Working Paper 97-50
Statistics and Econometrics Series 21
September 1997

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34)91-624-9849

THRESHOLD UNIT ROOT MODELS.

Martín González and Jesús Gonzalo*

Abstract

One of the main criticisms of unit root models is based on the theoretical fact that economic variables measured in rates cannot have unit roots. Nevertheless, standard unit root tests do not reject the existence of unit roots in many of those variables. In this paper we present a class of threshold models capable of replicating the behavior of economic variables such as unemployment, inflation and interest rates. Depending on the values of a threshold variable these models can have either a unit root or a stable root. However, despite the presence of the unit root, we prove they are stationary and geometrically ergodic. Least squares estimates of the parameters of these models are shown to be consistent and asymptotically normal. We propose the supremum of a t^2 test in order to test the null of no threshold against the alternative of threshold when the threshold value is unknown. The limiting distribution is derived under the null of I (0) as well as under the null of I(1). Critical values for both asymptotic distributions are computed and a finite sample study of the performance (size and power) of the tests developed in this paper is made. The paper concludes with an application to interest rates.

Keywords:
Brownian motion; Brownian sheet; geometric ergodicity; hypothesis testing; threshold models; unit root processes.

*González, Boston University and CEDES, Argentina, e-mail: martingr@arnet.com.ar; fax: 541-8620805; Gonzalo, Departamento de Estadística y Econometría, Universidad Carlos III de Madrid. C/ Madrid, 126, 28903 Madrid. Spain. Ph: 34-91-624.98.53; Fax: 34-91-624.98.49, e-mail: jgonzalo@est-econ.uc3m.es; We thank Bruce Hansen and Pierre Perron for estimulating comments on an earlier draft. This version has also benefited from comments of an associate editor and two anonymous referees. Jesús Gonzalo gratefully acknowledges financial aid from the Spanish Secretary of Education (PB 950298 DGICYT).

1 Introduction

For more than a decade, threshold time series models have been used to capture several nonlinear phenomena commonly observed in practice such as time-irreversibility, asymmetries, etc. The idea of these models is to approximate the behavior of certain time series using a threshold autoregression with a small number of regimes. In particular, consider the multiple thresholds first-order autoregressive model (TAR),

$$X_t = \begin{cases} \alpha_1 X_{t-1} + e_t & \text{if } Z_{t-d} \leq r_1, \\ \alpha_2 X_{t-1} + e_t & \text{if } r_1 < Z_{t-d} \leq r_2, \\ \dots & \\ \alpha_n X_{t-1} + e_t & \text{if } Z_{t-d} > r_n. \end{cases} \tag{1}$$

Where e_t is a white noise process, $r_1 < r_2 < ... < r_n$ are the threshold values, and α_i is the autoregressive coefficient in regime i. Z_{t-d} is the threshold variable and d is a fixed positive integer usually referred to as the delay parameter of Z_t.

A particular case of model (1), extensively analyzed in the literature, is the so-called Self Exciting Threshold Autoregressive (SETAR) model due to Tong (1983). In this model the regime switching is determined by the value of the variable's own past, i.e. $Z_{t-d} = X_{t-d}$.

In economics, the fact that the regime switching is determined by the same variable that generates the process may not be very appealing. In some situations a more realistic case is one in which another variable determines the regime switching in X_t. Some examples of economic variables whose behavior is affected by threshold variables (TV) are interest rates, GNP and unemployment. In the first case a candidate for a TV could be the inflation rate. For the last two variables a candidate for a TV could be a leading indicator.

Since the work of Beveridge and Nelson (1982) and Nelson and Plosser (1982), a widely believed fact is that most macroeconomic time series are best represented by models with unit roots. However, in theory, some of the economic time series mentioned above and, in general variables measured in rates, cannot have all the characteristics of a unit root process. This is so even though standard unit root tests applied to actual data do not reject the null hypothesis of unit root. In this paper we present a new type of model, a threshold unit root model (TUR) that is a combination of TAR (1) and unit root models. TUR models, while maintaining the structure and properties of the stationary TAR models, allow for unit roots in some of the regimes. This makes TUR models very good candidates to replicate the behavior of economic variables measured in rates.

The rest of the paper is organized as follows. In Section 2, we derive the conditions under which model (1) with unit roots is covariance stationary and geometrically ergodic. In Section 3 we show some results needed in order to obtain the asymptotic distributions of the tests developed in the following sections. Section 4 shows how to test a TAR model when the threshold value is assumed to be known. Asymptotic distributions of the proposed tests are derived under two types of different null hypotheses: the null of I(0) and the null of I(1). This section also shows the consistency and asymptotic normality of the OLS estimators of the coefficients of the TAR model (1) and therefore provides a test for the TUR model. Section 5 analyzes the same aspects of Section 4 but under the situation of an unknown threshold value. In this section we present and derive the asymptotic distribution of a supremum t^2 type of test for testing a TAR model. The finite sample performance (size and power) of the tests developed in this paper is analyzed in Section 6. In Section 7 we estimate a TUR model for interest rates finding evidence to support this type of model. The conclusions are found in Section 8. Proofs are provided in the Appendix.

A word on notation. We use the symbol "$\Longrightarrow$" to denote weak convergence. All limits are taken as $T \to \infty$. $|\cdot|$ means absolute value. Finally, Δ means the usual difference operator.

2 TUR Models

A more compact way of representing model (1) is,

$$\begin{aligned} X_t &= [\alpha_1 I(Z_{t-d} \leq r_1) + \alpha_2 I(r_1 < Z_{t-d} \leq r_2) + \dots + \alpha_n I(Z_{t-d} > r_n)]X_{t-1} + e_t \\ &= \delta_t X_{t-1} + e_t, \end{aligned} \tag{2}$$

where $I(\cdot)$ is an indicator function, $\delta_t = [\alpha_1 I(Z_{t-d} \leq r_1) + \alpha_2 I(r_1 < Z_{t-d} \leq r_2) + \dots + \alpha_n I(Z_{t-d} > r_n)]$, and e_t and Z_{t-d} satisfy the following assumptions,

Assumptions:

(A.1) (e_t, Z_{t-d}) *is strictly stationary and ergodic and adapted to the sigma-field* $\Im_t$.

(A.2) $E(e_t|\Im_{t-1}) = 0.$

(A.3) $E(e_t^2|\Im_{t-1}) = \sigma^2.$

(A.4) *for some* $\tau > 1$, $E(e_t^{2\tau}|\Im_{t-1}) \leq B < \infty.$

(A.5) $E(max(0, log|e_1|)) < \infty.$

(A.6) *essential supremum*$|e_1| < \infty.$

(A.7) e_1 *admits positive and continuous probability density function.*

Assumptions (A.1) and (A.5) are needed for strict stationarity. (A.6) is necessary for covariance stationarity of the TAR. (A.7) is required for geometric ergodicity. (A.2) and (A.3) are the standard assumptions specifying that the error is a conditionally homoskedastic martingale difference sequence. These two assumptions together with (A.4), that bound the extent of heterogeneity in the conditional distribution of e_t, are used to obtain the asymptotic results in Section 3. The conditions under which the TAR model (2) is covariance stationary are given in the following theorem.

