can be estimated using OLS. Then substitute $\hat{\alpha}$ into the log-likelihood function which only is now a function of only β which is tantamount to reduce-rank regression to solve the k eigenvalue, λ_i for $0 \leq \lambda_i \leq 1$

$$|\lambda_i S_{11} - S_{01}' S_{00}^{-1} S_{01}| = 0 \tag{18}$$

such that $\hat{\lambda}_1 \geq \ldots \geq \hat{\lambda}_k \geq 0$ and the remaining $k_1 - k$ eigenvalue of the cofeature equal to zero. and it associated eigenvector, b_i of $V = [b_1, \ldots, b_k]$ of normalized

$$\lambda_i S_{11} b_i = S_{01}' S_{00}^{-1} S_{01} b_i \tag{19}$$

such that $V'S_{11}V = I_K$. If estimator of β is obtained from the eigenvectors $\hat{\beta} = [b_1, \ldots, b_r] \in b_k$ becomes the true value MLE $\hat{\beta}$ of β, then true value of $\hat{\alpha}$ will be given by $\hat{\alpha} = \alpha(\hat{\beta})$ below:

$$\hat{\alpha} = \triangle XMX'_{-1}\hat{\beta}(\hat{\beta}'X_{-1}MX'_{-1}\hat{\beta})^{-1} \tag{20}$$

and finally the short-run parameter has a feasible estimator

$$\hat{\Gamma} = (\triangle X - \hat{\alpha}\hat{\beta}'X_{-1})X'(XX')^{-1} \tag{21}$$

The consistency and joint asymptotic normal are follow a more general assumptions than the Gaussianity condition.

3.2.1 Testing for cointegration and the estimation of VECM

We employ two kinds of tests to verify presence of long-run movement in series namely: the Enders [1998] two-step method; and the Johansen [1988] maximum likelihood ratio test for cointegration. Before we perform the estimate of the restricted parameters of the reduce form VECM using Johansen [1988] ML method in Appendix A1, the first assumption must be met first. Thus, the Johansen procedure jointly estimate for the $\Pi = \alpha\beta'$ (but not separating for only α or β); the rank of Π, the most significant r co-integrating vectors within β; and hence α.[8]

The hypothesis test for cointegration rank is more defined under the null H_0: $rank = k$ of full rank for $x_t \sim I(0)$ against an alternative H_a: $rank = r_0 < r \leq k$. This may be compared to the test statistic based on likelihood ratio (LR) test for finding r-cointegrating vectors

$$\Lambda_{LR}(H_a|H_0) = -TIn[(1-\lambda_{r+1})\ldots(1-\lambda_k)] = -T\sum_{i=r+1}^{k} In(1-\lambda_i) \tag{22}$$

where $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_k$. The LR test discriminates empirically between the large eigenvalues, λ_i, for $i = 1, 2, \ldots, r$ corresponding to stationary $\bar{\beta}'X_t \sim CI(1,1)$ and smaller roots, λ_i, for $i = r+1, r+2, \ldots, k$. LR is maximized by solving the generalized eigenvalue problem of the matrix

$$|\lambda S_{11} - S_{10}S_{00}^{-1}S_{01}| \tag{23}$$

[8]In practice, it obligatory that a stable system have $|\alpha_k| = [0, 1)$ for all k-dimension, since α explains the rate of the equilibrium error excluded for each variable movement for each period, t. The system may possibly diverge if α_k lies outside the boundary.

where $S_{ij} = T^{-1}Y_i^{'}Y_j$, for $i, j = 0, 1$. The significance of the r largest canonical correlation for the reduced ranked of the Π matrix, which is established by the ML estimator of β is further confirmed by Johansen [1988]'s two different LR test for the cointegrating ranks.

3.3 Threshold Vector Error Correction Model (TVECM)

A m-regime TVECM for time series data, X_{it} of lag length p is generalized as

$$\triangle X_t = \alpha^{(\tau)}\beta^{'}X_{t-1} + \sum_{i=1}^{p-1}\Gamma_i^{(\tau)} \triangle X_{t-i} + \epsilon_t^{(\tau)} \quad ; \quad \gamma^{(\tau-1)} \leq \beta^{'}X_{t-d} \leq \gamma^{\tau} \tag{24}$$

Thus, if the cointegrating vector representing the threshold variable at d lag-delay, $\beta^{'}X_{t-d}$ is bounded and closed within some transition bandwidths of value, γ such that $\gamma^{(m-1)} \leq \beta^{'}X_{t-d} \leq \gamma^{(m)}$ for $\tau = 1, \dots, m$ regimes of $m-1$ threshold effects. The notation $\Gamma_i^{(m)}$ is m different matrices of short-run coefficient lagged differences of each regime, and $\Pi = \alpha^{(m)}\beta'$ is also m different matrices of long-run lagged level coefficients, and $\epsilon_t^{(m)}$ is m different innovation disturbance such that each $\epsilon_t \sim i.i.d.(0, \Sigma)$.
The threshold effect is active if only $0 < P(\beta^{'}X_{t-d} \leq \gamma) < 1$, otherwise TVECM reduce to a regular VECM with $\Pi^{(1)} = \Pi^{(2)}$. Thus, allowing for 'threshold level stationary' equilibrium adjustments of linearly combine variables in different regimes. [9] The estimation of multivariate TVECM developed by Tsay [1998] pose dimensionality problem in executing it, so we limit our analysis for $\tau = 1, m = 2$ regime of one threshold effect, TAR(1)

$$\triangle X_t = \alpha^{(1)}\beta^{'}X_{t-1}.I_{(\beta^{'}x_{t-d} \leq \gamma)} + \alpha^{(2)}\beta^{'}X_{t-1}.I_{(\beta^{'}x_{t-d} > \gamma)}$$
$$+ \sum_{i=1}^{p-1}\Gamma^{(1)} \triangle X_{t-i}.I_{(\beta^{'}x_{t-d} \leq \gamma)} \qquad + \sum_{i=1}^{p-1}\Gamma^{(2)} \triangle X_{t-i}.I_{(\beta^{'}x_{t-d} > \gamma)} + \epsilon_t$$

where $I_{\{.\}}$ is a dummy of the heavy-side indicator function given by

$$I_{\{.\}} = \begin{cases} 1 & \text{for} \quad \beta^{'}X_{t-d} \leq \gamma \\ 0 & \text{for} \quad \beta^{'}X_{t-d} > \gamma \end{cases} \tag{25}$$

defines the different regimes encompassing both past changes in variables, $\triangle X_{it-p}$ and past deviation from equilibrating vector, $\beta^{'}X_{t-1}$ likely to induce varying effects on current changes. This discrete adjustment explains that deviation from equilibrium relationship would not exist for system exceeding critical threshold

[9] It obvious that TVECM has more than 1 linear equilibrium correction relation.

at which point in time the benefits of adjustment are higher than the costs where economic agents move the system back to equilibrium. Thus, it captures the 'deepness' of asymmetric responses of the equilibrium error correction.
An alternative adjustment of the threshold specification is momentum TAR (M-TAR) model allows a variable to display differing amounts of AR decay depending on whether it is increasing or decreasing in $\beta' X_{t-d}$

$$I_{\{.\}} = \begin{cases} 1 & \text{for} \quad \triangle\,\hat{\varepsilon}_{t-d} \leq \gamma \\ 0 & \text{for} \quad \triangle\,\hat{\varepsilon}_{t-d} > \gamma \end{cases} \tag{26}$$

the regime switch is initiated by the lagged changes in $\beta' X_{t-d}$. M-TAR describes the 'steepness' of the variation below or above the threshold value, γ for which the variable x_t follows two different stepwise stochastic processes.
In compact notation we obtain

$$\triangle X_t = \begin{cases} A_1' X_{t-1}(\beta) + \epsilon_t & \text{for} \quad \omega_{t-1}(\beta) \leq \gamma \\ A_2' X_{t-1}(\beta) + \epsilon_t & \text{for} \quad \omega_{t-1}(\beta) > \gamma \end{cases} \tag{27}$$

where $X_{t-1}(\beta) = (1, \omega_{t-1}(\beta), \triangle x_{t-i})'$ is a $K \times 1$ regressors, $A = [A_1 \quad A_2]$ is $k \times 2 = (2p+2) \times 2$ coefficient matrices which influences regime dynamics.
In an empirical estimation of the parameter $(\beta, A_1, A_2, \Sigma)$ TVECM, we may impose some form of constrains on the coefficients $A = [A_1 \quad A_2]$ to be permanently fixed but allows only the constant and residual deviation to switch across regimes. The probability of existence of a threshold effect on the error correction is constrained in a balanced parameter, π_0 such that $\pi_0 < P(\omega_{t-1} \leq \gamma) < 1 - \pi_0$. The ML estimation of the TVECM which is given in appendix - A2 satisfies the assumption that the errors in each regime, $\epsilon_{it} \sim i.i.d$'s and the combined is vector martingale difference sequence that has a finite covariance matrix, $\Sigma = [\epsilon_t \epsilon_t']$. Thus, the normalization imposed on β and the threshold restriction band $\pi_0 \leq \frac{1}{n}\sum_{t=1}^{n} 1(x_t'\beta) \leq 1 - \pi_0$ is minimized with each regime possessing such minimum fraction, p_0 of the entire sample.

