UNIVERSITY OF
Southampton
School of Social Sciences

School of Social Sciences
Economics Division
University of Southampton
Southampton SO17 1BJ, UK

Discussion Papers in
Economics and Econometrics

Threshold Autoregressions Under Near Integratedness

Jean-Yves Pitarakis

No. 1007

This paper is available on our website
http://www.southampton.ac.uk/socsci/economics/research/papers

ISSN 0966-4246

Threshold Autoregressions under Near Integratedness*

Jean-Yves Pitarakis
University of Southampton
Economics Division
Southampton SO17 1BJ, U.K

March 5, 2010

Abstract

We explore the properties of a Wald type test statistic for detecting the presence of threshold effects in time series when the underlying process could be nearly integrated as opposed to having an exact unit root. We derive its limiting distribution and establish its equivalence to a normalised squared Brownian Bridge process. More importantly we show that the limiting random variable no longer depends on the noncentrality parameter characterising the nearly integrated DGP. This is an unusual occurrence which is in stark contrast with the existing literature on conducting inferences under persistent regressors where it is well known that the noncentrality parameter appears in the limiting distribution of test statistics, making them impractical for inference purposes.

Keywords: Threshold Autoregressive Models, Near Unit Root, Noncentrality Parameter, Nonlinear time series

JEL: C22

*Address for correspondence: Jean-Yves Pitarakis, University of Southampton, School of Social Sciences, Economics Division, Southampton SO17 1BJ, UK. Email: j.pitarakis@soton.ac.uk

1 Introduction

This paper is concerned with conducting inferences within an environment that combines threshold type nonlinearities with the presence of a highly persistent variable that is close to a unit root process. We are interested in obtaining the limiting distribution of a Wald type test statistic for testing the null of linearity against a threshold autoregressive alternative when the linear process under the null is local to a unit root.

From the early literature on threshold models (see Tong (1983), Chan (1990), Hansen (1996) and references therein) it is well known that within stationary and ergodic environments the limiting distributions of test statistics designed to test for the presence of threshold effects are typically not free of nuisance parameters. More recently and moving away from a stationary and ergodic setting Caner and Hansen (2001) explored the asymptotics of a Wald type test statistic for threshold effects in an environment in which nonlinearities and exact unit root type nonstationarities could coexist. They developed its asymptotics for testing the null of linearity againts a threshold type alternative when the underlying null model contains a unit root. Pitarakis (2008) extended their results by showing that under the exact unit root setting and particular scenarios the nuisance parameter problem vanishes and the distribution of the Wald statistic for testing for the presence of threshold effects takes the well known form of a normalised quadratic form in Brownian Bridges. These contrasting results between purely stationary and exact unit root settings are the key motivations of this paper. More specifically, we are interested in assessing the asymptotics of the same Wald type test statistic for detecting threshold effects when the null model is local to a unit root process as opposed to having an exact unit root. This will also allow us to assess how accurate the asymptotic approximations developed in Caner and Hansen (2001) and Pitarakis (2008) are when we deviate from the exact unit root within the null model.

The plan of the paper is as follows. Section 2 introduces our model and test statistic and presents its limiting distribution. Section 3 illustrates our theoretical results through a simple simulation based exercise. Proofs are relegated to the appendix.

2 The Model and Asymptotic Inference

We are interested in testing $H_0 : \beta_1 = \beta_2$ in

$$\Delta y_t \quad = \quad \beta_1' x_{t-1} I(Z_{t-1} \leq \gamma) + \beta_2' x_{t-1} I(Z_{t-1} > \gamma) + e_t \tag{1}$$

with $x_{t-1} = (1 \ t \ y_{t-1})'$ and $\beta_i = (\mu_i \ \delta_i \ \rho_i)$ for $i = 1, 2$. Z_{t-1} is our threshold variable and the threshold parameter γ is assumed unknown with $\gamma \in \Gamma = [\gamma_1, \gamma_2]$. As in Caner and Hansen (2001) γ_1 and γ_2 are

selected such that $P(Z_t \leq \gamma_1) = \pi_1 > 0$ and $P(Z_t \leq \gamma_2) = \pi_2 < 1$ for some reasonable $\pi_i's$. In the discussion that follows we also replace the threshold variable Z_{t-1} by a uniformly distributed random variable making use of the equality $I(Z_{t-1} \leq \gamma) = I(G(Z_{t-1}) \leq G(\gamma)) \equiv I(U_{t-1} \leq u)$ where $G(.)$ is the marginal distribution of Z_t and U_t denotes a uniformly distributed random variable on $[0, 1]$. For notational simplicity we use $I_{1t-1} \equiv I(U_{t-1} \leq u)$ and $I_{2t-1} \equiv I(U_{t-1} > u)$ throughout.