Theorem 1. *Let X_t be generated by the TAR model (2), where the error term e_t satisfies assumptions (A.1), (A.5) and (A.6) and, Z_{t-d} is a threshold variable satisfying assumption (A.1). If $E|\delta_t^2| < 1$, then, the process is strictly stationary. Moreover, if $\sum_{j=1}^{\infty}(E|\prod_{n=1}^{j}\delta_n^2|)^{1/2} < \infty$, the process is also weakly stationary.*

Theorem 1 establishes sufficient conditions for strict stationarity of X_t given that (e_t, Z_{t-d}) is an strictly stationary and ergodic sequence and assumption (A.5) holds. However, this does not ensures the existence of moments. Assuming $\{e_t\}$ satisfies (A.6), Theorem 1 gives the condition for the process X_t to be covariance stationary. Notice that Theorem 1 does not either require the threshold variable Z_{t-d} to be an independent sequence or to be independent of e_t. For example, Z_{t-d} could be e_{t-d} and still conditions of Theorem 1 would be satisfied. On the other hand, the standard SETAR model can only be included in the theorem if $\alpha_1 < 1$, $\alpha_n < 1$ and $\alpha_1\alpha_n < 1$, otherwise the threshold variable will not be ergodic and assumption (A.1) will be violated (see Chan et al. (1985)).

The following three corollaries give particular examples of interesting processes satisfying the conditions of Theorem 1. These examples are presented for the simplest of the TAR models. A generalization to a model with more than two regimes is straightforward.

Corollary 1. *Consider the first-order threshold autoregressive model,*

$$X_t = [\alpha_1 I(Z_{t-d} \leq r) + \alpha_2 I(Z_{t-d} > r)]X_{t-1} + e_t = \delta_t X_{t-1} + e_t, \tag{3}$$

where $\{Z_{t-d}\}$ *is independent and identically distributed (iid) and mutually independent of* $\{e_t\}$ *with* $p = Pr(Z_{t-d} \leq r)$. *Then, if* $E|\delta_t^2| < 1$ *the process is covariance stationary.*

Corollary 2. *Consider the first-order threshold autoregressive model,*

$$X_t = [\alpha_1 I(Z_{t-d} \leq r) + \alpha_2 I(Z_{t-d} > r)]X_{t-1} + e_t = \delta_t X_{t-1} + e_t, \tag{4}$$

where Z_{t-d} *is an N-order Markov process. Then, if* $E|\delta_t^2| < 1$ *and* $p_{1/2...2} \geq p_{1/1...1}$ *the process is covariance stationary. Here,* $p_{1/j...j}$ *is the probability of being in state 1 given that during the previous N periods the process has been in state j (j=1,2).*

When $N = 1$ in Corollary 2, sufficient conditions for the process to be covariance stationary are $E|\delta_t^2| < 1$ and $p_{1/2} \geq p_{1/1}$ (see Appendix). Comparing both corollaries we see that, since in Corollary 2 we allow for some dependence in the structure of the threshold variable we need an additional condition to hold in order to get covariance stationarity of the TAR. This additional condition comes in the form of a restriction on the probabilities of the Markov process of the threshold variable.

These two corollaries set conditions that may allow for explosive roots in some of the regimes while maintaining the stationarity property. As an example, consider model (3) with $\alpha_1 = 1.3$, $\alpha_2 = 0.5$, and $p = (1-p) = 0.5$. In this case conditions of Corollary 1 are satisfied since $|\alpha_2| < 1$ and $1.3^2 = 1.69 < \frac{1-0.5^2 0.5}{0.5} = 1.75$, implying $E(\delta_t^2) = 0.97$, and the process is covariance stationary. This particular case of an *iid* threshold variable independent of the error process is generally used in random coefficient models (see Nicholls and Quinn (1982)).

Corollary 3. *Consider the first-order threshold autoregressive model,*

$$X_t = [\alpha_1 I(Z_{t-d} \leq r) + \alpha_2 I(Z_{t-d} > r)]X_{t-1} + e_t = \delta_t X_{t-1} + e_t, \tag{5}$$

where Z_{t-d} is an N-order Markov process, and $\alpha_i = 1$ and $|\alpha_j| < 1$, $i \neq j$, $(i,j) = (1,2)$. Then, the process is covariance stationary.

Corollary 3 introduces the TUR models and states sufficient conditions for these models to be stationary. Opposite to Corollary 2, no restrictions on the probabilities of the Markov process are needed, because the TUR coefficients are not allowed to be greater than one.

The processes introduced in corollaries 1 to 3 are more capable of replicating some of the characteristics of standard unit root models than simple autoregressive models. In particular, they show larger variance than autoregressive models. To see this, consider the analytical expression for the variance of the TAR model (3),

$$\frac{1}{1-(\alpha_1^2 p + \alpha_2^2(1-p))}, \tag{6}$$

and compare it with the variance of a first order autoregressive process with coefficient equal to the expected value of the TAR parameter,

$$\frac{1}{1-\rho^2} = \frac{1}{1-(\alpha_1^2 p + \alpha_2^2(1-p)) + p(1-p)(\alpha_1 - \alpha_2)^2}, \tag{7}$$

where $\rho = E(\delta_t)$. Since the denominator of (7) is greater than the denominator of (6) the variance of the TAR process is greater than the variance of a first-order autoregressive process with an autoregressive coefficient equal to the expected value of the TAR parameter.

Models (3) to (5) will also be difficult to differentiate from a pure unit root process in finite samples with standard unit root tests. To show this, we present some Monte-Carlo simulations.

Table 1 shows the empirical power of the Dickey-Fuller (DF) unit root test when the alternative hypothesis is the TUR model (5) with $\alpha_1 = 1$ and $\alpha_2 = 0.99$ (rows (4)-(6)), $\alpha_2 = 0.95$ (rows (7)-(9)), $\alpha_2 = 0.90$ (rows (10)-(12)), $\alpha_2 = 0.80$ (rows (13)-(15)), $\alpha_2 = 0.70$ (rows (16)-(18)), $\alpha_2 = 0.60$ (rows (19)-(21)), and $\alpha_2 = 0.50$ (rows (22)-(24)). To keep things simple, the threshold variable is an *iid* $U[0, 1]$ and the threshold value is $r = 0.5$. For comparison, the first three rows shows a case where $\alpha_1 = 1.1$ and $\alpha_2 = 0.90$, that is $E(\delta_t) = 1$ and the model is non-stationary. It is clear from Table 1 that the power of the unit root test between a TUR model and a first order autoregressive model with a coefficient equal to the expected value of δ_t is very similar.

TABLE 1 ABOUT HERE

Table 2 shows the empirical power of the DF test when the alternative hypothesis is the TAR model (3) with $\alpha_1 = 1.3$ and $\alpha_2 = 0.50$ (rows (4)-(6)), $\alpha_2 = 0.40$ (rows (7)-(9)), $\alpha_2 = 0.30$ (rows (10)-(12)), and $\alpha_2 = 0.20$ (rows (13)-(15)). The first three rows of the table are equal to Table 1 and they are only shown for reasons of comparison. Unlike Table 1, the power of the DF unit root test is larger here for the TAR model than for the autoregressive case with coefficient equal to the expected value of δ_t.

TABLE 2 ABOUT HERE

Notice that when the alternative hypothesis is a TAR model with $E(\delta_t) = 1$ (a non-stationary process), the power of the DF test is similar to its nominal size (rows (1)-(3) in both tables).

Stationarity and ergodicity suffice to obtain consistency and asymptotic normality of $\hat{\alpha}_i$, but in order to get consistency (rate T) for $\hat{r}$ we need geometric ergodicity (see Chan (1993)). Although this paper does not focus on the estimation of r (for this issue see Hansen (1996)), a related result will be needed in sections 4 and 5. For that reason the next theorem shows that TUR models are geometrically ergodic.