3.4 Estimation of the Parameters of Reduced form TVECM

From the technicalities of the Hansen and Seo [2002], and Hansen [2005], we rewrite the TVECM of the compact form as

$$\triangle x_t = M_1' X_{t-1}(\beta).d_{1t} + M_1' X_{t-1}(\beta).d_{2t} + \epsilon_t \tag{28}$$

where $d_{1t} = I_{\omega_{t-1}(\beta) \leq \gamma}$ and $d_{2t} = I_{\omega_{t-1}(\beta) > \gamma}$ and $\epsilon_t \sim i.i.d$ Gaussian applying a normalization on β. They proposed the quasi-MLE of the parameters $(\beta, M_1, M_2, \Sigma)$

denoted by $(\tilde{\beta}, \tilde{M_1}, \tilde{M_2}, \tilde{\Sigma})$ from the residual vector defined on the parameters of estimate$(\beta, M_1, M_2, \Sigma, \gamma)$ will be given as
$\tilde{\epsilon}_t(\beta, M_1, M_2, \Sigma, \gamma) = \triangle x_t - \tilde{M}' X_{t-1}(\tilde{\beta}) = \triangle x_{t-1} - seoM_1' X_{t-1}(\beta) d_1(\beta, \gamma) - M_2' X_{t-1}(\beta) d_2(\beta, \gamma)$. Then the Gaussian likelihood of the residual is

$$InL(\beta, M_1, M_2, \Sigma, \gamma) = -\frac{n}{2} \log |\Sigma|$$
$$-\frac{1}{2} \sum_{t=1}^{n} (\beta, M_1, M_2, \Sigma, \gamma)' \Sigma^{-1} \epsilon_t(\beta, M_1, M_2, \Sigma, \gamma)$$

Estimation of the ML on the above $InL(\beta, M_1, M_2, \Sigma, \gamma)$ focus partially first on the (M_1, M_2, Σ) parameters by holding fixed (β, γ). This yields estimates of the coefficient matrices, $\hat{M}_1(\beta, \gamma)$ and $\hat{M}_2(\beta, \gamma)$, which are obtained by conditional sequential LS regression of $\triangle x_t$ on $X_{t-1}(\beta)$ for each regime subsamples given below

$$\hat{M}_1(\beta, \gamma) = (\sum_{t=1}^{n} X_{t-1}(\beta) X_{t-1}(\beta)' d_1(\beta, \gamma))' (\sum_{t=1}^{n} X_{t-1}(\beta) \triangle x_{t-1}' d_1(\beta, \gamma)) \tag{29}$$

$$\text{for} \quad \omega_{t-1} \leq \gamma \tag{30}$$

$$\hat{M}_2(\beta, \gamma) = (\sum_{t=1}^{n} X_{t-1}(\beta) X_{t-1}(\beta)' d_2(\beta, \gamma))' (\sum_{t=1}^{n} X_{t-1}(\beta) \triangle x_{t-1}' d_2(\beta, \gamma)) \tag{31}$$

$$\text{for} \quad \omega_{t-1} > \gamma \tag{32}$$

Substituting into the residual vector and covariance matrix defined on (β, γ) given as

$$\hat{\epsilon}_t(\beta, \gamma) = \epsilon_t(\beta, \hat{M}_1'(\beta, \gamma), \hat{M}_2'(\beta, \gamma), \gamma) \text{and} \tag{33}$$

$$\hat{\Sigma}(\beta, \gamma) = \frac{1}{n} \sum_{t=1}^{n} \hat{\epsilon}_t(\beta, \gamma) \hat{\epsilon}_t(\beta, \gamma)' \tag{34}$$

By so doing it produces a concentrated Gaussian likelihood function

$$L_n(\beta, \gamma) = (\beta, \hat{M}_1'(\beta, \gamma), \hat{M}_2'(\beta, \gamma), \Sigma(\beta, \gamma), \gamma) \tag{35}$$

$$= -\frac{n}{2} \log |\hat{\Sigma}(\beta, \gamma)| - \frac{np}{2} \tag{36}$$

The MLE of the $L_n(\beta, \gamma)$ is seen to act as a minimizer of $\log |\hat{\Sigma}(\beta, \gamma)|$ subject to the normalization imposed of β discussed in earlier in previous paragraph and the constraint imposed within the probability of threshold effect of the trim parameter, π_0

$$\pi_0 \leq n^{-1} \sum_{t=1}^{n} 1(x_t^{'}\beta \leq \gamma) \leq 1 - \pi_0 \tag{37}$$

The MLE solution for M_1 and M_2 are $\hat{M}_1 = \hat{M}_1(\beta, \gamma)$ and $\hat{M}_1 = \hat{M}_1(\beta, \gamma)$

3.5 The m-SETAR(p) of the cointegrating rletionship

Following Enders & Granger Enders [1998] generalization of the Dickey & Fuller Dickey [1979] of long-run process to assume the TAR(m, p) consisting of m-AR(p) parts with variation occurring within an *'embedding dimension'* of $-\infty = \gamma_0 < \ldots < \gamma_\tau = \infty$ for $\tau = 1, \ldots, m$-regimes as one AR(p) process switches to another according to the specific value of a threshold parameter, γ_τ

$$\omega_t = (\sum_{j=1}^{p} \phi_j \omega_{t-j} + \sum_{j=1}^{p} \sigma_{\tau j} \epsilon_t) I_\tau\{\gamma_{j-1} \leq z_{t-d} < \gamma_j\} \tag{38}$$

where $I_\tau\{\gamma_{j-1} \leq z_{t-d} < \gamma_j\}$ is the heavy-side indicator function that determines the observed or partially unobserved properties under different or discontinuous decay nonlinear AR(p) states as series crosses the threshold value to the next $m-1$-regime that quantify the magnitude(speed) of decay adjustment to new trend path. The term $\sum_{j=1}^{p} \sigma_{\tau,i} \epsilon_t$ allows different error terms variances across regimes.

Tong Tong [1983] shows that the self-exciting threshold autoregressive (SETAR(m)) is a special kind of TAR(m, p) with a dynamic threshold variable, $z_{t-1} \approx z_{t-p-d-1}$ where $d \in \tau$ is threshold time-delay well-defined in threshold range, $\gamma_\tau \in [\gamma_1, \gamma_m]$. TAR(2) or SETAR(1) is a simplified piece-wise linear approximation of AR($p - d - 1$) expressed in general form as

$$\omega_t = \begin{cases} \phi_{10} + \sum_{j=1}^{p} \phi_{1j} \omega_{t-j} + \sigma_1 \epsilon_t & \text{for } z_t \leq \gamma_1 \\ \phi_{20} + \sum_{j=1}^{p} \phi_{2j} \omega_{t-j} + \sigma_2 \epsilon_t & \text{for } \gamma_1 < z_t \leq \gamma_2 \end{cases} \tag{39}$$

where ϕ_i are the constant speed parameters.

Based on an ordinary least square estimation of TAR-model, we employ the Hansen [1997] and Hansen and Johansen [1999] to test a H_0 : linearity 1-regime model or TAR(1) against $H_a : m > 1$-regime model, TAR(m) subject to the $\sup_{F-type}$ test statistics of the nonstandard F_{jk} distribution

$$F_{jk} = T \frac{SSR_j - SSR_k}{SSR_k} \quad ; \quad j < k \tag{40}$$

where SSR_j the minimum sum of squared residuals when fitting a j-regime SETAR against a k-regime SETAR model. A special estimation of the H_0 model is the SSR_1. The threshold value for k-regime which minimizes the SSR_k is computed by performing a grid search conditional on the value of the previous threshold that minimizes SSR_j. Under the H_0, vigorous re-sampling of the residuals is done using replicating bootstrap estimation of the threshold parameters and computation F_{jk}.

The residuals analysis of the models are

$$\hat{\epsilon}_t = \omega_t - \{\hat{\phi_{10}} - \sum_{j=1}^{p} \hat{\phi}_{1j}\omega_{t-j}\}I_{\omega_{t-d}\leq\hat{\gamma}} - \{\hat{\phi_{20}} - \sum_{j=1}^{p} \hat{\phi}_{2j}\omega_{t-j}\}I_{\omega_{t-d}>\hat{\gamma}} \quad (41)$$

The study employs the following performance criteria to compare the best Root mean square error

$$RMSE = \sqrt{\frac{\sum_{t=1}^{T}(\omega_t - f_t)^2}{T}} \quad (42)$$

Mean absolute error

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|\omega_t - f_t| \quad (43)$$

Mean absolute percentage error

$$MAPE = 100 \times \frac{1}{n}\sum_{t=1}^{n}|\frac{\omega_t - f_t}{\omega_t}| \quad (44)$$

where f_t and x_t denote the forecast and actual values at t period respectively.

The RMSE and MAE are a measure of fit, which indicates how well model fits with the historical data. It should be noted that unlike the MAPE which is scale invariant, both RMSE and MAE criteria rely on scale of the dependent variable. This require that the criteria are implemented as relative measures in forecast comparison for similar series across varying models. A smaller value indicates stronger forecasting power for the respective model.

4 Analysis and Discussion of Results

4.1 Data Assimilation and Exploration

Since cointegrating relationship (ω_t) in a 2(3)-regime TVECM could be specified as 2(3)-regime TAR model, so we proceed to estimate the ω_t from our APT models. Each computed vector of ω_t extraction comprises of 431 biweekly samples. An initial step to modeling these nonlinear price adjustment relationship is to understand graphically. Figure 1 displays pattern of the estimated mean-reversion at each stage. It indicates undoubtedly almost non-stationary strong upward trend for both initial and overall stages with respect to Brent crude oil in nexuses with

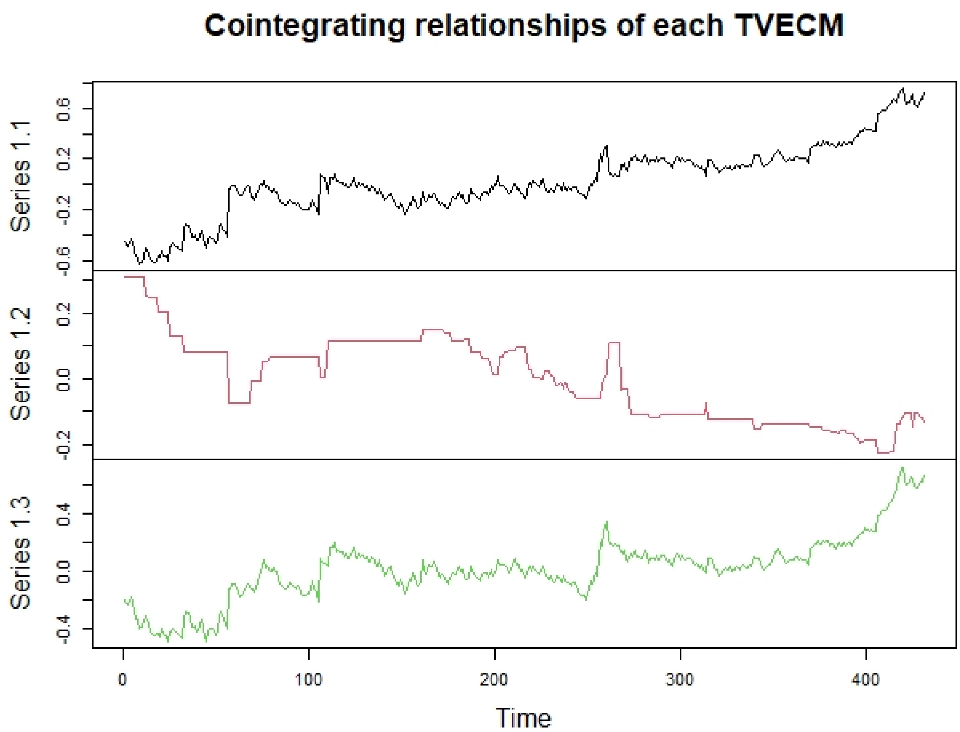

Figure 1: Estimated cointegrating relationship (ω_t) in a TVECM at each stage of price transmission 02/01/1999 - 27/06/2015.

both ex-pump and ex-refinery respectively, whereas the second stage show steady downward trend between paired localized co-movement in the sense of BDCs and OMCs. Figure 2 depicts the empirical densities for all three estimated cointegrating relationships in a TVECM.

