THRESHOLD STOCHASTIC UNIT ROOT MODELS

By Jesus Gonzalo* and Raquel Montesinos
Dept. of Statistics and Econometrics
Universidad Carlos III de Madrid

Last Version: September, 2002

Abstract

This paper introduces a new class of stochastic unit root (STUR) processes, where the randomness of the autorregresive unit root is driven by a threshold variable. These new models, the threshold autorregresive stochastic unit root (TARSUR) models, are stationary in some regimes and mildly explosive in others. TARSUR models are not only an alternative to fixed unit root models but present interpretation, estimation and testing advantages with respect to the existent STUR models. The paper analyzes the stationarity properties of the TARSUR models and proposes a simple t-statistic for testing the null hypothesis of a fixed unit root versus a stochastic unit root hypothesis. It is shown that its asymptotic distribution (AD) depends on the knowledge we have about the threshold values: known, unknown but identified, and unknown and unidentified. In the first two cases the AD is a standard Normal distribution, while in the last one the AD is a functional of Brownian Motions and Brownian Sheets. Monte Carlo simulations show that the proposed tests behave very well in finite samples and that the Dickey-Fuller test cannot easily distinguish between an exact unit root and a threshold stochastic unit root. The paper concludes with applications to stock prices and interest rates where the hypothesis of a fixed unit root is rejected in favor of the threshold stochastic unit root.

Key Words: Dickey-Fuller test; Difference stationary; Nonstationary time series; Stochastic difference equations; Stochastic unit roots; Time varying coefficients; Threshold models; Unit roots.

---

*Corresponding address: http://www.uc3m.es/uc3m/dpto/DEE/profesorado/personal/jgonzalo - Email:jgonzalo@est-econ.uc3m.es. Financial support DGCYT Grant (SEC01-0890) is gratefully acknowledged. We thank Karim Abadir, Juan Dolado, Francesc Marmol, Hashem Pesaran, Ruey Tsay, Mike Wickens and Taner Yigit for stimulating comments on earlier drafts. Special thanks go to two diligent referees whose careful reading of the first draft of this paper has seriously helped us to improve this version.

1. INTRODUCTION

It is well established that many economic series contain dominant, smooth components, even after removal of simple deterministic trends. Since the influential work of Nelson and Plosser (1982), this characteristic has been adequately captured by unit root (UR) models and unit roots have become a "stylized fact" for most of the macroeconomic and financial time series. This has produced an extensive literature on econometric issues related to unit root models (see Phillips and Xiao (1998) for a recent survey).

Trying to get away from the very tight constraints that an exact unit root imposes in a process, and to be able to generate more flexible models and of a more realistic kind, the research has recently evolved in two directions. The first one generalizes UR models by allowing for fractional roots: $ARFIMA$ models (see Granger and Joujeux (1980), Beran (1994), Robinson (1994), and Baillie (1996)). The second one makes the UR models more flexible by allowing the unit root to be stochastic (see Leybourne, McCabe and Tremayne (1996), Leybourne, McCabe and Mills (1996), McCabe and Tremayne (1995), and Granger and Swanson (1997)) instead of a fixed parameter. With both extensions more general forms of nonstationarity are allowed than the ones implied by the standard exact unit root autoregresive models. This paper forms part of the second line of research.

Stochastic unit root models ($STUR$) are seen to arise naturally in economic theory, as well as in many macroeconomic applications (see Leybourne, McCabe and Mills (1996), Granger and Swanson (1997)). $STUR$ models can be stationary for some periods or regimes, and mildly explosive for others. This characteristic makes them not to be difference stationary. If a series shows evidence of a nonstationarity not removable by differencing, it is inappropriate to estimate conventional $ARIMA$ or cointegration/error-correction models because the properties of the estimators and the tests involved are not the same as those in the standard difference-stationary case. For instance, two series generated by two independent $STUR$ models will be wrongly detected to be cointegrated according to some of the most commonly used cointegration tests (see Gonzalo and Lee (1998)). This problem is not detected with standard unit root tests, such as the Dickey-Fuller test, because they cannot easily distinguish between exact unit roots and stochastic unit roots. In order to obtain a better statistical distinction between these two type of unit roots, McCabe and Tremayne (1995) proposes a locally best invariant test (assuming gaussianity) for the null hypothesis of difference stationary versus a stochastic unit root. The application of this constancy parameter test to the macroeconomic variables analyzed in Nelson and Plosser (1982) suggests that about half of them are not difference stationary, opposite to what it has been widely believed (see Leybourne, McCabe and Tremayne (1996)). Hence, the notion that some economic time series are nonstationary in a rather more general way needs to be considered and, consequently, more elaborate techniques of modelling and estimation need to be explored.

From a statistical point of view, a suitable justification for using time varying parameter models to approximate or represent nonstationary processes is provided by Cramer's (1961) extension of Wold's theorem (see Granger and Newbold (1986), page 38). This extension implies that any nonstationary stochastic process, with finite second order moments, may be written as an *ARMA* process with coefficients that are allowed to vary with time. Most of the literature previously cited above considers that the time varying unit root varies as a sequence of independent and identically distributed (i.i.d.) random variables. This assumption is not necessarily the most appropriate in economics, because it implies that the model structure will change too often between states corresponding to stationary and explosive roots, whereas in reality, we might suppose that the transition between these two states occurs in a more gradual fashion. One way of introducing this gradual behavior is by allowing the unit autoregressive root itself to follow a random walk (see Leybourne, McCabe and Mills (1996)). In this case the change is smoother than in the i.i.d. case, but it has again the inconvenience that it occurs regularly at every moment of time. In this paper it is assumed that the economy stays in a "good" or "bad" state for a number of periods of time until certain determining variables overpass some key values. When this occurs the economy jumps from one state to the other type of state. This assumption is perfectly captured by modelling the evolution of economic variables via threshold models. In particular to model the random behavior of the largest root of an ARMA process, we propose a threshold autoregressive (TAR) model where the largest root is less than one in some regimes, and larger than one in others, in such a way that on average is equal to one. These threshold autorregresive stochastic unit root (*TARSUR*) models present several advantages with respect the previously mentioned approaches. First, its computational simplicity. The estimation of all the parameters is done by least squares (LS) regressions. Second, the t-statistic used to test the hypothesis of exact unit root versus stochastic unit root, in some cases follows asymptotically a standard distribution and therefore there is not need to generate new critical values. Third, we are able to introduce deterministic components with threshold effects. Fourth, the threshold variable is suggested by economic theory and it will be providing a possible explanation or cause for the existence of a unit root, something that to the best of our knowledge it is still absent in the econometric literature. And fifth, in many situations threshold models are easier to use for forecasting than random coefficient models. This is the case when the threshold variable is an observable variable with past time dependency.

The rest of the paper is organized as follows. In Section 2 we define the *TARSUR* model and examine its properties: strict stationarity, covariance stationarity and impulse response function. In Section 3, we present a t-test for testing the null hypothesis of an exact unit root versus a stochastic unit root. The asymptotic distribution of this test is developed under three different situations: when the threshold value is known, when the threshold value is unknown but identified and when the threshold value is unknown and unidentified. The finite sample performance (size and power) of the tests developed in this paper is analyzed in Section 4. Section 5 presents two empirical applications of our model: U.S. stock prices and international interest rates. The conclusions are found in Section 6. Proofs are provided in the Appendix.

2. *TARSUR* MODEL

Consider the following threshold first order autoregressive model

$$\begin{aligned} Y_t &= \left[\rho_1 I(Z_{t-d} \leq r_1) + \cdots + \rho_n I(Z_{t-d} > r_{n-1})\right] Y_{t-1} + \varepsilon_t \\ &= \delta_t Y_{t-1} + \varepsilon_t, \qquad t = 1, 2, \cdots, \end{aligned} \tag{1}$$

where $\delta_t = \rho_1 I(Z_{t-d} \leq r_1) + \cdots + \rho_n I(Z_{t-d} > r_{n-1})$, $I(\cdot)$ is an indicator function, and ε_t is an innovation term. Z_t is the threshold variable and in this paper will be a predetermined variable $(E(\varepsilon_{t+j}|Z_t) = 0, \forall j \geq 0)$. d is the delay parameter, and $r_1 < r_2 < \cdots < r_{n-1}$ are the threshold values determining the n different regimes.

Definition 1 *A first order TARSUR process is defined by equation (1) with* $E(\delta_t) = \sum_{i=1}^{n} \rho_i p_i = 1$, *where* p_i *is the probability of* Z_{t-d} *being in regime* i, *and* $V(\delta_t) > 0$.

For simplicity and without loss of generality, in this section where the properties of the *TARSUR* model are analyzed, no deterministic terms are included. They will be taken into account in the testing section, that is where they can really make a difference.

The variables $\{\varepsilon_t\}$ and $\{Z_t\}$ satisfy the following assumptions.

Assumptions

(A.1) $\{\varepsilon_t, Z_t\}$ is strictly stationary, ergodic, adapted to the *sigma-field* $\Im_t \overset{def}{=} \{(\varepsilon_j, Z_j), j \leq t\}$.

(A.2) $\{\varepsilon_t, Z_t\}$ is strong mixing with mixing coefficients α_m satisfying $\sum_{m=1}^{\infty} \alpha_m^{1/2-1/r} < \infty$ for some $r > 2$.

(A.3) ε_t is independent of $\Im_{t-1}$, $E(\varepsilon_t) = 0$ and $E\left|\varepsilon_t\right|^4 = k < \infty$.

(A.4) Z_t has a continuous and increasing distribution function.

(A.5) $E(\max(0, \log|\varepsilon_1|)) < \infty$.

(A.6) $ess.\sup\varepsilon_1 < \infty$[1].

Assumptions (A.1) and (A.3) specify that the error term is a conditionally homoskedastic martingale difference sequence. (A.3) also bounds the extent of heterogeneity in the conditional distribution of ε_t. (A.1), (A.2), (A.3), and (A.4) are needed to obtain the asymptotic distributions of the statistics proposed in this paper. Assumptions (A.1) and (A.5) are required for strict stationarity of Y_t, and (A.6) is needed for weak stationarity of Y_t. In many cases, (A.6) can be relaxed. For instance, if $\{\varepsilon_t\}$ and $\{Z_t\}$ are mutually independent, (A.6) can be replaced by $||\varepsilon_1||_p = [E|\varepsilon_1|^p]^{1/p} < \infty$, $\forall p < \infty$ (see Karlsen (1990)).

