Threshold Effects in Cointegrating Relationships

Jesús Gonzalo
Universidad Carlos III de Madrid
Department of Economics
and
Jean-Yves Pitarakis
University of Southampton
Department of Economics

New Version: www.economics.soton.ac.uk/staff/pitarakis
Introduction: Context

- Two or more variables that are expected to be in equilibrium must be cointegrated.
- The concept of cointegration as originally defined referred solely to linear relationships.

\[ z_t = y_t - \beta x_t \]

- Economic relationships are often nonlinear.
- We explore the possibility that some variables may be linked through nonlinear equilibrium relationships (threshold type).

\[ z_t = y_t - \beta_1 x_t I(A) - \beta_2 x_t (1 - I(A)) \]
Introduction: Context

- No consensus on a formal definition of threshold or nonlinear cointegration.
- Existing literature looked mainly at threshold effects in $z_t$, the adjustment process to equilibrium while forcing the cointegrating regression to be linear. This has been commonly called threshold cointegration (Balke and Fomby (1997)).
- But: Threshold cointegration defined in the above sense still falls within the standard Engle-Granger definition of Linear cointegration as an I(0) linear combination of I(1)’s.

**Novelty Here:** we are introducing nonlinearities in the form of threshold effects directly within the long run equilibrium relationship and this is what we call **threshold cointegration**.
Introduction: Objectives

- The type of nonlinear cointegration we are dealing with: piecewise linear cointegrating regressions separated according to the magnitude of a threshold variable, say $q_{t-d}$, which triggers the regime switches

$$y_t = \beta_1 x_t I(q_{t-d} \leq \gamma) + \beta_2 x_t I(q_{t-d} > \gamma) + u_t$$

- Equivalently: $y_t = \beta x_t + \lambda x_t I(q_{t-d} > \gamma) + u_t$

- Key Objective: Propose a test for detecting the presence of threshold nonlinearities within cointegrating regressions.

- $H_0: \beta_1 = \beta_2$ or $H_0: \lambda = 0$. 
Introduction: Motivations

- \( y_t = \beta x_t + \lambda x_t I(q_t-d > \gamma) + u_t \): Theory is often silent about the specific nature of nonlinearities that may be involved. The presence of threshold effects has been widely documented in the recent economics literature (Kanbur (2005)). Our model may be appropriate for modelling phenomena such as multiple or switching equilibria. The literature on animal spirits for instance suggests instances where some extraneous variable such as a confidence indicator may cause macro variables to switch across different equilibrium paths (Howitt and McAfee (1992)).

- Omitting nonlinearities within long run equilibrium relationships will lead to misleading interpretations of equilibrium relationships since the cointegrating vector will no longer be consistently estimated (unlike the case where nonlinearities enter the error correction term).
The Model: Defining Threshold Cointegration

- We consider the following cointegrating regression with a threshold nonlinearity

\[ y_t = \beta' x_t + \lambda' x_t I(q_{t-d} > \gamma) + u_t \]  
(1)

\[ x_t = x_{t-1} + v_t \]  
(2)

- Cointegration in the context of the specification in (1) is understood in the sense that although both \( y_t \) and \( x_t \) have variances that grow with \( t \), the threshold combination given by \( u_t \) is stationary \( \rightarrow \) we will formalise this alternative view of cointegration.

- The concept of I(.)’ness can no longer apply in a setup such as (1)-(2) and we need an alternative definition of the concept of cointegration. I(1)/I(0) are \textit{linear} concepts.
The Model: Defining Threshold Cointegration

- **Linear World:** \( y_t = \beta x_t + u_t \) with \( x_t = x_{t-1} + v_t \). Clearly \( \Delta y_t = \beta v_t + \Delta u_t \) and with \( x_t \) an I(1), \( y_t \) is also I(1) while \( \Delta y_t \) is I(0).

