Department of Economics

Testing for Cointegration in Nonlinear STAR Error Correction Models

George Kapetanios, Yongcheol Shin and Andy Snell

Working Paper No. 497 July 2003 ISSN 1473-0278

Queen Mary
University of London

Testing for Cointegration in Nonlinear STAR Error Correction Models*

George Kapetanios
Queen Mary, University of London

Yongcheol Shin
School of Economics, University of Edinburgh

Andy Snell
School of Economics, University of Edinburgh

April 2003

Abstract

In this paper we propose a new testing procedure to detect the presence of a cointegrating relationship that follows a globally stationary smooth transition autoregressive (STAR) process. We start from a general VAR model, embed the STAR error correction mechanism (ECM) and then derive the generalised nonlinear STAR error correction model. We provide two operational versions of the tests. Firstly, we obtain the associated nonlinear ECM-based test. Secondly, we generalise the well-known residual-based test for cointegration in linear models by Engle and Granger (1987) and obtain its nonlinear analogue. We derive the relevant asymptotic distributions of the proposed tests. We find via Monte Carlo simulation exercises that our proposed tests have much better power than the Engle and Granger test against the alternative of a globally stationary STAR cointegrating process. In an application to the price-dividend relationship, we also find that our test is able to find cointegration, whereas the linear-based tests fail to do so. Further analysis of impulse response functions of error correction terms (under the alternative) shows that the time taken to recover one half of a one standard deviation shock varies between five and twenty years, whereas the time taken to recover one half of a large shock varies between just 4 to 18 months. This clearly implies that data periods dominated by extreme volatility may display substantial mean reversion of the price-dividend relationship. By contrast this relationship may well look like a unit root when the underlying shocks take on smaller values.

JEL Classification: C12, C13, C32.
Key Words: Unit Roots, Globally Stationary Cointegrating Processes, Nonlinear Exponential Smooth Transition Autoregressive Error Correction Models, Monte Carlo Simulations, Prices and Dividends.

*We are grateful to seminar participants at University of Edinburgh and Heriott-Watt for their helpful comments. Partial financial support from the ESRC (grant No. R000223399) is gratefully acknowledged. The usual disclaimer applies.

1 Introduction

The investigation of nonstationarity in economics has assumed great significance over the past two decades. There has been increasing concern in macroeconomics that the information revealed by the analysis of a linear time series model may be insufficient to give definitive inference on important economic hypotheses. In particular, the power of tests such as the Dickey-Fuller (1979, DF) unit root test or the Engle-Granger (1987, EG) test for cointegration has been called into question. As a response to these problems, macroeconomists are increasingly turning to nonlinear dynamics to improve estimation and inference.

Balke and Fomby (1997) have recently popularised a joint analysis of nonstationarity and nonlinearity in the context of threshold cointegration, the case where a process may follow a unit root in a middle regime whilst at the same time being globally geometrically ergodic in outer regimes. They have also shown via Monte Carlo experiments that the power of both the DF unit root test and the EG cointegration test falls dramatically with threshold parameters. For a growing literature to address the joint issues of nonstationarity and nonlinearity see Michael *et al.* (1997), Enders and Granger (1998), Berben and van Dijk (1999), Caner and Hansen (2001), Lo and Zivot (2001), van Dijk *et al.* (2001), Saikkonen and Choi (2001), Kapetanios and Shin (2002) and Kapetanios *et al.* (2003).

In particular, Kapetanios *et al.* (2003) provide an alternative univariate framework by analysing a test of the null of a unit root against an alternative of a nonlinear exponential smooth transition autoregressive (ESTAR) process, develop the test that is specifically designed to have power against the globally stationary ESTAR process, and find via Monte Carlo simulations that the proposed test has better power than the DF test in cases where the nonlinear adjustment is locally prominent under the alternative of a globally stationary ESTAR process. However, these type of tests could also be regarded as cointegration tests in principle with (possibly) cointegrating parameters being known, since it can be applied to equilibrium relationships such as real exchange rates or real interest rates. In this regard, it would be more fruitful to derive a direct cointegration test under the STAR framework.

To bridge the two areas of nonstationarity and nonlinearity in the context of cointegration, we consider a single equation cointegrating relationship with nonlinearity in the speed of adjustment back to equilibrium. Following Kapetanios *et al.* (2003), we propose here to model a process where correction to cointegration is slower when the cointegrating residual is close to zero, and where the change in speed of this adjustment process is assumed smooth rather than sharp (as with threshold autoregressive models). Therefore, we focus on the case in which the cointegrating relationship or the error correction term follows a globally stationary ESTAR process under the alternative. Our approach is theoretically sensible in terms of speed of convergence arguments made in the literature, when considering economic hypothesis such as purchasing power parity in particular and hypotheses consisting of asset arbitrage conditions in general. See Sercu *et al.* (1995) and Michael *et al.* (1997) for a further discussion. More importantly, our framework is quite general. We start from a general VAR model and embed the ESTAR error correction mechanism in the resulting single equation conditional VAR model, and then develop the generalised nonlinear ESTAR error correction model in which we allow for the presence of nonlinear adjustment to the error correction mechanism under the alternative hypothesis.

Using this general nonlinear STAR error correction model framework and following a pragmatic residual-based two step approach advanced by Engle and Granger (1987) and Balke and Fomby (1997), we propose that a null hypothesis of no cointegration against an alternative of a globally stationary ESTAR cointegration be tested by examining the significance of the parameter controlling the degree of nonlinearity in the speed of adjustment. We derive the two operational test statistics, denoted t_{NLECM} and t_{NLEG}, respectively. The t_{NLECM} test refers to the t-type statistic obtained directly from the nonlinear ESTAR error correction regression, whereas the t_{NLEG} test is the nonlinear analogue to the Engle and Granger statistic for linear cointegration.

The small sample performance of the suggested tests is compared to that of the EG test via Monte Carlo experiments. We find that our newly proposed tests have good size properties and superior power properties compared to the EG test. In particular, the t_{NLECM} test is clearly superior to either linear or nonlinear EG tests when the regressors are weakly endogenous in a cointegrating regression. This clearly supports similar findings made in linear models by Kremers *et al.* (1992) that the EG test loses power relative to ECM-based cointegration tests because of losing potentially valuable information from the correlation between the regressors and the underlying disturbances.

Finally, we provide an application to investigating the presence of cointegration of asset prices and dividends for eleven stock portfolios (Germany, Belgium, Canada, Denmark, France, Ireland, Italy, Japan, Netherlands, UK and US) allowing for nonlinear ESTAR adjustment to equilibrium. The motivation for nonlinearity is that asset market models in the presence of transactions costs imply a nonlinear adjustment process toward an equilibrium through arbitrage. The test results clearly demonstrate the empirical worth of our approach. In particular, our proposed tests are able to reject the null of no cointegration in majority cases whereas the linear EG test rejects only twice. Given the strength of evidence in favor of the ESTAR cointegration we also estimate adjustment parameters under the alternative, and find that these estimates are significant in all cases. We further evaluate the impulse response functions of the error correction term with respect to initial impulses of 1 - 4 standard deviation shocks, respectively. The striking finding is that the time taken to recover one half of a one standard deviation shock varies between five and twenty years, whereas the time taken to recover one half of a large shock varies between just 4 to 18 months. This implies that data periods dominated by extreme volatility may display substantial reversion of prices towards their NPV relationship, while in "calmer" times where the error in the NPV relationship takes on smaller values, the process driving it may well look like a unit root.

The plan of the paper is as follows: Section 2 discusses general modelling issues and derives the nonlinear ESTAR error correction models. Section 3 develops the proposed test statistics and derives their asymptotic distributions. Section 4 evaluates the small sample performance of the proposed tests that take account of the specific ESTAR nonlinear nature of the alternative. Section 5 presents an empirical application to price-dividend relationships. Section 6 contains some concluding remarks. Mathematical proofs are collected in an appendix.

2 Nonlinear STAR Error Correction Models

We begin our analysis by reviewing cointegration testing in the linear case. We start with the following (possibly cointegrating) linear regression:

$$y_t = \boldsymbol{\beta}'\mathbf{x}_t + u_t, \tag{2.1}$$

$$\Delta\mathbf{x}_t = \mathbf{v}_t, \tag{2.2}$$

where y_t is a scalar I(1) variable, $\mathbf{x}_t$ is a $k \times 1$ vector of I(1) variables (not cointegrated among themselves), and the disturbances $(u_t, \mathbf{v}_t')'$ are assumed to follow the general stationary processes. We also suppose that the u_t follow an AR(1) process:

$$\Delta u_t = \rho u_{t-1} + \varepsilon_t. \tag{2.3}$$

Initially, we assume that the ε_t are *iid* processes with zero mean and finite variance, σ_ε^2, and the $\mathbf{x}_t$ are weakly exogenous with respect to ε_t. In this linear case, if $\rho = 0$, then y_t and $\mathbf{x}_t$ are not cointegrated, while if $\rho < 0$, there is a cointegrating relationship between y_t and $\mathbf{x}_t$.

The most popular approach to testing the presence of cointegration is the Engle and Granger (1987) two-step residual based test (hereafter EG). In the first stage one estimates $\boldsymbol{\beta}$ by OLS in (2.1) whereas in the second stage one carries out the Dickey-Fuller unit root t-test for the null of $\rho = 0$ (no cointegration) against the alternative of $\rho < 0$ (cointegration) using the following auxiliary regression:

$$\Delta\hat{u}_t = \rho\hat{u}_{t-1} + \varepsilon_t, \tag{2.4}$$

where $\hat{u}_t = y_t - \hat{\boldsymbol{\beta}}'\mathbf{x}_t$ are residuals obtained from (2.1) and $\hat{\boldsymbol{\beta}}$ is the OLS estimator of $\boldsymbol{\beta}$.