Theorem 2: *Let X_t be generated by the TUR model (5) and satisfy the assumptions of Theorem 1 plus (A.7). Then, X_t is geometrically ergodic.*

Even though Theorem 2 is only presented for TUR models, it is easy to see that models (3) and (4) are also geometrically ergodic.

3 Basic Results

In this section we show basic results necessary to obtain the limiting distribution of some of the tests developed in the next sections. In order to simplify, throughout the rest of the paper it is assumed that the relevant TUR model is given by

$$X_t = [\alpha_1 I(Z_{t-d} \leq r) + \alpha_2 I(Z_{t-d} > r)]X_{t-1} + e_t, \tag{8}$$

with $Pr(Z_{t-d} \leq r) = p(r)$. As it was mentioned in last section, the key assumptions used to obtain the following asymptotic results are (A.1) to (A.4).

Basic Result 1 (BR1):
Under the null of unit root ($X_t = X_{t-1} + e_t$) and under the assumptions that make $X_T(s) = \frac{1}{\sqrt{T}} \sum_{t=1}^{[Ts]} e_t \Longrightarrow \sigma W(s)$,

$$T^{-3/2} \sum_{t=1}^{T} I(Z_{t-d} \leq r)X_t \Longrightarrow p(r)\sigma \int_0^1 W(s)\, ds,$$

where $W(\cdot)$ is a standard Brownian motion.

BR1 implies that,

$$T^{-2} \sum_{t=1}^{T} I(Z_{t-d} \leq r)X_t^2 \Longrightarrow p(r)\sigma^2 \int_0^1 W^2(s)\, ds. \tag{9}$$

Next two results are taken from Caner and Hansen (1997).

Basic Result 2 (BR2):

$$X_T(s, p(r)) = \frac{1}{\sqrt{T}} \sum_{t=1}^{[Ts]} I(Z_{t-d} \leq r)e_t \Longrightarrow \sigma W(s, p(r)),$$

where $W(s, p(r))$ is a standard Brownian sheet on $[0,\ 1]^2$.

Definition: *A standard Brownian sheet S indexed by $R^+ \times [0,\ 1]$ is a zero-mean Gaussian process with continuous sample paths and covariance function,*

$$Cov[S(s,u), S(t,v)] = (s \wedge t)(u \wedge v).$$

Basic Result 3 (BR3):

$$\int_0^1 X_T(s)\,dX_T(s,p(r)) \Longrightarrow \sigma^2 \int_0^1 W(s)\,dW(s,p(r)).$$

BR2 and BR3 imply that,

$$\frac{1}{T}\sum_{t=1}^{T} I(Z_{t-d} \leq r) X_{t-1} e_t \Longrightarrow \sigma^2 \int_0^1 W(s)\,dW(s,p(r)). \tag{10}$$

Decomposing $e_t(p(r)) = I(Z_{t-d} \leq r)e_t$ into two orthogonal components (e_t, $v_t(r)$),

$$e_t(p(r)) = a(r)e_t + v_t(p(r)), \tag{11}$$

where $a(r) = E[e_t(p(r))e_t]/E[e_t^2] = p(r)$, we obtain two additional results.

Basic Result 4 (BR4):

$$V_T(s,p(r)) = \frac{1}{\sqrt{T}}\sum_{t=1}^{[Ts]} v_t(p(r)) \Longrightarrow \sigma V(s,p(r)),$$

where $V(s,p(r))$ is a standard Kiefer-Müller process on $[0,\ 1]^2$.

Definition: *A standard Kiefer-Müller process Z on $[0,\ 1]^2$ is given by,*

$$Z(t_1,t_2) = S(t_1,t_2) - t_2 S(t_1,1),$$

where $S(t_1,t_2)$ is a standard Brownian sheet. Then Z has covariance function,

$$Cov[Z(s_1,t_1), Z(s_2,t_2)] = (s_1 \wedge s_2)(t_1 \wedge t_2 - t_1 t_2).$$

Basic Result 5 (BR5):

$$\frac{1}{T}\sum_{t=1}^{[Ts]} X_{t-1} v_t(p(r)) \Longrightarrow \sigma^2 \int_0^1 W(s)\,dV(s,p(r)).$$

Since $W(s)$ is independent of $V(s,p(r))$ it can be proved that for a fixed r,

$$\{\int_0^1 W(s)^2\,ds\}^{-1/2}\int_0^1 W(s)\,dV(s,p(r)) \equiv N(0,\sigma_\nu^2), \tag{12}$$

where $\sigma_\nu^2 = Var(v_t(p(r))/\sigma) = p(r)(1-p(r))$.

4 Testing and Estimation when the Threshold Value is Known

4.1 Testing for Threshold

The case of a threshold value r known becomes relevant for pedagogical reasons as well as for cases where the regimes are determined by the sign of the threshold variable.

Equation (8) can be rearranged as follows,

$$X_t = \alpha_2 X_{t-1} + (\alpha_1-\alpha_2)I(Z_{t-d}\le r)X_{t-1} + e_t = \alpha_2 X_{t-1} + \gamma I(Z_{t-d}\le r)X_{t-1} + e_t. \tag{13}$$

Equation (13) can be re-written as

$$\Delta X_t = \rho X_{t-1} + \gamma I(Z_{t-d}\le r)X_{t-1} + e_t, \tag{14}$$

where $\rho = \alpha_2 - 1$, $\gamma = \alpha_1 - \alpha_2$.

Testing the null hypothesis of no threshold is equivalent to test

$$H_0 : \gamma = 0. \tag{15}$$

The test statistic used for testing (15) will be the t-statistic for $\hat{\gamma}$. Next proposition shows the limiting distribution of this test statistic under the null hypothesis of X_t being $I(0)$, as well as under the null of being $I(1)$.

Proposition 1: *Assume the threshold value is known. Under the null of no threshold plus X_t being I(0) or I(1), the $t_{\gamma=0}$ statistic in equation (14) has the following asymptotic distribution*

$$t_{\gamma=0} \Longrightarrow N(0,1).$$

Proposition 1 states an important result. If the threshold value is known it does not matter if the null hypothesis is a random walk or a stationary process, the limiting distribution for the t-statistic of $\hat{\gamma}$ will follow a standard normal distribution. However, in practice (finite samples), this test statistic is sensitive to X_0. In order to correct this problem we recommend introducing a threshold constant term in the regression model used to compute the t-test,

$$\Delta X_t = \mu_2 + (\mu_1 - \mu_2) I(Z_{t-d} \leq r) + \rho X_{t-1} + \gamma I(Z_{t-d} \leq r) X_{t-1} + e_t. \tag{16}$$

It can be easily shown that in this case $t_{\gamma=0}$ asymptotically also follows a $N(0,1)$ under both null hypotheses.

A special case for testing the null of no threshold against a TUR model is given by imposing the constraint $\rho = 0$ in (14) and testing the null hypothesis of $\gamma = 0$. This is the model used in Section 3 of Caner and Hansen (1997). In this case, it can be shown that the asymptotic distribution of the $t_{\gamma=0}$ statistic is a linear combination of the asymptotic distribution of the Dickey-Fuller unit root t-test and a standard normal distribution.

4.2 Estimation and Testing of the TUR Parameters

When the threshold value is known and X_t satisfies the conditions of Theorem 2, it can be proved (see the central limit theorems for stationary and ergodic time series in Hannan (1973)), that least squares estimation of model (8) produces consistent and asymptotically normal estimators. Basically,

$$D_i^{-1/2} T^{1/2} (\hat{\alpha}_i - \alpha_i) \Longrightarrow N(0,1),$$

where $D_i = \sigma^2 (p_i M)^{-1}$, $i = 1, 2$, $p_1 = p(r)$ and $p_2 = 1 - p(r)$, and $M = E(X_{t-1}^2)$. Therefore, is straightforward to test that one of the $\alpha's$ coefficients is equal to one (the TUR model).

