OXFORD BULLETIN OF ECONOMICS AND STATISTICS, 68, SUPPLEMENT (2006) 0305-9049

Threshold Effects in Cointegrating Relationships*

JESÚS GONZALO† and JEAN-YVES PITARAKIS‡

†*Department of Economics, Universidad Carlos III de Madrid, Madrid, Spain (e-mail: jesus.gonzalo@uc3m.es)*
‡*Economics Division, School of Social Sciences, University of Southampton, Southampton, UK*
(e-mail: j.pitarakis@soton.ac.uk)

Abstract

In this paper, we introduce threshold-type nonlinearities within a single-equation cointegrating regression model and propose a testing procedure for testing the null hypothesis of linear cointegration vs. cointegration with threshold effects. Our framework allows the modelling of long-run equilibrium relationships that may change according to the magnitude of a threshold variable assumed to be stationary and ergodic, and thus constitutes an attempt to deal econometrically with the potential presence of multiple equilibria. The framework is flexible enough to accommodate regressor endogeneity and serial correlation.

I. Introduction

Economic theory often predicts that some economic variables must be linked through long-run equilibrium relationships. From an applied viewpoint, a methodology for empirically assessing the existence of such relationships, when linear, is given by the concept of cointegration. Indeed, two or more economic variables that are expected to be in equilibrium must be cointegrated. Examples include linkages between wealth and consumption recently investigated in Lettau and Ludvigson (2004), variables which theory

*We wish to thank the Spanish Ministry of Education for supporting this research under grant SEJ2004-0401ECON. We also greatefully ackowledge the comments of two anonymous referees.
JEL Classification numbers: C22, C50.

© Blackwell Publishing Ltd, 2006. Published by Blackwell Publishing Ltd, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.

suggests are tied via present value relationships such as prices and dividends (Campbell and Shiller, 1987), amongst numerous others (see Ogaki, Jang and Lim, 2005, for a comprehensive overview). This *statistical* notion of an equilibrium relationship or long-run attractor linking two or more variables, together with Granger's representation theorem ensuring the existence of an error correction representation, provided researchers with valuable tools for jointly modelling and conducting inferences on the long-run equilibrium characterizing two or more variables, together with the short-run adjustment process towards such an equilibrium.

Although economic relationships are often nonlinear, characterized for instance by the presence of regime-specific behaviour such as the display of different dynamics depending on the phases of the business cycle, the statistical concept of cointegration as originally defined aimed to refer solely to *linear* combinations of variables linked through a long-run equilibrium relationship. In this paper our objective is to propose a test for assessing the presence of regime-specific nonlinearities within cointegrating relationships. Although the idea of extending the concept of cointegration to a nonlinear framework is not new (see Granger and Terasvirta, 1993; Granger, 1995; Balke and Fomby, 1997; Hansen and Seo, 2002; Seo, 2004; Kapetanios, Shin and Snell, 2004), nonlinearities in this context have been only understood to affect the adjustment process to equilibrium while the equilibrium relationship itself has typically been taken to be represented by a single and linear regression model.

The type of nonlinearity we propose to work with instead involves introducing regime-specific behaviour within the long-run equilibrium relationship itself. More specifically, we propose introducing threshold-type effects to model the possibility that the relationship linking the nonstationary variables undergoes regime switches. Such threshold effects lead to cointegrating regressions that are piecewise linear, separated according to the magnitude of a threshold variable which triggers the regime switches. Unlike Markov-switching type of nonlinearities where an unobservable state governs regime switches, our analysis exclusively focuses on nonlinearities induced by observable factors, possibly dictated by economic theory and that are assumed to be stationary throughout (e.g. growth rate in the economy, or the term structure of interest rates). In addition to offering an intuitive way of capturing economically meaningful nonlinearities, models with threshold effects are also straightforward to estimate by simple least squares methods or their variants, which is a considerable advantage in the present context. Although the multitude of alternative nonlinear specifications designed to capture regime-switching behaviour may suggest that the threshold family of models is only a narrow subset, especially given the fact that economic theory is often silent about the specific type of nonlinearity that may characterize the linkages among economic variables, Petruccelli (1992) has also shown that threshold

© Blackwell Publishing Ltd 2006

specifications may be viewed as an approximation to a more general class of nonlinear models.

Having inference tools designed to assess the presence of nonlinearities within a cointegrating regression is particularly important in applied work, unlike the case where the nonlinearities are introduced solely within the adjustment process to equilibrium. For instance, omitting the presence of nonlinear components such as threshold effects in long-run equilibrium, relationships themselves will lead to misleading interpretations of equilibrium relationships because the cointegrating vector will no longer be consistently estimated. Given the importance of threshold effects documented in the economics literature (see, for instance, Kanbur, 2005, and references therein) and the fact that cointegrating regressions can be viewed as a statistical counterpart to the notion of a long-run equilibrium linking two or more variables, we also conjecture that the econometric framework we focus on in this paper may provide a useful statistical environment for modelling the notion of switching or multiple long-run equilibria. The literature on animal spirits, for instance (Howitt and McAffee, 1992) suggests instances where some extraneous random variable such as a confidence indicator may cause macroeconomic variables to switch across different paths.

Our analytical framework involves working within a triangular type of representation of a set of variables known to be $I(1)$. Our goal is then to explore the properties of a test statistic for testing the null hypothesis of linear cointegration vs. nonlinear threshold-type cointegration, maintaining the presence of cointegration under both the null and alternative hypotheses.

In related work, Choi and Saikkonen (2004) have also explored the properties of statistical tests for detecting the presence of nonlinearities within cointegrating regressions. Unlike the present framework, however, they concentrated on a smooth transition type of functional form whose argument was taken to be one of the unit-root variables within the cointegrating regression. The idea of introducing nonlinear dynamics directly within cointegrating regressions has also been recently addressed in Park and Phillips (1999, 2001) and Chang, Park and Phillips (2001), where the authors' key concern was parameter estimation rather than testing in the context of regression models containing nonlinear transformations of $I(1)$ variables. Earlier research that explored the possibility of regime switches in cointegrating vectors can also be found in Hansen (1992) and Gregory and Hansen (1996), where the motivation of the authors was to test for the presence of structural break-type parameter instability within cointegrating relationships.

The plan of the paper is as follows. Section II introduces the formal statistical model, our operating assumptions, together with the resulting limiting distributions. Section III assesses the finite sample properties of the test for threshold cointegration. Section IV concludes.

© Blackwell Publishing Ltd 2006

II. Testing for threshold effects

2.1. The model and test statistic

We consider the following cointegrating regression with a threshold nonlinearity

$$y_t = \boldsymbol{\beta}'\mathbf{x}_t + \boldsymbol{\lambda}'\mathbf{x}_t I(q_{t-d} > \gamma) + u_t \tag{1}$$

$$\mathbf{x}_t = \mathbf{x}_{t-1} + \mathbf{v}_t \tag{2}$$

where u_t and $\mathbf{v}_t$ are scalar and p-vector-valued stationary disturbance terms, q_{t-d} with $d \geq 1$ is the stationary threshold variable and $I(q_{t-d} > \gamma)$ is the usual indicator function taking the value one when $q_{t-d} > \gamma$ and zero otherwise. The particular choice of d is not essential for our analysis and d will be taken as known throughout. For simplicity of exposition, we temporarily exclude the presence of deterministic components from equation (1) and relegate their inclusion to the end of this section.

Note that cointegration in the context of the specification in equation (1) is understood in the sense that although both y_t and $\mathbf{x}_t$ have variances that grow with t, the threshold combination given by u_t is stationary. Because of the lack of a precise definition of concepts such as nonlinear or threshold cointegration in the existing literature, it is not unusual to find the same terms referring to fundamentally different concepts. Unlike our specification in equation (1), for instance, the notion of threshold cointegration has often been used to refer to cases where the cointegrating relationship linking two or more variables is linear while the adjustment process to the long-run equilibrium contains threshold effects. Differently put, a cointegrating relationship that is linear has commonly been referred to as 'threshold cointegration' because of the presence of threshold effects in the adjustment mechanism.