There has been an alternative testing approach based on error correction models. Applying the first difference transformation to (2.1) and combining with (2.3), we obtain

$$\Delta y_t = \boldsymbol{\beta}'\Delta\mathbf{x}_t + \rho u_{t-1} + \varepsilon_t, \tag{2.5}$$

If $\rho = 0$, then (2.5) reduces to the linear regression involving only first differences, thus implying that there is no cointegration between the levels of y_t and $\mathbf{x}_t$. Therefore, in this case, the cointegration test can be carried out using the one-sided t-statistic for $\rho = 0$ (no cointegration) against $\rho < 0$ (cointegration) in the following regression:

$$\Delta y_t = \boldsymbol{\beta}'\Delta\mathbf{x}_t + \rho\hat{u}_{t-1} + \varepsilon_t. \tag{2.6}$$

There have also been some attempts to derive the ECM-based test directly from (2.5) without using residuals. See Banerjee *et al.* (1998) and Pesaran *et al.* (2001). One main motivation behind the use of ECM-based tests is their superior power performance over the EG tests. The plausible explanation for the much better power of the ECM-based test as compared to the EG test centers on an implicit common factor restriction imposed when using the EG procedure. If that restriction is invalid, the EG test remains consistent but loses power relative to ECM-based cointegration tests because of losing potentially valuable information from $\Delta \mathbf{x}_t$ (see (2.13) below). For some details see Kremers *et al.* (1992) and Hansen (1995).

In this paper we aim to develop an alternative modelling approach in which we allow for the presence of nonlinear adjustment to the error correction mechanism under the alternative hypothesis. To do this we consider the following general model:

$$\Delta u_t = F(u_{t-1}) + \varepsilon_t, \tag{2.7}$$

where various functional forms of $F(\bullet)$ can be analysed such that $F(u_{t-1})$ includes a linear model (2.3) as a special case. Here we focus on a special case where $F(\bullet)$ follows the exponential smooth transition autoregressive (ESTAR) functional form,[1]

$$F(u_{t-1}) = \gamma u_{t-1}\left(1 - e^{-\theta u_{t-1}^2}\right). \tag{2.8}$$

In this case (2.7) becomes

$$\Delta u_t = \gamma u_{t-1}\left(1 - e^{-\theta u_{t-1}^2}\right) + \varepsilon_t. \tag{2.9}$$

Kapetanios, Shin and Snell (2003, henceforth KSS) show that the u_t in (2.9) are geometrically ergodic or globally stationary as long as $\theta > 0$ and $-2 < \gamma < 0$. Combining (2.6) and (2.9) together,

$$\Delta y_t = \gamma u_{t-1}\left(1 - e^{-\theta u_{t-1}^2}\right) + \boldsymbol{\beta}' \Delta \mathbf{x}_t + \varepsilon_t, \tag{2.10}$$

which we call the nonlinear STAR error correction model. The representation (2.10) makes economic sense in that many economic models predict that the underlying system tends to display a dampened behavior towards an attractor when it is (sufficiently far) away from it, but shows some instability within the locality of that attractor. For a growing literature on the joint analysis of the cointegration and the STAR models see Sercu *et al.* (1995), Michael *et al.* (1997), Saikkonen and Choi (2001) and KSS.

Following KSS, it is straightforward to show that the test of the null of no cointegration against the alternative of globally stationary cointegration can be based on the single parameter θ. More particularly, we set the null hypothesis of no cointegration as

$$H_0 : \theta = 0, \tag{2.11}$$

against the alternative of nonlinear ESTAR cointegration of

$$H_1 : \theta > 0. \tag{2.12}$$

The positive value of θ effectively determines the stationary property of u_{t-1}. In general, the u_t are unobserved unless $\boldsymbol{\beta}$ are unknown, whereas the unknown STAR parameter γ is

[1] The exponential transition function is bounded between zero and 1, *i.e.* $F : \mathbb{R} \to [0, 1]$ has the properties: $F(0) = 0$; $\lim_{x\to\pm\infty} F(x) = 1$, and is symmetrically U-shaped around zero. An alternative nonlinear adjustment scheme is given by the first- or second-order logistic functions. In other application the threshold autoregressive models can also be considered.

not identified under the null $\theta = 0$. Furthermore, when considering the nonlinear STAR error correction model (2.10), both (nuisance) parameters $\boldsymbol{\beta}$ and γ are not identified under the null. Most solutions to this problem involve some sort of integrating out unidentified parameters, and this is usually achieved by calculating the summary test statistics obtained for a grid of possible values of $\boldsymbol{\beta}$ and γ, *e.g.*, Hansen (1996). The direct approach to dealing with this double Davies problem would be desirable, but the construction of a grid of possible values for $\boldsymbol{\beta}$ would be formidable as the dimension of $\boldsymbol{\beta}$ increases. There has been a limited approach, see for example Hansen and Seo (2001) for a two-regime TAR cointegration model. In next section we will develop the operational cointegration tests under the nonlinear STAR framework.

However, the above modelling approach may be restrictive in time series modelling in the sense that any tests developed from them are not expected to be robust to possible weak endogeneity of the regressors and serial correlation of the errors ε_t. We first consider relaxing the weak exogeneity assumption. For simplicity we consider the nonlinear STAR framework given by (2.1), (2.2), (2.9), but at first only allow for the contemporaneous correlation between $\mathbf{v}_t$ and ε_t as follow:

$$\varepsilon_t = \boldsymbol{\pi}'\mathbf{v}_t + e_t = \boldsymbol{\pi}'\Delta\mathbf{x}_t + e_t, \tag{2.13}$$

where the e_t's are *iid* variates with zero mean and finite variance. In this case, (2.1), (2.9) and (2.10) can be modified respectively to

$$y_t = \boldsymbol{\alpha}'\mathbf{x}_t + u_t^*, \tag{2.14}$$

$$\Delta u_t^* = \gamma u_{t-1}^* \left(1 - e^{-\theta u_{t-1}^{*2}}\right) + e_t, \tag{2.15}$$

$$\Delta y_t = \gamma u_{t-1}^* \left(1 - e^{-\theta u_{t-1}^{*2}}\right) + \boldsymbol{\alpha}'\Delta\mathbf{x}_t + e_t, \tag{2.16}$$

where $\boldsymbol{\alpha} = \boldsymbol{\beta} + \boldsymbol{\pi}$. This clearly shows that so far as the estimation of (2.14) and (2.16) is concerned under the null, (2.14), (2.15) and (2.16) are observationally equivalent to (2.1), (2.9) and (2.10). Therefore, the asymptotic null distributions of the (nonlinear) cointegration statistics obtained from (2.1), (2.9) and (2.10) would be equivalent to those obtained from (2.14), (2.15) and (2.16).

Next, turning to the case where the ε_t in (2.10) are serially correlated, and assuming that these serial correlations enter in a linear autoregressive fashion with finite lag order p, then[2]

$$\Delta u_t = \gamma u_{t-1} \left(1 - e^{-\theta u_{t-1}^2}\right) + \sum_{j=1}^{p} \varphi_j \Delta u_{t-j} + \eta_t, \tag{2.17}$$

[2]Of course in reality the augmentations may actually enter in a nonlinear way. In such cases, we would view the use of linear augmentation terms as a first order approximation to the underlying dynamics rather than a strict view about the exact nature of the dynamic process itself. Alternatively, we would follow the semi-parametric correction method advanced by Phillips and Perron (1988).

where η_t's are *iid* variates with zero mean and finite variance. Now, combining (2.17) with (2.10), we have

$$\begin{aligned} \Delta y_t &= \boldsymbol{\beta}'\Delta\mathbf{x}_t + \gamma u_{t-1}\left(1-e^{-\theta u_{t-1}^2}\right) + \sum_{j=1}^{p}\varphi_j\Delta u_{t-j} + \eta_t \\ &= \boldsymbol{\beta}'\Delta\mathbf{x}_t + \gamma u_{t-1}\left(1-e^{-\theta u_{t-1}^2}\right) + \sum_{j=1}^{p}\varphi_j\Delta y_{t-j} + \sum_{j=1}^{p}\boldsymbol{\lambda}_j'\Delta\mathbf{x}_{t-j} + \eta_t, \end{aligned} \quad (2.18)$$

where $\boldsymbol{\lambda}_j = -\varphi_j\boldsymbol{\beta}$. Following Said and Dickey (1984) and KSS, it can be easily seen that under the null the statistics testing $\theta = 0$ in (2.18) would have the same asymptotic distribution as obtained under non-serially correlated errors.

Finally, combining both contemporaneous correlation (2.13) and serial correlation (2.17) within (2.10), we obtain the following general nonlinear ESTAR error correction model:[3]

$$\Delta y_t = \gamma u_{t-1}\left(1-e^{-\theta u_{t-1}^2}\right) + \boldsymbol{\alpha}'\Delta\mathbf{x}_t + \sum_{j=1}^{p}\varphi_j\Delta y_{t-j} + \sum_{j=1}^{p}\boldsymbol{\lambda}_j'\Delta\mathbf{x}_{t-j} + e_t. \quad (2.19)$$

where the e_t are *iid* processes and all the regressors in (2.19) are weakly exogenous by construction.