- This reasoning no longer works in a nonlinear setup. Example:
  \( y_t = x_t I(q_{t-d} > \gamma) \) with \( x_t = x_{t-1} + v_t \). \( \Delta y_t \) does not make any sense. Strictly speaking \( y_t \) is I(\( \infty \)). We introduce a new concept called summability

**Definition (Summability)** A time series \( y_t \) is said to be summable of order \( \delta \), symbolically represented as \( S_y(\delta) \) if the sum \( S_y = \sum_{t=1}^{T}(y_t - E[y_t]) \) is such that \( S_y/T^{1/2+\delta} = O_p(1) \) as \( T \to \infty \).

- 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 \( S_y(1) \).
The Model: Defining Threshold Cointegration

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 $\delta_0 < \min(\delta_1, \delta_2)$.

- it is now clear that within our specification in (1), $y_t$ and $x_t$ 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 $(1, -\beta')$ and $(1, -(\beta + \lambda)')$ depending on whether $q_{t-d}$ crosses the unknown threshold level given by $\gamma$. 
Model and Testing

- The model: \( y_t = \beta' x_t + \lambda' x_t I(q_{t-d} > \gamma) + u_t \) with \( x_t = x_{t-1} + v_t \).
  In matrix form: \( y = X \beta + X \gamma \lambda + u \)

- We are interested in testing the null hypothesis of linear cointegration versus the alternative of threshold cointegration.

- \( H_0 : \lambda = 0 \) against \( H_1 : \lambda \neq 0 \)

- We explore the properties of an LM type test statistic for testing this null hypothesis. The test statistic: \( SupLM = \sup_{\gamma} LM_T(\gamma) \)

  where

  \[
  LM_T(\gamma) = \frac{1}{\hat{\sigma}_0^2} y'M X_\gamma (X'_\gamma M X_\gamma)^{-1} X'_\gamma M y
  \]  

  and \( M = I - X(X'X)^{-1}X' \)

(3)
Assumptions: \( y_t = \beta' x_t + \lambda' x_t I(q_{t-d} > \gamma) + u_t, \Delta x_t = v_t \)

(A1) The sequence \( \{u_t, v_t, q_t\} \) is strictly stationary and ergodic and strong mixing with mixing coefficients \( \alpha_n \) satisfying \( \sum_{n=1}^{\infty} \alpha_n^{\frac{1}{r}} < \infty \) for some \( r > 2 \); The threshold variable \( q_t \) has a distribution function \( F(.) \) that is continuous and strictly increasing.

(A2) \( \sum_{t=1}^{[Tr]} w_t / \sqrt{T} \Rightarrow B(r) = (B_u(r), B_v(r)')' \) where \( B(r) \) is a \((p + 1)\) dimensional Brownian Motion with a long run covariance matrix \( \Omega \) such that \( \sigma_u^2 = E[u_t^2] \), \( \Omega_{uv} = \Sigma_{uv} + \Lambda_{uv} + \Lambda_{vu}' \), \( \Omega_{vu} = \Sigma_{vu} + \Lambda_{vu} + \Lambda_{uv}' \) and \( \Omega_{vv} = \Sigma_{vv} + \Lambda_{vv} + \Lambda_{vv}' \) where \( \Sigma_{uv} = E[u_t v_t'], \Sigma_{vv} = E[v_t v_t'], \Lambda_{uv} = \sum_{k=1}^{\infty} E[u_{t-k} v_t], \Lambda_{vu} = \sum_{k=1}^{\infty} E[v_{t-k} u_t], \Lambda_{vv} = \sum_{k=1}^{\infty} E[v_{t-k} v_t'] \).

(A3) \( E(u_t) = 0, E|u_t|^4 = \kappa < \infty \) and \( u_t \) is independent of \( \mathcal{F}_{t-1}^{qu} \) where \( \mathcal{F}_{t}^{qu} = \sigma(q_{t-j}, u_{t-j}; j \geq 0) \).
Assumptions: Remarks

- We need (A1) and (A3) for establishing a limit theory for \( \sum_{t=1}^{[Tr]} u_t I(q_{t-1} \leq \gamma) / \sqrt{T} \).