An important source of difficulty when attempting to define a concept such as threshold cointegration comes from the fact that the concept of *integratedness* commonly used to refer to series as $I(1)$ or $I(0)$ is a linear concept and not helpful when trying to establish the stationarity properties of nonlinear processes. If x_t is an $I(1)$ variable, the degree of integration of powers of x_t is not obvious. Similar problems occur when dealing with stochastic unit-root models or models containing threshold effects with terms such as $y_t = x_t I(q_t > \gamma)$ where taking first differences does not make the series $I(0)$ (strictly speaking, they would be $I(\infty)$ processes). Because of these difficulties and before introducing a formal definition of the concept of threshold cointegration, we propose using the following alternative to the concept of $I(\cdot)$'ness.

Definition 1. A time series y_t is said to be summable of order δ, symbolically represented as $S_y(\delta)$ if the sum

© Blackwell Publishing Ltd 2006

$$S_y = \sum_{t=1}^{T} (y_t - E[y_t])$$

is such that

$$\frac{S_y}{T^{\frac{1}{2}+\delta}} = O_p(1) \quad \text{as } T \to \infty.$$

Note that in the context of the above definition, a process that is $I(d)$ can be referred to as $S_y(d)$ and the threshold process introduced in (1) is clearly $S_y(1)$. Using this concept of summability of order δ we can now provide a formal definition of the concept of threshold cointegration as follows.

Definition (threshold cointegration). Let y_t and x_t be $S_y(\delta_1)$ and $S_y(\delta_2)$, respectively. They are threshold cointegrated if there exists a threshold combination $(1, -(\beta_1 I(q_t \leq \gamma) + \beta_2 I(q_t > \gamma)))$ such that $z_t = y_t - \beta_1 x_t I(q_t \leq \gamma) - \beta_2 x_t I(q_t > \gamma)$ is $S_y(\delta_0)$ with $\delta_0 < \min(\delta_1, \delta_2)$.

Given the formal definition of threshold cointegration presented above, and its obvious extension to the multivariate case, it is now clear that within our specification in equation (1), y_t and $\mathbf{x}_t$ are threshold cointegrated with $\delta_0 = 0$ and $\delta_1 = \delta_2 = 1$. The threshold variable q_{t-d} allows the cointegrating vector to switch between $(1, -\boldsymbol{\beta}')$ and $(1, -(\boldsymbol{\beta} + \boldsymbol{\lambda})')$ depending on whether q_{t-d} crosses the unknown threshold level given by γ. This is also what we denote as two long-run equilibria.

We next turn to the properties of a test statistic for assessing the presence of threshold cointegration. For this purpose, it is convenient to reformulate equation (1) in matrix format as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{X}_\gamma \boldsymbol{\lambda} + \mathbf{u} \tag{3}$$

where $\mathbf{X}$ stacks $\mathbf{x}_t$ while $\mathbf{X}_\gamma$ stacks $\mathbf{x}_t I(q_{t-d} > \gamma)$. For later use, we also introduce the $(p + 1)$-dimensional random disturbance vector $\mathbf{w}_t = (u_t, \mathbf{v}_t')'$. In the context of the model in equation (1) or (3), a natural test of the hypothesis of linear cointegration vs. the alternative of threshold cointegration takes the form of testing $H_0 : \boldsymbol{\lambda} = 0$ against $H_1 : \boldsymbol{\lambda} \neq 0$ and here we propose exploring the properties of an LM-type test statistic for testing this null hypothesis. Note that under H_0 the threshold parameter given by γ remains unidentified. This is now a well-known problem in the literature on testing for regime-switching type of nonlinearities and is usually handled by viewing the test statistic as a random function of the nuisance parameter and basing inferences on a particular functional of the test statistic such as its supremum over γ, for instance (see Davies, 1977, 1987; Andrews and Ploberger, 1994; Hansen, 1996).

Letting $\text{LM}_T(\gamma)$ denote the Lagrange multiplier-type test statistic obtained for each γ, we will base our inferences on the quantity given by

© Blackwell Publishing Ltd 2006

$\text{supLM} = \sup_{\gamma\in\Gamma}\text{LM}_T(\gamma)$. At this stage, it is also important to recall that given $\gamma \in \Gamma$ the model in equation (3) is linear in its parameters and can therefore be easily estimated as it would have been the case in a purely linear framework. Letting $\mathbf{M} = \mathbf{I}-\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$, the standard form of the $\text{LM}_T(\gamma)$ statistic can be formulated as

$$\text{LM}_T(\gamma) = \frac{1}{\tilde{\sigma}_0^2}\mathbf{y}'\mathbf{M}\mathbf{X}_\gamma(\mathbf{X}_\gamma'\mathbf{M}\mathbf{X}_\gamma)^{-1}\mathbf{X}_\gamma'\mathbf{M}\mathbf{y} \tag{4}$$

where $\tilde{\sigma}_0^2$ denotes the residual variance estimated from equation (3) under the null hypothesis of linear cointegration. As, under the null hypothesis, we have $\mathbf{y}'\mathbf{M} = \mathbf{u}'\mathbf{M}$, the test statistic in equation (4) can also be written as

$$\text{LM}_T(\gamma) = \frac{1}{\tilde{\sigma}_0^2}\mathbf{u}'\mathbf{M}\mathbf{X}_\gamma(\mathbf{X}_\gamma'\mathbf{M}\mathbf{X}_\gamma)^{-1}\mathbf{X}_\gamma'\mathbf{M}\mathbf{u}. \tag{5}$$

2.2. Assumptions and limiting distributions

Throughout this paper, we will be operating under the following set of assumptions:

(A1) The sequence $\{u_t, \mathbf{v}_t, q_t\}$ is strictly stationary and ergodic and strong mixing with mixing coefficients α_n satisfying

$$\sum_{n=1}^{\infty}\alpha_n^{\frac{1}{2}-\frac{1}{r}} < \infty \quad \text{for some } r > 2;$$

the threshold variable q_t has a distribution function $F(\cdot)$ that is continuous and strictly increasing.

(A2)

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{[\text{Tr}]}\mathbf{w}_t \Rightarrow \mathbf{B}(r) = (B_u(r), \mathbf{B}_v(r)')'$$

where $\mathbf{B}(r)$ is a $(p + 1)$-dimensional Brownian motion with a long-run covariance matrix given by

$$\mathbf{\Omega} = \begin{pmatrix} \sigma_u^2 & \mathbf{\Omega}_{uv} \\ \mathbf{\Omega}_{vu} & \mathbf{\Omega}_{vv} \end{pmatrix} > 0,$$

where

$$\sigma_u^2 = E[u_t^2], \quad \mathbf{\Omega}_{uv} = \mathbf{\Sigma}_{uv} + \mathbf{\Lambda}_{uv} + \mathbf{\Lambda}_{vu}', \quad \mathbf{\Omega}_{vu} = \mathbf{\Sigma}_{vu} + \mathbf{\Lambda}_{vu} + \mathbf{\Lambda}_{uv}'$$
and $\mathbf{\Omega}_{vv} = \mathbf{\Sigma}_{vv} + \mathbf{\Lambda}_{vv} + \mathbf{\Lambda}_{vv}'$;

© Blackwell Publishing Ltd 2006

where

$$\mathbf{\Sigma}_{uv} = E[u_t \mathbf{v}_t'], \quad \mathbf{\Sigma}_{vv} = E[\mathbf{v}_t \mathbf{v}_t'], \quad \mathbf{\Lambda}_{uv} = \sum_{k=1}^{\infty} E[u_{t-k} \mathbf{v}_t],$$

$$\mathbf{\Lambda}_{vu} = \sum_{k=1}^{\infty} E[\mathbf{v}_{t-k} u_t] \quad \text{and} \quad \mathbf{\Lambda}_{vv} = \sum_{k=1}^{\infty} E[\mathbf{v}_t \mathbf{v}_{t-k}'].$$

(A3) $E(u_t) = 0$, $E|u_t|^4 = \kappa < \infty$ and u_t is independent of $\mathcal{F}_{t-1}^{qu}$ where $\mathcal{F}_t^{qu} = \sigma(q_{t-j}, u_{t-j}; j \geq 0)$.
(A4) The threshold parameter γ is such that $\gamma \in \Gamma = [\gamma_L, \gamma_U]$ a closed and bounded subset of the sample space of the threshold variable.
(A5) The p-dimensional $I(1)$ vector $\mathbf{x}_t$ is not cointegrated.