In Caner and Hansen (2001) the authors derived the limiting behaviour of a Wald type test statistic for testing $H_0 : \beta_1 = \beta_2$ in (1) when the underlying process was known to contain an exact unit root and given by $\Delta y_t = e_t$. Proceeding under the same probabilitic assumptions as in Caner and Hansen (2001, pp 1558-1559) our goal here is to instead explore the limiting behaviour of the same Wald statistic when the underlying process is nearly integrated as in

$$\Delta y_t = \frac{c}{T} y_{t-1} + e_t \tag{2}$$

with $c < 0$. Although not repeated here for space considerations, the assumptions of Caner and Hansen (2001) can be succintly stated as requiring $\{e_t, U_t\}$ to be strictly stationary and ergodic and strong mixing with appropriately defined mixing numbers. Furthermore e_t is also restricted to be zero mean i.i.d with a finite fourth moment. As it is standard in the unit root literature it is also understood that there are no deterministic trend components in the data generating process.

As highlighted in Hansen (1996) it is well known that limiting distributions of statistics designed to test $H_0 : \beta_1 = \beta_2$ are not free of nuisance parameters when operating under stationarity and ergodicity. They typically depend on model specific moments such as the variance of regressors amongst other quantities. At the same time, Pitarakis (2008) showed that under particular scenarios these limiting distributions become free of nuisance parameters and take familiar forms that are already tabulated in the literature when the model under the null contains an exact unit root. These two observations are what motivates our new setup in which the linear null model is now assumed to be local to a unit root process as in (2) above.

Given our parameterisation it is now straightforward to obtain the expression of the Wald statistic of interest. For greater convenience we write (1) in matrix form as $\Delta Y = X_1\beta_1 + X_2\beta_2 + e$. We let $X = X_1 + X_2$, $\beta = (\beta_1 \ \beta_2)'$ and define $R = [I_3 \ -I_3]$ so that our null hypothesis can be rewritten as $H_0 : R\beta = 0$ with the corresponding Wald statistic for given u or γ given by $W_T(u) = (\hat{\beta}_1 - \hat{\beta}_2)'[X_1'X_1 - X_1'X_1(X'X)^{-1}X_1'X_1](\hat{\beta}_1 - \hat{\beta}_2)/\hat{\sigma}_e^2$. The following proposition summarises the limiting behaviour of $W_T(u)$ under the null hypothesis of interest. Note that our inferences fall within the well known problem of an unidentified nuisance parameter under the null (here the threshold parameter) which we handle by following common practice and focusing on the Supremum Wald statistic.

Proposition 1. *Under the same assumptions as in Caner and Hansen (2001),* $H_0 : \beta_1 = \beta_2$ *and the parameter restriction* $\rho_1 = \rho_2 = c/T$ *we have* $\sup_u W_T(u) \Rightarrow \sup_u BB(u)/u(1-u)$ *as* $T \to \infty$ *with* $BB(u)$ *denoting a standard Brownian Bridge process of the same dimension as* β_i.

The above Proposition establishes that the limiting distribution of the Wald statistic for detecting the presence of threshold effects under near integratedness is identical to the one that occurs under the presence of an exact unit root. More importantly, our result in Proposition 1 highlights an environment in which the noncentrality parameter c no longer appears in the limiting distribution despite being present in the DGP. This is a highly unusual occurrence and appears to be unique to this threshold based framework. The difficulties with using parameterisations such as (2) are well known in the unit root, cointegration or predictive regression literature where inferences have always been plagued by the appearance of the noncentrality parameter c in the limiting distributions. It is only very recently that a solution to this problem has been proposed through the use of an IV based technique developed in Phillips and Magdalinos (2009). Our result in Proposition 1 also suggests that findings in Caner and Hansen (2001) continute to hold under deviations from the exact unit root scenario.