It is important to notice that if we limit the analysis to self exciting threshold autorregresive models ($Z_t = Y_t$), then it is not possible to handle the issue of stochastic unit roots (unless we introduce deterministic components with size and sign contraints). This is so because if any of the parameters ρ_i is larger than one, the process Y_t will not be stationary and ergodic (see Petrucelly and Woolford (1984)) and therefore assumption (A.1) will not hold.

Equation (1) represents a particular case of a stochastic difference equation, where δ_t is a discrete random variable that takes different values depending on the location of the threshold variable Z_{t-d}. In the next subsection we present the results from the theory of stochastic difference equations, that are useful to analyze the stationary properties and the impulse response function of a *TARSUR* process. The section concludes examining the consequences of differencing a *TARSUR* process.

2.2. Some preliminary results

Consider the following general first order stochastic difference equation

$$Y_t = \omega_t Y_{t-1} + \varepsilon_t, \qquad t = 1, 2, \cdots, \tag{2}$$

where $\{(\omega_t, \varepsilon_t)\}$ is a R^2-valued stochastic process on a probability space $(\Omega, \Im, P)$.

Iterating backwards the stochastic difference equation (2), we obtain

$$Y_t = \varepsilon_t + \sum_{j=1}^{n-1}\left(\prod_{i=0}^{j-1}\omega_{t-i}\right)\varepsilon_{t-j} + \left(\prod_{i=0}^{n-1}\omega_{t-i}\right)Y_{t-n}$$

[1] The essential supremum of X is *ess sup* $X = \inf\{x : P(|X| > x) = 0\} = ||x||_\infty$.

$$= \quad C_{1,t}(n) + C_{2,t}(n), \tag{3}$$

where $C_{1,t}(n) = \varepsilon_t + \sum_{j=1}^{n-1} \left(\prod_{i=0}^{j-1} \omega_{t-i} \right) \varepsilon_{t-j}$, and $C_{2,t}(n) = \left(\prod_{i=0}^{n-1} \omega_{t-i} \right) Y_{t-n}$.

From (2) and (3) the following results are obtained:

(a) if $C_{1,t}(n)$ converges, as $n \to \infty$ in L^p for $p \in [0, \infty]$ [2], then $C_{1,t} = \varepsilon_t + \sum_{j=1}^{\infty} \left(\prod_{i=0}^{j-1} \omega_{t-i} \right) \varepsilon_{t-j}$ is a strictly stationary solution of the stochastic difference equation defined by (2).

(b) if $C_{2,t}(n)$ converges in probability to zero, then the above solution is unique.

(c) if $p > 0$ in result (a), then $\{Y_t\}$ has a finite pth order moment.

The problem of finding conditions on $\{(\omega_t, \varepsilon_t)\}$ such that $\{Y_t\}$ has a strictly or second-order stationary solution has been studied by several authors. Vervaat (1979) and Nicholls and Quinn (1982) assume $\{(\omega_t, \varepsilon_t)\}$ to be *i.i.d.* and mutually independent. Pourahmadi (1986, 1988) and Tjϕstheim (1986) allow $\{\omega_t\}$ to be a dependent process. More general conditions are given in the following theorem based on Brandt (1986) and Karlsen (1990).

Theorem 1 *If the sequence $\{\varepsilon_t, Z_t\}$ satisfies assumptions (A.1), (A.5), and*

$$-\infty < E \log |\omega_1| < 0 \tag{4}$$

holds, then process (2) is strictly stationary. Moreover, if (A.6) is satisfied and

$$\sum_{j=0}^{\infty} \left(E \left| \psi_{t,j} \right|^2 \right)^{\frac{1}{2}} < \infty, \tag{5}$$

where $\psi_{t,0} = 1$ and $\psi_{t,j} = \prod_{i=0}^{j-1} \omega_{t-i}$ for $j \geq 1$, then process (2) is second-order stationary.

Theorem 1 provides sufficient conditions for (a) and (b) to hold when $p = 0, 1,$ or 2. It shows that strict and covariance stationarity will depend on the type of convergence of the infinite sequences $\left\{ \psi_{t,j} \right\}_{j=0}^{\infty}$. In fact, if condition (4) is satisfied, $\left\{ \psi_{t,j} \right\}$ will converge absolutely almost sure to zero as j goes to infinity, and this implies the strict stationarity of process (2) (see Brandt (1986)). Mean square convergence of $\left\{ \psi_{t,j} \right\}_{j=0}^{\infty}$ is obtained provided condition (5) holds, and in this case, process (2) is also second order stationary.

Note that there is a trade off between (A.6) and (5). For instance, assumption (A.6) can be relaxed by imposing $||\varepsilon_1||_p < \infty$, $\forall p < \infty$; but in this case, we need to modify (5) requiring a stronger condition

$$\sum_{j=0}^{\infty} \left(E \left| \psi_{t,j} \right|^{2+\delta} \right)^{\frac{1}{2+\delta}} < \infty, \text{ for a } \delta > 0. \tag{6}$$

[2] L^0 is equivalent to converge in probability.

Also, as it is mentioned before, if it is assumed that $\{\varepsilon_t\}$ and $\{Z_t\}$ are mutually independent with $||\varepsilon_1||_p < \infty$, $\forall p < \infty$, then condition (5) is a sufficient condition for second-order stationarity.

For the impulse response function (IRF) of Y_t, we need to derive its $MA(\infty)$ representation. This is possible from the conditions of the first part of Theorem 1 and it can be written as

$$Y_t = \varepsilon_t + \sum_{j=1}^{\infty} \left(\prod_{i=0}^{j-1} \omega_{t-i} \right) \varepsilon_{t-j} = \sum_{j=0}^{\infty} \psi_{t,j} \varepsilon_{t-j}. \tag{7}$$

From this representation, it is seen that the response of Y_t to a shock, $\frac{\delta Y_{t+h}}{\delta \varepsilon_t} = \psi_{t,h}$, becomes now stochastic in contrary to the fixed root case. For this reason, we define the impulse response function *(IRF)* as

$$\xi_h = E\left(\frac{\delta Y_{t+h}}{\delta \varepsilon_t}\right) = E\left(\psi_{t,h}\right) = E\left(\prod_{i=0}^{h-1} \omega_{t-i}\right), \quad h = 0, 1, 2, \cdots. \tag{8}$$

Theorem 1 produces explicit conditions for strict stationarity. However, no moments need to exist and to the best of our knowledge, there are not explicit conditions for second-order stationarity or for the convergence of the *IRF* (8), and therefore we must study each particular case. In order to obtain explicit expressions, in this Section 2, we work with the following representative case:

ω_t is a 1^{st}-order stationary Markov Chain with two regimes or states (v_1 and v_2).

This case can be generalized to an *N-order stationary Markov Chain* with $N > 1$, and to more than two regimes, but nothing is gained on the understanding of the process and the algebra become very tedious.

Sufficient conditions for second-order stationarity are presented in the following proposition.

Proposition 1 *Let ω_t be a 1^{st}-order stationary Markov Chain with two regimes (v_1 and v_2). Define the following 2×2 matrix*

$$F_2 = \begin{pmatrix} v_1^2 p_{11} & v_1^2 p_{21} \\ v_2^2 p_{12} & v_2^2 p_{22} \end{pmatrix},$$

where p_{ji} denotes the conditional probability $P(\omega_t = v_i \mid \omega_{t-1} = v_j)$, $i, j = 1, 2$. If the spectral radius of F_2, $\rho(F_2)$, is less than one, the process is covariance stationary.

Notice that if we consider ω_t to be an *i.i.d.* process, the sufficient condition for covariance stationarity can be formulated in terms of the marginal probabilities:

$$\rho(F_2) < 1 \Longleftrightarrow E(\omega_t^2) = v_1^2 p_1 + v_2^2 p_2 < 1.$$

This is the necessary and sufficient condition used in Nicholls and Quinn (1982) for the stationarity of random coefficient autoregressive models (RCA).

Proposition 2 *Under conditions of Proposition 1, the IRF of the process Y_t is given by*

$$\xi_h = \begin{pmatrix} 1 & 1 \end{pmatrix} F_1^h \begin{pmatrix} v_1 p_1 \\ v_2 p_2 \end{pmatrix}, \quad h = 1, 2, \cdots,$$

where $F_1 = \begin{pmatrix} v_1 p_{11} & v_1 p_{21} \\ v_2 p_{12} & v_2 p_{22} \end{pmatrix}$. *Shocks have transitory effects* ($\lim_{h\to\infty} \xi_h = 0$) *if and only if the spectral radius of* F_1, $\rho(F_1)$ *is less than one.*

Proposition 1 together with Proposition 2 establish that the covariance structure and the convergence of the IRF depend on the transition probabilities p_{ji}, and on the regime parameter values v_i.

2.3 Stationarity properties, covariance structure and impulse response function of a *TARSUR* model

As we mentioned before, a $TARSUR$ process is a particular case of a stochastic difference equation. In order to present its properties we will make constant use of the results obtained in the previous subsection.

Corollary 1 *A TARSUR process with $\rho_i \geq 0$, $\forall i$, is strictly stationary.*

Corollary 1 follows from Theorem 1, and it establishes sufficient conditions, easy to check, for a $TARSUR$ process to be strictly stationary. It covers the most appealing $TARSUR$ model from an empirical point of view, that is to say, the model with ρ_i values around unity: stationary for some regimes and mildly explosive for others. Notice that fixed unit root models are not stationary, but if we allow the root to be stochastic around unity we can achieve at least strict stationarity.

In order to present the second-order properties of a $TARSUR$ process, we adapt the particular representative case previously considered to the threshold framework. More concrete

$$
\begin{aligned}
Y_t &= [\rho_1 I(Z_{t-d} \leq r) + \rho_2 I(Z_{t-d} > r)]\, Y_{t-1} + \varepsilon_t \\
&= \delta_t Y_{t-1} + \varepsilon_t, \qquad (9)
\end{aligned}
$$

where $E(\delta_t) = 1$, and δ_t is a two regimes *1*st-order stationary Markov Chain.