We now turn to a generalization which includes all the above modelling approaches as a special case. We begin with the data generating process for $\mathbf{z}_t = (y_t, \mathbf{x}_t')'$ that follows the general VAR model of lag order $p+1$:

$$\mathbf{z}_t = \sum_{i=1}^{p+1}\Phi_i\mathbf{z}_{t-i} + \boldsymbol{\varepsilon}_t,\ t = 1, 2, ..., T \quad (2.20)$$

where Φ_i, $i = 1, ..., p+1$, are $(k+1)\times(k+1)$ matrices of unknown coefficients, the vector error processes $\boldsymbol{\varepsilon}_t$ are $iid(\mathbf{0}, \Sigma)$ with Σ being a $(k+1)\times(k+1)$ positive definite matrix, and the initial observations $\mathbf{Z}_0 \equiv (\mathbf{z}_{-p}, ..., \mathbf{z}_0)$ are given. The VAR$(p+1)$ model (2.20) may be rewritten in vector ECM form as

$$\Delta\mathbf{z}_t = \Pi\mathbf{z}_{t-1} + \sum_{i=1}^{p}\Gamma_i\Delta\mathbf{z}_{t-i} + \boldsymbol{\varepsilon}_t,\ t = 1, 2, ..., T, \quad (2.21)$$

The focus of this paper is on the conditional modelling of the variable y_t given the k-vector $\mathbf{x}_t$ and the past values of $\mathbf{z}_t$ and $\mathbf{Z}_0$. Partitioning the error term $\boldsymbol{\varepsilon}_t$ conformably with $\mathbf{z}_t$ as $\boldsymbol{\varepsilon}_t = (\varepsilon_{yt}, \boldsymbol{\varepsilon}_{xt}')'$ and its variance matrix as $\Sigma = \begin{pmatrix} \sigma_{yy} & \boldsymbol{\sigma}_{yx} \\ \boldsymbol{\sigma}_{xy} & \Sigma_{xx} \end{pmatrix}$, we may express ε_{yt} conditionally in terms of $\boldsymbol{\varepsilon}_{xt}$ as

$$\varepsilon_{yt} = \boldsymbol{\sigma}_{yx}\Sigma_{xx}^{-1}\boldsymbol{\varepsilon}_{xt} + e_t, \quad (2.22)$$

[3]The extension to a more general case that the $\mathbf{v}_t$ follow the VAR(q) processes will be straightforward. Then, the resulting model has essentially the same structure as (2.19). See the general discussion below.

where $e_t \sim iid(0, \sigma_e^2)$, $\sigma_e^2 \equiv \sigma_{yy} - \boldsymbol{\sigma}_{yx}\Sigma_{xx}^{-1}\boldsymbol{\sigma}_{xy}$ and e_t is independent of $\boldsymbol{\varepsilon}_{xt}$ by construction. Substituting (2.22) into (2.21), partitioning $\Pi = (\boldsymbol{\pi}_y', \Pi_x')'$, $\Gamma_i = (\boldsymbol{\gamma}_{yi}', \Gamma_{xi}')'$, $i = 1, ..., p$, and further assuming $\Pi_x = \mathbf{0}$, we obtain the following conditional ECM model for Δy_t:

$$\Delta y_t = \boldsymbol{\pi}_y \mathbf{z}_{t-1} + \boldsymbol{\omega}' \Delta \mathbf{x}_t + \sum_{i=1}^{p} \boldsymbol{\psi}_i' \Delta \mathbf{z}_{t-i} + e_t, \tag{2.23}$$

and the marginal VAR model for $\Delta \mathbf{x}_t$,

$$\Delta \mathbf{x}_t = \sum_{i=1}^{p} \Gamma_{xi} \Delta \mathbf{z}_{t-i} + \boldsymbol{\varepsilon}_{xt}, \tag{2.24}$$

where $\boldsymbol{\omega} \equiv \Sigma_{xx}^{-1}\boldsymbol{\sigma}_{xy}$ and $\boldsymbol{\psi}_i' \equiv \boldsymbol{\gamma}_{yi} - \boldsymbol{\omega}'\Gamma_{xi}$, $i = 1, ..., p$. Notice that the assumption $\Pi_x = \mathbf{0}$ implies that the process $\mathbf{x}_t$ are weakly exogenous for the parameters of (2.23) and therefore the parameters in (2.23) are variation-free from the parameters in (2.24), but also restricts consideration to cases in which there exists *at most* one conditional cointegrating relationship between y_t and $\mathbf{x}_t$ which includes *both* y_t *and* $\mathbf{x}_t$. See Johansen (1995), and Pesaran *et al.* (2001) for a more general discussion. Now, rewriting $\boldsymbol{\pi}_y \mathbf{z}_{t-1} = \rho\left(y_{t-1} - \boldsymbol{\beta}' x_{t-1}\right) = \rho u_{t-1}$ in (2.23), and embedding the STAR error correction mechanism (2.9), we obtain the following generalised nonlinear ESTAR error correction model:

$$\Delta y_t = \gamma u_{t-1}\left(1 - e^{-\theta u_{t-1}^2}\right) + \boldsymbol{\omega}' \Delta \mathbf{x}_t + \sum_{i=1}^{p} \boldsymbol{\psi}_i' \Delta \mathbf{z}_{t-i} + e_t. \tag{2.25}$$

In practice different lag orders for Δy_t and $\Delta \mathbf{x}_t$ can be selected in a data dependent way using standard model selection criteria or significance testing procedure without loss of generality or change in the asymptotic analysis. See Ng and Perron (1995).

3 Testing for Nonlinear STAR Cointegration

In this section we will develop the two operational versions of the cointegration test under the general nonlinear STAR-ECM framework given by (2.25). Here we follow Engle and Granger (1987) and Balke and Fomby (1997), and take a pragmatic residual-based two step approach. In the first stage, we obtain the residuals, $\hat{u}_t = y_t - \hat{\boldsymbol{\beta}}' \mathbf{x}_t$ from (2.1) with $\hat{\boldsymbol{\beta}}$ being the OLS estimate of $\boldsymbol{\beta}$. In the second stage, in order to overcome the Davies problem that γ is not identified under the null, we follow KSS and approximate (2.25) by

$$\Delta y_t = \delta u_{t-1}^3 + \boldsymbol{\omega}' \Delta \mathbf{x}_t + \sum_{i=1}^{p} \psi_i' \Delta \mathbf{z}_{t-i} + e_t, \tag{3.1}$$

where we use a first-order Talyor series approximation to $\left(1 - e^{-\theta u_{t-1}^2}\right)$ under the null $\theta = 0$ following Luukkonen *et al.* (1988).

This suggests that we could obtain the t-type statistic for $\delta = 0$ (no cointegration) against $\delta < 0$ (nonlinear ESTAR cointegration) by using $\hat{u}_t$ in (3.1). We then obtain the t-statistic for $\delta = 0$, denoted t_{NLECM}, by

$$t_{NLECM} = \frac{\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_1\Delta\mathbf{y}}{\sqrt{\hat{\sigma}^2\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_1\hat{\mathbf{u}}_{-1}^{3}}}, \tag{3.2}$$

$$\hat{\sigma}^2 = T^{-1}\sum_{t=1}^{T}\left(\Delta y_t - \hat{\delta}\hat{u}_{t-1}^3 - \hat{\boldsymbol{\omega}}'\Delta\mathbf{x}_t + \sum_{i=1}^{p}\hat{\boldsymbol{\psi}}_i'\Delta\mathbf{z}_{t-i}\right)^2, \tag{3.3}$$

where $\hat{\mathbf{u}}_{-1}^3 = \left(\hat{u}_0^3, ..., \hat{u}_{T-1}^3\right)'$, $\mathbf{Q}_1 = \mathbf{I}_T - \mathbf{S}\left(\mathbf{S}'\mathbf{S}\right)^{-1}\mathbf{S}'$, $\mathbf{S} = (\Delta\mathbf{X}, \Delta\mathbf{Z}_{-1}, ..., \Delta\mathbf{Z}_{-p})$, $\Delta\mathbf{X} = (\Delta\mathbf{x}_1, ..., \Delta\mathbf{x}_T)'$, $\Delta\mathbf{Z}_{-i} = (\Delta\mathbf{z}_{1-i}, ..., \Delta\mathbf{z}_{T-i})'$, $i = 1, ..., p$, $\Delta\mathbf{y} = (\Delta y_1, ..., \Delta y_T)'$, and $\hat{\delta}$, $\hat{\boldsymbol{\omega}}$, $\hat{\boldsymbol{\psi}}_i$, $i = 1, ..., p$, are the OLS estimates of δ, $\boldsymbol{\omega}$, $\boldsymbol{\psi}_i$, $i = 1, ..., p$.

Alternatively and in keeping with the tradition in linear cointegration, we propose a companion test, namely a test which is the analogue to the Engle and Granger statistic for linear cointegration. This statistic, denoted t_{NLEG}, is obtained by

$$t_{NLEG} = \frac{\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_2\Delta\hat{\mathbf{u}}}{\sqrt{\hat{\sigma}^2\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_2\hat{\mathbf{u}}_{-1}^{3}}}, \tag{3.4}$$

through the following regression:

$$\Delta\hat{u}_t = \delta\hat{u}_{t-1}^3 + \sum_{i=1}^{p}\varphi_i\Delta\hat{u}_{t-i} + \xi_t, \tag{3.5}$$

where

$$\hat{\sigma}_{EG}^2 = T^{-1}\sum_{t=1}^{T}\left(\Delta\hat{u}_t - \hat{\delta}\hat{u}_{t-1}^3 - \sum_{i=1}^{p}\hat{\varphi}_i\Delta\hat{u}_{t-i}\right)^2, \tag{3.6}$$

$\Delta\hat{\mathbf{u}} = (\Delta\hat{u}_1, ..., \Delta\hat{u}_T)'$, $\mathbf{Q}_2 = \mathbf{I}_T - \Delta\mathbf{U}_p\left(\Delta\mathbf{U}_p'\Delta\mathbf{U}_p\right)^{-1}\Delta\mathbf{U}_p'$, $\Delta\mathbf{U}_p = (\Delta\hat{\mathbf{u}}_{-1}, ..., \Delta\hat{\mathbf{u}}_{-p})$, $\Delta\hat{\mathbf{u}}_{-i} = (\Delta\hat{u}_{1-i}, ..., \Delta\hat{u}_{T-i})'$, $i = 1, ..., p$ and $\hat{\delta}$, $\hat{\varphi}_i$ are the OLS estimates of δ and φ_i.