3 Simulation Based Analysis

To gain further insights into the above large sample based result we simulate data taking (2) as our DGP and explore the finite sample size properties of our SupWald statistic across different magnitudes of c and sample sizes. The idea is to numerically demonstrate the phenomenon presented in Proposition 1 and document its robustness to alternative choices of c since from Hansen (1996) we know that the limiting distribution of the SupWald statistic for testing threshold effects is not free of nuisance parameters when we operate under pure stationarity.

For our size based analysis we use Hansen's (1997) p-value approximations which aim to provide pvalues for the supremum of a normalised squared Brownian Bridge type of limit as in Proposition 1 (see also Andrews (1992)). Once we have generated our samples from our null DGP we count the number of times $H_0 : \beta_1 = \beta_2$ is rejected using $c \in \{0, -1, -5, -10\}$ and $T \in \{200, 500, 1000\}$. Results are presented in Table 1 below which uses a nominal size of 2.5% and $N = 5000$ replications. All random variables are taken as NID(0,1).

Table 1. Empirical Size across Noncentrality Parameter Magnitudes (2.5% Nominal)

	$c = 0$	$c = -1$	$c = -5$	$c = -10$
$T = 200$	2.32	2.20	2.06	2.00
$T = 500$	2.30	2.32	2.44	2.44
$T = 1000$	2.46	2.47	2.48	2.45

As expected we first note the inadequacy of the limiting distribution in Proposition 1 when T is small and c is far away from 0. Under $c = -10$ and $T = 200$ for instance we have an empirical size of 2.00% when the nominal one is 2.50%. This discrepency highlights the fact that the above limiting approximation is inadequate when we move away from integratedness towards stationarity as discussed above. Looking at the results under $T = 1000$ on the other hand we note that the empirical sizes match their nominal counterparts very closely regardless of the magnitude of c, illustrating very clearly our theoretical result in Proposition 1.

APPENDIX

PROOF OF PROPOSITION 1: We write $W_T^A(u)$ under $H_0 : \beta_1 = \beta_2$ as

$$\begin{aligned} W_T^A(u) &= [e'X_1 - e'X(X'X)^{-1}X_1'X_1][X_1'X_1 - X_1'X_1(X'X)^{-1}X_1'X_1]^{-1} \\ &\quad [X_1'e - (X_1'X_1)(X'X)^{-1}X'e]/\hat{\sigma}_e^2. \end{aligned} \tag{3}$$

With $D_T = diag(T^{1/2}, T^{3/2}, T)$ we have

$$D_T^{-1}X_1{}'X_1D_T^{-1} = \begin{pmatrix} \frac{\sum I_{1t-1}}{T} & \frac{\sum tI_{1t-1}}{T^2} & \frac{\sum y_{t-1}I_{1t-1}}{T^{3/2}} \\ \frac{\sum tI_{1t-1}}{T^2} & \frac{\sum t^2I_{1t-1}}{T^3} & \frac{\sum y_{t-1}tI_{1t-1}}{T^{\frac{5}{2}}} \\ \frac{\sum y_{t-1}I_{1t-1}}{T^{3/2}} & \frac{\sum y_{t-1}tI_{1t-1}}{T^{\frac{5}{2}}} & \frac{\sum y_{t-1}^2I_{1t-1}}{T^2} \end{pmatrix} \tag{4}$$

from which we obtain the following weak convergence results

$$\begin{aligned} D_T^{-1}X_1{}'X_1D_T^{-1} &\Rightarrow \begin{pmatrix} u & \frac{1}{2}u & u\int_0^1 K_c(r)dr \\ \frac{1}{2}u & \frac{1}{3}u & u\int_0^1 rK_c(r) \\ u\int_0^1 rK_c(r) & u\int_0^1 rK_c(r) & \int_0^1 K_c^2(r) \end{pmatrix} \\ &\equiv u\int_0^1 \overline{K}_c(r)\overline{K}_c(r)' \end{aligned} \tag{5}$$

and

$$D_T^{-1}X'XD_T^{-1} \Rightarrow \int_0^1 \overline{K}_c(r)\overline{K}_c(r)' \tag{6}$$

where $\overline{K}_c(r) = (1, r, K_c(r))$. The above results follow from Theorem 3 in Caner and Hansen (2001) and Lemma 3.1 in Phillips (1988). The diffusion process $K_c(r)$ is an Ornstein-Uhlenbeck process defined as $K_c(r) = \int_0^r e^{(r-s)c}dB_e(s)$ with $K_c(r)$ such that $dK_c(r) = cK_c(r) + dB_e(r)$, $K_c(0) = 0$ and $B_e(r)$ is the Brownian Motion associated with e_t. Note that we can also write $K_c(r) = B_e(r) + c\int_0^r e^{(r-s)c}B_e(s)ds$. It now follows from the continuous mapping theorem that