5 Testing and Estimation when the Threshold Value is Unknown

5.1 Testing for Threshold

In general, the threshold value r can be considered to be unknown. When this is the case, it is usually assumed that r lies in a bounded interval, $\tilde{R}$. The null hypotheses continue to be the same but now the test statistic proposed is the supremum of the square of the t-statistic for $\gamma = 0$ over the whole range $\tilde{R}$. Again, the testing procedure is developed for two cases, one considering a stationary null and another considering a unit root null. The next proposition shows the limiting distribution for the test statistic when the null hypothesis is a stationary process.

Proposition 2: *Assume that the threshold value is unknown. Under the null hypothesis of no threshold and I(0), the asymptotic distribution of*

$$S_1 = Sup_{r \in \tilde{R}}\ t^2_{\gamma=0}(r),$$

in equation (14) is,

$$Sup_{r \in \tilde{R}}\ \frac{B^2_{p(r)}}{p(r)(1-p(r))},$$

where $\{B_{p(r)} : 0 \leq p(r) \leq 1\}$ *is a standard Brownian Bridge.*

The following proposition shows the asymptotic distribution of the supremum of the t^2-statistic of $\gamma = 0$ under the null hypothesis of unit root and no threshold.

Proposition 3: *Assume that the threshold value is unknown. Under the null hypothesis of no threshold and I(1), the asymptotic distribution of*

$$S_2 = Sup_{r \in \tilde{R}}\ t^2_{\gamma=0}(r),$$

in equation (14) is,

$$Sup_{r\in\bar{R}} \frac{\{\int_0^1 W(s)\, dV(s,p(r))\}^2}{p(r)(1-p(r))\int_0^1 W(s)^2\, ds},$$

where $V(s,p(r))$ is a standard Kiefer-Müller process on $[0, 1]^2$.

Notice that using (12) in Proposition 3, for a fixed r,

$$t_{\gamma=0}(r) \Longrightarrow \frac{\int_0^1 W(s)\, dV(s,p(r))}{\{p(r)(1-p(r))\int_0^1 W(s)^2\, ds\}^{1/2}} \equiv W(1),$$

as it was showed in Proposition 1.

In practice, distributions of the test statistics introduced in propositions 2 and 3 depend on the initial value X_0. To avoid this dependence the same tests should be derived from equation (16). Their asymptotic distributions are the same ones that appear in propositions 2 and 3 but with variables demeaned by the mean in each regime (μ_1 and μ_2). Asymptotic critical values corresponding to the supremum type of tests derived from equation (16) (S_1^* and S_2^*) are tabulated in Table 3. Critical values were computed using an *iid* $U[0, 1]$ threshold variable.

TABLE 3 ABOUT HERE

5.2 Estimation and Testing of the TUR Parameters

If the threshold value, r, is unknown, it can be shown that it can be estimated consistently (rate T) by the value of r that

$$min_{r\in\bar{R}}\ \hat{\sigma}^2(r), \tag{17}$$

in model (14). Its asymptotic distribution is independent on $\hat{\alpha}_i$ (see Chan (1993)). Therefore, Theorem 1 and propositions 2 and 3 still hold using $\hat{r}$ as the true threshold parameter. In order to test the TUR model, the results in section 4.2 can still be used.

6 Finite Sample Performance of Tests

Ignoring the existence of a threshold when the true model is given by equation (14) leads to an inconsistent estimator of ρ, by the standard argument of the omission of a relevant variable.

To avoid this inconsistency, we need to test for the presence of a threshold before estimating any parameter. In this section we analyze the finite sample performance of the tests developed in the previous section.

In order to calculate the empirical size of S_1 and S_2 for $T = 100$, we generated critical values from 10,000 replications for $T = 500$ using equation (14) under H_0. The threshold variable used in this study follows an *iid* $U[0, 1]$. Table 4 presents these results for S_1 and S_2.

TABLE 4 ABOUT HERE

It is clear that empirical sizes of these tests coincide with their nominal sizes.

To illustrate the power of S_1 and S_2 we specify four different alternative data generating processes for each test. The power is tabulated in Tables 5 and 6 using four nominal sizes, 2, 5, 10 and 20%, and two different sample sizes, $T = 100$ and $T = 500$.

TABLE 5 ABOUT HERE

As expected, the power increases with T, as well as with $|\gamma|$ $(= |\alpha_1 - \alpha_2|)$. For S_1 our experiment follows Hansen's (1996) experiment for an endogenous threshold variable with similar values for γ ($\gamma = -0.6$ and $\gamma = -1$). Comparing the power in both experiments, it can be seen that the power for $T = 100$, $\gamma = -0.6$ and a 5% size, more than duplicates the power for the endogenous threshold case. When $\gamma = -1$, the power increases from 0.7 to almost 1.

TABLE 6 ABOUT HERE

Similarly, for S_2 the power increases with the sample size as well as with $|\gamma|$. In this case our results are similar to the ones in the literature (see Gonzalo and Lee (1996)) on the power of the DF test when the alternative is a stationary AR(1) process. The S_2 test does remarkably well even for an alternative of $\gamma = -0.2$ and $T = 100$.

7 An Application to Interest Rates

It is well known that there exists a relationship between interest rates and inflation since nominal interest rates changes when inflation rate changes. In this section we analyze this relationship using the proposed TUR model with X_t = interest rates and Z_{t-d} = lagged inflation changes.

In this way, the autoregressive model for interest rates can have either a stationary root or a unit root depending on which regime is Z_{t-d}.

The data are 3-month bill U.S. government securities (FYGM3) from Citibase for the period March 1947 to May 1996. We performed DF tests on both variables, X_t and Z_{t-d}. In the interest rates series we are unable to reject the null hypothesis of a unit root at usual significance levels ($\tau_{ADF} = -2.33$). For the increments of the lagged inflation rate the null is clearly rejected.

The value of the supremum t^2-test, over $\tilde{R}$, for $\gamma = 0$ in equation (16) is 54.014 indicating a regime change at a lagged inflation increment ($\Delta\pi_{t-1}$) value of $\hat{r} = -0.0036$. The interval $\tilde{R}$ was set to exclude the top and bottom 15% of the threshold variable. The corresponding probabilities of being in regime 1 ($\Delta\pi_{t-1} \leq \hat{r}$) and 2 ($\Delta\pi_{t-1} > \hat{r}$) are 0.14 and 0.86 respectively. The estimated model is,

$$\begin{aligned} X_t &= \underset{(0.033)}{0.081} + \underset{(0.011)}{0.926}\; I(\Delta\pi_{t-1} \leq \hat{r})X_{t-1} + \underset{(0.006)}{0.992}\; I(\Delta\pi_{t-1} > \hat{r})X_{t-1} \\ &+ \underset{(0.039)}{0.370}\; \Delta X_{t-1} - \underset{(0.039)}{0.191}\; \Delta X_{t-2}, \end{aligned} \tag{18}$$

where standard errors are shown in parentheses. The estimated value of $E(\delta_t^2)$ is 0.966, and a 95% bootstrap confidence interval from 1000 bootstrapping samples is (0.927, 0.979). Therefore, we cannot reject that $E(\delta_t^2)$ is less than one. From Corollary 3, in order to test that interest rates follow a TUR model, is enough to test that one of the threshold coefficients is equal to one. In doing so, the null hypothesis of $\alpha_2 = 1$ is not rejected by a Wald test at any conventional significance level.