4.2 Test for 2-regime against 3-regime TAR model

The study further employs the Hansen's testHansen [1999] procedure to confirm the linearity of the cointegrating relationships extracted from the TVECM with 1000 bootstrap replication and trimmed off at 10% sample observations. Table 1 and Figure 3-5 indicate the test results and plots respectively, for H_0 of linear

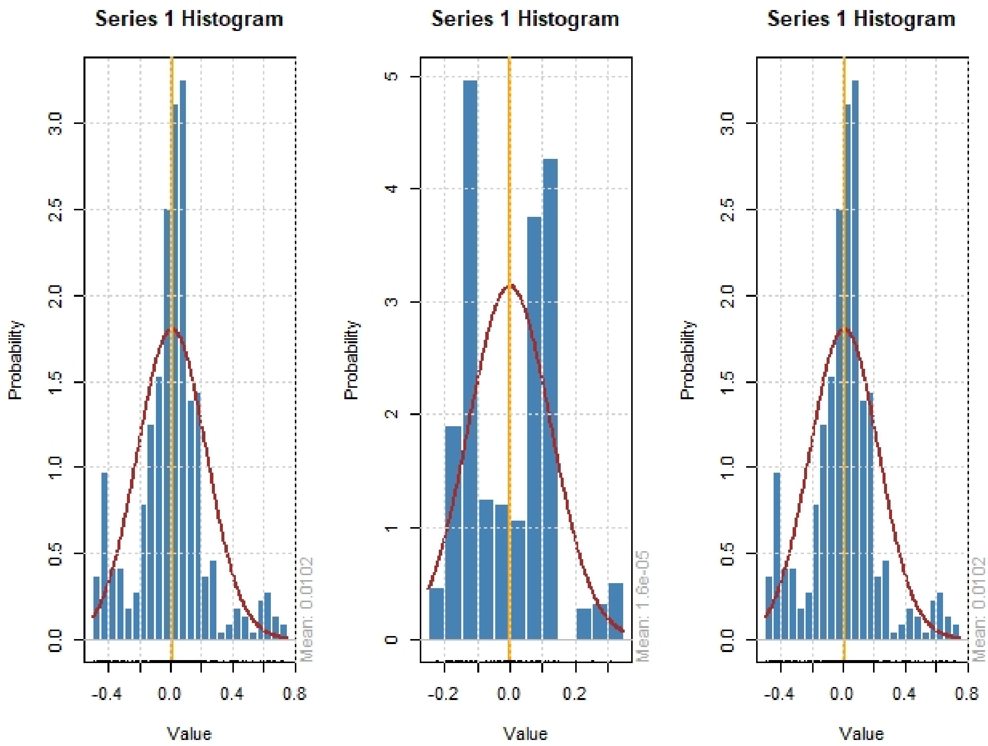

Figure 2: Plot of the empirical densities for estimated cointegrating relationship (ω_t) in a TVECM at each stage of price transmission 02/01/1999 - 27/06/2015.

cointegration relationship against an alternative of 2-regime threshold-type non-linear effect, with option for testing 2-regime versus 3-regime threshold-type AR model at each stage. We observe that $H_{0_1 vs2}$ of 1-regime tested against 2-regime is rejected at 1% and 5% significance levels for both initial and overall stages respectively. Whereas, on the contrary, the test for $H_{0_2 vs3}$ of 2-regime versus 3-regime cannot be rejected at any significance level for any of the modeling stage.

4.3 Modeling of SETAR model

As an initial model checking procedure, the study identifies appropriate SETAR models by employing the graphical diagnostic plots of the sample ACF and PACF of the residuals addressing degree of correlation between neighboring observations, and then test for normality which facilitates to represent ARMA(p, d) modeling. A forecasting exercise of the term spread for each level paired PBU from their respective TVECMs following that of Di Narzo Di Narzo [2008] procedure of out-of-sample forecasting of the crack spread

Table 1: Hansen's Test for 1vs2, 1vs3 and/or 2vs3-regime TAR of the contegrating relationship at each stage of APT

	SETAR Test					
	Test statistics		**Critical Value**			
	F-test	p_v	90%	95%	97.5%	99%
Stage 1:						
H_{1vs2}	40.86	0.02	24.82	31.12	37.93	46.85
H_{1vs3}	50.41	0.09	49.12	58.28	66.24	74.84
H_{2vs3}	8.71	0.85	20.91	25.92	30.92	40.26
Stage 2:						
H_{1vs2}	9.32	0.71	25.65	33.31	43.42	55.02
H_{1vs3}	31.27	0.48	62.10	74.61	87.77	103.41
H_{2vs3}	21.48	0.28	34.60	44.96	55.85	71.07
Overall:						
H_{1vs2}	27.42	0.05	22.97	27.12	30.59	40.14
H_{1vs3}	43.99	0.09	42.65	49.32	55.16	66.41
H_{2vs3}	15.57	0.38	22.24	26.72	30.13	34.24

Note:
sup*LM* is the supremum of the heteroskedastic-consistent Lagrange Multiplier test statistics, and p_{value} is the reasonable significant level for the reject of H_0.

The term spread is modeled by a SETAR(3) model. When testing one regime against two regimes, the null hypothesis of one regime, i.e., a linear model, is strongly rejected, but when testing 2 regimes against 3 regimes, the null hypothesis of 2 regimes cannot be rejected at any reasonable significance level. Therefore, we choose a two-regime SETAR(3) model for the term spread. In the out-of-sample forecasting of the term spread, the SETAR(3) model with two regimes gives much better prediction of the term spread than the linear models.

4.4 Forecast Comparison for ARMA and SETAR of Mean Reversion in Price Transmission

Further estimation of coefficients of the linear and nonlinear mean-reversion in the univariate sense are made based on the result of Hansen and Johansen [1999] method which test parameter constancy for number of regimes in TAR model. The model considers only 2-regime threshold for price transmission at initial and the overall stages. A search order of linear model using the Box and Jenkins [1994]

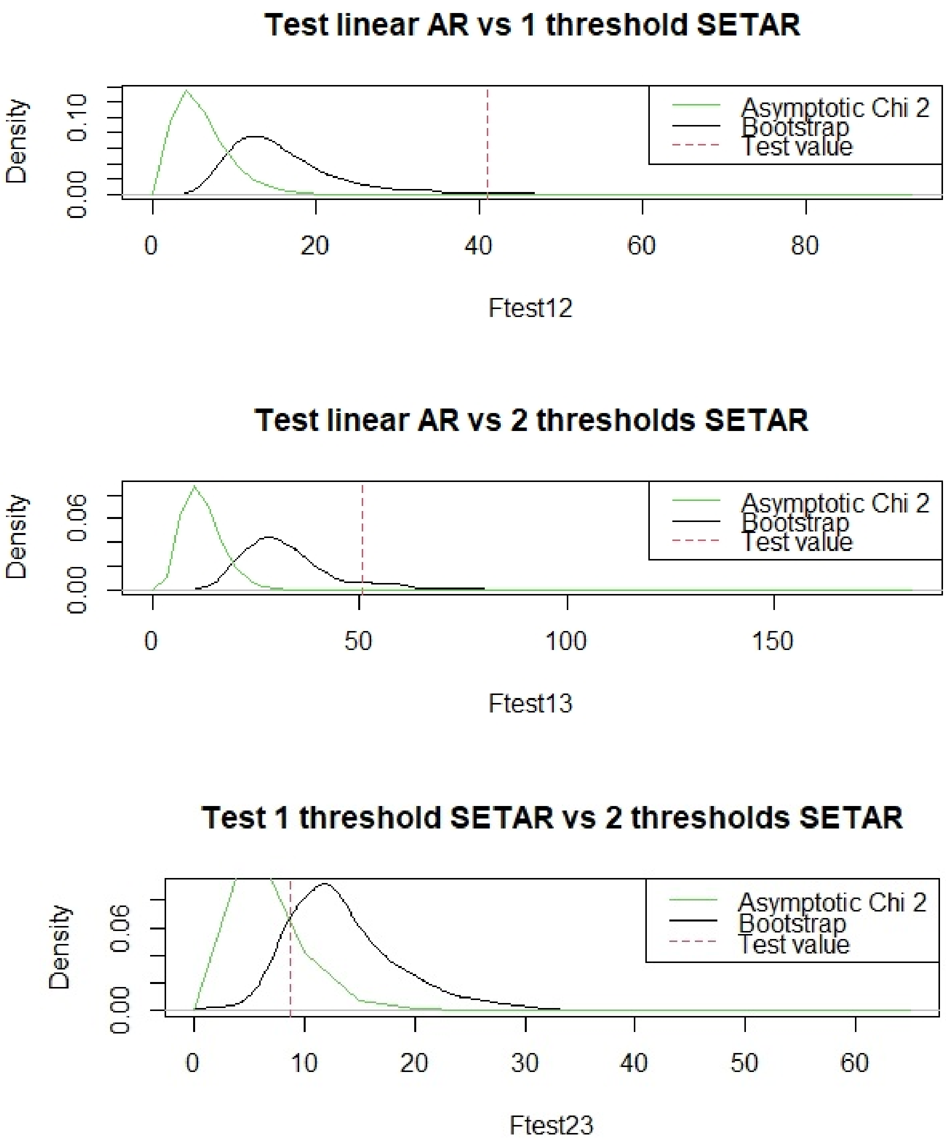

Figure 3: Biweekly Cointegrating relationship of the spread or mean-reversion at each stage of price transmission 02/01/1999 - 27/06/2015.