- (A2) is a multivariate invariance principle. It is implied by (A1) provided that we impose some moment restrictions on \( u_t \), \( v_t \) and their product.

- The fourth moment assumption in (A3) is not needed for (A2). We need it for establishing the tightness of the above empirical process.

- Structure of \( \Omega \): We disallow serial correlation in \( u_t \).

- Note that (A1) also requires \( q_t \) to be stationary and ergodic.
Additional Assumptions

(A4) The threshold parameter $\gamma$ 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 $x_t$ is not cointegrated.
• $y_t = \beta' x_t + \lambda' x_t I(q_{t-d} > \gamma) + u_t$, $\Delta x_t = v_t$. We obtain the limiting distribution of the SupLM statistic under various assumptions on the long run covariance matrix of $w_t = (u_t, v_t')' \text{ i.e. exogeneity, endogeneity etc.}$ Recall

$$\Omega = \begin{pmatrix} \sigma^2_u & \Omega_{uv} \\ \Omega_{vu} & \Omega_{vv} \end{pmatrix} > 0,$$

• 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 $
abla \Theta = [\theta_1, 1 - \theta_1]$ and throughout this paper we will be using $\theta$ and $F(\gamma)$ interchangeably.
**Proposition 1** Under the null hypothesis $H_0 : \lambda = 0$ and assumptions (A1)-(A5) with $\Omega_{uv} = \Omega_{vu} = 0$, we have

$$\text{SupLM} \Rightarrow \sup_{\theta} \frac{1}{\theta(1 - \theta)} \left[ \int_0^1 W_v(r) dK_u(r, \theta) \right]' \left[ \int_0^1 W_v(r) W_v'(r) dr \right]^{-1}$$

$$\left[ \int_0^1 W_v(r) dK_u(r, \theta) \right]$$

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.

- We can show that $K_u(r, \theta)$ and $W_v(r)$ are independent and $K_u(r, \theta)/\sqrt{\theta(1 - \theta)}$ is itself a standard Brownian Motion given $\theta$, say $\tilde{W}_u(r)$. It is then easy to establish that the random variable in Proposition 1 is equivalent to a normalised Brownian Bridge process $[W(\theta) - \theta W(1)]'[W(\theta) - \theta W(1)]/\theta(1 - \theta)$ with $W(.)$ denoting a p-dimensional standard Brownian Motion.
Proposition 2 Under the null hypothesis $H_0 : \lambda = 0$ and assumptions (A1)-(A5) we have

$$\text{SupLM} \Rightarrow \sup_{\theta} \frac{1}{\theta(1-\theta)} \left[ \int_0^1 B_v(r)d\tilde{K}_u(r, \theta) \right]' \left[ \int_0^1 B_v(r)B_v(r)'dr \right]^{-1} \left[ \int_0^1 B_v(r)d\tilde{K}_u(r, \theta) \right]$$

(5)

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 $\sigma_u^2 r\theta$.

- Standard algebra: The two distributions in Propositions 1 and 2 are the same!
  - Intuition: Although $\Omega_{uv} \neq 0$, $\tilde{K}_u(r, \theta)$ and $B_v(r)$ are independent.
- Summary: We have a test statistic that allows us to test for the presence of threshold effects in cointegrating regressions. Its limiting distribution is free of nuisance parameters even under endogeneity and serial correlation in $v_t$ and even more importantly it is already widely tabulated in the literature.
Adequacy of the Asymptotics

• \( y_t = \beta' x_t + \lambda' x_t I(q_{t-1} > \gamma) + u_t \) with \( x_t = x_{t-1} + v_t \)

• Some illustrations: (DGP) is given by \( y_t = \beta_0 + \beta_1 x_t + u_t \) with \( \Delta x_t = v_t \) and \( v_t = \rho v_{t-1} + \epsilon_t \). Fitted model: \( y_t = \beta_0 + \beta_1 x_t + (\lambda_0 + \lambda_1 x_t) I(q_{t-1} > \gamma) + u_t \). We take \( q_t \) to follow the AR(1) process \( q_t = \phi q_{t-1} + \nu_t \) and \( u_t = NID(0, 1) \).