8 Conclusion

The models introduced in this paper allow for unit roots without losing the stationarity property. This can be very useful for the analysis of those economic variables measured in rates. For these variables, in finite samples, standard unit root tests do not reject the null hypothesis of a unit root component, although theoretically they cannot have a random walk component (for instance, the variance cannot grow with t). Extensions to a multivariate framework are under current investigation.

Appendix

Proof of Theorem 1: Following Brandt (1986), Theorem 1, page 212, in order to prove strict stationarity we need $E(log|\delta_1|) < 0$. Notice that $E|\delta_1^2| < 1$ implies $E|\delta_1| < 1$, then, applying logarithms and using Jensen's inequality we get,

$$E(log|\delta_1|) \leq logE|\delta_1| < 0.$$

Once the strict stationarity is obtained, the covariance stationary solution follows from Corollary 1, page 132 of Karlsen (1990).

Proof of Corollary 1. In order to prove stationarity we have to compute the unconditional mean and covariance structure of X_t. The variance of δ_t is:

$$Var(\delta_t) = \sum_{i=1}^{2} \alpha_i^2 p_i - [\sum_{i=1}^{2} \alpha_i p_i]^2, \tag{19}$$

where $p_1 = p$ and $p_2 = 1 - p$. Then, assuming $X_0 = 0$ the process is characterized by the following moments:

$$E(X_t) = 0, \tag{20}$$

$$\begin{aligned} Var(X_t) &= Var(\delta_t X_{t-1}) + Var(e_t) + 2Cov(\delta_t X_{t-1}, e_t) \\ &= E(\delta_t^2 X_{t-1}^2) + \sigma^2 \\ &= E(\delta_t^2) Var(X_{t-1}) + \sigma^2 \\ &= \sigma^2[1 - (\sum_{i=1}^{2} \alpha_i^2 p_i)^t]/[1 - (\sum_{i=1}^{2} \alpha_i^2 p_i)], \end{aligned} \tag{21}$$

where the third equality follows from the independence between $\{Z_{t-d}\}$ and $\{e_t\}$.

$$\begin{aligned} Cov(X_t, X_{t-s}) &= E[(\delta_t X_{t-1} + e_t) X_{t-s}] \\ &= (\sum_{i=1}^{2} \alpha_i p_i)^s Var(X_{t-s}) \\ &= \sigma^2 (\sum_{i=1}^{2} \alpha_i p_i)^s [1 - (\sum_{i=1}^{2} \alpha_i^2 p_i)^{t-s}]/[1 - (\sum_{i=1}^{2} \alpha_i^2 p_i)]. \end{aligned} \tag{22}$$

Since $E|\delta_t^2| < 1$, it is clear from (21) and (22) that, when t goes to infinity, the covariance structure of X_t is given by the following expressions:

$$Var(X_t) = \sigma^2/[1 - (\sum_{i=1}^{2} \alpha_i^2 p_i)], \tag{23}$$

and,

$$Cov(X_t, X_{t-s}) = \sigma^2(\sum_{i=1}^{2} \alpha_i p_i)^s/[1 - (\sum_{i=1}^{2} \alpha_i^2 p_i)]. \tag{24}$$

Therefore the autocorrelation of order s is

$$\rho_s = Cov(X_t, X_{t-s})/Var(X_t) = (\sum_{i=1}^{2} \alpha_i p_i)^s, \tag{25}$$

and the process is covariance stationary.

Proof of Corollary 2. In order to simplify the proof we will assume that Z_{t-d} follows a first-order Markov chain. The proof for $N > 1$ is similar. From Theorem 1, the condition for covariance stationarity is given by,

$$\sum_{j=1}^{\infty}(E|\prod_{n=1}^{j} \delta_n^2|)^{1/2} = [(\ 1,\ \ 1\ )\sum_{j=1}^{\infty} F_2^j A_\delta]^{1/2} < \infty, \tag{26}$$

where $A'_\delta = (\alpha_1^2 p,\ \ \alpha_2^2(1-p))$ and $F_2 = \begin{bmatrix} \alpha_1^2 p_{1/1} & \alpha_1^2 p_{1/2} \\ \alpha_2^2 p_{2/1} & \alpha_2^2 p_{2/2} \end{bmatrix}$. For equation (26) to hold we need the spectral radius of F_2 being less than one. The characteristic equation associated with F_2 is, $\lambda^2 - (\alpha_1^2 p_{1/1} + \alpha_2^2 p_{2/2})\lambda + \alpha_1^2\alpha_2^2 p_{1/1}p_{2/2}$. Necessary and sufficient conditions for the largest eigenvalue in absolute value to be less than one are,

$$\alpha_1^2 p_{1/1} + \alpha_2^2 p_{2/2} - \alpha_1^2\alpha_2^2 p_{1/1}p_{2/2} < 1, \tag{27}$$

$$-\alpha_1^2 p_{1/1} - \alpha_2^2 p_{2/2} - \alpha_1^2\alpha_2^2 p_{1/1}p_{2/2} < 1, \tag{28}$$

and,

$$|\alpha_1^2\alpha_2^2 p_{1/1}p_{2/2}| < 1. \tag{29}$$

We need to show that $E(\delta_t^2) < 1$ implies (27), (28) and (29). *Sufficient* conditions for $E(\delta_t^2) < 1$ are,

$$|\alpha_2^2| < 1, \ and \ \alpha_2^2 \leq \alpha_1^2 < \alpha_2^2 + \frac{1-\alpha_2^2}{p}, \tag{30}$$

or,

$$|\alpha_1^2| < 1, \ and \ 1 < \alpha_2^2 < \frac{1-\alpha_1^2 p}{1-p}. \tag{31}$$

It is straightforward to show that conditions (27), (28) and (29) are satisfied when $|\alpha_2^2| < 1$ and $\alpha_2^2 \leq \alpha_1^2 \leq 1$. When $|\alpha_2^2| < 1$ but $1 < \alpha_1^2 < \alpha_2^2 + \frac{1-\alpha_2^2}{p}$ condition (27) holds when $p_{1/2} = 1 - p_{2/2} \geq p_{1/1}$, while (28) and (29) are satisfied for any value of the probabilities and any value of α_2^2. Therefore, (30) satisfies conditions (27), (28) and (29). The rest of the proof, using (31), follows similar steps and it is omitted here.

When the threshold variable follows an N-order Markov process it can be shown after some tedious algebra that the spectral radius of F_2 is less than one if $E|\delta_t^2| < 1$ and $p_{1/2\ldots2} = (1 - p_{2/2\ldots2}) \geq p_{1/1\ldots1} > 0$. The intuition behind these conditions comes from the fact that the characteristic equation of F_2 has $N-2$ roots equal to zero, and the other two roots given by,

$$\lambda^2 - (\alpha_1^2 p_{1/1\ldots1} + \alpha_2^2 p_{2/2\ldots2})\lambda + \alpha_1^2\alpha_2^2 p_{1/1\ldots1}p_{2/2\ldots2}, \tag{32}$$

which has a structure similar to the characteristic equation for the first-order Markov chain.

Proof of Corollary 3. The proof of Corollary 3 follows the same steps as the proof of Corollary 2 but in this case conditions (27), (28) and (29) hold if $\alpha_1 = 1$ and $|\alpha_2| < 1$.