Theorem 3.1 *Consider the generalised nonlinear ESTAR error correction model (2.25). Under the null (2.11), the t_{NLECM} and t_{NLEG} statistics defined by (3.2) and (3.4) have the following asymptotic distributions, respectively:*

$$t_{NLECM} \Rightarrow \frac{\int B(r)^3 dW(r)}{\sqrt{\int B(r)^6 dr}}, \tag{3.7}$$

$$t_{NLEG} \Rightarrow \frac{\int B(r)^3 dB(r)}{\left(1 + \boldsymbol{\tau}'\boldsymbol{\tau}\right)\sqrt{\int B(r)^6 dr}}, \tag{3.8}$$

where '⇒' denotes a weak convergence,

$$B(r) = W(r) - \mathbf{W}(r)' \left(\int_0^1 \mathbf{W}(r)\mathbf{W}(r)' \right)^{-1} \left(\int_0^1 \mathbf{W}(r)W(r) \right), \tag{3.9}$$

$$\boldsymbol{\tau} = \left[\int_0^1 \mathbf{W}'(r)\,\mathbf{W}(r)\,dr \right]^{-1} \left[\int_0^1 \mathbf{W}'(r)\,W(r) \right], \tag{3.10}$$

and $W(r)$ and $\mathbf{W}(r)$ are independent scalar and k-vector standard Brownian motions, respectively, defined on $r \in [0,1]$. Under the alternative (2.12), both t_{NLECM} and t_{NLEG} statistics diverge to negative infinity.

To accommodate deterministic components in the regression (2.1), we extend to consider the regression with an intercept

$$y_t = a_0 + \boldsymbol{\beta}'\mathbf{x}_t + u_t, \tag{3.11}$$

and the regression with an intercept and a linear deterministic time trend,

$$y_t = a_0 + a_1 t + \boldsymbol{\beta}'\mathbf{x}_t + u_t. \tag{3.12}$$

Notice that the u_t are still of the same form given by (2.9). Alternatively, we take a simpler but equivalent approach, in which we re-express (3.11) and (3.12) as

$$y_t^* = \boldsymbol{\beta}'\mathbf{x}_t^* + u_t^*, \tag{3.13}$$

$$y_t^+ = \boldsymbol{\beta}'\mathbf{x}_t^+ + u_t^+, \tag{3.14}$$

where superscripts '$*$' and '$+$' indicate the demeaned data and the demeaned and detrended data, respectively.

The respective t_{NLECM} and t_{NLEG} statistics are then obtained as follows: First, the appropriate residuals are obtained from either (3.13) or (3.14), and then the corresponding regressions are constructed. More specifically, we have

$$\Delta y_t^* = \delta \hat{u}_{t-1}^* + \boldsymbol{\omega}'\Delta\mathbf{x}_t^* + \sum_{i=1}^{p-1} \boldsymbol{\psi}_i'\Delta\mathbf{z}_{t-i}^* + error, \tag{3.15}$$

$$\Delta y_t^+ = \delta \hat{u}_{t-1}^+ + \boldsymbol{\omega}'\Delta\mathbf{x}_t^+ + \sum_{i=1}^{p-1} \boldsymbol{\psi}_i'\Delta\mathbf{z}_{t-i}^+ + error, \tag{3.16}$$

where $\hat{u}_t^* = y_t^* - \boldsymbol{\beta}'\mathbf{x}_t^*$, and $\hat{u}_t^+ = y_t^+ - \boldsymbol{\beta}'\mathbf{x}_t^+$. The appropriate t_{NLECM} statistics are now obtained as the t-statistic for $\delta = 0$ in (3.15) or (3.16), respectively. The corresponding t_{NLEG} statistics are also similarly obtained. For the regression with a non-zero intercept (3.11), it is easily seen that the asymptotic distribution of both t_{NLECM} and t_{NLEG} statistics is the same as in (3.7) and (3.8), except that $W(r)$ and $\mathbf{W}(r)$ is replaced by the de-meaned Brownian motions, denoted $\widetilde{W}(r)$ and $\widetilde{\mathbf{W}}(r)$, defined on $r \in [0, 1]$. Similarly, for the regression with non-zero intercept and non-zero linear trend coefficient (3.12), the associated asymptotic distributions are such that $W(r)$ and $\mathbf{W}(r)$ are replaced by the de-meaned and de-trended Brownian motions, denoted $\widehat{W}(r)$ and $\widehat{\mathbf{W}}(r)$. Asymptotic critical values of the t_{NLEGDF} and t_{NLECM} statistics for the above three cases have been tabulated for $k = 1, ..., 5$, via stochastic simulations with $T = 1,000$ and 50,000 replications in Table 1.

Table 1 about here

4 Finite Sample Properties

In this section we undertake a small-scale Monte Carlo investigation of the finite sample size and power performance of our proposed t_{NLECM} and t_{NLEG} tests in conjunction with the linear EG test, denoted t_{EG}.

We consider experiments based on the bivariate regression. In the first set of experiments (Experiment 1) we focus on the size of the tests after constructing the null model by

$$y_t = \beta x_t + u_t, \tag{4.1}$$

$$\Delta x_t = v_t, \tag{4.2}$$

$$\Delta u_t = \varepsilon_t, \tag{4.3}$$

$$\varepsilon_t = \lambda v_t + \eta_t, \tag{4.4}$$

where we fix $\beta = 1$, $v_t \sim iidN(0, 1)$, $\eta_t \sim iidN(0, 1)$ and η_t is independently distributed of v_t. We consider the two cases: $\lambda = 0$ (Experiment 1A with exogenous regressor) and $\lambda = 1$ (Experiment 1B with endogenous regressor). Since the tests are similar with respect to intercepts and/or time trends, we will set all intercepts and trend coefficients to zero in what follows. In order to accommodate the possible serial correlation of the errors, we also consider the case of AR(1) errors given by

$$\eta_t = \phi_1 \eta_{t-1} + e_t, \tag{4.5}$$

$$v_t = \phi_2 v_{t-1} + \varsigma, \tag{4.6}$$

where we set $\phi_1 = \phi_2 = 0.4$, $e_t \sim iidN(0, 1)$, $\varsigma_t \sim iidN(0, 1)$, and e_t, ς_t are independently generated. We also consider the two cases: $\lambda = 0$ (Experiment 1C) and $\lambda = 1$ (Experiment 1D).

Secondly, in order to evaluate the power of alternative tests, we now generate (4.3) by

$$\Delta u_t = \gamma u_{t-1} \left[1 - \exp\left(-\theta u_{t-1}^2\right)\right] + \varepsilon_t, \tag{4.7}$$

where ε_t is still generated by (4.4) with $v_t \sim iidN(0,1)$ and $\eta_t \sim iidN(0,1)$. To make a general power comparison we choose a broad range of parameter values for $\gamma = \{-1.5, -1, -0.5\}$ and $\theta = \{0.01, 0.1\}$. We also consider the two cases, $\lambda = 0$ (Experiment 2A) and $\lambda = 1$ (Experiment 2B). Here we also consider the special case of the linear alternative given by

$$\Delta u_t = \gamma u_{t-1} + \varepsilon_t, \tag{4.8}$$

where we simply set $\gamma = -0.1$.

Table 2 presents results on the size of the alternative tests. Notice that for Experiments 1A and 1B in which the ε_t are not serially correlated, all the test statistics are obtained using the saponification with no augmentations, whereas we use the correct specification with one augmentation for Experiments 1C and 1D with AR(1) errors. As expected, sizes for the t_{NLECM}, t_{NLEG} and t_{EG} tests are all close to the nominal level of 5% even in the presence of serially correlated errors.

Table 2 about here

Turning to the power performance of the tests, which are summarised in Table 3, a general finding is that our suggested nonlinear-based t_{NLECM} and t_{NLEG} tests are more powerful than the linear t_{EG} test for almost all cases considered. In particular, the power gain of the t_{NLECM} and t_{NLEG} tests over the t_{EG} test becomes substantial, when θ is relatively small, e.g. $\theta = 0.01$.[4] This result is consistent with the univariate finding in KSS. For example, when looking at Case 2 for Experiment 2A (with exogenous regressor) with $(\gamma, \theta) = (-1.5, 0.01)$ and $T = 100$, the powers of the t_{NLECM} and t_{NLEG} tests are .534 and .514, whereas the power of t_{EG} is only .345. In Experiment 2A, the t_{NLECM} test is only marginally more powerful than the t_{NLEG} test. Interestingly and as expected, however, the power dominance of the t_{NLECM} test becomes more pronounced for Experiment 2B in which the regressor is endogenous. For example, when looking at Case 2 for Experiment 2B (with endogenous regressor) with $(\gamma, \theta) = (-1, 0.01)$ and $T = 100$, the powers of the t_{NLECM}, t_{NLEG} and t_{EG} tests are .792, .458 and .288, respectively. As mentioned earlier, the plausible explanation is that the t_{NLECM} test explicitly employs potentially valuable information from the correlation between Δx_t and ε_t, whereas both t_{NLEG} and t_{EG} tests fail to do so. Finally, looking at the results for the linear alternative (see the rows with $(\gamma, \theta) = (-0.1, \infty)$), we find that the linear t_{EG} test is slightly more powerful than both t_{NLECM} and t_{NLEG} tests for Experiment 2A, whereas the t_{NLECM} is still most powerful for Experiment 2B. This clearly demonstrates the superior power performance of the ECM-based test over the EG-type tests in general when the regressors are likely to be endogenous.

Table 3 about here

[4]Notice in our application below that the estimates of θ (which we obtain under the constraint that $\gamma = -1$) are indeed quite small, ranging as they do between 0.007 and 0.017.

5 Empirical Application to Asset Pricing in the Presence of Transactions Costs

In a seminal paper, Campbell and Shiller (1987) investigate the existence of linear cointegration between aggregate US stock prices and US dividends, as predicted by a simple equilibrium model of constant expected asset returns. Their results were ambiguous. A null hypothesis of no linear cointegration was marginally rejected in their data but the implied estimates of long run asset returns was implausible. Imposing a more credible long run return caused non rejection of the null of no cointegration. Subsequent literature has met with similar mixed results.