Proposition 1 determines that the covariance stationarity of a *TARSUR* process depends on the transition probabilities p_{21} and p_{11}, and on the parameter values ρ_1 and ρ_2. For instance, for values of the parameters $\rho_1 = 1.1$, $\rho_2 = 0.9$, $p_{21} = 0.8$, and $p_{11} = 0.2$ the *TARSUR* process is covariance stationary. More general, it is straightforward to show that a necessary condition for $\rho(F_2) < 1$ is $p_{21} > p_{11}$ (or equivalently $p_{12} > p_{22}$). In other words, the transition probability of being in the same regime has to be strictly smaller than the probability of changing regimes. The idea behind this condition is to avoid staying in the explosive regime for too long.

It is worthwhile to mention that a *TARSUR* process with an *i.i.d.* threshold variable is not covariance stationary, since $E(\delta_t^2) > 1$.

With respect to the IRF, Proposition 2 establishes that depending on the transition probabilities, shocks can have transitory or permanent effects. It is easy to check that for a *TARSUR* process, the following implications hold:

1. If $p_{11} > p_{21} : \lim\limits_{h \to \infty} \xi_h = \infty$, as it happens in an explosive model.
2. If $p_{11} = p_{21} : \xi_h = 1, \quad \forall h$, as it happens in a random walk model. Note that in this case Z_t is an *i.i.d.* process.
3. If $p_{11} < p_{21} : \lim\limits_{h \to \infty} \xi_h = 0$, as it happens in a stationary model.

Proposition 1 together with Proposition 2 show that *TARSUR* processes are more flexible than fixed unit root models, in the sense of being able to produce a richer set of plausible scenarios. If $p_{11} \geq p_{21}$ the process is not covariance stationary and shocks have permanent and even increasing effects in mean; but if $p_{11} < p_{21}$, shocks will have only transitory effects in mean and depending on the parameter values, it can be stationary or not. This latter case of non covariance stationarity but transitory effects resembles, in this sense, the *ARFIMA* models with a long memory parameter between 0.5 and 1 (see Dolado, Gonzalo and Mayoral (2002)).

Figure 1, a-c, displays simulated realizations from *TARSUR* and Random Walk (RW) models. The *TARSUR* series are generated by model (9), for $t = 1, ..., 550$, with ε_t as *i.i.d. Normal* (0,1) and Z_t as a standard stationary AR(1) process. The random walk series is generated from the same set of innovations. The first 50 observations of each series have been disregarded to avoid any initial conditions dependency. For comparison reasons, each figure shows a random walk versus three different types of *TARSUR* processes, that depends on the relationship between the conditional probabilities: $p_{11} > p_{21}$, $p_{11} = p_{21}$ and $p_{11} < p_{21}$. Each figure differs by the value taken by the variance of the stochastic unit root coefficient. More specifically, in figure 1a $\rho_1 = 1.01$ and $\rho_2 = 0.99$ ($V(\rho_t) = 0.0001$), in figure 1b $\rho_1 = 1.03$ and $\rho_2 = 0.97$ ($V(\rho_t) = 0.0009$), and in figure 1c $\rho_1 = 1.1$ and $\rho_2 = 0.9$ ($V(\rho_t) = 0.001$). It can be seen that for small values of $V(\rho_t)$ the RW and *TARSUR* series are indistinguishable. As $V(\rho_t)$ increases the *TARSUR* series becomes more volatile than its corresponding RW. It is worth to mention that even in the most unstable case (see figure 1c) the "explosive" *TARSUR* series ($p_{11} > p_{21}$) does not look like a standard AR(1) with a fixed explosive root.

2.4 Differencing a *TARSUR* process

Differencing model (2) we obtain

$$\triangle Y_t = (\omega_t - 1)Y_{t-1} + \varepsilon_t. \tag{10}$$

Proposition 3 *Assume that Y_t follows model (2). If ω_t has a strictly positive variance, ΔY_t is strictly (covariance) stationary if and only if Y_t is strictly (covariance) stationary.*

In contrast to fixed unit root models, stochastic unit root models are not difference stationary, in the sense that if the process is not stationary in levels, its differences will not be stationary either. Alternatively, if the process is strictly stationary (i.e., conditions of the first part of Theorem 1 are satisfied), its difference will also be strictly stationary. In this case we can express model (10) as a *MA*(∞)

$$\Delta Y_t = \sum_{j=0}^{\infty} \Psi_{t,j}\varepsilon_{t-j} \tag{11}$$

where $\Psi_{t,0} = 1$ and $\Psi_{t,j} = (\omega_t - 1)\psi_{t-1,j-1}$, $j \geq 1$.

In order to obtain the covariance structure and the *IRF* of ΔY_t, the representative case of subsection 2.2 needs to be assumed again.

Proposition 4 *Under the conditions of Proposition 1, ΔY_t is covariance stationary if the spectral radius of F_2 is less than one. Moreover, the IRF of ΔY_t , $\Upsilon_j = E\left(\frac{\delta \Delta Y_{t+j}}{\delta \varepsilon_t}\right)$, is given by*

$$\Upsilon_j = \begin{pmatrix} 1 & 1 \end{pmatrix} G_1 F_1^{j-2} \begin{pmatrix} v_1 p_1 \\ v_2 p_2 \end{pmatrix}, \; j \geq 2, \tag{12}$$

with $G_1 = \begin{pmatrix} (v_1-1)p_{11} & (v_1-1)p_{21} \\ (v_2-1)p_{12} & (v_2-1)p_{22} \end{pmatrix}$.

From these general results it is straightforward to conclude that a *TARSUR* process is not difference stationary. Nevertheless, under the scenario of Corollary 1 (i.e., stochastic root around unity), the *TARSUR* model is strictly stationary and therefore by Proposition 3 its difference will be too.

With respect the IRF, as it is expected, the long-run effect of the shocks on ΔY_t depends on the transition probabilities:

1. If $p_{11} > p_{21} : \lim_{j\to\infty} \Upsilon_j = \infty$, and ΔY_t is not covariance stationary.
2. If $p_{11} = p_{21} : \Upsilon_j = 0, \forall j \geq 1$, and ΔY_t is not covariance stationary.
3. If $p_{11} < p_{21} : \lim_{j\to\infty} \Upsilon_j = 0$, and ΔY_t could be covariance stationary.

The expression of the covariance function is omitted since we consider that its contribution to the analysis does not compensate its complexity. In any case, the covariance function of ΔY_t does not always exist, it depends on the spectral radius of F_2. In spite of this, for many of the cases ΔY_t resembles a white noise process.

3. STOCHASTIC UNIT ROOT TESTS

The goal of this section is to construct a test for the null hypothesis of an exact unit root versus the alternative of an stochastic unit root. It is worthwhile to emphasize that under both hypotheses $E(\delta_t) = 1$.

The data generating process (DGP) considered is the following model:

$$Y_t = [\mu_1 + \rho_1 Y_{t-1}]\, I(Z_{t-d} \leq r) + [\mu_2 + \rho_2 Y_{t-1}]\, I(Z_{t-d} > r) + \varepsilon_t. \tag{13}$$

Rearranging terms this DGP can be rewritten as

$$\begin{aligned} \Delta Y_t \quad = \quad & (\mu_1 I(Z_{t-d} \leq r) + \mu_2 I(Z_{t-d} > r)) \\ & + ((\rho_1 - \rho_2) I(Z_{t-d} \leq r) + (\rho_2 - 1))\, Y_{t-1} + \varepsilon_t, \end{aligned} \tag{14}$$

and imposing $E(\delta_t) = 1$, it is obtained

$$\begin{aligned} \Delta Y_t \quad = \quad & (\mu_1 I(Z_{t-d} \leq r) + \mu_2 I(Z_{t-d} > r)) \\ & + \gamma U_t(r) Y_{t-1} + \varepsilon_t, \end{aligned} \tag{15}$$

where $\gamma = (\rho_1 - \rho_2)$ and $U_t(r) = I(Z_{t-d} \leq r) - p(r)$, with $p(r) = P(Z_{t-d} \leq r)$.

Assuming $0 < p(r) < 1$ and given that $V(\delta_t) = \gamma^2 p(r)(1 - p(r))$, the null hypothesis of an exact unit root ($V(\delta_t) = 0$) versus the alternative of a stochastic unit root ($V(\delta_t) \neq 0$) is tested by testing

$$H_0 : \gamma = 0 \tag{16}$$

against

$$H_1 : \gamma \neq 0 \tag{17}$$

in model (15).

As it occurs with the Dickey-Fuller (DF) t-test, in order to obtain asymptotic distributions that are invariant to the deterministic terms contained in the DGP, the regression model (RM) used to implement our tests will contain a threshold constant term as well as a threshold deterministic trend:

$$\begin{aligned} \Delta Y_t \;=\; & (\mu_1 I(Z_{t-d} \leq r) + \mu_2 I(Z_{t-d} > r)) \\ & + (\beta_1 t I(Z_{t-d} \leq r) + \beta_2 t I(Z_{t-d} > r)) \\ & + \gamma U_t(r) Y_{t-1} + \varepsilon_t. \end{aligned} \tag{18}$$

The asymptotic distribution of our tests will basically depend on whether the threshold value is known or unknown, and in the latter case on whether is identified or unidentified. In the rest of the paper " $\Rightarrow$ " denotes weak convergence as $T \to \infty$ with respect to the uniform metric on $[0,1]^2$.

3.1 Threshold value known

The case of a known threshold value becomes relevant for pedagogical or explanatory reasons as well as for cases where the regimes are determined by the sign of the threshold variable (see Enders and Granger (1998) momentum TAR model). In this situation the test proposed is the $t-$statistic for $\widehat{\gamma}$, $t_{\gamma=0}$, in regression model (18), and its asymptotic distribution is shown in the next proposition.