Assumptions (A1) and (A3) above are needed for establishing a limit theory for an empirical process such as $\sum_{t=1}^{T} u_t I(q_{t-1} \leq \gamma)/\sqrt{T}$ and are similar to the assumptions in Caner and Hansen (2001). Assumption (A2) is a multivariate invariance principle for the partial sum process constructed from $\mathbf{w}_t$ and is standard in the literature. Note, for instance, that (A2) is implied by (A1) provided that some moment requirements are imposed on u_t and $\mathbf{v}_t$. More specifically, under $E(u_t) = 0$, $E(\mathbf{v}_t) = 0$ and the requirement that $E|u_t|^{2+\rho} < \infty$, $E|v_{it}|^{2+\rho} < \infty$ and $E|u_t v_{it}|^{2+\rho} < \infty$ for some $\rho > 0$ and $i = 1, \ldots, p$, the multivariate invariance principle follows directly from (A1). Although the finite fourth moment assumption we are making in (A3) is not strictly necessary for the invariance principle, it is needed here for establishing the tightness properties of the above-mentioned empirical process. It is important to note at this stage that the particular structure of the long-run covariance matrix Ω restricts the nature of the temporal dependence displayed by u_t. Although it allows for the endogeneity of $\mathbf{x}_t$ and serial correlation in $\mathbf{v}_t$, it rules out the possibility of serial correlation in u_t. The treatment of serial correlation in u_t or more generally the presence of dependent errors in nonlinear models with $I(1)$ variables is a new and technically challenging area for which few results are available in the literature (see De Jong, 2002). We relegate the treatment of this case to the end of this section, where we present a new set of results covering a range of relevant scenarios allowing u_t to be a serially correlated process. It is also important to note at this stage that even under the more restrictive framework described by (A2), the properties of tests on the parameters underlying equation (1) turn out to be substantially different from what is commonly documented in the linear cointegration literature. An assumption similar to (A2)–(A3) above is also made in Chang *et al.* (2001), for instance, where the author's goal was to explore the asymptotic properties of parameter estimators obtained from a nonlinear cointegrating regression.

© Blackwell Publishing Ltd 2006

It is also important to highlight the fact that assumption (A1) requires the variable triggering the regime switches to be stationary, while at the same time allowing it to follow a very rich class of stochastic processes. It rules out, however, the possibility of q_t being $I(1)$ itself. Assumption (A4) is standard in the literature on threshold models (see, for instance, Hansen, 1996, 1999, 2000; Gonzalo and Pitarakis, 2002, 2006a, b). The threshold variable sample space Γ is typically taken to be $[\gamma_L, \gamma_U]$, with γ_L and γ_U chosen such that $P(q_{t-d} \leq \gamma_L) = \theta_1 > 0$ and $P(q_{t-d} \leq \gamma_U) = 1 - \theta_1$. The choice of θ_1 is commonly taken to be 10% or 15%. Restricting the parameter space of γ in this fashion ensures that there are enough observations in each regime and also guarantees the existence of nondegenerate limits for the test statistics of interest. Assumption (A5) rules out the possibility that the components of the p-vector $\mathbf{x}_t$ being themselves cointegrated. It can be seen as equivalent to requiring that the long-run covariance matrix Ω_{vv} be positive definite.

Given the above assumptions, our key objective is to next obtain the limiting distribution of the SupLM statistic under the null hypothesis $H_0 : \boldsymbol{\lambda} = 0$. In analogy with the linear cointegration framework in the context of testing hypotheses on cointegrating vectors, our analysis will distinguish across various scenarios about the possible endogeneity of regressors. Taking Ω_{uv} and Ω_{vu} different from zero in Ω, for instance, allows for $\mathbf{x}_t$ to be endogenous in equation (1), while setting them equal to zero forces strict exogeneity of regressors. In what follows we make use of the equality $I(q_{t-d} \leq \gamma) = I(F(q_{t-d}) \leq F(\gamma))$, which allows us to use uniformly distributed random variables (see Caner and Hansen, 2001, p. 1586). In this context, we let $\theta \equiv F(\gamma) \in \Theta$ with $\Theta = [\theta_1, 1 - \theta_1]$ and throughout this paper we will be using θ and $F(\gamma)$ interchangeably.

The next propositions present the limiting distributions of the test statistic of interest under alternative assumptions on Ω. In what follows, the quantity $W_u(r, \theta)$ will denote a scalar *standard* Brownian sheet (two parameter standard Brownian motion) indexed by $[0, 1]^2$, which is a zero-mean Gaussian process with covariance $(r_1 \wedge r_2)(\theta_1 \wedge \theta_2)$, while $B_u(r, \theta)$ denotes a Brownian sheet with covariance $\sigma_u^2(r_1 \wedge r_2)(\theta_1 \wedge \theta_2)$. Similarly, the quantity $K_u(r, \theta)$ will denote a *standard* Kiefer process defined as $K_u(r, \theta) = W_u(r, \theta) - \theta W_u(r, 1)$. A Kiefer process on $[0, 1]^2$ is also a Gaussian process with zero mean and covariance function $(r_1 \wedge r_2)$ $(\theta_1 \wedge \theta_2 - \theta_1\theta_2)$. Finally, in analogy to $\mathbf{B}_v(r)$ introduced in (A2), we also let $\mathbf{W}_v(r)$ denote a p-dimensional standard Brownian motion.

Proposition 1 initially concentrates on the case where there is no regressor endogeneity in equation (1) by setting $\Omega_{uv} = \Omega_{vu} = 0$ in the expression of Ω in (A2).

© Blackwell Publishing Ltd 2006

Proposition 1. Under the null hypothesis $H_0 : \boldsymbol{\lambda} = 0$ and Assumptions (A1)–(A5) with $\boldsymbol{\Omega}_{uv} = \boldsymbol{\Omega}_{vu} = 0$, we have

$$\text{SupLM} \Rightarrow \sup_{\theta} \frac{1}{\theta(1-\theta)} \left[\int_0^1 \mathbf{W}_v(r) \mathrm{d}K_u(r,\theta) \right]' \left[\int_0^1 \mathbf{W}_v(r)\mathbf{W}_v(r)' \mathrm{d}r \right]^{-1} \times \left[\int_0^1 \mathbf{W}_v(r) \mathrm{d}K_u(r,\theta) \right] \tag{6}$$

where $K_u(r, \theta) = W_u(r, \theta) - \theta W_u(r, 1)$ is a standard Kiefer process and $W_u(r, \theta)$ a standard Brownian sheet.