test of autocorrelation were conducted and concluded on the lag orders using ACF and PACF for both ARMA(p,q) and 2-SETAR(m,d) as in Larsen [2012]. Figure 6 - 8 represent the plot of volatile residuals, ACF and PACF for SETAR of ARMA model for cointegrating series of price transmission stage. The estimated ACFs and PACFs points to a SETAR of embedded dimensional orders, $m = 5$ and

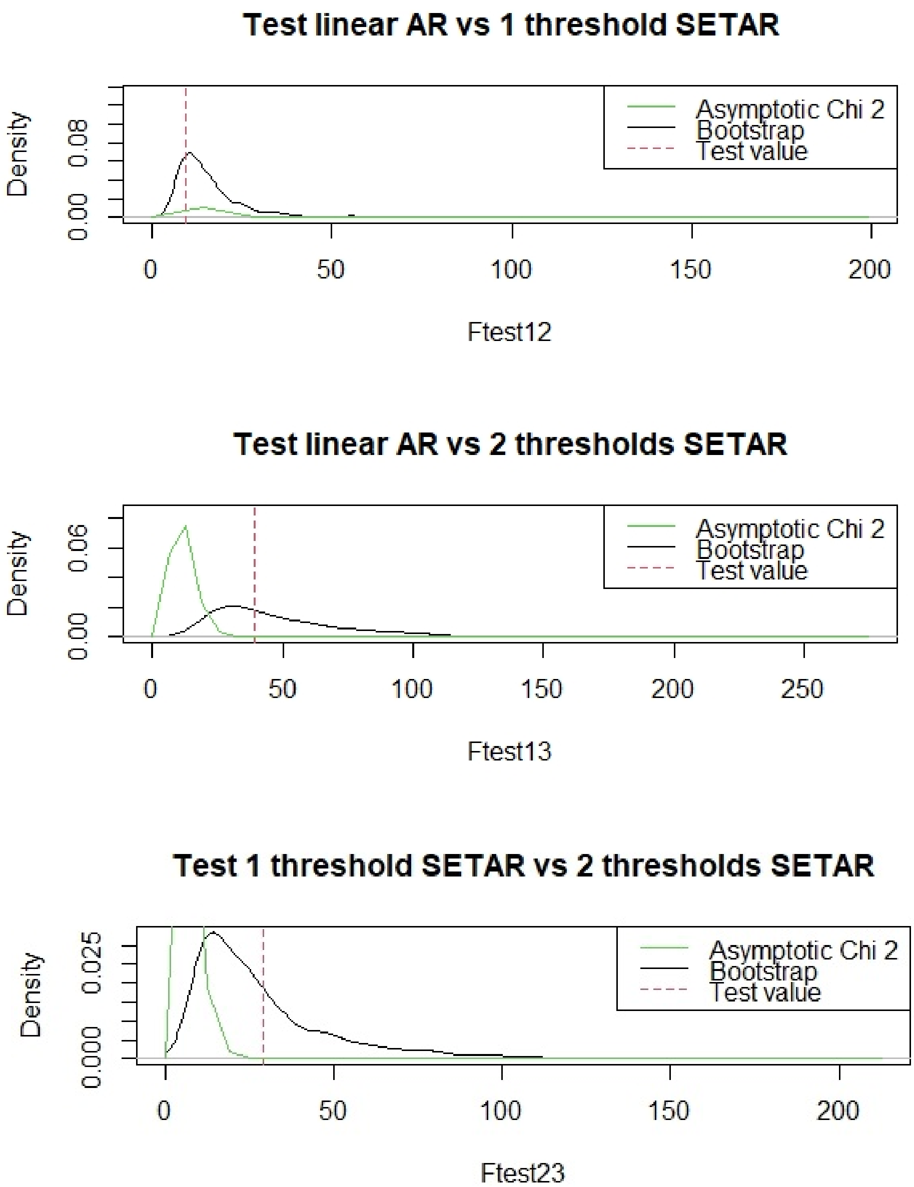

Figure 4: Biweekly Cointegrating relationship of the spread or mean-reversion at each stage of price transmission 02/01/1999 - 27/06/2015.

$m = 4$ for stage one and overall stage transmission respectively in Figure 8 & 10. Also from Figure 9, the estimated ACF and PACF confirmed a linear cointegrating relationship, as it points to ARMA of lag order $p = 4$ and $q = 2$ as the most suitable to estimate the second stage .

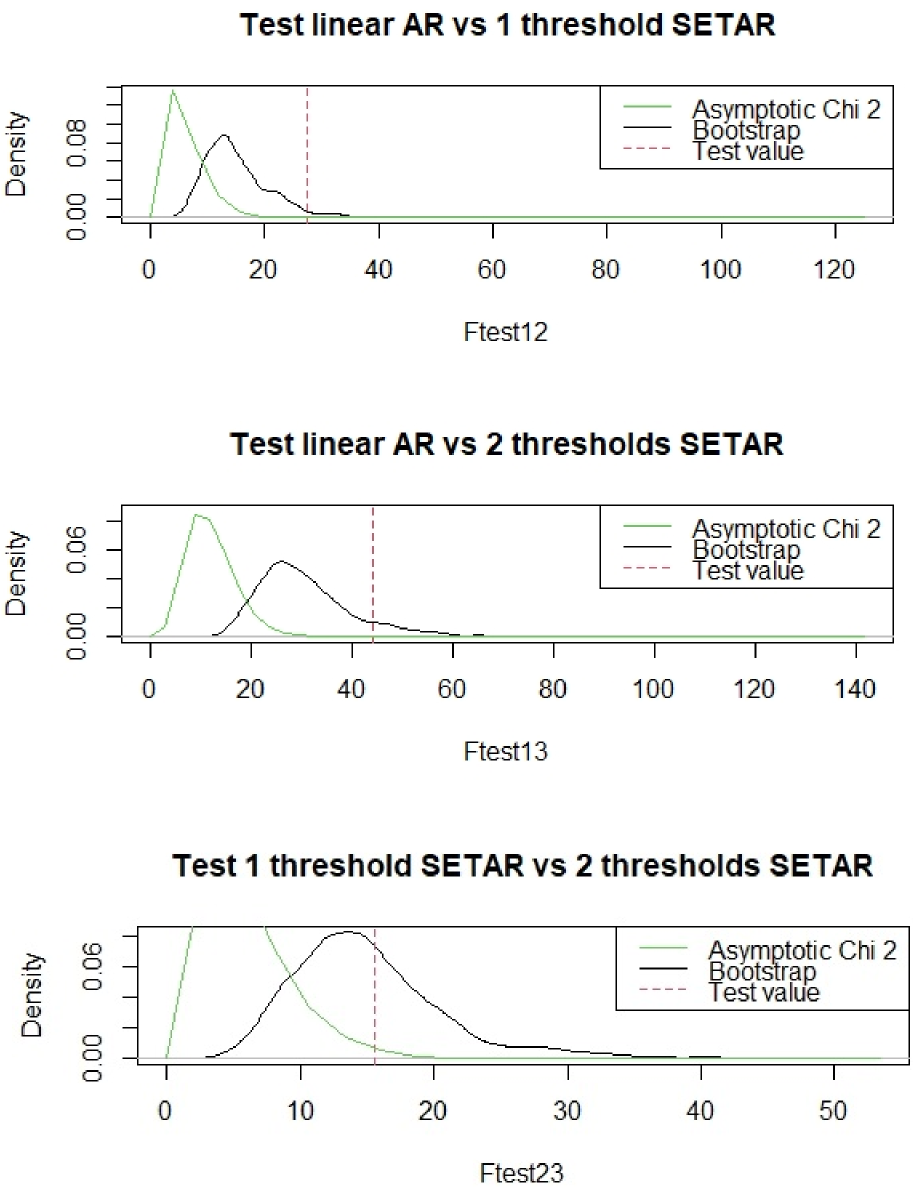

Figure 5: Biweekly Cointegrating relationship of the spread or mean-reversion at each stage of price transmission 02/01/1999 - 27/06/2015.