Proposition 5 *Suppose that the threshold value is known and that assumptions (A.1), (A.2) (A.3) and (A.4) hold. Under the null of no threshold the $t_{\gamma=0}$ statistic in the regression model (18) has the following asymptotic distribution*

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

3.2 Threshold value unknown

When the threshold value r is unknown it is assumed that this parameter lies in a bounded interval R^*. The LS estimate of r is the value that

$$\min_{r \in R*} \widehat{\sigma}^2(r),$$

where $\widehat{\sigma}^2(r) = T^{-1} \sum_{t=1}^{T} \widehat{\varepsilon}_t^2$ denotes the residual variance from the LS estimation of model (18) for a fixed r. This estimator $\widehat{r}$ coincides with the one obtained by maximizing the Wald statistic

$$W_T = \sup_{r \in R^*} W_T(r)$$

of the null hypothesis of no threshold ($\mu_1 = \mu_2$ and $\gamma = 0$).

For statistical reasons (different AD) when r is unknown we need to distinguish whether this threshold parameter is identified or not under the null hypothesis. In the first case we assume that there exists a threshold effect under the null hypothesis of exact unit root, that is, $\gamma = 0$ but $\mu_1 \neq \mu_2$ in DGP (15). In the second case we assume that $\mu_1 = \mu_2$, so the test statistic of the null hypothesis of exact unit root is the same as the test statistic of no threshold at all.

3.2.1 Threshold value unknown but identified

When the DGP has a threshold effect in the drift term we can identify this threshold value before testing for a stochastic unit root. In this case it is enough to use the $t-$statistic for $\gamma = 0$ evaluated at $\widehat{r}$, $t_{\gamma=0}(\widehat{r})$. This is so because r is estimated super-consistently ($T-$consistent) by LS in a first step, and it can be taken as known, getting back into Proposition 2.

Proposition 6 *Suppose that assumptions (A.1), (A.2), (A.3) and (A.4) hold. Under $H_0 : \gamma = 0$, $\mu_1 \neq \mu_2$, the $t_{\gamma=0}$ statistic in regression model (18) has the following asymptotic distribution*

$$t_{\gamma=0}(\widehat{r}) \Rightarrow N(0,1).$$

3.2.2 Threshold value unknown and unindentified

In this subsection we consider models with no threshold effect in the constant term under the null ($\mu_1 = \mu_2 = \mu$). The appropriate test statistic is the supremum of the Wald statistic W_T introduced in section 3.2,

$$W_T = \sup_{r \in R^*} W_T(r), \text{ where } W_T(r) = t^2_{\gamma=0}(r).$$

The asymptotic distribution of W_T turns out to be different, as it happens with the DF tests, depending on the deterministic components introduced in the regression model and whether the DGP is characterized by a nonzero drift or not.

Proposition 7 *Suppose that assumptions (A.1), (A.2) (A.3) and (A.4) hold.*

1. *Consider DGP (15) with* $\mu_1 = \mu_2 = 0$, *and regression model (18) with no deterministic terms. Then under* $H_0 : \gamma = 0$

$$W_T \Rightarrow \sup_{r \in R^*} \frac{\left(\int B(s)dV(s, p(r))\right)^2}{p(r)(1-p(r)) \int B(s)^2 ds},$$

where $B(\cdot)$ *is a standard Brownian motion and* $V(s, p(r))$ *is a Kiefer-Müller process*[3] *on* $[0,1]^2$ *.*

2. *Consider DGP (15) with* $\mu_1 = \mu_2 = 0$, *and regression model (18) with a threshold constant term. Then under* $H_0 : \gamma = 0$

$$W_T \Rightarrow \sup_{r \in R^*} \frac{\left(\int B^*(s)dV(s, p(r))\right)^2}{p(r)(1-p(r)) \int B^*(s)^2 ds},$$

where $B^*(\cdot) = B(\cdot) - \int_0^1 B(s)ds$.

3. *Consider DGP (15) with* $\mu_1 = \mu_2 = \mu \neq 0$, *and regression model (18) with a threshold constant term. Then under* $H_0 : \gamma = 0$

$$W_T \Rightarrow \sup_{r \in R^*} \frac{\left(\int_0^1 f(s)dV(s, p(r))\right)^2}{p(r)(1-p(r))\frac{\mu^2}{12}},$$

where $f(s) = \mu s - \frac{\mu}{2}$.

4. *Consider DGP (15) with* $\mu_1 = \mu_2 = \mu$, *and regression model (18) with a threshold constant term and a threshold deterministic trend. Then under* $H_0 : \gamma = 0$

$$W_T \Rightarrow \sup_{r \in R^*} \frac{\left(\int B^{**}(s)dV(s, p(r))\right)^2}{p(r)(1-p(r)) \int B^{**}(s)^2 ds},$$

where $B^*(s) = B(s) - \int_0^1 B(a)g(a)'da \left(\int_0^1 g(a)g(a)'da\right)^{-1} g(s)$ *and* $g(s) = (1\ s)'$.

[3]A *Kieffer-Müller* process V on $[0,1]^2$ is given by $V(t_1, t_2) = B(t_1, t_2) - t_2 B(t_1, 1)$ where $B(t_1, t_2)$ is a *standard Brownian sheet.* The *standard Brownian sheet* $B(t_1, t_2)$ is a zero-mean Gaussian process indexed by $T = [0,1]^2$ and covariance function $Cov[B(s,t), B(u,v)] = (s \bigwedge t)(u \bigwedge v)$.

In general it will not be known whether $\mu_1 = \mu_2 = 0$ or not under H_0. For this reason, as in the standard DF test, it is recommended for practical purposes to use the regression model and critical values corresponding to case 4 of last proposition. Critical values (5% significant level) for cases 1, 2 and 4 are tabulated in Table 1. The data are generated under DGP (13), with $\rho_1 = \rho_2 = 1$, and ε_t as *i.i.d.* $N(0,1)$. The regression model considered is model (18) with Z_t an *i.i.d.* $U(0,1)$.

4. A MONTE CARLO EXPERIMENT AND A TESTING STRATEGY

Using Monte Carlo methods we now examine the performance of the proposed stochastic unit root tests. The power of the Dickey Fuller t test against $TARSUR$ alternatives is also analyzed. The Monte Carlo experiment consists on 10,000 replications with sample sizes $T = 100$, 250 and 500. The error term ε_t is generated as *i.i.d.* $N(0,1)$ and the threshold variable Z_t follows, without loss of generality, an $U(0,1)$ independent of Y_t .

Tables 2 and 3 show the empirical size of the proposed test for different values of the threshold effect in the drift term, $\Delta\mu = \mu_1 - \mu_2$, under the null hypothesis. From Table 2 it can be seen that, when the threshold parameter r is known, the empirical and nominal sizes coincide. Table 3 shows the empirical size for r unknown. As it is expected the asymptotic Normal approximation to the finite sample distribution improves with the sample size as well as with the size of $\Delta\mu$. When $\Delta\mu = 0$, the threshold parameter is not identified and the Normal distribution is not the correct asymptotic distribution. In brackets we report the empirical size based on the critical values of the supremum Wald statistic.

In order to study the power we analyze several $TARSUR$ alternatives that allow for different values of $|\gamma| = |\rho_1 - \rho_2| = (0.02,\ 0.06,\ 0.2)$, and $|\Delta\mu| = (0,\ 0.3,\ 0.6,\ 1,\ 2)$. Results are presented in Table 4 for r known and in Table 5 for r unknown. In both tables it is observed that the power increases with the sample size as well as with the size of the threshold effect $|\gamma| = |\rho_1 - \rho_2|$.

Table 6 shows the power of the DF t-test against the same $TARSUR$ alternatives previously considered. The t-statistic is calculated from the regression

$$\Delta Y_t = \pi_1 + \pi_2 t + \pi_3 Y_{t-1} + v_t. \tag{19}$$

The conclusion is that the DF unit root tests can not easily distinguish between a pure unit root and a threshold stochastic unit root.

Finally, we have also studied the power against alternative models with different values of

$E(\delta_t) = 0.3$, 0.5, 0.7, 0.9. For each value of $E(\delta_t)$, the $V(\delta_t)$ is allowed to vary from 0 to 0.3, and the threshold effect in the drift term ($\Delta\mu = \mu_1 - \mu_2$) goes between -2 and 2. A summary of these results are available upon request. As it is expected the power increases with the $V(\delta_t)$, the size of $\Delta\mu$ and with the value of $E(\delta_t)$.

In the light of these results we propose the following two steps strategy for empirical work:

(1) Test for a fixed unit root using a standard test like the DF t-test.

(2) If the null hypothesis of fixed unit root is not rejected then test for a *TARSUR* model.

5. EMPIRICAL APPLICATIONS

In order to provide an empirical illustration of how the estimation and testing of a *TARSUR* model can be applied in practice, we present two applications where there exists some theoretical and/or empirical controversy about the randomness of the unit root in the AR representation. The first example is the modelling of the U.S. stock prices and the second one analyzes interest rates from Japan, UK, U.S. and West Germany.

5.1 U.S. stock prices

In this application we investigate via our *TARSUR* model the link between asset prices and real activity, as well as the predictability in stock returns. The data analyzed is the quarterly series of Standard and Poor Composite Stock Price Index from 1947:1 to 1999:4. The threshold variable representing the real activity is the increment of GDP. Both variables are deflated by the implicit GDP price deflator (1996=100). More information about the data on stock prices can be found in Shiller (www.econ.yale.edu/~shiller) and about the GDP (S.A.) series in the U.S. Department of Commerce, Bureau of Economic Analysis (www.bea.doc.gov).

Since the work by Samuelson (1965) asset prices have been modeled as a martingale process considering returns to be unpredictable. Following LeRoy (1973) and Lucas (1978) the martingale property is obtained from the Euler equation that describes the optimal behaviour of the representative consumer:

$$p_t U'_t = (1+\rho)^{-1} E_t(p_{t+1} + d_{t+1}) U'_{t+1}, \tag{20}$$

where p_t is stock price at time t, d_t the dividends, ρ a discount factor, and U' the marginal utility. Assuming risk neutrality, $\rho = 0$, and removing the dividends from equation (20), the martingale model holds. Relaxation of these strong restrictions, for instance, assuming risk aversion, will lead to a departure from the martingale model. Note that the random walk specification is still more restrictive and it can not be derived within the framework of competitive price theory.