• Covariance structure: \( z_t = (u_t, \epsilon_t, \nu_t)' \). For \( \Sigma_z = E[z_t z_t'] \), we use

\[
\Sigma_z = \begin{pmatrix}
1 & \sigma_{u\epsilon} & \sigma_{uv} \\
\sigma_{\epsilon u} & 1 & \sigma_{\epsilon\nu} \\
\sigma_{uv} & \sigma_{\epsilon\nu} & 1
\end{pmatrix}
\]

• \( DGP_1 : \{\sigma_{u\epsilon}, \sigma_{uv}, \sigma_{\epsilon\nu}\} = \{0.3, 0.7, 0.6\} \). \( DGP_2 \) has \{0.3, 0.0, 0.6\} and \( DGP_3 \) has \{0.0, 0.0, 0.0\}.

• \( DGP_1 \) and \( DGP_2 \) allow for the endogeneity of \( x_t \) while \( DGP_3 \) takes it as strictly exogenous. The implementation of the SupLM test also assumes 10% trimming at each end of the sample.
# Adequacy of the Asymptotics

## Table 1: Asymptotic Critical Values

<table>
<thead>
<tr>
<th></th>
<th>10%</th>
<th>5%</th>
<th>2.5%</th>
<th>1%</th>
<th>10%</th>
<th>5%</th>
<th>2.5%</th>
<th>1%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\phi = 0.5$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$DGP_1, \rho = 0.0$</td>
<td>10.28</td>
<td>12.12</td>
<td>13.85</td>
<td>15.97</td>
<td>10.75</td>
<td>12.35</td>
<td>14.04</td>
<td>15.43</td>
</tr>
<tr>
<td>$DGP_2, \rho = 0.0$</td>
<td>10.15</td>
<td>12.09</td>
<td>13.74</td>
<td>15.60</td>
<td>10.66</td>
<td>12.37</td>
<td>13.95</td>
<td>15.94</td>
</tr>
<tr>
<td>$DGP_3, \rho = 0.0$</td>
<td>10.38</td>
<td>12.10</td>
<td>13.38</td>
<td>15.78</td>
<td>10.34</td>
<td>12.58</td>
<td>14.00</td>
<td>15.99</td>
</tr>
<tr>
<td>$DGP_1, \rho = 0.4$</td>
<td>10.34</td>
<td>12.00</td>
<td>13.81</td>
<td>15.91</td>
<td>10.33</td>
<td>12.08</td>
<td>13.51</td>
<td>15.05</td>
</tr>
<tr>
<td>$DGP_2, \rho = 0.4$</td>
<td>10.19</td>
<td>11.73</td>
<td>13.70</td>
<td>15.46</td>
<td>10.64</td>
<td>12.36</td>
<td>13.94</td>
<td>15.98</td>
</tr>
<tr>
<td>$DGP_3, \rho = 0.4$</td>
<td>10.34</td>
<td>11.95</td>
<td>13.36</td>
<td>15.76</td>
<td>10.37</td>
<td>12.60</td>
<td>13.98</td>
<td>16.02</td>
</tr>
<tr>
<td>Andrews</td>
<td>10.50</td>
<td>12.27</td>
<td>NA</td>
<td>16.04</td>
<td>10.50</td>
<td>12.27</td>
<td>NA</td>
<td>16.04</td>
</tr>
</tbody>
</table>
## Adequacy of the Asymptotics