In this section we test for cointegration of asset prices and dividends for eleven stock portfolios allowing for nonlinear adjustment to equilibrium of the STAR variety. The motivation for nonlinearity is the existence of transactions costs via a bid ask spread that varies over stocks. At first we might expect that transactions costs which arise from a fixed spread might motivate the consideration of price adjustment mechanism of the TAR variety. However our data consists of prices and dividends averaged over a widely diversified portfolio of stocks and it has been shown in numerical simulations that this aggregation process leads to a specification that is better approximated by an STAR rather than a TAR model, see for example Taylor *et al.* (2003).

We collected monthly data from January 1974 to November 2002 on end period real prices and within period real dividend yields for value weighted market portfolio indices of stocks traded on the main exchanges of the following eleven countries: Germany, Belgium, Canada, Denmark, France, Ireland, Italy, Japan, Netherlands, UK and US. A dividend series was constructed as the product of dividend yield and prices. Although not presented here, simple ADF tests from an initial data analysis give overwhelming support to the hypothesis that all variates are I(1).

As alluded to above we test for the existence of a linear cointegrating relationship between dividends and prices of the form,

$$p_t = \beta d_t + u_t. \tag{5.9}$$

The existence of bid ask spreads discussed above motivate the specification of a nonlinear dynamic adjustment mechanism such as the $ESTAR$ model giving the following nonlinear STAR-ECM model:

$$\Delta p_t = \gamma \left(1 - e^{-\theta u_{t-1}^2}\right) u_{t-1} + \alpha \Delta x_t + \varepsilon_t, \tag{5.10}$$

where $u_{t-1} = p_{t-1} - \beta d_{t-1}$ and ε_t is a (possibly autocorrelated) error term which captures other microstructure effects such as specific kinds of noise trading and dealer inventory control mechanisms whereby the adjustment of prices to elicit inventory-correcting trades generates autocorrelated price movements, see for example Snell and Tonks (1998).

We computed three tests. The first two, t_{EG} and t_{NLEG} are the linear Engel-Granger test and its nonlinear counterpart. The third, t_{NLECM} is the t-ratio on $\hat{u}_{t-1}^3$ in the STAR-ECM formulation where $\hat{u}_{t-1}$ is the residual from the first stage (spurious under the null) regression of p_t on d_t. The price and dividend series appeared to have an upward trend so that all series were demeaned and detrended before use.[5] We estimated the appropriate auxiliary regressions for $p = 12$ and then dropped all insignificant lags in a single round of general to specific modelling.[6]

The results for the tests are in columns 2 to 4 in Table 4. Looking at the results we see that viewed through the "eyes" of linear cointegration tests there is little support for the hypothesis that dividends and prices move together in the long run with only 2 of the t_{EG} tests rejecting the null, albeit at the 1% level. Furthermore, none of the remaining 9 t_{EG} statistics are significant even at the 10% level. By contrast the nonlinear t_{NLEG} test rejects in 7 out of 11 stock markets with four of these rejections occurring at the 1% level. A further three statistics are quite close to the 10% critical value. The success in rejecting the null of no cointegration is less marked for t_{NLECM} with only 5 rejections at standard significance levels although three of these reject also at the 1%. A further two t_{NLECM} statistics are quite close to the 10% critical value.[7]

Table 4 about here

Given the strength of evidence against the null and support for the alternative we could obtain estimates of adjustment parameters under the alternative. Focusing on the univariate model we obtained nonlinear least squares estimates of θ from the alternative $ESTAR$ model,

$$\Delta\hat{u}_t = -\left\{1 - \exp\left(-\theta\hat{u}_{t-1}^2\right)\right\}\hat{u}_{t-1} + \sum_{i=1}^{12}\varphi_i\Delta\hat{u}_{t-i} + \xi_t, \tag{5.11}$$

where the model has been specialised compared with the general $ESTAR$ considered above by imposing a unit coefficient on γ. Early attempts to estimate γ jointly with θ foundered on severe identification problems and our nonlinear algorithm failed to converge in most cases - hence the specialisation. Under the alternative (and estimation of (5.11) only makes sense if the alternative is true), θ is scale dependent. To clarify its interpretation and to facilitate numerical convergence, we normalised the $\hat{u}_t$ series to have unit sample variance (a procedure which only makes sense under the alternative). The results for $\hat{\theta}$ and its t-statistic are given in Table 4. Although we cannot interpret the t-statistic as a significance from zero test (for obvious reasons) we refer to it as "significant" if an asymptotic 95% confidence interval around the estimate excludes zero. We see that $\hat{\theta}$ is "significant" in all cases and varies between .007 and .017.

To get a feel for what such values imply, Figure 1 below plots impulse response functions (irfs) for the error correction term for initial impulses of 1, 2, 3, 4 standard deviation shocks,

[5] The issue of whether or not stock prices and dividends contain a deterministic time trend in the long run is contentious. However there is a clearly discernible trend in both dividends and prices in our data hence we detrend and demean. It is comforting to note that if we do not detrend but only demean, the results are qualitatively almost identical.

[6] We should note that although further exploration revealed some significant lags beyond 12th order, the test statistics were not in general very sensitive to the choice of lag length.

[7] If the alternative is really true then we could interpret this finding as being somewhat at odds with the Monte Carlo evidence, which generally shows that t_{NLECM} has more power than t_{NLEG}. However, there is good prior reason to believe that dividends are weakly exogenous in the system. If true, a bivariate ECM would lack parsimony compared with the univariate specification and this may have lead to a loss in power.

respectively. For completeness and comparison we compare the corresponding irf with that obtained from the estimated linear models. The striking thing about the graphs is the length of time taken to recover from small shocks. In particular the time taken to recover one half of a one standard deviation shock varies between five and twenty years. By contrast, the time taken to recover one half of a large shock (such as 3 or 4 standard deviations) is comparable to that of the linear case and varies between just 4 to 18 months. This implies that data periods dominated by extreme volatility may display substantial reversion of prices towards their NPV relationship but in "calmer" times, where the error in the NPV relationship takes on smaller values, the process driving it may well look like a unit root. This suggests that in practice the $ESTAR$ and $SETAR$ models may not be too dissimilar in terms of overall inference in any given (finite) sample.

Figure 1 about here

6 Concluding Remarks

The investigation of nonstationarity in conjunction with nonlinear autoregressive modelling has recently assumed a prominent role in econometric study. It is clear that misclassifying a stable nonlinear process as nonstationary can be misleading both in impulse response and forecasting analysis. Similarly not allowing for the presence of cointegration when the speed of adjustment varies with the position of the system as in the case of nonlinear cointegration can be equally misleading. In this paper we have proposed a direct cointegration test that is designed to have power against nonlinear error correction specifications. Our proposed tests are shown to have better power than the Engle and Granger test that ignores the nonlinear nature of the alternative. An empirical application clearly shows the potential of our approach. Unlike linear cointegration tests, the nonlinear tests find substantial evidence of cointegration in stock price and dividend systems. Further research to develop similar tests for alternative nonlinear models such as threshold autoregressive models is currently under way.

Table 1. Asymptotic Critical Values of the t_{NLEGDF} and t_{NLECM} Statistics

t_{NLEGDF}									
	Case 1			Case 2			Case 3		
k	90%	95%	99%	90%	95%	99%	90%	95%	99%
1	-2.59	-2.85	-3.38	-2.98	-3.28	-3.84	-3.41	-3.71	-4.26
2	-3.01	-3.30	-3.89	-3.36	-3.67	-4.23	-3.64	-3.99	-4.53
3	-3.34	-3.66	-4.23	-3.63	-3.93	-4.50	-3.90	-4.18	-4.76
4	-3.65	-3.95	-4.56	-3.90	-4.19	-4.68	-4.09	-4.39	-4.95
5	-3.88	-4.13	-4.75	-4.10	-4.42	-4.97	-4.36	-4.67	-5.23

t_{NLECM}									
	Case 1			Case 2			Case 3		
k	90%	95%	99%	90%	95%	99%	90%	95%	99%
1	-2.38	-2.66	-3.35	-2.92	-3.22	-3.78	-3.30	-3.59	-4.17
2	-2.67	-3.01	-3.59	-3.12	-3.43	-4.00	-3.46	-3.79	-4.40
3	-2.95	-3.28	-3.93	-3.32	-3.61	-4.19	-3.62	-3.96	-4.54
4	-3.15	-3.47	-4.14	-3.46	-3.77	-4.38	-3.75	-4.07	-4.70
5	-3.33	-3.67	-4.31	-3.58	-3.92	-4.53	-3.87	-4.20	-4.85

Table 2. Size of Alternative Tests

	Experiment 1A			Experiment 1B		
	t_{EGDF}	t_{NLEGDF}	t_{NLECM}	t_{EGDF}	t_{NLEGDF}	t_{NLECM}
Case 1						
$T = 100$	.056	.045	.052	.042	.042	.048
$T = 200$	.047	.041	.053	.060	.051	.045
Case 2						
$T = 100$	.056	.047	.041	.049	.053	.054
$T = 100$	.049	.046	.046	.038	.051	.047
Case 3						
$T = 100$	.070	.041	.046	.061	.045	.053
$T = 100$	.051	.044	.047	.044	.046	.051

	Experiment 1C			Experiment 1D		
	t_{EGDF}	t_{NLEGDF}	t_{NLECM}	t_{EGDF}	t_{NLEGDF}	t_{NLECM}
Case 1						
$T = 100$	.068	.062	.054	.046	.054	.052
$T = 200$	.046	.054	.054	.056	.047	.055
Case 2						
$T = 100$	.048	.053	.036	.056	.053	.031
$T = 100$	.047	.054	.049	.056	.061	.039
Case 3						
$T = 100$	.047	.040	.040	.055	.047	.040
$T = 100$	.051	.051	.037	.059	.045	.036