In order to generalize the martingale model we propose a stochastic unit root specification. The stochastic unit root model could be a martingale or not depending on the type of process followed by the stochastic coefficient δ_t. It will be a martingale if and only if $E(\delta_t|\Im_{t-1}) = 1$.

The estimated model for the stock prices is the $TARSUR$ model

$$\begin{aligned} \Delta Y_t &= (\mu_1 I(Z_{t-d} \leq r) + \mu_2 I(Z_{t-d} > r)) \\ &\quad + (\beta_1 t I(Z_{t-d} \leq r) + \beta_2 t I(Z_{t-d} > r)) \\ &\quad + \gamma \left(I(Z_{t-d} \leq r) - p(r) \right) Y_{t-1} + \varepsilon_t, \end{aligned}$$

where Y_t is the real stock price index and Z_t corresponds to the changes in the real GDP ($\Delta rgdp_t$). Dickey-Fuller unit root tests suggest that real stock prices as well as the real GDP contain a unit root, while Z_t clearly rejects the null hypothesis of a unit root.

Figure 2 presents a plot of both variables and Table 7 summarizes the estimation results of the $TARSUR$ model. The null hypothesis of exact unit root versus the alternative of a threshold stochastic unit root is clearly rejected at 5% significant level, $W_T = 13.61$ versus the critical value of 7.41 (see Table 1). For comparative purposes we have also estimated the following linear model, where the returns ΔY_t, are explained in terms of their own lagged values and lagged values of changes in GDP:

$$\Delta Y_t = \mu + \beta t + \rho \Delta Y_{t-1} + \alpha_1 \Delta rgdp_{t-1} + \alpha_2 \Delta rgdp_{t-2} + \alpha_3 \Delta rgdp_{t-3} + \nu_t.$$

From Table 8 it can be seen that our simple $TARSUR$ model is superior to the linear model. The $TARSUR$ model does not only capture a clearly positive relationship between the stock market and the real activity, but it does find a candidate (Z_t) to explain the cause of why stock prices may have a unit root. From the maintained hypothesis of unit root ($\rho_1 p(r) + \rho_2(1 - p(r)) = 1$) and the estimated parameters, $\widehat{\gamma}$ and $\widehat{p(r)}$, it is straightforward to obtain the estimates of ρ_1, ρ_2 and conditional probabilities p_{22} and p_{12} (see Table 9).

The results in Tables 7 and 9 show that when the increment of real GDP is less than 0.71 (corresponding approximately to a growth rate of 1.6%), the stock price index is in the stationary and mean reverting regime (autoregresive parameter equal to 0.98). The estimated probability of being in this regime is 0.8. On the other hand, when the increments of real GDP are larger than 0.71, prices follow a mildly explosive model (autoregresive parameter equal to 1.05), and this occurs with probability 0.2. On the overall, the stochastic root of the autoregresive representation is on average unity.

Moreover, looking at the transition probabilities, the stochastic coefficient δ_t seems to follow an *AR(1)* process with positive parameter (the correlogram of $\Delta rgdp_t$ also suggests this result). Therefore stock prices will not be a martingale process with respect the information set formed by past values of Y_t and $\Delta rgdp_t$. In other words, if $\Delta rgdp_t$ is considered a plausible explanation of the stochastic unit root, future returns could be predictable in the sense that

$$E_{t-1}\left(\frac{Y_t - Y_{t-1}}{Y_{t-1}}\right) = E_{t-1}\left(\delta_t - 1\right) \neq 0. \tag{21}$$

From (21) and the results in Tables 7 and 9 we conclude that if we were in a "recession" state at time $t-1$ ($\Delta rgdp_{t-1} < 0.71$), the expected value of the returns at time t would be negative. On the contrary, if we were in an "expansion" state ($\Delta rgdp_t > 0.71$) the expected value would be positive. In that way, we find that there exists a positive non-linear relationship between the expected stock returns and the real activity of the economy. Linear links between stock returns and macroeconomic variables have already been found in the finance literature although with a clear smaller R^2 (Chen et al. (1986), Fama (1990)) than the one in our TARSUR model.

5.2 Interest rates from different countries

The data analyzed is the same as in Leybourne, McCabe and Mills (1996). This data set corresponds to international U.S. (BUS), U.K. (BUK), Japan (BJP) and West Germany (BWG) bond yield data. The series are daily close of trade observations from April 1^{st} 1986 to December 29^{th} 1989 and can be obtained from Mills (1993). The four variables are plotted in Figure 3.

Leybourne, McCabe and Mills (1996) find that the null hypothesis of a fixed unit root versus the alternative of a stochastic unit root is clearly not rejected for U.S. bond yields. However, the fixed unit root null is mildly rejected for the U.K. and West Germany bond yield data, and strongly rejected for the Japanese series.

In order to apply our *TARSUR* model we need a candidate for a threshold variable. The fact that there is evidence of U.S. bond yields Granger causing the other yields, but not the other way around, together with the fact that U.S. bond yields do not reject the null hypothesis of exact unit root, makes the changes in the U.S. bond yields a perfect candidate for threshold variable.

The results obtained for the U.K., Japan and West Germany bond yield series are in Table

10. The values of the Wald test for testing the null of a unit root against a *TARSUR* alternative suggest that the null of a fixed unit root is not rejected for the U.K. and West Germany, while for Japan it is rejected in favor of the alternative hypothesis of a stochastic unit root. These results are similar to the ones obtained by Leybourne, McCabe and Mills (1996) with a different methodology. The advantages of the *TARSUR* model is that in the case of Japan we find a possible cause for the existence of a stochastic unit root, changes on the U.S. bond yields. For an alternative threshold model of interest rates, see Gonzalez and Gonzalo (1998).

6. CONCLUSION

This paper introduces a new class of stochastic unit root models (*TARSUR*) where the random behavior of the unit root is driven by an economic threshold variable. By doing that, we do not only make the unit root models more flexible but we find an explanation for the existence of unit roots. Flexibility is obtained because depending on the values of certain parameters, *TARSUR* processes can behave like an explosive process, like an exact unit root process, or like a stationary process. Explanatory power is gained because *TARSUR* models, by identifying an economic variable as a threshold variable, can provide a cause for the existence of unit roots.

Empirical applications show that estimation and testing of *TARSUR* models is not more complex than the estimation and testing involved in fixed unit root models. This is a clear advantage of *TARSUR* models with respect to other stochastic unit root methodologies available in the literature.

Extension of these models to the cointegration framework is undergoing research by the authors.

8. APPENDIX

Proof of Theorem 1. The condition of strict stationarity follows from Brandt (1986), and the weak stationarity from Karlsen (1990).

Proof of Proposition 1. The condition for covariance stationarity is given by,

$$\sum_{j=0}^{\infty} E\left(\left|\psi_{t,j}\right|^2\right)^{\frac{1}{2}} = \left[\begin{pmatrix} 1 & 1 \end{pmatrix} \sum_{j=1}^{\infty} F_2^j \begin{pmatrix} v_1^2 p_1 \\ v_2^2 p_2 \end{pmatrix}\right] < \infty, \tag{22}$$

with $F_2 = \begin{pmatrix} v_1^2 p_{11} & v_1^2 p_{21} \\ v_2^2 p_{12} & v_2^2 p_{22} \end{pmatrix}$. This infinite sum converges if the spectral radius of F_2 is less than one.

Proof of Proposition 2. The *IRF* can be expressed as

$$\xi_h = \begin{pmatrix} 1 & 1 \end{pmatrix} F_1^h \begin{pmatrix} v_1 p_1 \\ v_2 p_2 \end{pmatrix}, \quad h = 1, 2, \cdots, \tag{23}$$

where $F_1 = \begin{pmatrix} v_1 p_{11} & v_1 p_{21} \\ v_2 p_{12} & v_2 p_{22} \end{pmatrix}$. Therefore $\lim_{h\to\infty} \xi_h$ converges to zero if and only if the spectral radius of F_1 is less than one.

Proof of Corollary 1. From $V(\delta_1) > 0$ and by Jensen's inequality we get

$$E \log |\delta_1| < \log E |\delta_1| = \log E\delta_1 = 0.$$

Therefore condition (4) holds.

Proof of Proposition 3. Iterating backwards (10),

$$\triangle Y_t = \varepsilon_t + (\omega_t - 1) \sum_{j=1}^{n-1} \left(\prod_{i=1}^{j-1} \omega_{t-i}\right) \varepsilon_{t-j} + (\omega_t - 1) \left(\prod_{i=1}^{n-1} \omega_{t-i}\right) Y_{t-n}. \tag{24}$$

Substracting (10) from (24)

$$\Delta Y_t(Y_{t-n}) - \triangle Y_t = (\omega_t - 1)\left(Y_{t-1}(Y_{t-n}) - Y_{t-1}\right),$$

where $\Delta Y_t(Y_{t-n})$ corresponds to equation (24), and $\triangle Y_t$ to equation (10). As long as $V(\omega_t) > 0$, $\Delta Y_t(Y_{t-n})$ converges almost sure (in mean square) to $\triangle Y_t$ as $n \to \infty$, if and only if $Y_{t-1}(Y_{t-n})$ converges almost sure (in mean square) to Y_{t-1}.

Proof of Proposition 4. Covariance stationary follows from Propositions 1 and 3. Expression (12) is easily obtained after some algebra.

Proof of Proposition 5. The proof is divided in three parts depending on the deterministic terms included in the regression model (18): (1) no deterministic terms ($\mu_1 = \mu_2 = 0$ and $\beta_1 = \beta_2 = 0$), (2) a threshold constant term ($\mu_1 \neq \mu_2$ and $\beta_1 = \beta_2 = 0$), and (3) a threshold constant term as well as a threshold deterministic trend ($\mu_1 \neq \mu_2$ and $\beta_1 \neq \beta_2$).

In order to derive the asymptotic distribution of the proposed statistic we need to use some of the asymptotic tools developed in Caner and Hansen (2001).