The results of long-run adjustment equation for each stage of price transmission displayed in Table 2 to Table 3 show the estimated coefficients and the vari-

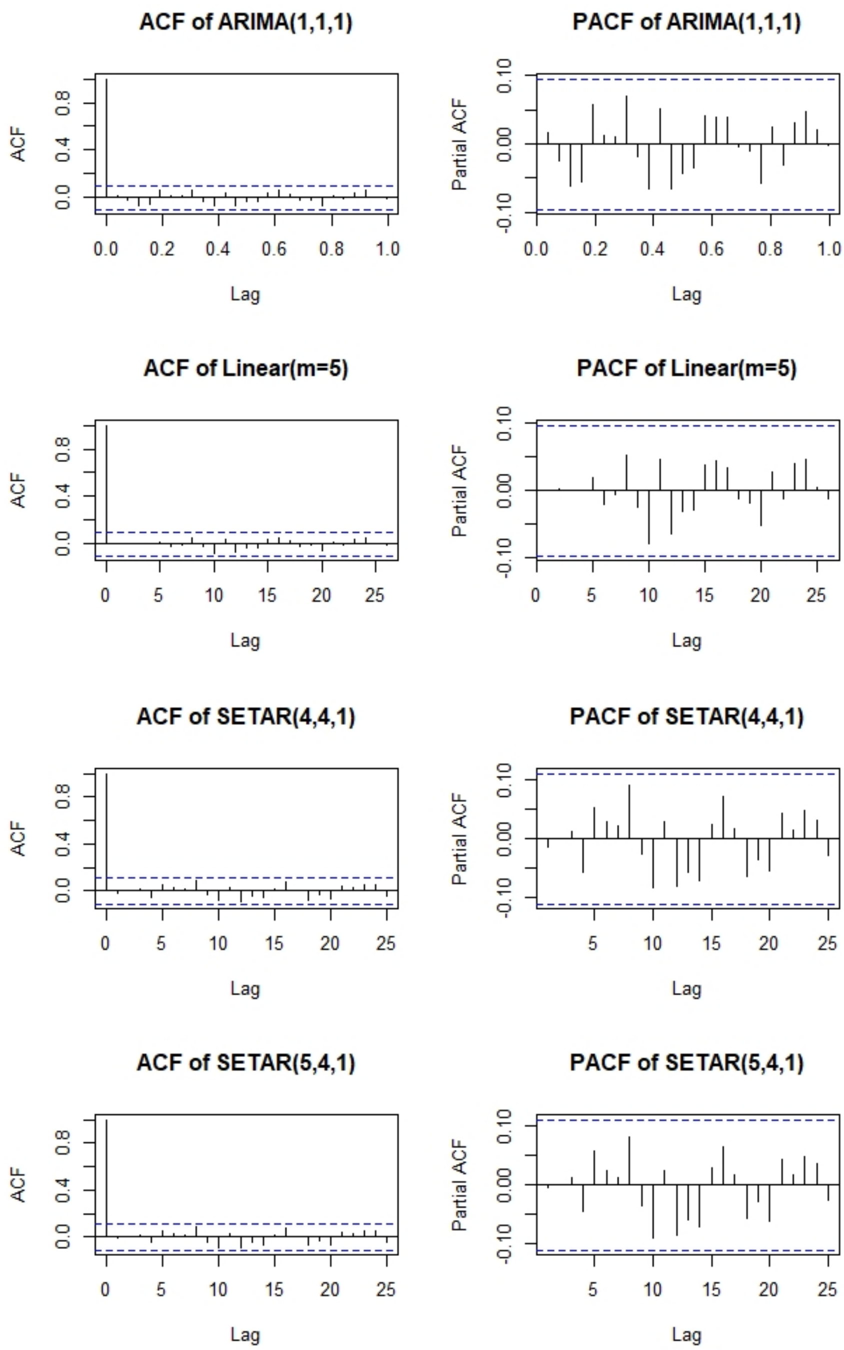

Figure 6: Plot of estimated ACF and PACF functions for the residuals of fitted AR(7), ARIMA(1,1,1), SETAR(5,5,1),STAR(11), NNET(4), AAR(4) in stage 1 PT mean-reversion

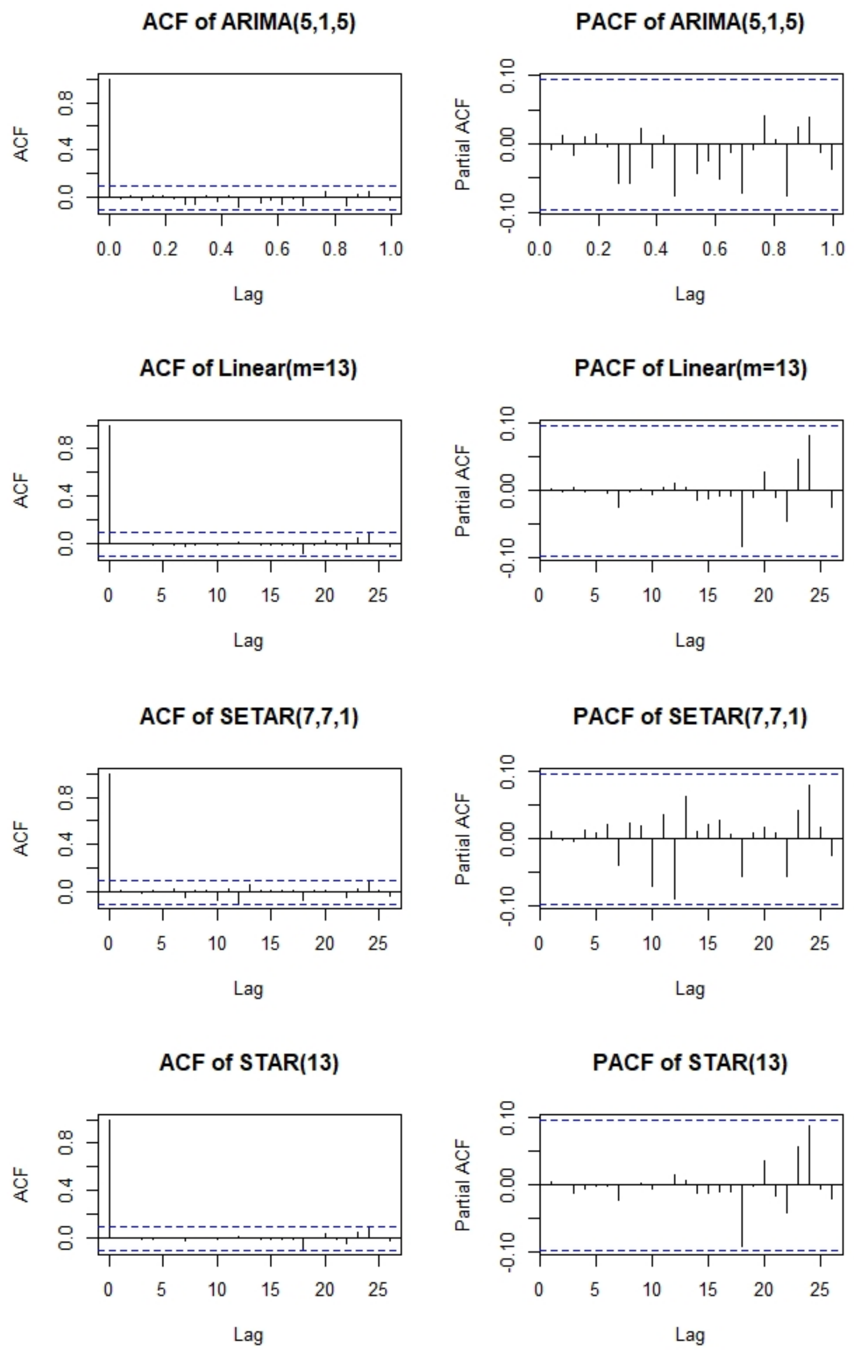

Figure 7: Plot of estimated ACF and PACF functions for the residuals of fitted AR(13), ARIMA(1,1,1), SETAR(7,7,1),STAR(13), NNET(13), AAR(13) in stage 1 PT mean-reversion

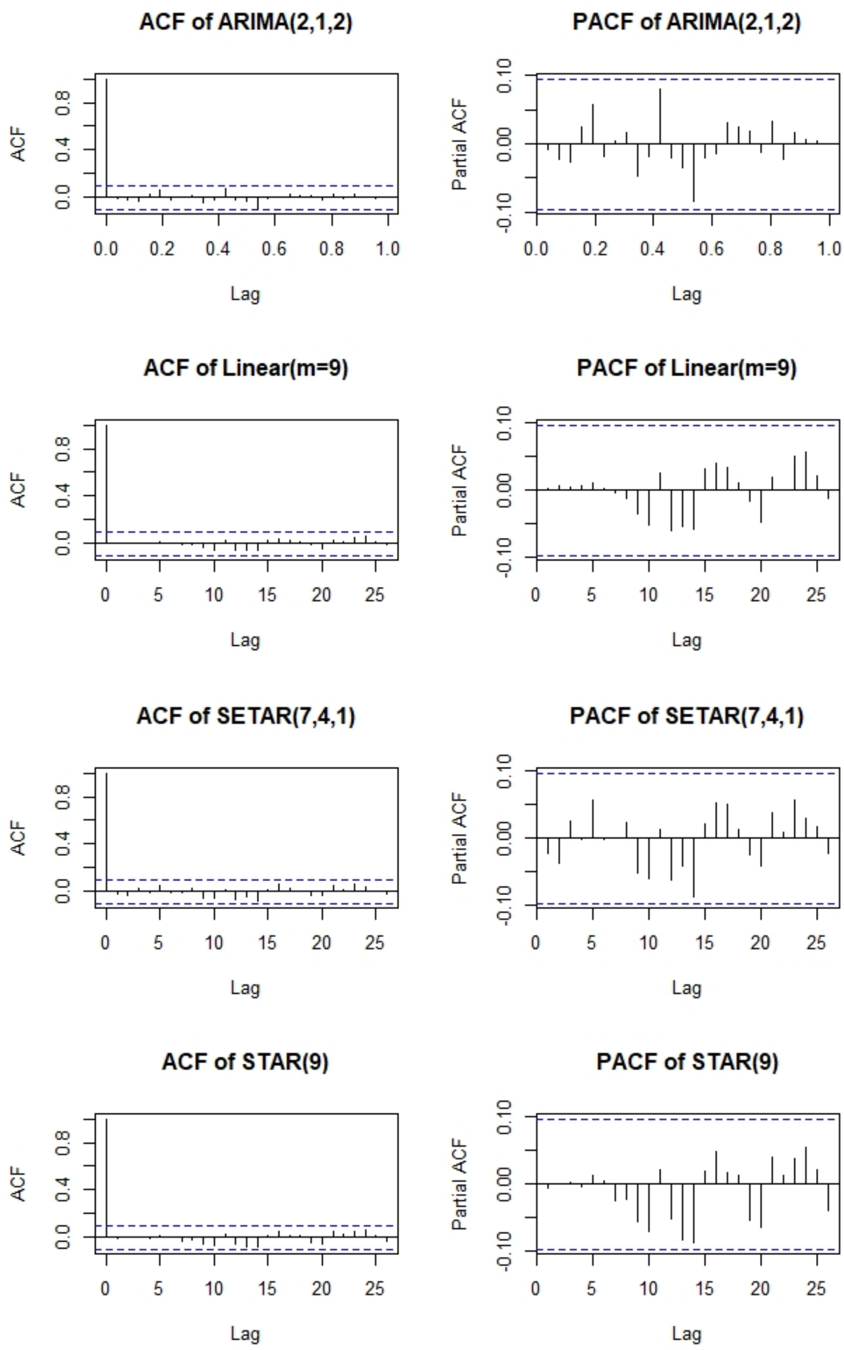

Figure 8: Plot of estimated ACF and PACF functions for the residuals of fitted AR(13), ARIMA(1,1,1), SETAR(7,4,1), STAR(9), NNET(13), AAR(13) in stage 1 PT mean-reversion

ance of fitted residual with it mean absolute prediction error and as displayed on Figures 11-14. The SETAR models are evaluated at the values of maximum threshold parameter attained using the Threshold Vector Error Correction Method. It is observe that the 2-regime SETAR model of mean-reversion for transmission from p_{co} to p_{exr} is best modeled at a maximum autoregressive order for both low and high regimes of embedding dimension $m = 5$ with the low regime order,$mL = 4$ and the high regime, $mH = 1$ to yield the minimum AIC and MAPE. The equation is significant for both constant and trend deterministic part in both regimes.
The linear cointegrating relationship of the transmission from p_{xrg} to p_{xpg} is modeled by ARMA of order $p = 4$ and $q = 2$ at the minimum AIC and MAPE, hence stand out to be best selection for the models for forecast.