Table 3. Power of Alternative Tests

Case 1		Experiment 2A			Experiment 2B		
(γ, θ)	T	t_{EGDF}	t_{NLEGDF}	t_{NLECM}	t_{EGDF}	t_{NLEGDF}	t_{NLECM}
(-0.5,0.01)	100	.172	.236	.317	.260	.327	.710
	200	.633	.778	.859	.932	.894	.990
(-0.5,0.1)	100	.982	.958	.976	1.0	.978	1.0
	200	1.0	1.0	1.0	1.0	1.0	1.0
(-1,0.01)	100	.345	.500	.619	.667	.695	.935
	200	.977	.977	.986	1.0	.992	.999
(-1,0.1)	100	1.0	.999	.999	1.0	1.0	1.0
	200	1.0	1.0	1.0	1.0	1.0	1.0
(-1.5,0.01)	100	.606	.737	.824	.915	.885	.982
	200	.999	.993	.995	1.0	1.0	1.0
(-1.5,0.1)	100	1.0	1.0	1.0	1.0	1.0	1.0
	200	1.0	1.0	1.0	1.0	1.0	1.0
(-0.1,∞)	100	.373	.236	.336	.286	.213	.538
	200	.922	.631	.720	.847	.576	.920

Case 2		Experiment 2A			Experiment 2B		
(γ, θ)	T	t_{EGDF}	t_{NLEGDF}	t_{NLECM}	t_{EGDF}	t_{NLEGDF}	t_{NLECM}
(-0.5,0.01)	100	.150	.183	.191	.123	.177	.463
	200	.376	.515	.541	.547	.664	.936
(-0.5,0.1)	100	.843	.826	.841	.963	.902	.995
	200	1.0	1.0	1.0	1.0	1.0	1.0
(-1,0.01)	100	.215	.293	.312	.288	.458	.792
	200	.786	.859	.869	.978	.963	.996
(-1,0.1)	100	1.0	.993	.994	1.0	1.0	1.0
	200	1.0	1.0	1.0	1.0	1.0	1.0
(-1.5,0.01)	100	.345	.514	.534	.558	.698	.919
	200	.977	.952	.954	1.0	.996	1.0
(-1.5,0.1)	100	1.0	1.0	1.0	1.0	1.0	1.0
	200	1.0	1.0	1.0	1.0	1.0	1.0
(-0.1,∞)	100	.205	.165	.174	.123	.135	.326
	200	.701	.456	.479	.547	.350	.758

Case 3		Experiment 2A			Experiment 2B		
(γ, θ)	T	t_{EGDF}	t_{NLEGDF}	t_{NLECM}	t_{EGDF}	t_{NLEGDF}	t_{NLECM}
(-0.5,0.01)	100	.131	.120	.141	.108	.095	.297
	200	.274	.335	.367	.205	.365	.802
(-0.5,0.1)	100	.664	.646	.691	.794	.708	.967
	200	1.0	.994	.995	1.0	.997	1.0
(-1,0.01)	100	.174	.192	.222	.211	.273	.630
	200	.613	.681	.720	.830	.830	.978
(-1,0.1)	100	.992	.976	.985	1.0	.994	1.0
	200	1.0	1.0	1.0	1.0	1.0	1.0
(-1.5,0.01)	100	.265	.313	.374	.332	.464	.825
	200	.854	.861	.883	.991	.968	1.0
(-1.5,0.1)	100	1.0	.998	.997	1.0	1.0	1.0
	200	1.0	1.0	1.0	1.0	1.0	1.0
(-0.1,∞)	100	.176	.115	.132	.108	.073	.190
	200	.553	.322	.352	.295	.196	.576

Table 4. Cointegration Tests and Estimates of the ESTAR Parameter for Asset Prices and Dividends

Country	t_{EG}	t_{NLEG}	t_{NLECM}	θ	t_θ
Germany	-2.62	-3.68*	-4.69**	.011	3.40
Belgium	-2.06	-4.69**	-4.89**	.007	3.79
Canada	-1.73	-3.05	-1.53	.008	3.26
Denmark	-5.04**	-4.82**	-4.73**	.009	3.75
France	-3.27	-3.07	-0.99	.007	2.83
Ireland	-2.72	-3.76*	-3.73*	.017	3.85
Italy	-1.81	-5.50**	-3.04	.010	3.51
Japan	-2.44	-2.46	-2.11	.009	2.86
Netherlands	-4.94**	-7.79**	-1.90	.015	6.18
UK	-2.43	-3.28	-3.22	.008	2.78
US	-2.10	-3.90*	-3.70*	.007	3.76

Notes: The sample period in all cases is March 1974 to November 2002. *(**) denotes significance at the 5% (1%) level.

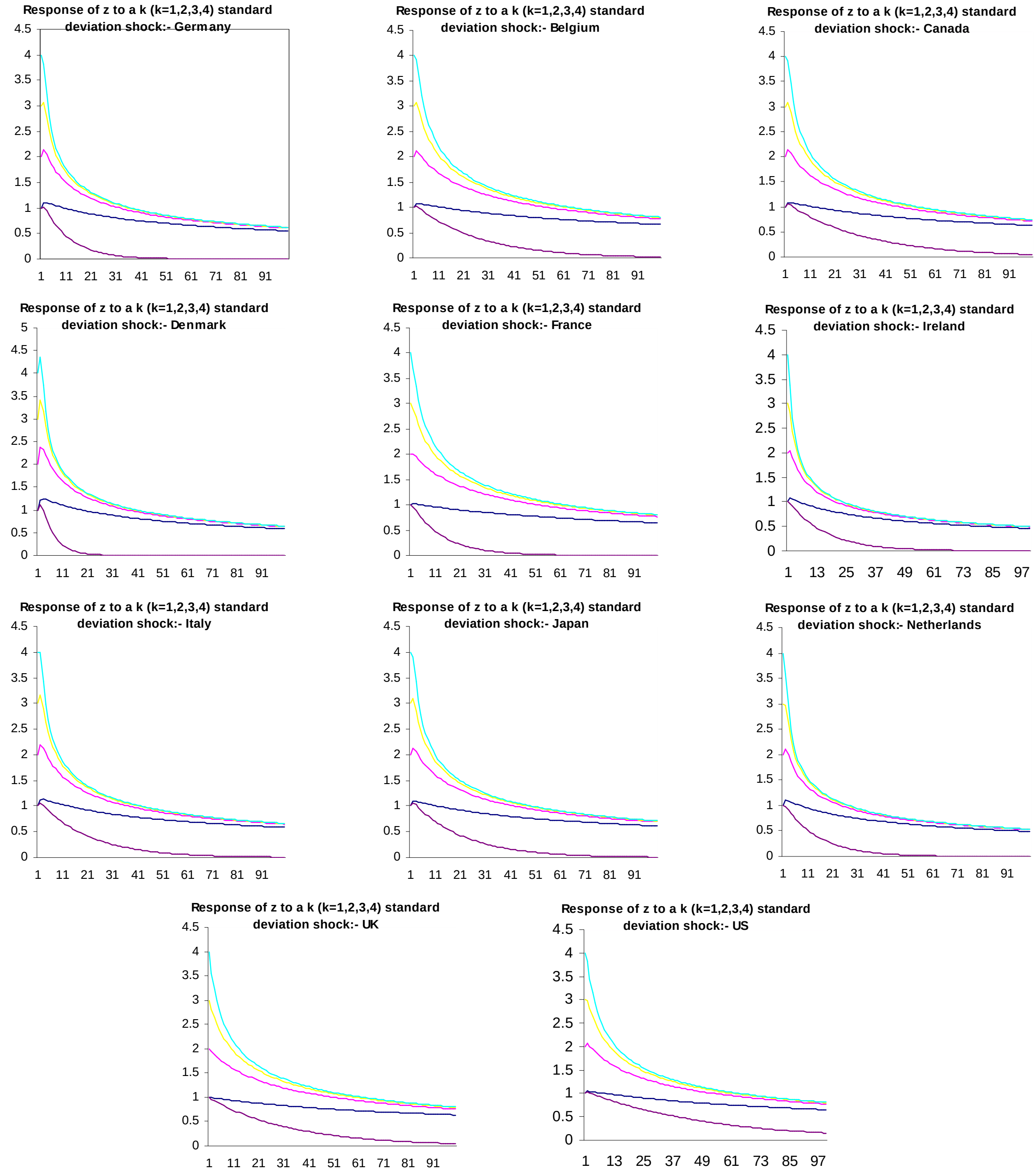

Figure 1. Impulse Response Functions of Error Coorection Terms

Notes: The lower line in each graph is the impulse response functions for the linear case whilst the lines above are impulse response functions the nonlinear case with $k = 1, 2, 3, 4$ standard deviations shocks (in ascending order), respectively.

A Appendix: Proof of Theorem 3.1

Under the null, $\delta = 0$ and thus t_{NLECM} given by (3.2) can be written as

$$t_{NLECM} = \frac{\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_1\mathbf{e}}{\sqrt{\hat{\sigma}_{ECM}^2 \hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_1\hat{\mathbf{u}}_{-1}^3}} = \frac{T^{-2}\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_1\mathbf{e}}{\sqrt{\hat{\sigma}_{ECM}^2 \left(T^{-4}\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_1\hat{\mathbf{u}}_{-1}^3\right)}}, \tag{A.1}$$

where $\mathbf{e} = (e_1, ..., e_T)'$. Next we write the residuals $\hat{u}_t$ obtained from (2.1) as

$$\hat{u}_t = y_t - \hat{\boldsymbol{\beta}}'\mathbf{x}_t = u_t - \mathbf{x}_t'\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right), \tag{A.2}$$

where $\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}$, $\mathbf{X} = (\mathbf{x}_1, ..., \mathbf{x}_T)'$ and $\mathbf{u} = (u_1, ..., u_T)'$. Notice under the null of no cointegration that the u_t are I(1). It is now well-established (*e.g.*, Following Phillips and Ouliaris (1990)) that