$$\begin{aligned} [D_T^{-1}X_1{}'X_1D_T^{-1} - D_T^{-1}X_1{}'X_1(X'X)^{-1}X_1{}'X_1D_T^{-1}]^{-1} &\Rightarrow \frac{1}{u(1-u)} \\ &\times \left(\int_0^1 \overline{K}_c(r)\overline{K}_c(r)'\right)^{-1}. \end{aligned} \tag{7}$$

We next focus on the limiting behaviour of $D_T^{-1}X'u$ and $D_T^{-1}X_1'u$. Looking at each component separately, setting $\sigma_e^2 = 1$ for simplicity and no loss of generality and using Theorem 2 in Caner and Hansen (2001), we have

$$D_T^{-1}X_1{}'u = \begin{pmatrix} \frac{\sum I_{1t-1}e_t}{\sqrt{T}} \\ \frac{\sum tI_{1t-1}e_t}{T^{3/2}} \\ \frac{\sum y_{t-1}I_{1t-1}e_t}{T} \end{pmatrix} \Rightarrow \begin{pmatrix} B_e(r,u) \\ \int_0^1 rdB_e(r,u) \\ \int_0^1 K_c(r)dB_e(r,u) \end{pmatrix} \tag{8}$$

and

$$D_T^{-1}X'u \;=\; \begin{pmatrix} \frac{\sum e_t}{\sqrt{T}} \\ \frac{\sum te_t}{T^{3/2}} \\ \frac{\sum y_{t-1}e_t}{T} \end{pmatrix} \Rightarrow \begin{pmatrix} B_e(r,1) \\ \int_0^1 r dB_e(r,1) \\ \int_0^1 K_c(r) dB_e(r,1) \end{pmatrix}. \tag{9}$$

where $B_e(r,u)$ is a two parameter Brownian Motion as defined in Theorem 1 of Caner and Hansen (2001). The above now allows us to formulate the limiting behaviour of $D_T^{-1}X_1'e - uD_T^{-1}X'e$ as

$$D_T^{-1}X_1'e - uD_T^{-1}X'e \;\Rightarrow\; \int_0^1 \overline{K}_c(r) dG_e(r,u) \tag{10}$$

where $G_e(r,u) = B_e(r,u) - uB_e(r,1)$. The result then follows straightforwardly through the use of the continuous mapping theorem, standard algebra, and crucially, by noting that the random variable in (A.8) is mixed normal $N(0, u(1-u)\int \overline{K}_c(r)\overline{K}_c(r)')$ due to the independence of $G_e(r,u)$ and $\overline{K}_c(r)$ since $E[G_e(r_1,u_1)K_c(r_2)] = 0$ and both processes are Gaussian. Thus combining (A.8) and (A.5) gives the result in Proposition 1.

REFERENCES

Andrews, D. W. K., 1993. Tests for Parameter Instability and Structural Change with Unknown Change Point, *Econometrica*, 61, 821-856.

Caner, M., AND B. E. Hansen, 2001. Threshold Autoregression with a Unit Root, Econometrica, 69, 1555-1596.

Chan, K. S., 1990. Testing for threshold autoregression, Annals of Statistics, 18, 1886-1894.

Hansen, B. E., 1996. Inference when a Nuisance Parameter is not Identified under the Null Hypothesis, Econometrica, 64, 413-430.

Hansen, B. E., 1997. Approximate Asymptotic p-Values for Structural Change Tests, Journal of Business and Economic Statistics, 15, 60-67.

Phillips, P.C. B and A. Magdalinos, 2009. Econometric Inference in the Vicinity of Unity, Singapore Management University, CoFie Working Paper No. 7.

Phillips, P. C. B., 1988. Regression Theory for Near-Integrated Time Series, Econometrica, Vol. 56, pp. 1021-1043.

Pitarakis, J., 2008. Threshold Autoregressions with a unit root: Comment, Econometrica, Vol. 76, pp. 1207-1217.

Tong, H., 1983. Threshold Models in Non-linear Time Series Analysis. Springer Verlag.