Define the partial-sum process

$$B_T(s,u) = \frac{1}{\sqrt{T}} \sum_{t=1}^{[Ts]} I(Z_{t-d} \leq r)\varepsilon_t,$$

with $u = P(Z_{t-d} \leq r) = p(r)$. Theorem 1 in Caner and Hansen (2001) establishes that

$$B_T(s,u) \Rightarrow \sigma B(s,u), \tag{25}$$

on $(s,u) \in [0,1]^2$ as $T \to \infty$, where $B(s,u)$ is a *standard Brownian sheet* on $[0,1]^2$, and $\sigma^2 = E(\varepsilon_t^2)$.

Following Theorem 2 in Caner and Hansen (2001) we have that if $X_t = X_{t-1} + \varepsilon_t$

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} X_t I(Z_{t-d} \leq r)\varepsilon_t \Rightarrow \sigma \int_0^1 B(s)dB(s,u), \tag{26}$$

where $B(s)$ is a standard Brownian Motion. Finally, from Theorem 3 in Caner and Hansen (2001) we have that

$$\frac{1}{T} \sum_{t=1}^{T} X_t I(Z_{t-d} \leq r) \Rightarrow p(r) \int_0^1 B(s)ds. \tag{27}$$

Case 1. The DGP is

$$\Delta Y_t = \varepsilon_t. \tag{28}$$

The regression model considered is (18) with $\mu_1 = \mu_2 = 0$ and $\beta_1 = \beta_2 = 0$,

$$\Delta Y_t = \gamma U_t(r) Y_{t-1} + \varepsilon_t, \tag{29}$$

where $U_t(r) = I(Z_{t-d} \leq r) - p(r)$.

The t-statistic for $\gamma = 0$ is

$$t_{\gamma=0}(r) = \frac{\frac{1}{T}\sum_2^T Y_{t-1}U_t(r)\varepsilon_t}{\sqrt{\frac{1}{T^2}\sum_2^T U_t^2(r)Y_{t-1}^2}}.$$

From (25) and (26)

$$\frac{1}{T}\sum_{2}^{T} U_t(r) Y_{t-1}\varepsilon_t \Rightarrow \sigma^2 \int_0^1 B(s) dV(s, p(r)),$$

where $V(s, p(r))$ is a Kiefer-Müller process on $[0, 1]^2$. On the other hand, (25) and (27) imply that

$$\frac{1}{T^2}\sum_{2}^{T} I(Z_{t-d} \leq r) Y_{t-1}^2 \Rightarrow \sigma^2 p(r) \int_0^1 B(s)^2 ds.$$

Therefore

$$\frac{1}{T^2}\sum_{2}^{T} U_t(r)^2 Y_{t-1}^2 \Rightarrow \sigma^2 p(r)(1-p(r)) \int_0^1 B(s)^2 ds,$$

and the asymptotic distribution of the t-statistic is,

$$t_{\gamma=0}(r) \Rightarrow \frac{\int_0^1 B(s) dV(s, p(r))}{\sqrt{p(r)(1-p(r))(\int_0^1 B(s)^2 ds)}}.$$

Since $V(s, p(r))$ and $B(s) \equiv B(s, 1)$ are independent, it can be proved that for a fixed r,

$$\frac{\int_0^1 B(s) dV(s, p(r))}{\sqrt{\int_0^1 B(s)^2 ds}} \equiv N(0, \sigma_\nu^2),$$

where $\sigma_\nu^2 = Var\left(U_t(r)\varepsilon_t/\sigma\right) = p(r)(1-p(r))$.

Case 2. Regression model contains a threshold constant term

$$\Delta Y_t = \mu_1 I(Z_{t-d} \leq r) + \mu_2 I(Z_{t-d} > r) + \gamma U_t(r) Y_{t-1} + \varepsilon_t. \tag{30}$$

Note that $U_t(r) = (1-p(r))I(Z_{t-d} \leq r) - p(r)I(Z_{t-d} > r)$. We can estimate γ from the following transformed model:

$$\Delta Y_t = \gamma\left[(1-p(r))I(Z_{t-d} \leq r)Y_{t-1}^I - p(r)I(Z_{t-d} > r)Y_{t-1}^{II}\right] + \varepsilon_t, \tag{31}$$

where $Y_{t-1}^I = Y_{t-1} - \frac{\sum I(Z_{t-d} \leq r)Y_{t-1}}{\sum I(Z_{t-d} \leq r)}$ and $Y_{t-1}^{II} = Y_{t-1} - \frac{\sum I(Z_{t-d} > r)Y_{t-1}}{\sum I(Z_{t-d} > r)}$.

The t-statistic for $\gamma = 0$ is

$$t_{\gamma=0}(r) = \frac{T^{-1}\left((1-p(r))\sum Y_{t-1}^I I(Z_{t-d} \leq r)\varepsilon_t - p(r)\sum Y_{t-1}^{II} I(Z_{t-d} > r)\varepsilon_t\right)}{\sqrt{\widehat{\sigma}^2 T^{-2}\left((1-p(r))^2 \sum I(Z_{t-d} \leq r)\left(Y_{t-1}^I\right)^2 + p(r)^2 \sum I(Z_{t-d} \leq r)\left(Y_{t-1}^I\right)^2\right)}}.$$

Its asymptotic distribution is obtained under different nulls of interest:

(i) $H_0 :\ \gamma = 0$ and $\mu_1 = \mu_2 = 0$ (DGP (28)).

By applying (25) and (27)

$$T^{-1}Y^i_{Tt} \Rightarrow B(s) - \int_0^1 B(s)ds = B^*(s) \quad , \quad i = I, II.$$

Then, by using (26) we obtain

$$\begin{aligned} t_{\gamma=0}(r) &\Rightarrow \frac{(1-p(r))\int B^*(s)dB(s,p(r)) - p(r)\int B^*(s)dB(s,1-p(r))}{\sqrt{(1-p(r))p(r)\int B^*(s)^2 ds}} \\ &\equiv \frac{\int B^*(s)dV(s,p(r))}{\sqrt{p(r)(1-p(r))(\int B^*(s)^2 ds)}}. \end{aligned}$$

Since $B^*(s)$ and $V(s,p(r))$ are independent, it can be proved that for fixed r

$$\frac{\int B^*(s)dV(s,p(r))}{\sqrt{p(r)(1-p(r))(\int B^*(s)^2 ds)}} \equiv N(0,1).$$

(ii) $H_0: \ \gamma = 0$ and $\mu_1 = \mu_2 = \mu \neq 0$. The DGP is

$$Y_t = \mu + Y_{t-1} + \varepsilon_t. \tag{32}$$

By applying (25) and (27)

$$T^{-1}Y^i_{[Ts]} \Rightarrow \mu s - \frac{\mu}{2} = f(s), \quad i = I, II.$$

Using the same reasoning as before,

$$t_{\gamma=0}(r) \Rightarrow \frac{\int f(s)dV(s,p(r)}{\sqrt{p(r)(1-p(r)\frac{\mu^2}{12}}}.$$

For a fixed r

$$\frac{\int f(s)dV(s,p(r)}{\sqrt{p(r)(1-p(r)\frac{\mu^2}{12}}} \equiv N(0,1).$$

(iii) $H_0: \ \gamma = 0$ and $\mu_1 \neq \mu_2$. The DGP is

$$\begin{aligned} Y_t &= \mu_1 I(Z_{t-d} \leq r) + \mu_2 I(Z_{t-d} > r) + Y_{t-1} + \varepsilon_t = \\ &= \mu + Y_{t-1} + \xi_t, \end{aligned} \tag{33}$$

where $\mu = \mu_1 p(r) + \mu_2(1-p(r))$ and $\xi_t = \mu_1(I(Z_{t-d} \leq r) - p(r)) + \mu_2(I(Z_{t-d} > r) - (1 - p(r))) + \varepsilon_t$ is a zero mean strictly stationary process. Then

$$\begin{pmatrix} T^{-\frac{1}{2}} \sum_{t=1}^{[Ts]} \varepsilon_t I(Z_{t-d} \leq r) \\ T^{-\frac{1}{2}} \sum_{t=1}^{[Ts]} \xi_t \end{pmatrix} \Rightarrow \begin{pmatrix} B(s,p(r)) \\ \sigma_\xi B(s) \end{pmatrix}.$$

Note that as in the previous case

$$T^{-1}Y^i_{[Ts]} \Rightarrow \mu s - \frac{\mu}{2} = f(s), \quad i = I, II.$$

Then,

$$t_{\gamma=0}(r) \Rightarrow \frac{\int f(s)dV(s,p(r))}{\sqrt{p(r)(1-p(r)\frac{\mu^2}{12}}}.$$

For a fixed r,

$$\frac{\int f(s)dV(s,p(r))}{\sqrt{p(r)(1-p(r)\frac{\mu^2}{12}}} \equiv N(0,1).$$

Case 3. The regression model considered is

$$\Delta Y_t = (\mu_1 + \beta_1 t)I(Z_{t-d} \leq r) + (\mu_2 + \beta_2 t)I(Z_{t-d} > r) + \gamma U_t(r)Y_{t-1} + \varepsilon_t. \tag{34}$$

Following the same logic as in case 2, γ can be estimated from the following transformed model:

$$\Delta Y_t = \gamma \left[(1-p(r))I(Z_{t-d} \leq r)Y^l_{t-1} - p(r)I(Z_{t-d} > r)Y^h_{t-1}\right] + \varepsilon_t, \tag{35}$$

where

$$Y^l_{t-1} = Y_{t-1} - \sum_{j=1}^{T} Y_{j-1}I(Z_{j-d} \leq r)\begin{pmatrix} 1 & j \end{pmatrix}\left(\sum \begin{pmatrix} 1 & j \end{pmatrix}\begin{pmatrix} 1 \\ j \end{pmatrix} I(Z_{j-d} \leq r)\right)^{-1}\begin{pmatrix} 1 \\ t \end{pmatrix}$$

and

$$Y^h_{t-1} = Y_{t-1} - \sum Y_{j-1}I(Z_{j-d} > r)\begin{pmatrix} 1 & j \end{pmatrix}\left(\sum \begin{pmatrix} 1 & j \end{pmatrix}\begin{pmatrix} 1 \\ j \end{pmatrix} I(Z_{j-d} > r)\right)^{-1}\begin{pmatrix} 1 \\ t \end{pmatrix}.$$

The $t-$statistic for $\gamma = 0$ is

$$t_{\gamma=0}(r) = \frac{T^{-1}\left((1-p(r))\sum Y^l_{t-1}I(Z_{t-d} \leq r)\varepsilon_t - p(r)\sum Y^h_{t-1}I(Z_{t-d} > r)\varepsilon_t\right)}{\sqrt{\widehat{\sigma}^2 T^{-2}\left((1-p(r))^2 \sum I(Z_{t-d} \leq r)\left(Y^l_{t-1}\right)^2 + p(r)^2 \sum I(Z_{t-d} \leq r)\left(Y^h_{t-1}\right)^2\right)}}.$$

Its asymptotic distribution is obtained under different nulls of interest:

(i) $H_0: \ \gamma = 0$ and $\mu_1 = \mu_2 = 0$ (DGP (28)).