The above proposition establishes the limiting distribution of the SupLM test statistic for testing the null hypothesis of linear cointegration against threshold cointegration. Although it is free of nuisance parameters, making it directly operational, it is important to emphasize the fact that the formulation in equation (6) rules out the presence of regressor endogeneity. It is very interesting to observe, however, that because of the independence of $W_u(r)$ and $\mathbf{W}_v(r)$, the Kiefer process given by $K_u(r, \theta)$ is also independent of $\mathbf{W}_v(r)$. As $K_u(r, \theta)/\sqrt{\theta(1-\theta)}$ is itself a standard Brownian motion (given θ), writing $\tilde{W}_u(r) \equiv K_u(r, \theta)/\sqrt{\theta(1-\theta)}$, we can reformulate the limiting random variable in the right-hand side of equation (6) as

$$\left[\int_0^1 \mathbf{W}_v(r) \mathrm{d}\tilde{W}_u(r,\theta) \right]' \left[\int_0^1 \mathbf{W}_v(r)\mathbf{W}_v(r)' \mathrm{d}r \right]^{-1} \left[\int_0^1 \mathbf{W}_v(r) \mathrm{d}\tilde{W}_u(r,\theta) \right]$$

from which it is easy to infer its equivalence to a $\chi^2(p)$ random variable (see Park and Phillips, 1988). In effect, the above result is directly analogous to the *standard asymptotic phenomenon* obtained in the linear cointegration literature when testing hypotheses about cointegrating vectors, except that here the analogy is with the distributional theory obtained in the stationary nonlinear framework as in Hansen (1996) and where the limiting random variable is typically referred to as a χ^2 *process*, indexed by a nuisance parameter such as θ. It is also interesting to note here that the limiting distribution of the SupLM test statistic presented in Proposition 1 is equivalent to a random variable given by the supremum of a squared normalized Brownian bridge process, say $[\mathbf{W}(\theta) - \theta\mathbf{W}(1)]' \times [\mathbf{W}(\theta) - \theta\mathbf{W}(1)]/\theta(1-\theta)$, with $\mathbf{W}(\cdot)$ denoting a p-dimensional standard Brownian motion. For inference purposes, critical values for the above distribution are readily available (see Andrews, 1993, Table 1, p. 840 or Hansen, 1997).

We next consider the case where regressor endogeneity is allowed within the cointegrating regression (1). The limiting distribution of the SupLM test statistic under this scenario is summarized in Proposition 2.

© Blackwell Publishing Ltd 2006

Proposition 2. Under the null hypothesis $H_0 : \boldsymbol{\lambda} = 0$ and Assumptions (A1)–(A5) we have

$$\text{SupLM} \Rightarrow \sup_{\theta} \frac{1}{\theta(1-\theta)} \left[\int_0^1 \mathbf{B}_v(r)\mathrm{d}\tilde{K}_u(r,\theta)\right]' \left[\int_0^1 \mathbf{B}_v(r)\mathbf{B}_v(r)'\mathrm{d}r\right]^{-1} \times \left[\int_0^1 \mathbf{B}_v(r)\mathrm{d}\tilde{K}_u(r,\theta)\right] \quad (7)$$

where $\tilde{K}_u(r, \theta) = (B_u(r, \theta) - \theta B_u(r, 1))/\sigma_u$ with $B_u(r, \theta)$ denoting a Brownian Sheet with variance $\sigma_u^2 r\theta$.

At this stage, it is interesting to note the key difference between the result in Proposition 2 and the typical result that is documented in the linear cointegration literature. An important difference in the present framework arises from the fact that although $B_u(r)$ and $\mathbf{B}_v(r)$ are not independent since $\Omega_{uv} \neq 0$, we have

$$\begin{aligned} E[\tilde{K}_u(r,\theta)\mathbf{B}_v(r)] &= (E[B_u(r,\theta)\mathbf{B}_v(r)] - \theta E[B_u(r,1)\mathbf{B}_v(r)])/\sigma_u \\ &= (r\theta\Omega_{uv} - \theta r\Omega_{uv})/\sigma_u = 0 \end{aligned}$$

and in effect

$$\int_0^1 \mathbf{B}_v(r)\mathrm{d}\tilde{K}_u(r,\theta) \approx N(0, \text{var}[\tilde{K}_u(r,\theta)]\mathbf{G})\mathrm{d}P(\mathbf{G}) \text{ with } \mathbf{G} = \int_0^1 \mathbf{B}_v(r)\mathbf{B}_v(r)'\mathrm{d}r.$$

Since $\text{var}[\tilde{K}_u(r, \theta)] = r\theta(1-\theta)$, it follows that the SupLM statistic will have a limiting distribution that is free of nuisance parameters despite the presence of endogeneity, and the limiting distributions obtained under Propositions 1 and 2 are the same.

2.3. Extensions

Deterministic components

So far, our results in Propositions 1 and 2 are based on the specification of a cointegrating regression without an intercept term. It is algebraically straightforward to generalize all our results to the case where the model in equation (1) is allowed to contain an intercept term by focusing on a specification given by $y_t = \beta_0 + \boldsymbol{\beta}'\mathbf{x}_t + (\lambda_0 + \boldsymbol{\lambda}'\mathbf{x}_t)I(q_{t-d} > \gamma) + u_t$. In this context, the regressor matrices $\mathbf{X}$ and $\mathbf{X}_\gamma$ in equation (3) would stack 1 and $\mathbf{x}_t$, and $I(q_{t-d} > \gamma)$ and $\mathbf{x}_t I(q_{t-d} > \gamma)$, respectively. It is then straightforward to establish that both our results in Propositions 1 and 2 continue to hold provided that the Brownian motions $\mathbf{W}_v(r)$ and $\mathbf{B}_v(r)$ are replaced with $\mathbf{W}_v^*(r) = (1, \mathbf{W}_v(r)')'$ and $\mathbf{B}_v^*(r) = (1, \mathbf{B}_v(r)')'$, respectively.

Serial correlation in u_t

Given the structure of the long-run covariance matrix we operated under, with our Assumption (A2), our results in Propositions 1 and 2 ruled out the

© Blackwell Publishing Ltd 2006

possibility of u_t being serially correlated. This assumption allowed us to appeal directly to the weak convergence results involving quantities such as

$$G_{uT}(r,\theta)=\sum_{t=1}^{[\mathrm{Tr}]}u_tI(F(q_{t-d})\leq\theta)/\sqrt{T}$$

established in Caner and Hansen (2001) and to evaluate the long-run variance and asymptotic covariance kernel of related quantities in a straightforward manner. In what follows and for notational simplicity we set $d=1$ and since $I(q_{t-1}\leq\gamma)=I(F(q_{t-1})\leq F(\gamma))$, with no loss of generality we proceed taking q_t as a uniform random variable and let $\theta\equiv F(\gamma)$. We also let $I_t(\theta)\equiv I(F(q_t)\leq\theta)$.

Although the treatment of a serially correlated u_t within a linear cointegration framework is straightforward, handling the same setup within our specification in equation (1) becomes much more involved. The main complications come from the need to obtain a functional central limit theorem (FCLT) for a marked empirical process such as $G_{uT}(r,\theta)$ in which *both* the marks u_t as well as q_t are possibly correlated general stationary processes (see Caner and Hansen, 2001, Theorem 1, for the case of u_t following an independent and identially distributed i.i.d process). To our knowledge, formal FCLTs for such a case are not readily available in either the econometrics or the statistics literature. The statistics literature on empirical processes for instance has considered such processes, typically under an i.i.d. or martingale difference sequence assumption for the marks (see Koul, 1996; Koul and Stute, 1999; Stute, 1997 and references therein). Although the literature on linear processes offers a valuable tool such as the Beveridge and Nelson (BN) decomposition for handling the case of weakly dependent errors in regression models via the choice of processes such as $u_t=C(L)e_t$ with e_t an i.i.d. or a martingale difference sequence and $C(L)=\sum_{j=0}^{\infty}c_jL^j$ (see Phillips and Solo, 1992), the same decomposition applied to a nonlinear process such as $G_{uT}(r,\theta)$ does not lead to a convenient formulation for establishing an FCLT type of result via the use of a CLT for martingale differences, for instance. Indeed, using the BN decomposition of u_t together with summation by parts, we can write

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}u_tI_{t-1}=\frac{1}{\sqrt{T}}C(1)\sum_{t=1}^{T}e_tI_{t-1}+\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\tilde{e}_t\Delta I_t+o_p(1),\qquad(8)$$

where

$$\tilde{e}_t=\tilde{C}(L)e_t=\sum_{j=0}^{\infty}\tilde{c}_je_{t-j}\quad\text{with }\tilde{c}_j=\sum_{k=j+1}^{\infty}c_k.$$