### Table 2: Empirical Size Estimates

<table>
<thead>
<tr>
<th>Nominal</th>
<th>5%</th>
<th>2.5%</th>
<th>1%</th>
<th>5%</th>
<th>2.5%</th>
<th>1%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>φ = 0.5</td>
<td></td>
<td>φ = 0.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$DGP_1, \rho = 0.4, T = 200$</td>
<td>5.00%</td>
<td>2.60%</td>
<td>0.95%</td>
<td>4.55%</td>
<td>2.15%</td>
<td>0.85%</td>
</tr>
<tr>
<td>$DGP_1, \rho = 0.4, T = 400$</td>
<td>4.95%</td>
<td>2.45%</td>
<td>0.96%</td>
<td>4.20%</td>
<td>2.00%</td>
<td>1.00%</td>
</tr>
<tr>
<td>$DGP_2, \rho = 0.4, T = 200$</td>
<td>4.65%</td>
<td>2.45%</td>
<td>0.80%</td>
<td>4.90%</td>
<td>2.35%</td>
<td>0.85%</td>
</tr>
<tr>
<td>$DGP_2, \rho = 0.4, T = 400$</td>
<td>4.75%</td>
<td>2.45%</td>
<td>0.75%</td>
<td>4.35%</td>
<td>1.70%</td>
<td>0.70%</td>
</tr>
<tr>
<td>$DGP_3, \rho = 0.4, T = 200$</td>
<td>4.50%</td>
<td>2.30%</td>
<td>0.75%</td>
<td>3.85%</td>
<td>2.10%</td>
<td>0.70%</td>
</tr>
<tr>
<td>$DGP_3, \rho = 0.4, T = 400$</td>
<td>4.70%</td>
<td>2.70%</td>
<td>1.20%</td>
<td>3.98%</td>
<td>1.90%</td>
<td>0.60%</td>
</tr>
</tbody>
</table>
### Adequacy of the Asymptotics

#### Table 3: Empirical Power Estimates

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

• Recall: \( y_t = \beta' x_t + \lambda' x_t I(q_{t-1} > \gamma) + u_t \) with \( x_t = x_{t-1} + v_t \)

• Inclusion of deterministic components → straightforward.

• Most Important: Allowing for serial correlation in \( u_t \) → very tricky and not a straightforward extension as in the linear case.

• Important object: \( G_{uT}(r, \theta) = \sum_{t=1}^{[Tr]} u_t I(F(q_{t-1}) \leq \theta)/\sqrt{T} \) where \( F(q_{t-1}) \equiv U_{t-1} \) is a uniformly distributed random variable. For simplicity I write \( G_{uT}(r, \theta) = \sum_{t=1}^{[Tr]} u_t I_{t-1}/\sqrt{T} \)

• Complications: Need an FCLT type of result for the marked empirical process \( G_{uT}(r, \theta) \) in which both the marks \( u_t \) and \( q_t \) are possibly correlated general stationary processes. For \( u_t \) iid see Caner and Hansen (2001). 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)).
• A Beveridge and Nelson type of approach or a martingale approximation do not work. Suppose that we assume \( u_t = C(L)e_t \) with \( e_t \) i.i.d or m.d.s. Using BN and summation by parts gives

\[
\frac{1}{\sqrt{T}} \sum_{t=1}^{T} u_t I_{t-1} = \frac{1}{\sqrt{T}} C(1) \sum_{t=1}^{T} e_t I_{t-1} + \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \tilde{e}_t \Delta I_t + o_p(1), \tag{6}
\]

where \( \tilde{e}_t = \tilde{C}(L)e_t = \sum_{j=0}^{\infty} \tilde{c}_j e_{t-j} \) with \( \tilde{c}_j = \sum_{k=j+1}^{\infty} c_k \)

• Had the second term in the right hand side vanished asymptotically, the use of the results in Caner and Hansen (2001) would have been sufficient to treat the case of serial correlation in \( u_t \). Note also that the two components in the right hand side of (6) are not independent and the second one clearly does not form a martingale difference sequence since \( \tilde{e}_t \) is another weakly dependent process.