Applying (25) and (27) we get

$$\begin{aligned} T^{-\frac{1}{2}}Y^i_{[Ts]} \quad &\Rightarrow \quad B(s) - \int_0^1 B(a)\begin{pmatrix} 1 & a \end{pmatrix} da \left(\int_0^1 \begin{pmatrix} 1 & a \end{pmatrix}\begin{pmatrix} 1 \\ a \end{pmatrix} da\right)^{-1}\begin{pmatrix} 1 \\ s \end{pmatrix} \\ &= \quad B^{**}(s), \quad i = l, h. \end{aligned}$$

Then, by (26)

$$
\begin{aligned}
t_{\gamma=0}(r) &\Rightarrow \frac{(1-p(r))\int B^{**}(s)dB(s,p(r)) - p(r)\int B^{**}(s)dB(s,1-p(r))}{\sqrt{(1-p(r))p(r)\int B^{**}(s)^2ds}} \\
&\equiv \frac{\int B^{**}(s)dV(s,p(r))}{\sqrt{(1-p(r))p(r)(\int B^{**}(s)^2ds}}.
\end{aligned}
$$

Since $V(s,p(r))$ and $B(s)$ are independent, for a fixed r

$$
\frac{\int B^{**}(s)dV(s,p(r))}{\sqrt{(1-p(r))p(r)(\int B^{**}(s)^2ds}} \equiv N(0,1).
$$

(ii) $H_0:\ \gamma = 0$ and $\mu_1 = \mu_2 = \mu \neq 0$ (DGP (32)).

Since the regression model contains a trend component, the test statistic is invariant to μ, so we can set $\mu = 0$. Then we are back into case (i).

(iii) $H_0:\ \gamma = 0$ and $\mu_1 \neq \mu_2$ (DGP (33)).

Since the regression model contain a trend component, it can be shown that the test statistic is invariant to μ, so we can set $\mu = 0$ in expression (33). Then, applying results (25), (26) and (27) we obtain the following asymptotic distribution

$$
\begin{aligned}
t_{\gamma=0}(r) &\Rightarrow \frac{(1-p(r))\int B_b^{**}(s)dB_a(s,p(r)) - p(r)\int B_b^{**}(s)dB_a(s,1-p(r))}{\sqrt{(1-p(r))p(r)\int B_b^{**}(s)^2ds}} \\
&\equiv \frac{\int B_b^{**}(s)dV_a(s,p(r))}{\sqrt{(1-p(r))p(r)\int B_b^{**}(s)^2ds}}.
\end{aligned}
$$

Again, for fixed r

$$
\frac{\int B_b^{**}(s)dV_a(s,p(r))}{\sqrt{(1-p(r))p(r)\int B_b^{**}(s)^2ds}} \equiv N(0,1).
$$

Proof of Proposition 6

Given that $\widehat{r}$ is T-consistent[4], it can be shown that

$$
t_{\gamma=0}(\widehat{r}) \Rightarrow t_{\gamma=0}(r).
$$

Then, results in Proposition 5 can be directly applied.

Proof of Proposition 7

Since the threshold value is unknown and unidentified, the test statistic proposed is

$$
W_T = \sup_{r\in R^*} t_{\gamma=0}(r)^2.
$$

[4] Caner and Hansen (2000) proof that $T(\widehat{r} - r_0) = O_p(1)$ in presence of nonstationary variables.

All the cases considered in Proposition 7 are examined in Proposition 5. Applying the continuous mapping theorem we have that

$$W_T \Rightarrow \sup_{r \in R^*} t(r)^2,$$

where $t(r)$ is the asymptotic distribution of the $t - statistic$ obtained in Proposition 5.

References

Baillie, R.T. (1996), "Long memory processes and fractional integration in econometrics" *Journal of Econometrics,* 73, 5-59.

Beran, J. (1994), *Statistics for long memory processes.* New York: Chapman & Hall.

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

Caner, M., and B. Hansen (2001), "Threshold autoregression with a near unit root", *Econometrica* 69, 1555-1596.

Chen, N., R. Roll and S. Ross (1986), "Economics forces and the stock market", *Journal of Business*, 56, 383-403.

Cramer, H. (1961), "On some classes of nonstationary stochastic processes", *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability* (Vol.2), Berkeley, University of California Press, 57-78.

Dolado, J., J. Gonzalo and L. Mayoral (2002), "A Fractional Dickey-Fuller Test", *Econometrica,* 70, 1963-2006.

Enders, W., and C.W.J. Granger (1998), "Unit root test and asymmetric adjustment with an example using the term structure of interest rates", *Journal of Business & Economic Statistics*, 16, 304-311.

Fama, F. E. (1990), "Stock returns, expected returns, and real activity", *Journal of Finance,* 45, 1089-1108.

Gonzalo, J., and T. Lee (1998), "Pitfalls in testing for long run relationships", *Journal of Econometrics,* 86, 129-154.

González, M., and J. Gonzalo (1998), "Threshold unit root models", Working Paper, U. Carlos III de Madrid.

Granger, C., and J. Jouyeux (1980), "An introduction to long memory time series models and fractional differencing" *Journal of Time Series Analysis,* 1, 15-39.

Granger, C.W.J, and Newbold, P. (1986), *Forecasting Economic Time Series* (2nd ed.). Orlando, FL: Academic Press.

Granger, C., and N.R. Swanson (1997), "An introduction to stochastic unit-root processes", *Journal of Econometrics,* 80, 35-62.

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

LeRoy, S.F. (1973), "Risk aversion and the martingale property of stock prices", *Interna-*

tional Economic Review, 14, 436-446.

Leybourne, S.J., B. McCabe and T. Mills (1996), "Randomized unit root processes for modelling and forecasting financial time series: Theory and Applications", *Journal of Forecasting,* 15, 253-270.

Leybourne, S.J., B. McCabe and A.R. Tremayne (1996), "Can economic time series be differenced to stationarity?", *Journal of Business and Economic Statistic,* 14, 435-446.

Lucas, R.E. (1978), "Asset prices in an exchange economy", *Econometrica,* 49, 1426-1445.

McCabe,B., and A.R. Tremayne (1995), "Testing a time series for difference stationarity", *The Annals of Statistics*, 23, 1015-1028.

Mills, T.C. (1993), *The econometric modelling of financial time series.* Cambridge University Press.

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

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

Petruccelli, J., and S. Woolford (1984), "A threshold AR(1) model", *Journal of Applied Probability*, 21, 270-286.

Phillips, P.C.B., and Z. Xiao (1998),"A primer in unit root testing", *Journal of Economic Surveys,* 12, 423-470.

Pourahmadi, M. (1986), "On stationary of the solution of a doubly stochastic model", *Journal of Time Series Analysis,* 7, 123-132.

—(1988), "Stationary of the solution of $X_t = A_t X_{t-1} + \varepsilon_t$ and analysis of non-gaussian dependent random variables", *Journal of Time Series Analysis,* 9, 225-239.

Robinson, P.M. (1994), "Time series with strong dependence"*, Advances in Econometrics,* Sixth World Congress. Edited by C. Sims, Cambridge University Press.

Samuelson, P.A. (1965), "Time series with strong dependence", *Industrial Management Review*, 6, 41-49.

Tjϕstheim, D. (1986), "Some doubly stochastic time series", *Journal of Time Series Analysis,* 7, 51-72.

Vervaat, W. (1979), "On a stochastic difference equation and a representation of non-negative infinitely divisible random variables", *Advances in Applied Probability,* 11, 750-783.

Table 1: 5% Critical Values for the *TARSUR* test when r is unknown and unidentified

DGP	RM	$T = 100$	$T = 250$	$T = 500$
$\mu_1 = \mu_2 = 0$	$\mu_1 = \mu_2 = 0$ $\beta_1 = \beta_2 = 0$	7.36	7.30	7.09
$\mu_1 = \mu_2 = 0$	$\mu_1 \neq \mu_2$ $\beta_1 = \beta_2 = 0$	7.54	7.34	7.11
$\mu_1 = \mu_2 \neq 0$	$\mu_1 \neq \mu_2$ $\beta_1 \neq \beta_2$	7.79	7.41	7.17

Note: 10,000 replications.

DGP: model (13); RM: model (18)

Table 2: Empirical size (%) of the *TARSUR* test for r known

DGP	RM	$T = 100$	$T = 250$	$T = 500$
$\mu_1 = \mu_2 = 0$	$\mu_1 = \mu_2 = 0$ $\beta_1 = \beta_2 = 0$	5.14	5.02	4.87
$\mu_1 = \mu_2 = 0$	$\mu_1 \neq \mu_2$ $\beta_1 = \beta_2 = 0$	5.77	5.47	5.26
$\Delta\mu = -2$	$\mu_1 \neq \mu_2$ $\beta_1 \neq \beta_2$	5.15	5.46	5.29
$\Delta\mu = -1$	$\mu_1 \neq \mu_2$ $\beta_1 \neq \beta_2$	5.04	5.07	4.96
$\Delta\mu = -0.6$	$\mu_1 \neq \mu_2$ $\beta_1 \neq \beta_2$	5.79	5.24	5.34
$\Delta\mu = -0.3$	$\mu_1 \neq \mu_2$ $\beta_1 \neq \beta_2$	5.80	4.90	5.24
$\Delta\mu = 0$	$\mu_1 \neq \mu_2$ $\beta_1 \neq \beta_2$	6.01	5.53	5.24
$\Delta\mu = 0.3$	$\mu_1 \neq \mu_2$ $\beta_1 \neq \beta_2$	5.68	5.50	4.82
$\Delta\mu = 0.6$	$\mu_1 \neq \mu_2$ $\beta_1 \neq \beta_2$	5.28	4.93	5.21
$\Delta\mu = 1$	$\mu_1 \neq \mu_2$ $\beta_1 \neq \beta_2$	5.53	5.64	5.13
$\Delta\mu = 2$	$\mu_1 \neq \mu_2$ $\beta_1 \neq \beta_2$	5.43	5.41	4.92

Note: Rejection rates (%) from 10,000 replications. Nominal size 5%.