Table 2: Estimated coefficients of 2-regime SETAR(m) models at each PT stage of the price build-up for Stage 1.

Equations		
Stage 1: 2-SETAR(5,4,1)	**Regime 1** $\omega_t \leq -0.1619$ (16.83% Obs.)	**Regime 2** $\omega_t > -0.1619$ (83.17% Obs.)
Intercept	- 4.2054e-02 (4.4317e-02)	- 1.8447e-02* (7.9242e-03)
Trend	4.2942e-04. (2.4712e-04)	9.2684e-05* (3.8082e-05)
ω_{t-1}:	8.3411e-01*** (1.2774e-01)	9.6617e-01*** (2.1341e-02)
ω_{t-2}:	- 1.9984e-02 (1.8212e-01)	
ω_{t-3}:	- 4.9994e-01** (1.7976e-01)	
ω_{t-4}:	6.0509e-01*** (1.2198e-01)	
Fitt. resid. variance: 0.001643	AIC: - 2673	MAPE: 87%

Note: ***(or **,*,.) denotes the 0.1% (or 1%, 5%, 10%) level of significant for rejecting H_0.

Table 3: Estimated coefficients of 2-regime SETAR(m) models at each PT stage of the price build-up for Stage 2.

Equations		
Stage 2: 2-SETAR(7,7,1)	**Regime 1 (52.06%)** $\omega_t \leq 0.01423$	**Regime 2(47.94%)** $\omega_t > 0.01423$
Intercept	1.2955e-02**	5.1566e-05
	(4.2977e-03)	(4.2908e-03)
Trend	- 3.9997e-05.	- 8.7269e-06
	(2.0805e-05)	(1.9142e-05)
ω_{t-1}	1.0279e+00***	9.7487e-01***
	(6.7593e-02)	(2.4246e-02)
ω_{t-2}	1.5027e-02	
	(8.6038e-02)	
ω_{t-3}	7.3650e-02	
	(8.5459e-02)	
ω_{t-4}	- 1.8934e-01*	
	(8.5242e-02)	
ω_{t-5}	2.2093e-01*	
	(8.9211e-02)	
ω_{t-6}	- 9.5453e-02	
	(9.2777e-02)	
ω_{t-7}	- 6.3025e-02	
	(6.7460e-02)	
Fitt. resid. variance: 0.0002829 AIC: - 3406 MAPE: 28.45%		
Overall: 2-SETAR(7,4,1)	$\omega \leq -0.08925$ (23.97% Obs.)	$\omega > -0.08925$ (76.03% Obs.)
Intercept :	- 1.6853e-02	- 1.1776e-02.
	(1.4534e-02)	(6.2736e-03)
Trend :	1.4458e-04.	4.9765e-05.
	(7.6403e-05)	(2.6596e-05)
ω_{t-1}:	9.1403e-01***	9.9883e-01***
	(8.9703e-02)	(2.0604e-02)
ω_{t-2}:	8.6990e-03	
	(1.2746e-01)	
ω_{t-3}:	- 5.1775e-01***	
	(1.4559e-01)	
ω_{t-4}:	5.4928e-01***	
	(1.0792e-01)	
Fitt. resid. variance: 0.001438 AIC: - 2729 MAPE: 121%		

Note: ***(or **,*) denotes 0.1% (or 1%, 5%, 10%) level of sig. for rejecting H_0.

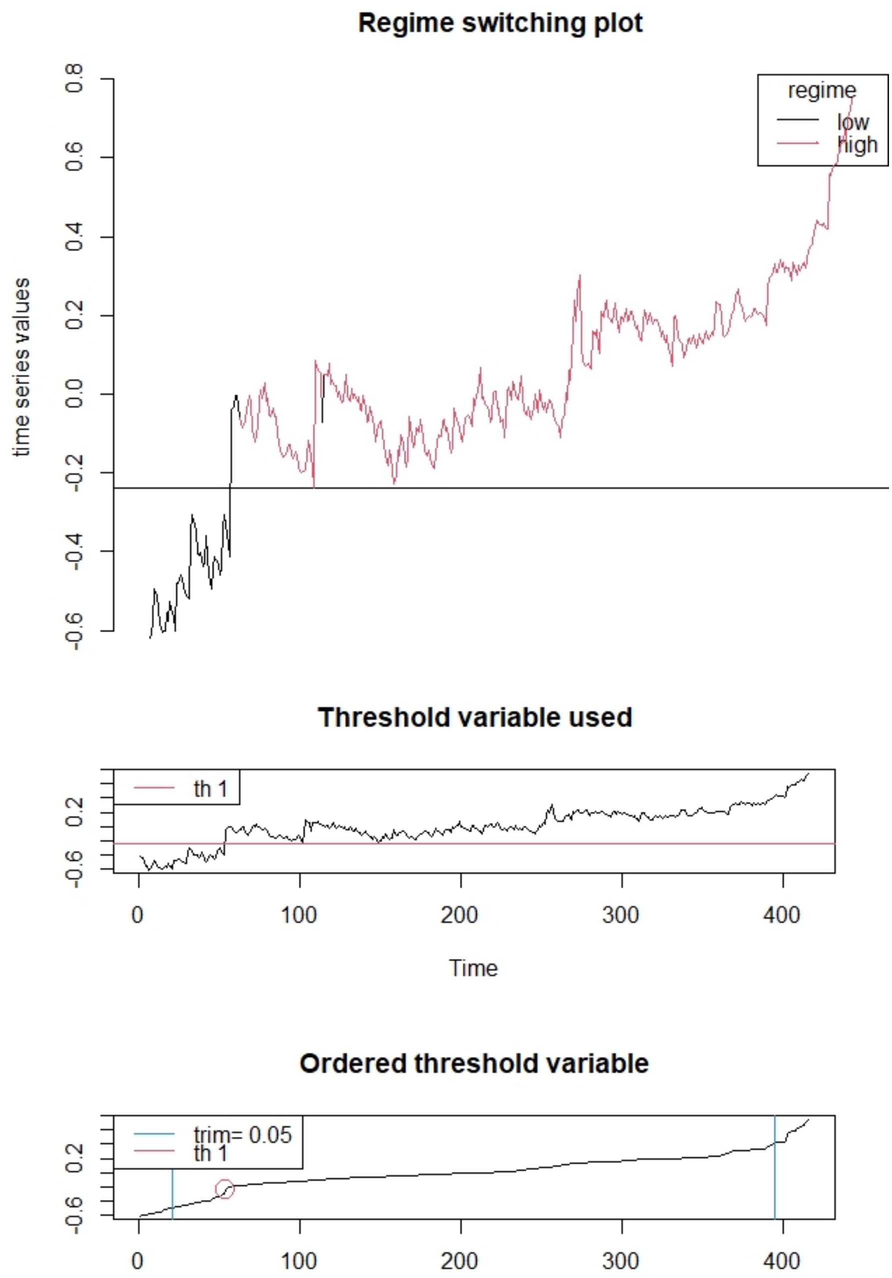

Figure 9: Plot of the regime switching level and threshold variable of the mean-reversion process in initial price transmission between crude oil and ex-refinery from 02/01/1999 - 27/06/2015.

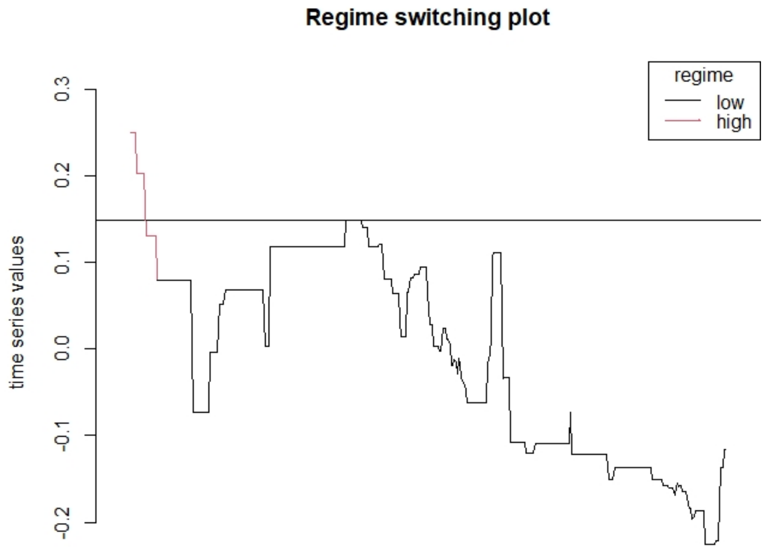

Figure 10: Plot of the regime switching level and threshold variable of the mean-reversion process in initial price transmission between ex-refinery and ex-pump from 02/01/1999 - 27/06/2015.