Proof of Theorem 2. Theorem 1 in Chan (1989) establishes that if (X_t) is an *aperiodic* and *irreducible* Markov chain and there exists a small set C, a non-negative measurable function g, and constants $\kappa > 1$, $\iota < 0$, and $B > 0$ such that

$$sup_{x \in C^c} E(\kappa g(X_{t+1}) - g(X_t)/X_t = x) = \iota < 0; \tag{33}$$

$$sup_{x \in C} E(g(X_{t+1}); X_{t+1} \in C^c / X_t = x) \equiv B < \infty; \tag{34}$$

$$g(x) \text{ is bounded away from 0 and } +\infty \text{ on } C. \tag{35}$$

Then (X_t) is geometrically ergodic.

To see that a process satisfying assumptions in Theorem 2 is both *irreducible* and *aperiodic* see Tong (1990) Appendix 1, Proposition A1.7.

Now, let $g(x) = |x| + 1$, then equation (33) reduces to,

$$\begin{aligned} E[\kappa(|\delta_{t+1}X_t + e_{t+1}| + 1) - (|X_t| + 1)/X_t = x] &= E[\kappa(|\delta_{t+1}x + e_{t+1}| + 1) - (|x| + 1)/X_t = x] \\ &= E[\kappa|\delta_{t+1}x + e_{t+1}| + \kappa - (|x| + 1)/X_t = x] \\ &\leq \kappa|x|E(|\delta_{t+1}|/X_t = x) + \kappa - (|x| + 1), \end{aligned} \tag{36}$$

where the last inequality follows from the fact that $|a + b| \leq |a| + |b|$ and $E(|e_{t+1}|/X_t = x) = 0$. Therefore, the supremum of equation (36) will be negative if $\kappa < \frac{1+|x|}{1+|x|E(|\delta_{t+1}|/X_t = x)}$, which is satisfied if $E(|\delta_{t+1}|/X_t = x)$ is in between zero and one. Now,

$$E(|\delta_{t+1}|/X_t = x) = |\alpha_1|P(Z_{t-d} \leq r/x) + |\alpha_2|P(Z_{t-d} > r/x). \tag{37}$$

Since $\alpha_1^2 = 1$ and $|\alpha_2^2| < 1$, then, $0 < E(|\delta_{t+1}|/X_t = x) < 1$.

Equation (34) is given by

$$\begin{aligned} E[|\delta_{t+1}X_t + e_{t+1}| + 1/X_t = x] &= E[|\delta_{t+1}x + e_{t+1}| + 1/X_t = x] \\ &\leq |x|E(|\delta_{t+1}|/X_t = x) + 1. \end{aligned} \tag{38}$$

Expression (38) is finite following the same type of argument as above, and $x \in C = [-c, c]$ which is small.

Therefore, equations (33) and (34) are satisfied by the TUR model and Theorem 2 holds.

Proof of BR1:

The limiting distribution of basic result 1 follows from,

$$\sum_{t=1}^{T} a_t B_t + \sum_{t=1}^{T} b_t A_{t-1} = A_T B_T - A_0 B_0, \tag{39}$$

where $a_t = (A_t - A_{t-1})$ and $b_t = (B_t - B_{t-1})$.

Assume $X_0 = 0$, and let $a_t = I(Z_{t-d} \le r)$ and $B_t = X_t = \sum_{i=1}^{t} e_i$. Then, by (39) we have,

$$\sum_{t=1}^{T} I(Z_{t-d} \le r) X_t = \sum_{t=1}^{T} I(Z_{t-d} \le r) \sum_{t=1}^{T} e_t - \sum_{t=1}^{T} \{e_t \sum_{s<t} I(Z_s \le r)\}. \tag{40}$$

Dividing (40) by $T^{3/2}$,

$$\begin{aligned} T^{-3/2} \sum_{t=1}^{T} I(Z_{t-d} \le r) X_t &= T^{-1} \sum_{t=1}^{T} I(Z_{t-d} \le r) (\frac{1}{\sqrt{T}} \sum_{t=1}^{T} e_t) \\ &- T^{-1/2} \sum_{t=1}^{T} \{e_t T^{-1} \sum_{s<t} I(Z_s \le r)\}. \end{aligned} \tag{41}$$

BR1 is obtained by applying the following results to the right hand side (RHS) of (41),

$$T^{-1} \sum_{t=1}^{T} I(Z_{t-d} \le r) - p(r) \xrightarrow{p} 0, \tag{42}$$

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} e_t \Longrightarrow \sigma W(1), \tag{43}$$

$$T^{-1} \sum_{s<t} I(Z_s \le r) - p(r) t/T \xrightarrow{p} 0, \tag{44}$$

and,

$$T^{-3/2} p(r) \sum_{t=1}^{T} t e_t \Longrightarrow p(r) \sigma [W(1) - \int_0^1 W(s)\, ds], \tag{45}$$

where $\xrightarrow{p}$ denote convergence in probability.

Proof of BR2:

BR2 follows from Theorem 2 in Caner and Hansen (1997).

Proof of BR3:

BR3 follows from Theorem 3 in Caner and Hansen (1997).

Proof of BR4:

Decomposing $e_t(p(r)) = I(Z_{t-d} \leq r)e_t$ into two orthogonal components (e_t and $v_t(p(r))$) we have,

$$e_t(p(r)) = a(r)e_t + v_t(p(r)), \tag{46}$$

where

$$a(r) = E(e_t(p(r))e_t)/E(e_t^2) = p(r). \tag{47}$$

Adding over t and dividing by $\sqrt{T}$, we can write (46) as,

$$T^{-1/2}\sum_{t=1}^{[Ts]} e_t(p(r)) = a(r)T^{-1/2}\sum_{t=1}^{[Ts]} e_t + T^{-1/2}\sum_{t=1}^{[Ts]} v_t(p(r)). \tag{48}$$

Rearranging (48),

$$T^{-1/2}\sum_{t=1}^{[Ts]} v_t(p(r)) = T^{-1/2}\sum_{t=1}^{[Ts]} e_t(p(r)) - a(r)T^{-1/2}\sum_{t=1}^{[Ts]} e_t. \tag{49}$$

Using BR2, the first term of the RHS in (49) converges to,

$$T^{-1/2}\sum_{t=1}^{[Ts]} e_t(p(r)) \Longrightarrow \sigma W(s, p(r)). \tag{50}$$

The second term of the RHS goes to,

$$a(r)T^{-1/2}\sum_{t=1}^{[Ts]} e_t \Longrightarrow p(r)\sigma W(s). \tag{51}$$

From (50) and (51) we get,

$$T^{-1/2}\sum_{t=1}^{[Ts]} v_t(p(r)) \Longrightarrow \sigma[W(s, p(r)) - p(r)W(s,1)], \tag{52}$$

where $W(s, p(r)) - p(r)W(s,1) = V(s, p(r))$ is a standard Kiefer-Müller process on $[0, 1]^2$.

Proof of BR5:

It follows from BR2 and BR3 replacing $e_t(p(r))$ by $v_t(p(r))$.

Proof of Proposition 1.

Case 1: Threshold Known, Stationary Root Case.

This is a standard result. For a formal proof see González and Gonzalo (1997).

Case 2: Threshold Known, Unit Root Case.