Clearly, had the second term in the right hand-side of equation (8) vanished asymptotically, the use of the results in Caner and Hansen (2001) would have

© Blackwell Publishing Ltd 2006

been sufficient to treat the case of serial correlation in u_t. Note also that the two components in the right-hand side of equation (8) are not independent and the second one clearly does not form a martingale difference sequence since $\tilde{e}_t$ is another weakly dependent process. It could also have been possible to take an alternative approach whereby one uses our Assumption (A1) to argue that $u_t I_{t-1}$ is also strong mixing. Finite dimensional convergence would then follow through the use of a CLT for strongly mixing sequences. Unfortunately, in this context, it becomes difficult to establish the tightness of $G_{uT}(r, \theta)$ because of the lack of an appropriate inequality equivalent to Rosenthal's inequality available for martingale differences. To our knowledge, the existing statistical toolkit and the literature on moment inequalities for strongly mixing processes, in particular, does not presently allow a full treatment of the asymptotics of objects such as $G_{uT}(r, \theta)$ under more general assumptions. Shao and Yu (1996) present a set of useful moment inequalities for strong mixing sequences but they are not useful for dealing with the special case of the above empirical process.

In order to deal with the case of serial correlation in u_t and to allow us to obtain a tractable limit theory for handling quantities such as $G_{uT}(r, \theta)$ above, our generalization will involve taking u_t to follow a linear moving average process of finite order ℓ. More formally, we introduce the following assumption about the behaviour of the u_t sequence which now replaces our earlier assumption (A3):

(B1) $u_t = \Psi_\ell(L)e_t$ where $\Psi_\ell(L) = \sum_{j=0}^{\ell} \Psi_j L^j$ with $\Psi_0 = 1$, $E(e_t) = 0$, $E(e_t^2) = \sigma_e^2$, $E|e_t|^4 < \infty$ and e_t is independent of $\mathcal{F}_{t-1}^{q_e} = \sigma(q_{t+\ell-j}, e_{t-j}; j \geq 1)$.

We also let ω_u^2 denote the long-run variance of u_t which given our chosen process can be written as $\omega_u^2 = \sigma_e^2 \Psi_\ell(1)^2$. In what follows we will also refer to Assumption (A2′) as (A2) but with the stochastic behaviour of u_t replaced with that in (B1). Although we allow for serial correlation in u_t as well as in q_t, in order for us to obtain a functional CLT for $G_{uT}(r, \theta)$ it is crucial to restrict the dependence structure between the error process driving the cointegrating regression and the threshold variable.

Our next goal is to obtain an FCLT type of result for our nonlinear process and subsequently generalize our earlier propositions to cover the present framework.

Proposition 3. Letting

$$G_{uT}(r, \theta) = \sum_{t=1}^{[\mathrm{Tr}]} u_t I_{t-1} / \sqrt{T}$$

and under Assumptions (A1), (B1) and (A4) we have $G_{uT}(r, \theta) \Rightarrow G_u(r, \theta)$ on $(r, \theta) \in [0, 1]^2$ as $T \to \infty$, where $G_u(r, \theta)$ is a zero-mean Gaussian process with covariance kernel

© Blackwell Publishing Ltd 2006

$$\omega_G^2(r_1, r_2, \theta_1, \theta_2) = (r_1 \wedge r_2)\sigma_e^2 E\left[\left(\sum_{j=0}^{\ell} \Psi_j I_{t-1+j}(\theta_1)\right)\left(\sum_{j=0}^{\ell} \Psi_j I_{t-1+j}(\theta_2)\right)\right].$$

Our result in Proposition 3 specializes to Theorem 1 of Caner and Hansen (2001) if we set $\Psi_j = 0$ for $j \geq 1$, because this corresponds to the case where the marks of the empirical process are i.i.d. Indeed, from the expression of $\omega_G^2(r_1, r_2, \theta_1, \theta_2)$ above, we obtain $\omega_G^2(r_1, r_2, \theta_1, \theta_2) = \sigma_e^2(r_1 \wedge r_2)(\theta_1 \wedge \theta_2)$ which can be recognized as the covariance kernel of a standard Brownian sheet. We now have Proposition 4.

Proposition 4. Under the null hypothesis $H_0 : \boldsymbol{\lambda} = 0$ and Assumptions (A1), (A2′), (B1), (A4) and (A5) we have

$$\text{SupLM} \Rightarrow \sup_{\theta} \frac{1}{\theta(1-\theta)}\left[\int_0^1 \mathbf{B}_v(r)\mathrm{d}\tilde{M}_u(r,\theta)\right]'\left[\int_0^1 \mathbf{B}_v(r)\mathbf{B}_v(r)'\mathrm{d}r\right]^{-1} \times \left[\int_0^1 \mathbf{B}_v(r)\mathrm{d}\tilde{M}_u(r,\theta)\right] \quad (9)$$

where $\tilde{M}_u(r,\theta) = (G_u(r,\theta) - \theta G_u(r,1))/\sigma_u$.

Unlike Propositions 1 and 2 where the limiting distributions were free of nuisance parameters, the limiting process in equation (9) clearly depends on model-specific parameters such as the $\Psi_j's$. This occurrence is in fact the norm in this literature where it is well documented that universal tabulations of distributions cannot be obtained. In this sense, it is remarkable that in the context of our Propositions 1 and 2, our particular specification led to tractable asymptotics free of nuisance parameters. In general, however, this is not the case and the common approach for conducting inferences involves using bootstrap methods to approximate the null distribution of the test statistic. It is beyond the scope of this paper to develop new bootstrapping techniques that can accomodate the above framework. It could be an interesting extension to follow the approach proposed in Caner and Hansen (2001) for this purpose; to our knowledge, however, the validity of the bootstrap has not been established in this literature.

In the context of our results in Proposition 4, however, an important simplification occurs under the additional assumption that q_t follows an i.i.d. process. Under this particular case, it is straightforward to write the long-run variance of $G_u(r,\theta)$ as $\omega_G^2(r,\theta) = r[\theta^2\omega_u^2 + \theta(1-\theta)\sigma_u^2]$. It is then easy to see that the process $\tilde{M}_u(r,\theta)$ in equation (9) is such that $\tilde{M}_u(r,\theta) \equiv K_u(r,\theta)$ and the limiting distribution in equation (9) reduces to that presented in Propositions 1 and 2. We summarize the results pertaining to this particular scenario in the following proposition.

© Blackwell Publishing Ltd 2006

Proposition 5. When q_t is an i.i.d. process and u_t is as in (B1), the limiting distribution in equation (9) is identical to the one obtained in Propositions 1 and 2.

III. Adequacy of asymptotic approximations and finite sample performance

We initially illustrate the key features of the distributions presented in Propositions 1 and 2 through the simulation of a wide range of models under the null hypothesis of linear cointegration. Our first goal is to highlight the robustness of the limiting distribution of the SupLM test statistic to regressor endogeneity using a large sample size such as $T = 1{,}000$ and comparing the resulting critical values.