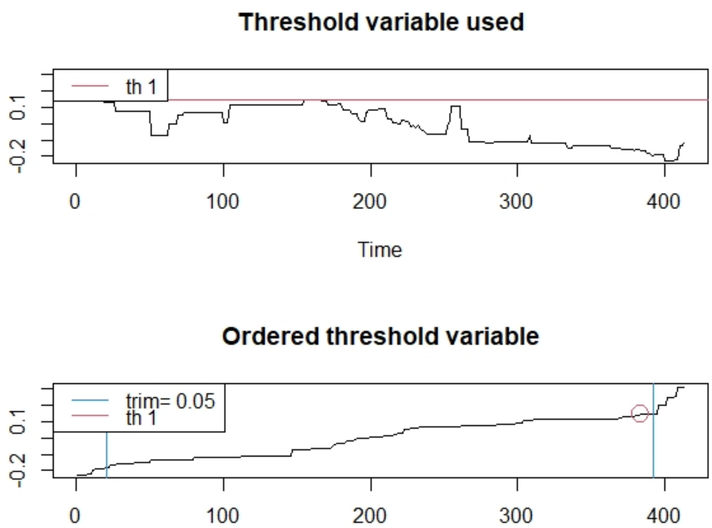

Figure 11: Ordered Threshold Plot of the regime switching level and threshold variable of the mean-reversion process in initial price transmission between ex-refinery and ex-pump from 02/01/1999 - 27/06/2015.

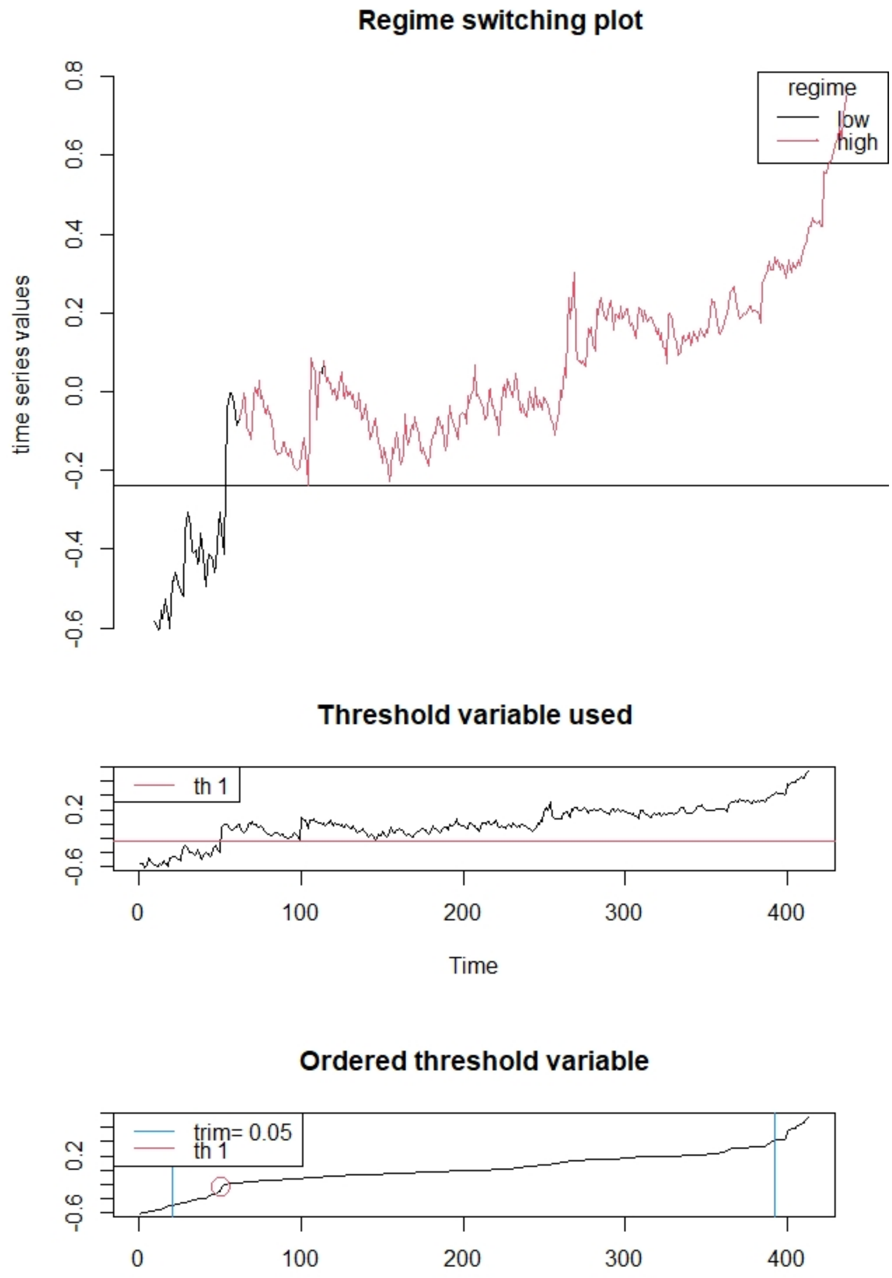

Figure 12: Plot of the regime switching level and threshold variable of the mean-reversion process in the overall PT between crude oil and ex-pump from 02/01/1999 - 27/06/2015.

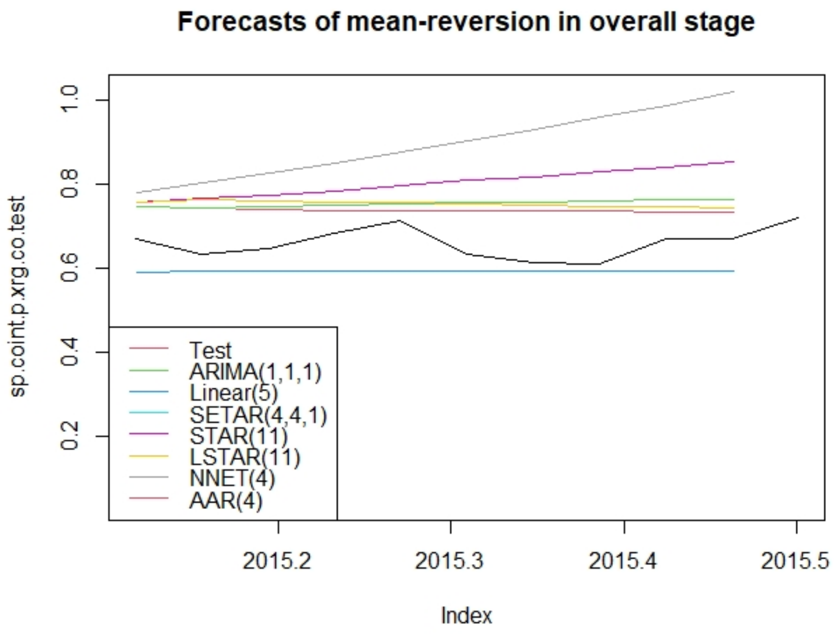

Figure 13: Plot of forecast comparison for nonlinear mean-reversion at stage 1

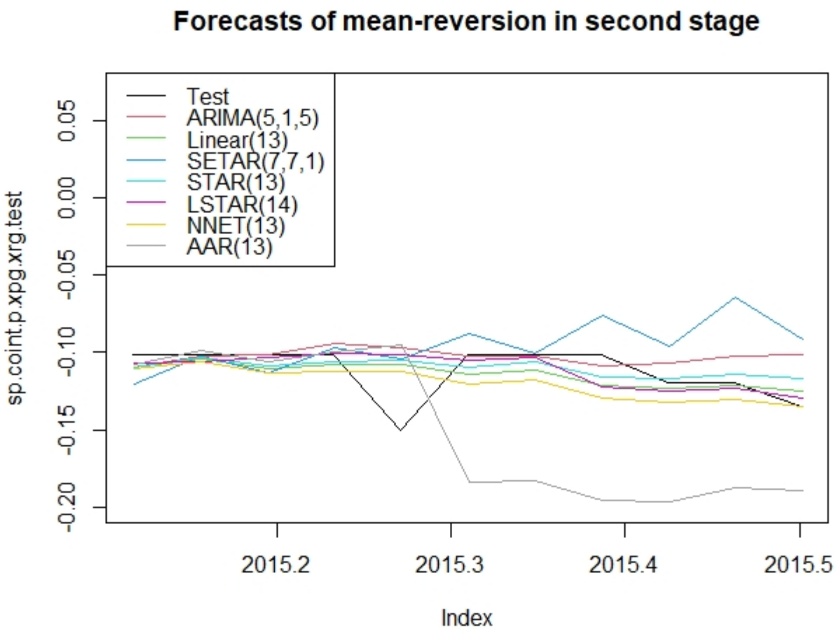

Figure 14: Plot of forecast comparison for nonlinear mean-reversion at stage 2

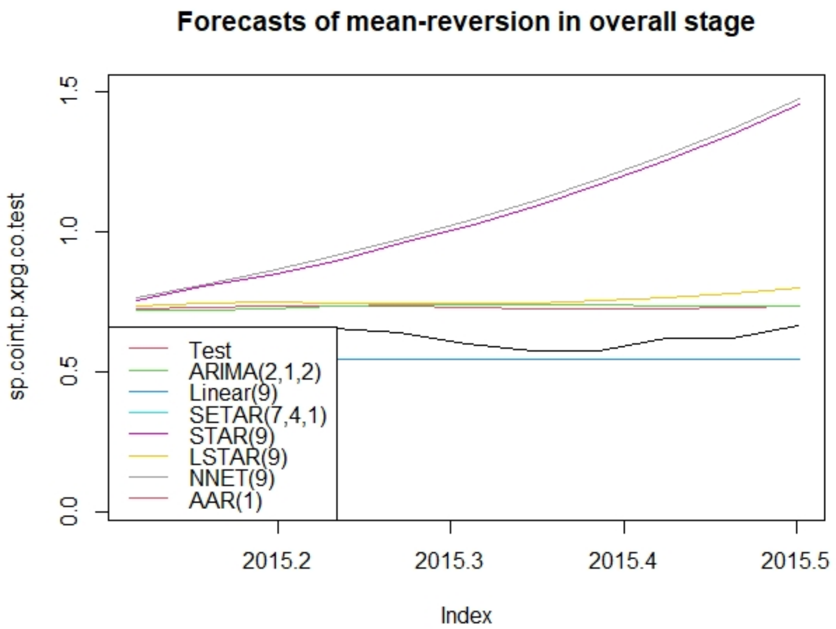

Figure 15: Plot of forecast comparison for nonlinear mean-reversion at overall stage

4.5 Evidence of Asymmetries price Transmission in Three Different Stages of Bivariate Time Series between Crude oil, ex-refinery and ex-pump Gasoline Prices in the Ghanaian petroleum market

The summary statistics of the mean return and variance coupled with graph of the indicates unstable or volatile residuals of the times series to random walk process. This observation enable the test for stationarity of the series using both the Dickey [1979] and Kwiatkowski et al. [1992] methods which indicate unit root processes $I(1)$ with significant trends and drifts deterministic forms in all three energy prices.
The choice of a suitable univariate technique took a naive approach to test linearity within the data. However, Hansen and Johansen [1999] tests for numbers of one regime AR versus two or three TAR model for univariate analysis pointed to threshold model of the initial and overall stage transmission and rather model the second stage with linear model. The major findings pertaining to the study, objectives and hypothesis are in all three stages of the price transmission of the price build-up formula are outlaid as follows:

The claim of presence of nonlinearities in terms of threshold effects is accepted evident from Hansen and Seo test under H_0 of linear cointegration. Additionally, the Hansen and Johansen test for 1-regime versus 2-regime and 2-regime versus 3-regime of the estimated univariate equations show that the initial price transmission is best characterized by two regimes.