Rewriting equation (14),

$$\Delta X_t = W_{t-1}\beta + e_t, \tag{53}$$

where $W_{t-1} = [X_{t-1} \quad I(Z_{t-d} \leq r)X_{t-1}]$ and $\beta = [\rho \quad \gamma]$. Then,

$$\hat{\beta} = (\sum_{t=1}^{T} W'_{t-1}W_{t-1})^{-1}\sum_{t=1}^{T} W'_{t-1}\Delta X_t, \tag{54}$$

or,

$$\hat{\beta} - \beta = (\sum_{t=1}^{T} W'_{t-1}W_{t-1})^{-1}\sum_{t=1}^{T} W'_{t-1}e_t. \tag{55}$$

Now,

$$\sum_{t=1}^{T} W'_{t-1}W_{t-1} = \begin{bmatrix} \sum_{t=1}^{T} X_{t-1}^2 & \sum_{t=1}^{T} I(Z_{t-d} \leq r)X_{t-1}^2 \\ \sum_{t=1}^{T} I(Z_{t-d} \leq r)X_{t-1}^2 & \sum_{t=1}^{T} I(Z_{t-d} \leq r)X_{t-1}^2 \end{bmatrix}. \tag{56}$$

Defining $\Upsilon_T = \begin{bmatrix} T & 0 \\ 0 & T \end{bmatrix}$ and multiplying both sides of (55) we get,

$$\Upsilon_T(\hat{\beta} - \beta) =$$

$$\begin{bmatrix} T^{-2}\sum_{t=1}^{T} X_{t-1}^2 & T^{-2}\sum_{t=1}^{T} I(Z_{t-d} \leq r)X_{t-1}^2 \\ T^{-2}\sum_{t=1}^{T} I(Z_{t-d} \leq r)X_{t-1}^2 & T^{-2}\sum_{t=1}^{T} I(Z_{t-d} \leq r)X_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} T^{-1}\sum_{t=1}^{T} X_{t-1}e_t \\ T^{-1}\sum_{t=1}^{T} I(Z_{t-d} \leq r)X_{t-1}e_t \end{bmatrix}. \tag{57}$$

Then, using BR1 and (9),

$$\left[\begin{array}{cc} T^{-2}\sum_{t=1}^{T} X_{t-1}^2 & T^{-2}\sum_{t=1}^{T} I(Z_{t-d}\leq r)X_{t-1}^2 \\ T^{-2}\sum_{t=1}^{T} I(Z_{t-d}\leq r)X_{t-1}^2 & T^{-2}\sum_{t=1}^{T} I(Z_{t-d}\leq r)X_{t-1}^2 \end{array}\right]^{-1} \Longrightarrow$$

$$(p(1-p)\sigma^2 \int_0^1 W(j)^2\, dj)^{-1} \left[\begin{array}{cc} p & -p \\ -p & 1 \end{array}\right]. \tag{58}$$

Decomposing $I(Z_{t-d} \leq r)e_t$ into two orthogonal components (e_t and v_t) we get $I(Z_{t-d} \leq r)e_t = a(r)e_t + v_t$. Then, we can write,

$$T^{-1}\sum_{t=1}^{T} I(Z_{t-d}\leq r)X_{t-1}e_t = a(r)T^{-1}\sum_{t=1}^{T} X_{t-1}e_t + T^{-1}\sum_{t=1}^{T} X_{t-1}v_t,$$

where using standard arguments, the first term of the RHS goes to $\frac{1}{2}p\sigma^2[W(1)^2-1]$, and the second term converges to $(p(1-p)\sigma^4 \int_0^1 W(j)^2\, dj)^{1/2}W(1)$. Therefore,

$$\left[\begin{array}{c} T^{-1}\sum_{t=1}^{T} X_{t-1}e_t \\ T^{-1}\sum_{t=1}^{T} I(Z_{t-d}\leq r)X_{t-1}e_t \end{array}\right] \Longrightarrow \left[\begin{array}{c} \frac{1}{2}\sigma^2[W(1)^2-1] \\ \frac{1}{2}p\sigma^2[W(1)^2-1] + (p(1-p)\sigma^4\int_0^1 W(j)^2\, dj)^{1/2}W(1) \end{array}\right]. \tag{59}$$

Then,

$$T(\hat{\gamma}-\gamma) \Longrightarrow [p(1-p)\int_0^1 W(j)^2\, dj]^{-1/2}W(1), \tag{60}$$

implying,

$$t_{\hat{\gamma}} \Longrightarrow W(1). \tag{61}$$

Proof of Proposition 2:

The asymptotic distribution of the supremum of the $t^2_{\gamma=0}(r)$ statistic is the same as the one developed by Chan and Tong (1990) for their LR test, special case (a), using $p(r) = \frac{E[I(Z_{t-d}\leq r)X_{t-1}^2]}{Var(X_t)}$ instead of their $s(r)$.

Proof of Proposition 3:

Using (9) we have,

$$\left[\begin{array}{cc} T^{-2}\sum_{t=1}^{T} X_{t-1}^2 & T^{-2}\sum_{t=1}^{T} I(Z_{t-d}\leq r)X_{t-1}^2 \\ T^{-2}\sum_{t=1}^{T} I(Z_{t-d}\leq r)X_{t-1}^2 & T^{-2}\sum_{t=1}^{T} I(Z_{t-d}\leq r)X_{t-1}^2 \end{array}\right]^{-1} \Longrightarrow$$

$$(p(r)(1-p(r))\sigma^2 \int_0^1 W(s)^2\,ds)^{-1} \left[\begin{array}{cc} p(r) & -p(r) \\ -p(r) & 1 \end{array}\right]. \tag{62}$$

Decomposing $e_t(p(r)) = I(Z_{t-d}\leq r)e_t$ into two orthogonal components (e_t and $v_t(p(r))$) we get,

$$e_t(p(r)) = a(r)e_t + v_t(p(r)), \tag{63}$$

where $a(r) = E[e_t(p(r))e_t]/E[e_t^2] = p(r)$. Then,

$$T^{-1}\sum_{t=1}^{T} I(Z_{t-d}\leq r)X_{t-1}e_t = a(r)T^{-1}\sum_{t=1}^{T} X_{t-1}e_t + T^{-1}\sum_{t=1}^{T} X_{t-1}v_t(p(r)),$$

where using standard arguments, the first term of the RHS goes to $\frac{1}{2}p(r)\sigma^2[W(1)^2-1]$. Using BR5, the second term of the RHS converges to $\sigma^2\int_0^1 W(s)\,dV(s,p(r))$. Therefore,

$$\left[\begin{array}{c} T^{-1}\sum_{t=1}^{T} X_{t-1}e_t \\ T^{-1}\sum_{t=1}^{T} I(Z_{t-d}\leq r)X_{t-1}e_t \end{array}\right] \Longrightarrow \left[\begin{array}{c} \frac{1}{2}\sigma^2[W(1)^2-1] \\ \frac{1}{2}p(r)\sigma^2[W(1)^2-1] + \sigma^2\int_0^1 W(s)\,dV(s,p(r)) \end{array}\right]. \tag{64}$$

Then,

$$T(\hat{\gamma}-\gamma) \Longrightarrow \frac{\int_0^1 W(s)\,dV(s,p(r))}{p(r)(1-p(r))\int_0^1 W(s)^2\,ds}, \tag{65}$$

implying,

$$t_{\hat{\gamma}} \Longrightarrow \frac{\int_0^1 W(s)\,dV(s,p(r))}{\{p(r)(1-p(r))\int_0^1 W(s)^2\,ds\}^{1/2}}. \tag{66}$$

Therefore,

$$t_{\hat{\gamma}}^2 \Longrightarrow \frac{\{\int_0^1 W(s)\,dV(s,p(r))\}^2}{p(r)(1-p(r))\int_0^1 W(s)^2\,ds}, \tag{67}$$

and the limiting distribution of the supremum of (67) follows from the continuous mapping theorem.

References

Beveridge, S. and C. Nelson (1981), "A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the business cycle". *Journal of Monetary Economics*, 7, pp. 151-174.