• Other ideas: Use (A1) to argue that \( u_t I_{t-1} \) is also strong mixing. Use CLT for strongly mixing sequences to establish fidi convergence. Problem: We don’t know how to establish tightness of \( G_{uT}(r, \theta) \) due to the lack of an inequality equivalent to Rosenthal’s inequality available for mds sequences.
• \( u_t \) follows a finite order \( MA \) process.

• We replace assumption (A3) with: \( u_t = \Psi_\ell(L)e_t \) 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 \( \mathcal{F}_{t-1} = \sigma(q_{t+\ell-j}, e_{t-j}; j \geq 1) \).

• The long run variance of \( u_t \) is denoted \( \omega_u^2 = \sigma_e^2 \Psi_\ell(1)^2 \).

• 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.

**Proposition 3** Letting \( G_{uT}(r, \theta) = \sum_{t=1}^{[Tr]} u_t I_{t-1} / \sqrt{T} \) 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 \( \omega_G^2(r_1, r_2, \theta_1, \theta_2) = (r_1 \wedge r_2)\sigma_e^2 E[(\sum_{j=0}^{\ell} \Psi_j I_{t-1+j}(\theta_1))(\sum_{j=0}^{\ell} \Psi_j I_{t-1+j}(\theta_2))]. \)
• Our result in Proposition 3 specialises to Theorem 1 of Caner and Hansen (2001) if we set $\Psi_j = 0$ for $j \geq 1$ since this corresponds to the case where the marks of the empirical process are i.i.d. Indeed, from the expression of $\omega^2_G(r_1, r_2, \theta_1, \theta_2)$ above we obtain $\omega^2_G(r_1, r_2, \theta_1, \theta_2) = \sigma^2_e(r_1 \wedge r_2)(\theta_1 \wedge \theta_2)$ which can be recognised as the covariance kernel of a standard Brownian Sheet
Proposition 4 Under the null hypothesis $H_0 : \lambda = 0$ and assumptions (A1), (A2'), (B1) and (A4)-(A5) we have

$$SupLM \Rightarrow \sup_{\theta} \frac{1}{\theta(1 - \theta)} \left[ \int_{0}^{1} B_v(r) \tilde{M}_u(r, \theta) \right]' \left[ \int_{0}^{1} B_v(r) B_v(r)' dr \right]^{-1} \left[ \int_{0}^{1} B_v(r) \tilde{M}_u(r, \theta) \right]$$

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

• The limiting process in (7) 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.
• In the context of our results in Proposition 4 however, an important simplification occurs under the additional assumption that $q_t$ follows an iid process. Under this particular case it is straightforward to write the long run variance of $G_u(r, \theta)$ as $\omega^2_G(r, \theta) = r[\theta^2 \omega^2_u + \theta(1 - \theta)\sigma^2_u]$. It is then easy to see that the process $\tilde{M}_u(r, \theta)$ in (7) is such that $\tilde{M}_u(r, \theta) \equiv K_u(r, \theta)$ and the limiting distribution in (7) reduces to that presented in Propositions 1 and 2. We summarise the results pertaining to this particular scenario in the following Proposition.

**Proposition 5** When $q_t$ is an iid process and $u_t$ is as in (B1) the limiting distribution in (7) is identical to the one obtained in Propositions 1 and 2.
Conclusions

• We proposed a test of linear versus threshold cointegration. Under quite a general setting (endogeneity and serial correlation in $v_t$) we introduced a test statistic whose tabulated distribution is readily available from the literature. We also obtained a new FCLT for a marked empirical process with weakly dependent marks.

• In progress: Allow for general forms of serial correlation in $u_t$.

• Economic Applications: Present Value Relationships with time varying discount rates.