Our general data-generating process (DGP) is given by $y_t = \beta_0 + \beta_1 x_t + u_t$ with $\Delta x_t = v_t$ and $v_t = \rho v_{t-1} + \epsilon_t$, while the fitted model is given by $y_t = \beta_0 + \beta_1 x_t + (\lambda_0 + \lambda_1 x_t) I(q_{t-1} > \gamma) + u_t$. We take q_t to follow the AR(1) process $q_t = \phi q_{t-1} + v_t$ and $u_t = \text{n.i.d.}(0, 1)$. The null hypothesis of interest is $H_0 : \lambda_0 = \lambda_1 = 0$. Regarding the covariance structure of the random disturbances, letting $z_t = (u_t, \epsilon_t, v_t)'$ and $\mathbf{\Sigma}_z = E[z_t z_t']$, we use

$$\mathbf{\Sigma}_z = \begin{pmatrix} 1 & \sigma_{u\epsilon} & \sigma_{uv} \\ \sigma_{\epsilon u} & 1 & \sigma_{\epsilon v} \\ \sigma_{uv} & \sigma_{\epsilon v} & 1 \end{pmatrix}$$

which allows for a sufficiently general covariance structure while imposing unit variances. Our covariance matrix parameterization allows the threshold variable to be contemporaneously correlated with the random shocks hitting the cointegrating regression and its regressors. All our experiments are based on $N = 2{,}000$ replications and set $\{\beta_0, \beta_1\} = \{1, 2\}$ throughout. In order to document the robustness of our results to the degree of persistence in the threshold variable q_t, our experiments are also conducted using $\phi = \{0.5, 0.9\}$. Since our initial motivation is to explore the robustness of the limiting distributions to the presence/absence of endogeneity, we consider the following DGPs. $\text{DGP}_1 : \{\sigma_{u\epsilon}, \sigma_{uv}, \sigma_{\epsilon v}\} = \{0.3, 0.7, 0.6\}$, $\text{DGP}_2 : \{\sigma_{u\epsilon}, \sigma_{uv}, \sigma_{\epsilon v}\} = \{0.3, 0.0, 0.6\}$ and $\text{DGP}_3 : \{\sigma_{u\epsilon}, \sigma_{uv}, \sigma_{\epsilon v}\} = \{0.0, 0.0, 0.0\}$. Note that DGP_1 and DGP_2 allow for the endogeneity of $\mathbf{x}_t$ while DGP_3 takes it as strictly exogenous. The implementation of the SupLM test also assumes 10% trimming at each end of the sample.

Table 1 presents a range of quantiles of the simulated limiting distribution approximated by the use of a sample of size $T = 1{,}000$.

The critical values tabulated in Table 1 support the asymptotic-based result that the limiting distributions are robust to the presence of endogeneity as well

© Blackwell Publishing Ltd 2006

TABLE 1

Asymptotic critical values

	$\phi = 0.5$				$\phi = 0.9$			
	10%	*5%*	*2.5%*	*1%*	*10%*	*5%*	*2.5%*	*1%*
DGP_1, $\rho = 0.0$, $T = 1{,}000$	10.28	12.12	13.85	15.97	10.75	12.35	14.04	15.43
DGP_2, $\rho = 0.0$, $T = 1{,}000$	10.15	12.09	13.74	15.60	10.66	12.37	13.95	15.94
DGP_3, $\rho = 0.0$, $T = 1{,}000$	10.38	12.10	13.38	15.78	10.34	12.58	14.00	15.99
DGP_1, $\rho = 0.4$, $T = 1{,}000$	10.34	12.00	13.81	15.91	10.33	12.08	13.51	15.05
DGP_2, $\rho = 0.4$, $T = 1{,}000$	10.19	11.73	13.70	15.46	10.64	12.36	13.94	15.98
DGP_3, $\rho = 0.4$, $T = 1{,}000$	10.34	11.95	13.36	15.76	10.37	12.60	13.98	16.02
Andrews	10.50	12.27	NA	16.04	10.50	12.27	NA	16.04

as serial correlation in $\mathbf{v}_t$. Comparing the tabulated critical values with the process tabulated in Andrews (1993), it is also clear that the documented equivalence of the limiting distributions is supported by our $T = 1{,}000$ based simulations. Looking at the tabulations across the two magnitudes of ϕ, it also appears that the limiting distributions are reasonably well approximated even when the threshold variable is highly persistent. It is also useful to recall that Hansen (1997) provided approximate asymptotic p-values for the same distribution as Andrews (1993) and that they can easily be used in the present context. In what follows, therefore, our inferences will be based on the p-values of the computed SupLM statistic obtained through the tabulations in Hansen (1997).

We next aim to document the size properties of the SupLM test statistic across smaller sample sizes. For this purpose, we simulate the same DGPs as in Table 1 using samples of size $T = 200$ and $T = 400$ and document the number of times the null hypothesis is rejected using the asymptotic-based critical values. Results for this experiment are presented in Table 2 below. Looking at the first column of Table 2, the reported empirical size estimates

TABLE 2

Empirical size estimates

	$\phi = 0.5$			$\phi = 0.9$		
Nominal	*5%*	*2.5%*	*1%*	*5%*	*2.5%*	*1%*
DGP_1, $\rho = 0.4$, $T = 200$	5.00%	2.60%	0.95%	4.55%	2.15%	0.85%
DGP_1, $\rho = 0.4$, $T = 400$	4.95%	2.45%	0.96%	4.20%	2.00%	1.00%
DGP_2, $\rho = 0.4$, $T = 200$	4.65%	2.45%	0.80%	4.90%	2.35%	0.85%
DGP_2, $\rho = 0.4$, $T = 400$	4.75%	2.45%	0.75%	4.35%	1.70%	0.70%
DGP_3, $\rho = 0.4$, $T = 200$	4.50%	2.30%	0.75%	3.85%	2.10%	0.70%
DGP_3, $\rho = 0.4$, $T = 400$	4.70%	2.70%	1.20%	3.98%	1.90%	0.60%

© Blackwell Publishing Ltd 2006

match their nominal counterparts reasonably well even for samples such as $T = 200$. Recall that our size estimates have been computed using Hansen's (1997) p-value approximations which are not exact p-values but obtained through simulations instead. As the degree of persistence of the threshold variable is allowed to increase, we note a slight deterioration in the size properties of the test in small to moderate samples. The reported figures suggest that the direction of the distortions depends on the DGP parameters (e.g. covariance structure) in a complicated manner with some tendency towards undersizeness. Overall, however, the documented size properties of our test compare favourably with those reported in the existing literature (see, for instance, Table II in Hansen, 1996).

Our next set of experiments are designed to highlight the finite sample power properties of the SupLM test statistic. For this purpose, we consider the DGP $y_t = \beta_0 + \beta_1 x_t + (\lambda_0 + \lambda_1 x_t)I(q_{t-1} > \gamma_0) + u_t$ with q_t, u_t and v_t maintained as in our size simulations. We also maintain the earlier parameterization for β_0, β_1 and ϕ as $\{\beta_0, \beta_1\} = \{1, 2\}$ and $\phi = \{0.5, 0.9\}$. Results are summarized in Table 3, which concentrates on the parameterization given by DGP_3 across alternative magnitudes for λ_0 and λ_1. Results for DGP_1 and DGP_2 were virtually identical and are therefore omitted.

The empirical power estimates presented in Table 3 suggest that the SupLM-based test has good power properties for detecting the presence of threshold effects in cointegrating regressions. Even under an alternative that is very close to the null hypothesis (e.g. $\lambda_0 = \lambda_1 = 0.05$), the test is able to reject the null approximately 85% of the times under $T = 400$. These figures also compare favourably with the power properties of the SupLM-type test statistics documented in the literature in the context of stationary regression frameworks such as testing for SETAR-type nonlinearities (see Hansen, 1996). Experiments with different magnitudes of γ_0 (e.g. $\gamma_0 = 0.5$) led to virtually identical results and are omitted.