More importantly, although the half-life of both p_{co} and $\triangle p_{xrg}$ to return mean-reversion is approximately 2 months in quick unusual regime and it takes 7 months for the later to return in the slow usual regime for every unit of p_{co}. Also within the asymmetric price transmission, $\triangle p_{xrg}$ does not respond spontaneously to itself in whether quick or slow direction although it responds positively to $\triangle p_{co}$. The change in downstream does not impact on the upstream.

The asymmetric behavior was still a dominant feature of our mean-reversion also in this stage with 2-SETAR(5,1,4) equation also exhibits significant constant and trending intervention properties of the price build-up mechanism. The threshold models outperformed the AR and ARMA process from the forecasts in Figure 8 although the MAPE is slightly high. That is, the mean reversion property depends significantly on what happens in the past 4 biweekly periods in the quick unusual directions, whereas in slow usual direction it depends on recurrence of the same in the just the immediate biweekly periods.

For a single equation of the mean-reversion, Hansen and Johansen test suggest we use a linear model of embedding dimension of $m \geq 2$ as seen Larsen [2012]. Both prices does not response spontaneously to dynamic external shocks from each other and within themselves.

This asymmetric behavior of the sample mean-reversion observation is best modeled with linear univariate process specifically the ARMA(4,0,2) since it forecasting performance slightly outweighs that of 2-regime SETAR or any order univariate method in Figure 10 according to the MAPE. We imply that domestic interaction through taxes, levies and margins added onto ex-pump price build-up is behave linearly upwards. Prices are not returning in the manner when even the upstream falls.

5 Conclusion

In the further investigation of whether this mean-reversion above for bivariate analysis is nonlinear, the Hansen and Seo test unexpectedly accepted the null hypothesis of linear cointegration over claim of asymmetric in paired mean-reversion in this stage. Additionally, the hansen and Johansen test for 1-regime versus 2-regime and 2-regime versus 3-regime of the univariate equation mean-reversion indicate we rather model by 2-regime or one threshold model. The 2-SETAR(5,1,4) equation, also showcases the robust and persistent intervention elements of the price build-up process. The presence of asymmetric behavior continued to be a dominant factor in our mean-reversion strategy, and it was also evident in this stage. While the Mean Absolute Percentage Error (MAPE) is somewhat greater for the threshold models compared to the Autoregressive (AR) and Autoregressive Moving Average (ARMA) processes, the threshold models outperformed the AR and ARMA processes. In other words, the mean reversion property is significantly influenced by the events that have taken place in the previous four biweekly periods when there have been sudden and unexpected changes. However, when there are gradual and expected changes, it is influenced by the events in the biweekly periods immediately preceding it.

Acknowledgements

The authors would like to acknowledge College of Science Liberian of the Kwame Nkrumah University of Science Technology for his contribution of making access to the digital library use for accessing literature for the study.

Funding

The authors did not receive any form of funding on this study, hence has no conflict of interest in this paper.

References

Sticky prices, inventories, and market power in wholesale gasoline markets. *The RAND Journal of Economics*, 33:116–139, 2002.

Rockets and feathers revisited: an international comparison on european gasoline market. *Energy Economics*, 25:175–90, 2002.

J. Bachmeier, L.J. & Griffin. New evidence on asymmetric gasoline price responses. *Review of Economics and Statistics*, 85:772–776, 2003.

R. Bacon. Rockets and feathers: the asymmetric speed of adjustment of uk retail gasoline prices to cost changes. *Energy economics*, 13:211–218, 1991.

S. . Y. M. Balke, N.S.; Brown. Crude oil and gasoline prices: an asymmetric relationship. *Federal Reserve Bank of Dallas Economic Review*, pages 2–11, 1998.

T. Balke, N.S. & Fomby. Threshold cointegratio. *International economic review*, pages 627–645, 1997.

A. . G. R. Borenstein, S.; Cameron. Do gasoline prices respond asymmetrically to crude oil price changes? *National Bureau of Economic Research*, 1997.

G. E. P. Box and G. M. Jenkins. Time series analysis, forecasting and control. *Prentice Hall, Englewood Clifs*, 1994.

S. P. Brown and M. K. Yucel. Gasoline and crude oil prices: why the asymmetry? *Economic and financial review*, page 23, 2000.

M. Caner and B. E. Hansen. Threshold autoregression with a unit root. *Econometrica*, 69(6):1555–1596, 2001.

A. Di Narzo. Nonlinear autoregressive time series models in r using tsdyn version 0.7. 2008.

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

C. Enders, W. & Granger. Unit-root tests and asymmetric adjustment with an example using the term structure of interest rates. *Journal of Business and Economic Statistics*, 16:304–311, 1998.

P. Enders, W. & Siklos. Cointegration and threshold adjustment. *Journal of Business and Economic Statistics*, 19:166–176, 2001.

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

J. Gonzalo, J. & Pitarakis. Threshold effects in cointegrating relationships. *Oxford Bulletin of Economics and Statistics*, 68:813–833, 2006.

B. Gregory, A.W. & Hansen. Residual-based tests for cointegration in models with regime shifts. *Journal of econometrics*, 70:99–126, 1996.

B. Hansen. Testing for linearity. *Journal of Economic Surveys*, 13, 1999.

B. E. Hansen. Inference in tar models. *Studies in nonlinear dynamics & econometrics*, 2(1), 1997.

B. E. Hansen and B. Seo. Testing for two-regime threshold cointegration in vector error-correction models. *Journal of econometrics*, 110(2):293–318, 2002.

H. Hansen and S. Johansen. Some tests for parameter constancy in cointegrated var-models. *The Econometrics Journal*, 2(2):306–333, 1999.

P. R. Hansen. A test for superior predictive ability. *Journal of Business & Economic Statistics*, 23(4):365–380, 2005.

B. N. Huang, C. W. Yang, and M. J. Hwang. The dynamics of a nonlinear relationship between crude oil spot and futures prices: A multivariate threshold regression approach. *Energy Economics*, 31 31:91 to 98, 2009.

S. Johansen. Statistical analysis of cointegration vectors. *Journal of economic dynamics and control*, 12(2-3):231–254, 1988.

S. Johansen. Identifying restrictions of linear equations with applications to simultaneous equations and cointegration. *Journal of econometrics*, 69(1):111–132, 1995.

S. Johansen and K. Juselius. Maximum likelihood estimation and inference on cointegrationwith appucations to the demand for money. *Oxford Bulletin of Economics and statistics*, 52(2):169–210, 1990.

J. D. Karrenbrock et al. The behavior of retail gasoline prices: symmetric or not? *Federal Reserve Bank of St. Louis Review*, 73(4):19–29, 1991.

D. Kwiatkowski, P. C. Phillips, P. Schmidt, and Y. Shin. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of econometrics*, 54(1-3): 159–178, 1992.

B. Larsen. A threshold cointegration analysis of norwegian interest rates. Master's thesis, Universitetet i Tromsø, 2012.

J. Meyer and S. von Cramon-Taubadel. Asymmetric price transmission: a survey. *Journal of agricultural economics*, 55(3):581–611, 2004.

V. Natanelov, M. J. Alam, A. M. McKenzie, and G. V. Huylenbroeck. Is there co-movement of agricultural commodities futures prices and crude oil? *Energy Policy*, 39:4971 to 4984, 2011.

S. Peltzman. Prices rise faster than they fall. *Journal of Political Economy*, 108: 466–502., 2000.

P. C. Phillips and S. Ouliaris. Testing for cointegration using principal components methods. *Journal of Economic Dynamics and Control*, 12(2):205 – 230, 1988.

M. Seo. Bootstrap testing for the null of no cointegration in a threshold vector error correction model. *Journal of Econometrics*, 134(1):129–150, 2006.

D. Shin. Do product prices repspond symmetrically to changes in crude oil prices? *OPEC Review*, 18:137–157, 1992.

J. H. Stock and M. W. Watson. Testing for common trends. *Journal of the American statistical Association*, 83(404):1097–1107, 1988.

M. Tappata. Rockets and feathers: Understanding asymmetric pricing. *The RAND Journal of Economics*, 40(4):673–687, 2009.

H. Tong. *Threshold models in non-linear time series analysis. Lecture notes in statistics, No. 21.* Springer-Verlag, 1983.

H. Tong and K. S. Lim. Threshold autoregression, limit cycles and cyclical data. *Journal of the Royal Statistical Society: Series B (Methodological)*, 42(3):245–268, 1980.

R. S. Tsay. Testing and modeling multivariate threshold models. *journal of the american statistical association*, 93(443):1188–1202, 1998.

H.-M. Zhu, S.-F. Li, and K. Yu. Crude oil shocks and stock markets: A panel threshold cointegration approach. *Energy Economics*, 33(5):987–994, 2011.