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{[Tr]} e_t \Rightarrow \sigma_e W(r),\ T^{-1/2}u_t \Rightarrow B_1(r), \tag{A.3}$$

$$\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} = (T^{-2}\mathbf{X}'\mathbf{X})^{-1}T^{-2}\mathbf{X}'\mathbf{u} \Rightarrow \left[\int_0^1 \mathbf{B}_2'(r)\,\mathbf{B}_2(r)\,dr\right]^{-1}\int_0^1 \mathbf{B}_2'(r)\,B_1(r), \tag{A.4}$$

where $[Tr]$ is the integer part of Tr, $W(r)$ is a scalar standard Brownian motion, $B_1(r)$ is a scalar Brownian motion with a long-run variance $\varpi_{11}^2 = \omega_{11} - \omega_{21}'\Omega_{22}^{-1}\omega_{21}$, and $\mathbf{B}_2(r)$ are the k-vector Brownian motions with a covariance matrix $\Omega_{22} = E(\mathbf{v}_t\mathbf{v}_t')$, and $\Omega = \begin{pmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \Omega_{22} \end{pmatrix}$ is the long run covariance matrix of $(\Delta y, \Delta\mathbf{x}')'$. Using (A.3) and (A.4) in (A.2), it is easily seen that

$$T^{-1/2}\hat{u}_t \Rightarrow \varpi_{11}B(r),$$

where $B(r)$ is defined in (3.9). Next, noting that the $\hat{u}_{t-1}^j$, $j = 1, 2, ...$, are a regular transformation of $\hat{u}_{t-1}$ in the sense of Park and Phillips (2001), we can apply Lemma 2.1 of their paper and obtain (see also KSS)

$$T^{-4}\sum_{t=1}^{T}\hat{u}_{t-1}^6 = T^{-1}\sum_{t=1}^{T}\left(T^{-1/2}\hat{u}_{t-1}\right)^6 \Rightarrow \varpi_{11}^6\int B(r)^6 dr,$$

$$T^{-2}\sum_{t=1}^{T}\hat{u}_{t-1}^3 e_t = T^{-\frac{1}{2}}\sum_{t=1}^{T}\left(T^{-1/2}\hat{u}_{t-1}\right)^3 e_t \Rightarrow \sigma_e\varpi_{11}^3\int_0^1 B(r)^3 dW(r),$$

$$T^{-2}\sum_{t=1}^{T}\hat{u}_{t-1}^3\Delta\mathbf{z}_t = T^{-\frac{1}{2}}\sum_{t=1}^{T}\left(T^{-1/2}\hat{u}_{t-1}\right)^3\mathbf{s}_t = O_p(1),$$

where the $\mathbf{s}_t$ are the t-th row of $\mathbf{S}$. Using the above results, it is now readily seen that

$$\begin{aligned} T^{-2}\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_1\mathbf{e} &= T^{-2}\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{e} + T^{-1/2}\left(T^{-2}\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{S}\right)\left(T^{-1}\mathbf{S}'\mathbf{S}\right)^{-1}\left(T^{-1/2}\mathbf{S}'\mathbf{e}\right) \\ &= T^{-2}\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{e} + o_p(1) \Rightarrow \sigma_e\varpi_{11}^3\int_0^1 B(r)^3 dW(r), \end{aligned} \tag{A.5}$$

$$\begin{aligned} T^{-4}\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_1\hat{\mathbf{u}}_{-1}^3 &= T^{-4}\hat{\mathbf{u}}_{-1}^{3\prime}\hat{\mathbf{u}}_{-1}^3 + T^{-1}\left(T^{-2}\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{S}\right)\left(T^{-1}\mathbf{S}'\mathbf{S}\right)^{-1}\left(T^{-2}\mathbf{S}'\hat{\mathbf{u}}_{-1}^3\right) \\ &= T^{-4}\hat{\mathbf{u}}_{-1}^{3\prime}\hat{\mathbf{u}}_{-1}^3 + o_p(1) \Rightarrow \varpi_{11}^6\int B(r)^6 dr. \end{aligned} \tag{A.6}$$

Finally, it is straightforward to show under the null that

$$\begin{aligned}
\hat{\sigma}^2 &= T^{-1}\sum_{t=1}^{T}\left(\Delta y_t - \hat{\delta}\hat{u}_{t-1}^3 - \hat{\boldsymbol{\omega}}'\Delta\mathbf{x}_t - \sum_{i=1}^{p}\hat{\boldsymbol{\psi}}_i'\Delta\mathbf{z}_{t-i}\right)^2 \qquad (A.7)\\
&= T^{-1}\sum_{t=1}^{T}\left(e_t - \hat{\delta}\hat{u}_{t-1}^3 - (\hat{\boldsymbol{\omega}}-\boldsymbol{\omega})'\Delta\mathbf{x}_t - \sum_{i=1}^{p}\left(\hat{\boldsymbol{\psi}}_i-\boldsymbol{\psi}_i\right)'\Delta\mathbf{z}_{t-i}\right)^2\\
&= T^{-1}\sum_{t=1}^{T}e_t^2 + o_p(1) \to \sigma_e^2,
\end{aligned}$$

where we used $T^2\hat{\delta} = O_p(1)$, $\sqrt{T}(\hat{\boldsymbol{\omega}}-\boldsymbol{\omega}) = O_p(1)$ and $\sqrt{T}\left(\hat{\boldsymbol{\psi}}_i-\boldsymbol{\psi}_i\right) = O_p(1)$. Using (A.5) - (A.7) in (A.1), we obtain the required results for the asymptotic distribution of the t_{NLECM} test.

We move on to prove results for the t_{NLEG} test, which under the null can be written as

$$t_{NLEG} = \frac{\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_2\boldsymbol{\xi}}{\sqrt{\hat{\sigma}_{EG}^2\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_2\hat{\mathbf{u}}_{-1}^3}} = \frac{T^{-2}\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_2\boldsymbol{\xi}}{\sqrt{\hat{\sigma}_{EG}^2\left(T^{-4}\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_2\hat{\mathbf{u}}_{-1}^3\right)}}, \qquad (A.8)$$

where $\boldsymbol{\xi} = (\xi_1, ..., \xi_T)'$. Notice under the null that the DGP for $\Delta\hat{u}_t$ is given by

$$\varphi(L)\Delta\hat{u}_t = \xi_t,$$

where $\varphi(L) = 1 - \sum_{i=1}^{p}\varphi_i L^i$, (see (3.5) and also discussion around (2.18)). Following Phillips and Ouliaris (1990), it is straightforward to show that

$$T^{-1/2}\sum_{t=1}^{[Tr]}\xi_t \Rightarrow \varphi(1)\varpi_{11}B(r), \qquad (A.9)$$

$$\frac{1}{T}\sum_{t=1}^{T}\Delta\hat{u}_t^2 \Rightarrow \varpi_{11}^2(1+\boldsymbol{\tau}'\boldsymbol{\tau}), \qquad (A.10)$$

$$\frac{1}{T}\sum_{t=1}^{T}\xi_t^2 \Rightarrow [\varphi(1)]^2\varpi_{11}^2(1+\boldsymbol{\tau}'\boldsymbol{\tau}), \qquad (A.11)$$

where $\boldsymbol{\tau}$ is defined in (3.10). As before we also have

$$T^{-4}\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_2\hat{\mathbf{u}}_{-1}^3 = T^{-4}\hat{\mathbf{u}}_{-1}^{3\prime}\hat{\mathbf{u}}_{-1}^3 + o_p(1) \Rightarrow \varpi_{11}^6\int_0^1 B(r)^6 dr, \qquad (A.12)$$

$$T^{-2}\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_2\boldsymbol{\xi} = T^{-2}\hat{\mathbf{u}}_{-1}^{3\prime}\boldsymbol{\xi} + o_p(1) \Rightarrow \varphi(1)\varpi_{11}^4\int_0^1 B(r)^3 dB(r), \qquad (A.13)$$

$$\begin{aligned}
\hat{\sigma}_{EG}^2 &= T^{-1}\sum_{t=1}^{T}\left(\Delta\hat{u}_t - \hat{\delta}\hat{u}_{t-1}^3 - \sum_{i=1}^{p}(\hat{\varphi}_i-\varphi_i)\Delta\hat{u}_{t-i}\right)^2 = T^{-1}\sum_{t=1}^{T}\xi_t^2 + o_p(1)\\
&\Rightarrow [\varphi(1)]^2\varpi_{11}^2(1+\boldsymbol{\tau}'\boldsymbol{\tau}),
\end{aligned}$$

where we used (A.11), $\hat{\delta} = O_p\left(T^{-2}\right)$ and $(\hat{\varphi}_i-\varphi_i) = O_p\left(T^{-1/2}\right)$, $i = 1, ..., p$. Using (A.12) - (A.14) in (A.8), we obtain the required results for the asymptotic distribution of the t_{NLEG} test.

We now prove that both t_{NLECM} and t_{NLEG} tests are consistent under the alternative. First, under the alternative, t_{NLECM} can be expressed as

$$t_{NLECM} = \delta\sqrt{\frac{\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_1\hat{\mathbf{u}}_{-1}^{3}}{\hat{\sigma}^2}} + \frac{\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_1\mathbf{e}^*}{\sqrt{\hat{\sigma}^2\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_1\mathbf{u}_{-1}^{3}}}, \tag{A.14}$$

where $\delta < 0$, $\mathbf{e}^* = (e_1^*, ..., e_T^*)'$ and $e_t^* = e_t + \delta\left(u_{-1}^3 - \hat{u}_{-1}^3\right)$. Since u_t, $\hat{u}_t = u_t - \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)' \mathbf{x}_t$, u_t^3, $\widehat{u}_{t-1}^3$ are all I(0) and $T\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) = O_p\left(T^{-1}\right)$ under the alternative, it is easily seen that

$$T^{-1}\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_1\hat{\mathbf{u}}_{-1}^{3} = O_p\left(1\right),\ T^{-1/2}\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_1\mathbf{e} = O_p\left(1\right), \tag{A.15}$$

$$\begin{aligned}
\hat{\sigma}^2 &= T^{-1}\sum_{t=1}^{T}\left(\Delta y_t - \hat{\delta}\hat{u}_{t-1}^3 - \hat{\boldsymbol{\omega}}'\Delta\mathbf{x}_t - \sum_{i=1}^{p}\hat{\boldsymbol{\psi}}_i'\Delta\mathbf{z}_{t-i}\right)^2 \\
&= T^{-1}\sum_{t=1}^{T}\left[e_t^* - \left(\hat{\delta} - \delta\right)\hat{u}_{t-1}^3 - \left(\hat{\boldsymbol{\omega}} - \boldsymbol{\omega}\right)'\Delta\mathbf{x}_t - \sum_{i=1}^{p}\left(\hat{\boldsymbol{\psi}}_i - \boldsymbol{\psi}_i\right)'\Delta\mathbf{z}_{t-i}\right]^2 \\
&= T^{-1}\sum_{t=1}^{T}\varepsilon_t^{*2} + o_p\left(1\right) = O_p\left(1\right),
\end{aligned}$$

where we used $\left(\hat{\delta} - \delta\right) = O_p\left(T^{-1/2}\right)$, $\left(\hat{\boldsymbol{\omega}} - \boldsymbol{\omega}\right) = O_p\left(T^{-1/2}\right)$ and $\left(\hat{\boldsymbol{\psi}}_i - \boldsymbol{\psi}_i\right) = O_p\left(T^{-1/2}\right)$. Notice in (A.15) that we also used that u_{t-1}, $\hat{u}_{t-1}$ and all lagged I(0) regressors in $\mathbf{S}$ are uncorrelated with e_t. Next, a tedious but straightforward computation shows that

$$T^{-1/2}\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_1\mathbf{e}^* = T^{-1/2}\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_1\mathbf{e} + o_p\left(1\right) = O_p\left(1\right). \tag{A.16}$$