TABLE 3

Empirical power estimates

	$\phi = 0.5,\ \gamma_0 = 0$			$\phi = 0.9,\ \gamma_0 = 0$		
Nominal	*5%*	*2.5%*	*1%*	*5%*	*2.5%*	*1%*
DGP_3, $\lambda_0 = \lambda_1 = 0.05$, $T = 200$	48.90	41.40	34.90	46.25	38.80	31.20
DGP_3, $\lambda_0 = \lambda_1 = 0.05$, $T = 400$	84.90	81.10	76.15	83.55	79.05	73.70
DGP_3, $\lambda_0 = \lambda_1 = 0.05$, $T = 800$	99.15	98.90	98.15	99.20	98.75	98.15
DGP_3, $\lambda_0 = \lambda_1 = 0.15$, $T = 200$	95.55	93.90	92.10	94.00	92.00	89.25
DGP_3, $\lambda_0 = \lambda_1 = 0.15$, $T = 400$	99.95	99.85	99.65	99.95	99.00	99.85
DGP_3, $\lambda_0 = \lambda_1 = 0.15$, $T = 800$	100.00	100.00	100.00	100.00	100.00	100.00

© Blackwell Publishing Ltd 2006

TABLE 4

Asymptotic critical values under MA cointegrating errors

	10%	*5%*	*2.5%*	*1%*
DGP_4, $T = 1{,}000$	10.35	11.74	13.41	15.07
DGP_5, $T = 1{,}000$	10.42	11.97	13.46	15.42

Our last set of experiments are designed to illustrate our analysis of the case where u_t follows an MA process. More specifically, we concentrate on the particular scenario where the limiting distributions of the SupLM statistic were shown to be unaffected by the presence of serial correlation in u_t. For this purpose, we let u_t follow an MA(1) process given by $u_t = e_t - 0.5e_{t-1}$ with $e_t =$ n.i.d.(0, 1). The process for v_t is chosen as an AR(1) with $\rho = 0.4$. Our chosen covariance structure allows for endogeneity while forcing q_t and e_t to be independent with $q_t =$ n.i.d.(0, 1). Replacing z_t above with $z_t = (e_t, \epsilon_t, q_t)$, the covariance matrix $\mathbf{\Sigma}_z$ now has $\sigma_{e\epsilon} = 0.5$, $\sigma_{eq} = 0$, and $\sigma_{\epsilon q} = 0.6$ with unit variances. Thus although we allow q_t to be correlated with the shock hitting the regressor x_t, it is independent of u_t. We refer to this parameterization as DGP_4 and also consider the case where $\mathbf{\Sigma}_z = I$, referring to it as DGP_5. Results for this set of experiments are presented in Table 4.

Comparing the critical values obtained in Table 4 with the ones in Table 1 clearly illustrates empirically our point in Proposition 5 where we established that even under a serially correlated error process there are instances where the limiting distribution of the LM statistic remains free of nuisance parameters and is identical to the one obtained in Propositions 1 and 2.

IV. Conclusions

The goal of this paper was to introduce a test statistic designed to detect the presence of threshold effects within equilibrium relationships. Operating within the framework of a cointegrating regression and allowing for the presence of endogeneity and some degree of temporal dependence, we obtained the asymptotic distribution of an LM-based test statistic designed to test the null hypothesis of linear cointegration against threshold cointegration.

Our work assumed cointegration under both the null and alternative hypotheses. This is common in this literature, but a very useful albeit challenging extension would be to develop a test for the presence of cointegration that is robust to the presence or absence of nonlinear components such as the threshold effects analysed here. Although the framework under which we operated is quite general, there is also considerable scope for

© Blackwell Publishing Ltd 2006

extension by exploring the effect of including deterministic trend components and additional stationary regressors within the specification in equation (1). Perhaps, more importantly, a general estimation and inference theory dealing with the presence of threshold-type nonlinearities in both the long-run equilibrium relationship(s) and the adjustment process to equilibrium is still not completed. These issues are currently being investigated by the authors.

References

Andrews, D. W. K. (1993). 'Tests for parameter instability and structural change with unknown change point', *Econometrica*, Vol. 61, pp. 821–856.

Andrews, D. W. K. and Ploberger, W. (1994). 'Optimal tests when a nuisance parameter is present only under the alternative', *Econometrica*, Vol. 62, pp. 1383–1414.

Balke, N. and Fomby, T. (1997). 'Threshold cointegration', *International Economic Review*, Vol. 38, pp. 627–645.

Campbell, J. Y. and Shiller, R. J. (1987). 'Cointegration and tests of present value models', *Journal of Political Economy*, Vol. 95, pp. 1062–1088.

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

Chang, Y., Park, J. Y. and Phillips, P. C. B. (2001). 'Nonlinear econometric models with cointegrated and deterministically trending regressors', *Econometrics Journal*, Vol. 4, pp. 1–36.

Choi, I. and Saikkonen, P. (2004). 'Testing linearity in cointegrating smooth transition regressions', *Econometrics Journal*, Vol. 7, pp. 341–346.

Davies, R. B. (1977). 'Hypothesis testing when a nuisance parameter is present only under the alternative', *Biometrika*, Vol. 64, pp. 247–254.

Davies, R. B. (1987). 'Hypothesis testing when a nuisance parameter is present only under the alternative', *Biometrika*, Vol. 74, pp. 33–43.

De Jong, R. (2002). *Nonlinear Estimators with Integrated Regressors but without Exogeneity*, Manuscript, Department of Economics, Michigan State University – East Lansing,

Gonzalo, J. and Pitarakis, J. (2002). 'Estimation and model selection based inference in single and multiple threshold models', *Journal of Econometrics*, Vol. 110, pp. 319–352.

Gonzalo, J. and Pitarakis, J. (2005). *Threshold Effects in Cointegrating Relationships*, University of Southampton Discussion Papers in Economics and Econometrics No. 0506.

Gonzalo, J. and Pitarakis, J. (2006a). 'Threshold effects in multivariate error correction models', in Mills T. C. and Patterson K. (eds), *Palgrave Handbook of Econometrics: Econometric Theory* Volume 1, Palgrave MacMillan, New York, NY, pp. 578–609.

Gonzalo, J. and Pitarakis, J. (2006b). 'Estimation of common long-memory components in threshold cointegrated systems', Manuscript, Department of Economics, University of Southampton, Southampton.

Granger, C. W. J. and Terasvirta, T. (1993). *Modelling Nonlinear Economic Relationships*, Oxford University Press, Oxford.

Granger, C. W. J. (1995). 'Modelling nonlinear relationships between extended memory variables', *Econometrica*, Vol. 63, pp. 265–279.

Gregory, A. W. and Hansen, B. E. (1996). 'Residual-based tests for cointegration in models with regime shifts', *Journal of Econometrics*, Vol. 70, pp. 99–126.

Hansen, B. E. (1992). 'Tests for parameter instability in regressions with I(1) processes', *Journal of Business and Economic Statistics*, Vol. 10, pp. 321–335.