Brandt, A. (1986), "The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients". *Advances in Applied Probability*, 18, pp. 211-220.

Caner, M. and B. Hansen (1997), "Threshold autoregressions with a unit root". Mimeo.

Chan, K.S. (1993), "Consistency and limiting distribution of the least squares estimator of a threshold autoregressive model". *The Annals of Statistics*, Vol 21, No. 1, pp. 520-533.

Chan, K.S. (1989), "A note on the geometric ergodicity of a Markov chain". *Advances in Applied Probability*, 21, pp. 702-704.

Chan, K.S. and H. Tong (1990), "On likelihood ratio tests for threshold autoregression". *Journal of the Royal Statistical Society*, Vol. 52, No. 3, pp. 469-476.

Chan, K.S., J.D. Petruccelli, H. Tong and S. W. Woolford (1985), "A multiple threshold AR(1) model". *Journal of Applied Probability*, 22, pp. 267-279.

González, M. and J. Gonzalo (1997) "Threshold unit root models". WP 97-50 (21), Statistics and Econometrics, U. Carlos III de Madrid.

Gonzalo, J. and T. Lee (1996), "Relative power of t type tests for stationary and unit root processes". *Journal of Time Series Analysis*, Vol 17, No. 1., pp. 37-47.

Hannan, E.J. (1973), "Central limit theorems for time series regression". *Z. Wahrsch. Verw.* Gebiete 26 pp. 157-170.

Hansen, B. (1996), "Inference when a nuisance parameter is not identified under the null hypothesis". *Econometrica*, Vol 64, No. 2, pp. 413-430.

Hansen, B. (1996), "Estimation of TAR models". Mimeo.

Karlsen, H.A. (1990), "Existence of moments in a stationary stochastic difference equation". *Advances in Applied Probability*, 22, pp. 129-146.

Nelson, C. and C. Plosser (1982), "Trends and random walks in macroeconomics time series: some evidence and implications". *Journal of Monetary Economics*, 10, pp. 130-162.

Nicholls, D.F. and B.G. Quinn (1982) "Random coefficient autoregressive models: an introduction". *Springer-Verlag, Lecture Notes in Statistics*, Vol. 11.

Tong, H. (1990), "Non-Linear time series: A dynamical system approach". *Oxford Statistical Science Series*, 6.

Tong, H. (1983), "Threshold models in non-linear time series analysis". *Springer and Verlag, Lectures Notes in Statistics*, vol 21.

Table 1: Empirical Power of Unit Root Tests
Model (5)

Critical Values	1%	5%	10%
(1) No Constant, No Trend	0.018	0.081	0.146
(2) Constant, No Trend	0.009	0.049	0.102
(3) Constant, Trend	0.011	0.047	0.100
(4) No Constant, No Trend	0.015	0.071	0.140
(5) Constant, No Trend	0.010	0.056	0.110
(6) Constant, Trend	0.009	0.051	0.103
(7) No Constant, No Trend	0.040	0.181	0.328
(8) Constant, No Trend	0.015	0.075	0.145
(9) Constant, Trend	0.012	0.059	0.115
(10) No Constant, No Trend	0.105	0.376	0.597
(11) Constant, No Trend	0.029	0.128	0.236
(12) Constant, Trend	0.019	0.082	0.158
(13) No Constant, No Trend	0.371	0.801	0.932
(14) Constant, No Trend	0.103	0.339	0.527
(15) Constant, Trend	0.051	0.197	0.336
(16) No Constant, No Trend	0.727	0.971	0.993
(17) Constant, No Trend	0.273	0.650	0.821
(18) Constant, Trend	0.143	0.406	0.598
(19) No Constant, No Trend	0.924	0.995	0.999
(20) Constant, No Trend	0.543	0.875	0.953
(21) Constant, Trend	0.319	0.665	0.824
(22) No Constant, No Trend	0.984	0.999	0.999
(23) Constant, No Trend	0.782	0.963	0.990
(24) Constant, Trend	0.549	0.851	0.940

Note: Simulations were computed using model (5), sample size equal to 100, and 10000 replications.

Table 2: Empirical Power of Unit Root Tests
Model (3)

Critical Values	1%	5%	10%
(1) No Constant, No Trend	0.018	0.081	0.146
(2) Constant, No Trend	0.009	0.049	0.102
(3) Constant, Trend	0.011	0.047	0.100
(4) No Constant, No Trend	0.692	0.970	0.982
(5) Constant, No Trend	0.266	0.631	0.812
(6) Constant, Trend	0.147	0.401	0.588
(7) No Constant, No Trend	0.915	0.988	0.991
(8) Constant, No Trend	0.533	0.861	0.955
(9) Constant, Trend	0.322	0.650	0.809
(10) No Constant, No Trend	0.983	0.992	0.994
(11) Constant, No Trend	0.769	0.964	0.986
(12) Constant, Trend	0.559	0.838	0.932
(13) No Constant, No Trend	0.990	0.994	0.995
(14) Constant, No Trend	0.905	0.988	0.991
(15) Constant, Trend	0.753	0.941	0.983

Note: Simulations were computed using model (3), sample size equal to 100, and 10000 replications.

Table 3: Critical Values

	S_1^*	S_2^*
1%	0.536	0.489
2.5%	0.678	0.654
5%	0.850	0.832
10%	1.092	1.089
25%	1.705	1.688
50%	2.729	2.722
75%	4.221	4.227
90%	6.165	6.106
95%	7.576	7.520
97.5%	8.897	8.959
99%	10.674	10.818

Note: Critical values were computed using a sample size equal to 500, and 10000 replications.

Table 4: Finite Sample Size of Tests

	S_1	S_2
2%	0.026	0.022
5%	0.055	0.055
10%	0.108	0.103
20%	0.202	0.202

Note: Critical values were computed using a sample size equal to 500, and 10000 replications.

Table 5: Power of S_1 Test

	$T = 100$, $\alpha_2 = 0.3$				$T = 500$, $\alpha_2 = 0.3$			
	$\alpha_1 = -0.7$	$\alpha_1 = -0.3$	$\alpha_1 = 0$	$\alpha_1 = 0.1$	$\alpha_1 = -0.7$	$\alpha_1 = -0.3$	$\alpha_1 = 0$	$\alpha_1 = 0.1$
2%	0.985	0.584	0.125	0.060	1.000	1.000	0.728	0.339
5%	0.994	0.693	0.202	0.110	1.000	1.000	0.821	0.474
10%	0.997	0.792	0.310	0.194	1.000	1.000	0.894	0.597
20%	0.999	0.881	0.463	0.328	1.000	1.000	0.940	0.719

Note: Percentages of rejection were computed using 10000 replications.

Table 6: Power of S_2 Test

	$T = 100$, $\alpha_2 = 1$				$T = 500$, $\alpha_2 = 1$			
	$\alpha_1 = 0.7$	$\alpha_1 = 0.8$	$\alpha_1 = 0.9$	$\alpha_1 = 0.95$	$\alpha_1 = 0.7$	$\alpha_1 = 0.8$	$\alpha_1 = 0.9$	$\alpha_1 = 0.95$
2%	0.488	0.297	0.129	0.061	0.999	0.974	0.748	0.364
5%	0.600	0.410	0.205	0.111	1.000	0.988	0.832	0.507
10%	0.713	0.529	0.306	0.191	1.000	0.996	0.898	0.623
20%	0.822	0.672	0.447	0.314	1.000	0.999	0.940	0.741

Note: Percentages of rejection were computed using 10000 replications.