Therefore, using (A.15) - (A.16) in (A.14), we obtain:

$$t_{NLECM} \Rightarrow \sqrt{T}\delta\sqrt{\frac{T^{-1}\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{M}\hat{\mathbf{u}}_{-1}^{3}}{\hat{\sigma}^2}} + O_p\left(1\right) = O_p\left(\sqrt{T}\right), \tag{A.17}$$

which clearly shows that the t_{NLECM} statistic diverges to negative infinity as $T \to \infty$.

Next, under the alternative, t_{NLEG} can be expressed as

$$t_{NLEG} = \delta\sqrt{\frac{\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_2\hat{\mathbf{u}}_{-1}^{3}}{\hat{\sigma}_{EG}^2}} + \frac{\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_2\boldsymbol{\xi}}{\sqrt{\hat{\sigma}_{EG}^2\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_2\hat{\mathbf{u}}_{-1}^{3}}}. \tag{A.18}$$

As before, it is straightforward to show that

$$T^{-1}\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_2\hat{\mathbf{u}}_{-1}^{3} = O_p\left(1\right),\ T^{-1/2}\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_1\boldsymbol{\xi} = O_p\left(1\right), \tag{A.19}$$

$$\hat{\sigma}_{EG}^2 = T^{-1}\sum_{t=1}^{T}\left(\Delta\hat{u}_t - \left(\hat{\delta} - \delta\right)\hat{u}_{t-1}^3 - \sum_{i=1}^{p}\left(\hat{\varphi}_i - \varphi_i\right)\Delta\hat{u}_{t-i}\right)^2 = T^{-1}\sum_{t=1}^{T}\xi_t^2 + o_p\left(1\right) = O_p\left(1\right), \tag{A.20}$$

where we used $\left(\hat{\delta} - \delta\right) = O_p\left(T^{-1/2}\right)$ and $\left(\hat{\varphi}_i - \varphi_i\right) = O_p\left(T^{-1/2}\right)$. Using (A.19) and (A.20) in (A.18), we find that

$$t_{NLEG} = \sqrt{T}\delta\sqrt{\frac{T^{-1}\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_2\hat{\mathbf{u}}_{-1}^{3}}{\hat{\sigma}_{EG}^2}} + O_p\left(1\right) = O_p\left(\sqrt{T}\right), \tag{A.21}$$

which also shows that the t_{NLEG} statistic diverges to negative infinity as $T \to \infty$.

We note in passing that Theorem 3.1 can be proved for the most general case of an infinite order VAR under the further condition that the truncated lag order is imposed by $p = O_p(T^{1/3})$. This upper bound follows from Berk (1974) who shows that a second moment matrix for the higher (than p) lagged regressors does not converge to its population moment in norm, and thus any coefficients on these stationary variables would not be estimated consistently. See also Ng and Perron (1995). Therefore, for consistent estimation we need to impose this upper bound. The detailed technical proof will be available upon request.

References

[1] Balke, N.S. and T.B. Fomby (1997), "Threshold Cointegration," *International Economic Review*, 38, 627-645.

[2] Banerjee A., J.J. Dolado, and R. Mestre (1998), "ECM Tests for Cointegration in a Single Equation Framework," *Journal of Time Series Analysis,* 19, 267-283.

[3] Berk, K.N. (1974), "Consistent Autoregressive Spectral Estimates," , *Annals of Statistics*, 2, 489–502.

[4] Berben, R. and D. van Dijk (1999), "Unit Root Tests and Asymmetric Adjustment: A Reassessment," unpublished manuscript, Erasmus University Rotterdam.

[5] Campbell J.Y. and R.J. Shiller (1987), "Cointegration and Tests of Present Value Models," *Journal of Political Economy,* 95, 1062-1088.

[6] Caner, M. and B.E. Hansen (2001), "Threshold Autoregression with a Unit Root," *Econometrica*, 69, 1555-1596.

[7] Davies, R.B. (1987), "Hypothesis Testing When a Nuisance Parameter is Present Under the Alternative," *Biometrika*, 74, 33-43.

[8] Dickey, D.A. and W.A. Fuller (1979), "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association*, 74, 427-431.

[9] Enders, W. and C.W.J. Granger (1998), "Unit Root Test and Asymmetric Adjustment with an Example Using the Term Structure of Interest Rates," *Journal of Business and Economic Statistics*, 16, 304-311.

[10] Engle, R. and C.W.J. Granger (1987), "Cointegration and Error Correction: Representation, Estimation and Testing," *Econometrica*, 55, 251-276.

[11] Hansen, B.E. (1995), "Rethinking the Univariate Approach to Unit Root Tests: How to Use Covariates to Increase Power," *Econometric Theory,* 11, 1148-1171.

[12] Hansen, B.E. (1996), "Inference when a Nuisance Parameter is not Identified under the Null Hypothesis," *Econometrica*, 64, 414–430.

[13] Hansen, B.E. and B. Seo (2001), "Testing for Threshold Cointegration in Vector Error Correction Models," unpublished manuscript, University of Wisconsin.

[14] Johansen, S. (1995), *Likelihood Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press: Oxford.

[15] Kapetanios, G. and Y. Shin (2002), "Unit Root Tests in Three-Regime SETAR Models," unpublished manuscript, University of Edinburgh.

[16] Kapetanios, G., Y. Shin and A. Snell, (2003), "Testing for a Unit Root in the Nonlinear STAR Framework," *Journal of Econometrics*, 112, 359-379.

[17] Kremers, J.J.M., K.R. Ericsson and J.J. Dolado (1992) "The Power of Cointegration Tests, " *Oxford Bulletin of Economics and Statistics,* 54, 325–348.

[18] Lo, M.C. and E. Zivot (2001), "Threshold Cointegration and Nonlinear Adjustment to the Law of One Price," *Macroeconomic Dynamics*, 5, 533-576.

[19] Luukkonen, R., P. Saikkonen and T. Teräsvirta (1988), "Testing Linearity against Smooth Transition Autoregressive Models," *Biometrika*, 75, 491-499.

[20] Michael, P., R.A. Nobay and D.A. Peel (1997), "Transactions Costs and Nonlinear Adjustment in Real Exchange Rates: an Empirical Investigation," *Journal of Political Economy*, 105, 862-879.

[21] Ng, S. and P. Perron (1995), "Unit Root Tests in ARMA models with Data-dependent Methods for the Selection of the Truncation Lag." *Journal of the American Statistical Association*, 90, 268-281.

[22] Pesaran M.H, R.J. Smith and Y. Shin (2001), "Bounds Testing Approaches to the Analysis of Level Relationships," *Journal of Applied Econometrics,* 16, 289-326.

[23] Phillips P.C.B. and S. Ouliaris (1990), "Asymptotic Properties of Residual Based Tests of Cointegration", *Econometrica*, 58, *165-193.*

[24] Phillips, P.C.B. and P. Perron (1988), "Testing for a Unit Root in Time Series Regression," *Biometrika*, 75, 335-346.

[25] Said, S.E. and D.A. Dickey (1984), "Testing for Unit roots in Autoregressive Moving Average Models of Unknown Order," *Biometrika*, 71, 599-607.

[26] Saikkonen, P. and I. Choi (2001), "Cointegrating Smooth Transition Regressions," unpublished manuscript, University of Helsinki.

[27] Sercu, P., R. Uppal and C. Van Hulle (1995), "The Exchange Rate in the Presence of Transaction Costs: Implications for Tests of Purchasing Power Parity," *Journal of Finance*, 50, 1309-1319.

[28] Snell A. and I. Tonks (1998), "Testing for Asymmetric Information and Inventory Control Effects in Market Maker Behaviour on the London Stock Exchange," *Journal of Empirical Finance*, 5, 1-25.

[29] Taylor, M.P., I. Chowdhury and L. Sarno (2003), "Nonlinear Dynamics in Deviations from the Law of One Price: A Broad-Based Empirical Study," forthcoming in *Journal of International Money and Finance.*

[30] van Dijk, D., T. Teräsvirta and P.H. Franses (2001), "Smooth Transition Autoregressive Models - A Survey of Recent Developments," unpublished manuscript, Erasmus University Rotterdam.

Queen Mary
University of London

This working paper has been produced by
the Department of Economics at
Queen Mary, University of London

Department of Economics
Queen Mary, University of London
Mile End Road
London E1 4NS
Tel: +44 (0)20 7882 5096 or Fax: +44 (0)20 8983 3580
Email: j.conner@qmul.ac.uk
Website: www.econ.qmul.ac.uk/papers/wp.htm