© Blackwell Publishing Ltd 2006

Hansen, B. E. (1996). 'Inference when a nuisance parameter is not identified under the null hypothesis', *Econometrica*, Vol. 64, pp. 413–430.
Hansen, B. E. (1997). 'Approximate asymptotic p-values for structural change tests', *Journal of Business and Economic Statistics*, Vol. 15, pp. 60–67.
Hansen, B. E. (1999). 'Testing for linearity', *Journal of Economic Surveys*, Vol. 13, pp. 551–576.
Hansen, B. E. (2000). 'Sample splitting and threshold estimation', *Econometrica*, Vol. 68, pp. 575–603.
Hansen, B. E. and Seo, B. (2002). 'Testing for two-regime threshold cointegration in vector error correction models', *Journal of Econometrics*, Vol. 110, pp. 293–318.
Howitt, P. W. and McAffee, P. R. (1992). 'Animal spirits', *American Economic Review*, Vol. 82, pp. 493–507.
Kanbur, R. (2005). *Programme of the IFPRI/Cornell Conference on Threshold Nonlinearities in Growth and Development*. http://www.people.cornell.edu/pages/sk145/.
Kapetanios, G., Shin, Y. and Snell, A. (2004). *Testing for Cointegration in Nonlinear STAR Error Correction Framework*, Working Paper No. 497, Queen Mary University of London.
Koul, H. (1996). 'Asymptotics of some estimators and sequential residual empiricals in non-linear time series', *Annals of Statistics*, Vol. 24, pp. 380–404.
Koul, H. and Stute, W. (1999). 'Nonparametric model checks for time series', *Annals of Statistics*, Vol. 27, pp. 204–236.
Lettau, M. and Ludvigson, S. C. (2004). 'Understanding trend and cycle in asset values: reevaluating the wealth effect on consumption', *American Economic Review*, Vol. 94, pp. 276–299.
Ogaki, M., Jang, K. and Lim, H. L. (2005). *Structural Macroeconometrics*, manuscript, Department of Economics, Ohio State University, Ohio, USA.
Park, J. Y. and Phillips, P. C. B. (1988). 'Statistical inference in regressions with integrated processes: Part1', *Econometric Theory*, Vol. 4, pp. 468–497.
Park, J. Y. and Phillips, P. C. B. (1999). 'Asymptotics for nonlinear transformations of integrated time series', *Econometric Theory*, Vol. 15, pp. 269–298.
Park, J. Y. and Phillips, P. C. B. (2001). 'Nonlinear regressions with integrated time series', *Econometrica*, Vol. 69, pp. 1452–1498.
Petruccelli, J. D. (1992). 'On the approximation of time series by threshold autoregressive models', *Sankhya Series B*, Vol. 54, pp. 54–61.
Phillips, P. C. B. and Durlauf, S. N. (1986). 'Multiple time series regression with integrated processes', *Review of Economic Studies*, Vol. 53, pp. 473–495.
Phillips, P. C. B. and Solo, V. (1992). 'Asymptotics for linear processes', *Annals of Statistics*, Vol. 20, pp. 971–1001.
Seo, M. (2004). *Cointegration Test in the Threshold Cointegration Model*, Manuscript, Department of Economics, University of Wisconsin-Madison.
Shao, Q. M. and Yu, H. (1996). 'Weak convergence for weighted empirical processes of dependent sequences', *Annals of Probability*, Vol. 24, pp. 2098–2127.
Stute, W. (1997). 'Nonparametric model checks for regression', *Annals of Statistics*, Vol. 25, pp. 613–641.

Appendix

A more detailed presentation of all our proofs can be found in Gonzalo and Pitarakis (2005).

© Blackwell Publishing Ltd 2006

Proof of Proposition 1

Follows as a special case of the proof of Proposition 2 writing $B_u(r, \theta) = \sigma_u W_u(r, \theta)$, $\mathbf{B}_v(r) = \mathbf{\Omega}_{vv}^{\frac{1}{2}} \mathbf{W}_v(r)$ and letting $\mathbf{\Omega}_{uv} = 0$. ■

Proof of Proposition 2

Lemma 3.1 in Phillips and Durlauf (1986), Lemma A3 in Gonzalo and Pitarakis (2006a) and the continuous mapping theorem give

$$\mathbf{X}_\gamma' \mathbf{M} \mathbf{X}_\gamma / T^2 \Rightarrow \theta(1-\theta) \int_0^1 \mathbf{B}_v \mathbf{B}_v' \mathrm{d}r.$$

From Lemma 3.1 in Phillips and Durlauf (1986) we have

$$\frac{\mathbf{X}'\mathbf{u}}{T} \Rightarrow \int_0^1 \mathbf{B}_v(r) \mathrm{d}B_u(r, 1) + \mathbf{\Lambda}_{vu} + \mathbf{\Sigma}_{vu}.$$

From Theorem 2 in Caner and Hansen (2001) and standard manipulations, it follows that

$$\frac{\mathbf{X}_\gamma'\mathbf{u}}{T} \Rightarrow \int_0^1 \mathbf{B}_v(r) \mathrm{d}B_u(r, 1) - \int_0^1 \mathbf{B}_v \mathrm{d}B_u(r, \theta) + (1-\theta)\mathbf{\Lambda}_{vu} + (1-\theta)\mathbf{\Sigma}_{vu},$$

leading to

$$\frac{\mathbf{X}_\gamma'\mathbf{M}\mathbf{u}}{T\tilde{\sigma}_0} \Rightarrow -\int_0^1 \mathbf{B}_v(r) \mathrm{d}\tilde{K}_u(r, \theta).$$

Combining the above into the expression of the SupLM test statistic and the use of the continuous mapping theorem leads to the desired result. ■

Proof of Proposition 3

From

$$\sum_{t=1}^T u_t I_{t-1}(\theta) = \sum_{t=1}^T (\sum_{j=0}^{\ell} \Psi_j e_{t-j}) I_{t-1}(\theta),$$

using summation by parts we obtain

$$\frac{\sum_{t=1}^{[\mathrm{Tr}]} u_t I_{t-1}(\theta)}{\sqrt{T}} = \frac{\sum_{t=1}^{[\mathrm{Tr}]} e_t \delta_{t-1}(\theta)}{\sqrt{T}} + o_p(1) \quad \text{where } \delta_{t-1}(\theta) = \sum_{j=0}^{\ell} \Psi_j I_{t-1+j}(\theta).$$

For all θ, $\{e_t \delta_{t-1}(\theta), \mathcal{F}_{t-1}^{q_e}\}$ with $\mathcal{F}_{t-1}^{q_e} = \sigma(q_{t+\ell-j}, e_{t-j}; j \geq 0)$ is a strictly stationary and ergodic martingale difference sequence with variance

© Blackwell Publishing Ltd 2006

$E[e_t^2 \delta_{t-1}(\theta)^2]$. By the CLT for martingale difference sequences, it follows that for any (r, θ), $G_T(r, \theta) \xrightarrow{d} N(0, \omega_G^2(r, \theta))$ with $\omega_G^2(r, \theta) = rE[e_t^2 \delta_{t-1}(\theta)^2]$. It is also straightforward to see that the limiting covariance kernel is given by $\omega_G^2(r_1, r_2, \theta_1, \theta_2) = (r_1 \wedge r_2)\sigma_e^2 E[e_t^2 \delta_{t-1}(\theta_1)\delta_{t-1}(\theta_2)]$. Combined with the Cramer–Wold device the fidi convergence follows. Given our assumptions on e_t, the stochastic equicontinuity of $G_T(r, \theta)$ follows directly from Theorem 1 in Caner and Hansen (2001). ■

Proof of Proposition 4

The difference between the setup of Proposition 2 and the present framework is the treatment of the quantity $u'\bar{\mathbf{X}}_\gamma/T$ where $\bar{\mathbf{X}}_\gamma = \mathbf{X} * I(q \leq \gamma)$. It is a direct consequence of Theorem 2 in Caner and Hansen (2001) that we have

$$\frac{u'\bar{\mathbf{X}}_\gamma}{T} \Rightarrow \int_0^1 \mathbf{B}_v(r)\mathrm{d}G_u(r, \theta) + \mathbf{\Lambda}_{vu}^G + \mathbf{\Sigma}_{vu}^G$$

on $\theta \in [0, 1]$ and the above limiting process is almost surely continuous. Here

$$\mathbf{\Lambda}_{vu}^G = \sum_{k=1}^{\infty} E[\mathbf{v}_{t-k} u_t I_{t-1}] + \sum_{k=1}^{\infty} E[u_{t-k} I_{t-k-1} \mathbf{v}_t], \quad \mathbf{\Sigma}_{vu}^G = E[\mathbf{v}_t u_t I_{t-1}]$$

and given our assumptions we have $\mathbf{\Lambda}_{vu}^G = \theta\mathbf{\Lambda}_{vu}$ and $\mathbf{\Sigma}_{vu}^G = \theta\mathbf{\Sigma}_{vu}$. The rest of the proof follows as the proof of Proposition 2. ■

Proof of Proposition 5

Under q_t i.i.d. it can be shown that $G_u(r, \theta)$ is equivalent to $\theta B_u(r) + \sigma_u K_u(r, \theta)$. Plugging this in the expression of $\tilde{M}_u(r, \theta)$ in equation (9) gives $\tilde{M}_u(r, \theta) = K_u(r, \theta)$ and the result follows. ■

© Blackwell Publishing Ltd 2006