DGP: model (13); RM: model (18)

Table 3: Empirical size (%) of the *TARSUR* test when r is unknown and identified

DGP	$T = 100$	$T = 250$	$T = 500$
$\Delta\mu = -2$	6.01	5.55	4.84
$\Delta\mu = -1$	8.98	5.98	5.30
$\Delta\mu = -0.6$	14.34	7.71	5.80
$\Delta\mu = -0.3$	20.13 (4.54)	14.65 (3.20)	10.96 (2.89)
$\Delta\mu = 0$	23.31 (5.10)	20.97 (4.59)	19.64 (5.27)
$\Delta\mu = 0.3$	20.28 (4.52)	14.57 (3.49)	10.74 (2.88)
$\Delta\mu = 0.6$	14.49	8.02	5.47
$\Delta\mu = 1$	8.97	5.58	5.41
$\Delta\mu = 2$	5.93	5.13	4.80

Note: Rejection rates (%) from 10,000 replications. Nominal size 5%.

In brackets the empirical size based on the a.d. of $sup\ t^2_{\gamma=0}$ instead of the Normal distribution.

DGP: model (13); RM: model (18)

Table 4: Empirical power (%) of the *TARSUR* test for r known

DGP	RM	$T = 100$			$T = 250$			$T = 500$		
$\lvert\gamma\rvert =$		0.02	0.06	0.2	0.02	0.06	0.2	0.02	0.06	0.2
$\mu_1 = \mu_2 = 0$	$\mu_1 = \mu_2 = 0$ $\beta_1 = \beta_2 = 0$	10.95	43.28	92.97	35.44	86.71	99.97	70.13	98.90	100
$\mu_1 = \mu_2 = 0$	$\mu_1 \neq \mu_2$ $\beta_1 = \beta_2 = 0$	6.45	21.71	85.30	16.98	70.95	99.91	45.78	96.79	100
$\Delta\mu = -2$	$\mu_1 \neq \mu_2$ $\beta_1 \neq \beta_2$	7.07	19.43	99.82	13.29	98.56	100	70.21	100	100
$\Delta\mu = -1$	$\mu_1 \neq \mu_2$ $\beta_1 \neq \beta_2$	6.19	13.81	95.21	11.36	84.66	100	41.00	100	100
$\Delta\mu = -0.6$	$\mu_1 \neq \mu_2$ $\beta_1 \neq \beta_2$	5.82	11.81	86.59	9.69	67.78	100	31.89	99.90	100
$\Delta\mu = -0.3$	$\mu_1 \neq \mu_2$ $\beta_1 \neq \beta_2$	5.44	11.55	77.13	9.70	54.50	99.96	26.79	97.29	100
$\Delta\mu = 0$	$\mu_1 \neq \mu_2$ $\beta_1 \neq \beta_2$	5.37	11.00	71.37	9.27	47.88	99.74	25.11	91.58	100
$\Delta\mu = 0.3$	$\mu_1 \neq \mu_2$ $\beta_1 \neq \beta_2$	5.78	11.68	76.85	9.83	53.49	99.94	26.52	97.41	100
$\Delta\mu = 0.6$	$\mu_1 \neq \mu_2$ $\beta_1 \neq \beta_2$	5..56	12.18	87.11	10.32	66.98	100	30.83	99.79	100
$\Delta\mu = 1$	$\mu_1 \neq \mu_2$ $\beta_1 \neq \beta_2$	5.75	12.81	95.78	10.31	84.59	100	41.49	100	100
$\Delta\mu = 2$	$\mu_1 \neq \mu_2$ $\beta_1 \neq \beta_2$	6.96	19.50	99.84	13.20	98.58	100	70.60	100	100

Note: Rejection rates (%) from 10,000 replications. Nominal size 5%.

DGP: model (13); RM: model (18)

Table 5: Empirical power (%) of the *TARSUR* test for r unknown

DGP	RM	T=100			T=250			T=500		
$\|\gamma\| =$		0.02	0.06	0.2	0.02	0.06	0.2	0.02	0.06	0.2
$\mu_1 = \mu_2 = 0$	$\mu_1 = \mu_2 = 0$ $\beta_1 = \beta_2 = 0$	8.87	35.05	90.66	27.49	80.70	99.94	63.67	98.15	100
$\mu_1 = \mu_2 = 0$	$\mu_1 \neq \mu_2$ $\beta_1 = \beta_2 = 0$	6.01	14.04	75.85	11.36	57.60	99.80	33.24	93.64	100
$\Delta\mu = -2$	$\mu_1 \neq \mu_2$ $\beta_1 \neq \beta_2$	8.17	19.37	99.92	12.34	98.69	100	72.04	100	100
$\Delta\mu = -1$	$\mu_1 \neq \mu_2$ $\beta_1 \neq \beta_2$	15.38	19.05	97.01	12.46	85.54	100	42.39	100	100
$\Delta\mu = -0.6$	$\mu_1 \neq \mu_2$ $\beta_1 \neq \beta_2$	19.62	21.85	89.65	15.06	69.30	99.99	32.67	99.84	100
$\Delta\mu = -0.3$	$\mu_1 \neq \mu_2$ $\beta_1 \neq \beta_2$	22.32	25.87	81.76	20.69	56.12	100	29.19	97.58	100
$\Delta\mu = 0$	$\mu_1 \neq \mu_2$ $\beta_1 \neq \beta_2$	23.71	27.46	78.34	23.61	54.19	99.86	33.68	92.45	100
$\Delta\mu = 0.3$	$\mu_1 \neq \mu_2$ $\beta_1 \neq \beta_2$	22.44	25.19	81.85	20.58	57.14	99.97	28.40	98.01	100
$\Delta\mu = 0.6$	$\mu_1 \neq \mu_2$ $\beta_1 \neq \beta_2$	20.32	21.85	89.57	16.03	69.02	100	32.32	99.89	100
$\Delta\mu = 1$	$\mu_1 \neq \mu_2$ $\beta_1 \neq \beta_2$	16.03	18.56	96.82	12.87	84.88	100	42.59	100	100
$\Delta\mu = 2$	$\mu_1 \neq \mu_2$ $\beta_1 \neq \beta_2$	8.51	19.54	99.91	13.49	98.75	100	70.87	100	100

Note: Rejection rates (%) from 10,000 replications. Nominal size 5%.

DGP: model (13); RM: model (18)

Table 6: Empirical Power of the DF t-test for TARSUR alternatives

DGP	RM	$T = 100$			$T = 250$			$T = 500$		
$\|\gamma\| =$		0.02	0.06	0.2	0.02	0.06	0.2	0.02	0.06	0.2
$\mu_1 = \mu_2 = 0$	$\mu = 0$ $\beta = 0$	4.70	4.69	5.70	4.87	4.99	4.73	4.99	4.80	4.32
$\mu_1 = \mu_2 = 0$	$\mu \neq 0$ $\beta = 0$	2.29	1.17	0.29	1.93	0.82	0.33	1.19	0.50	0.33
$\Delta\mu = -2$	$\mu \neq 0$ $\beta \neq 0$	0.00	0.00	0.05	0.00	0.00	0.12	0.00	0.03	0.22
$\Delta\mu = -1$	$\mu \neq 0$ $\beta \neq 0$	0.00	0.00	0.03	0.00	0.01	0.26	0.00	0.04	0.25
$\Delta\mu = -0.6$	$\mu \neq 0$ $\beta \neq 0$	0.00	0.00	0.05	0.00	0.01	0.10	0.00	0.08	0.22
$\Delta\mu = -0.3$	$\mu \neq 0$ $\beta \neq 0$	0.00	0.00	0.04	0.00	0.00	0.15	0.00	0.04	0.23
$\Delta\mu = 0$	$\mu \neq 0$ $\beta \neq 0$	0.00	0.00	0.04	0.00	0.00	0.14	0.00	0.04	0.22
$\Delta\mu = 0.3$	$\mu \neq 0$ $\beta \neq 0$	0.00	0.00	0.06	0.00	0.00	0.18	0.00	0.06	0.22
$\Delta\mu = 0.6$	$\mu \neq 0$ $\beta \neq 0$	0.00	0.00	0.03	0.00	0.00	0.21	0.00	0.01	0.16
$\Delta\mu = 1$	$\mu \neq 0$ $\beta \neq 0$	0.00	0.00	0.01	0.00	0.00	0.26	0.00	0.04	0.26
$\Delta\mu = 2$	$\mu \neq 0$ $\beta \neq 0$	0.00	0.00	0.01	0.00	0.00	0.16	0.00	0.05	0.25

Note: Rejection rates (%) from 10,000 replications. Nominal size 5%.

DGP: model (13); RM: $\triangle Y_t = \pi_1 + \pi_2 t + \pi_3 Y_{t-1} + v_t$

Table 7: *TARSUR* model for U.S. Stock Prices

$\widehat{\mu}_1$	$\widehat{\mu}_2$	$\widehat{\beta}_1$	$\widehat{\beta}_2$	$\widehat{\gamma}$	$\widehat{r}$	$\widehat{d}$	$\widehat{p}(r)$	W_T	R^2	AIC
0.007 (0.028)	− 0.439 (0.14)	− 0.0006 (0.0002)	0.002 (0.0011)	− 0.063 (0.017)	0.71	1	0.8	13.61	*0.25*	−3.22

Table 8: Linear regression model for U.S. Stock Prices

$\widehat{\mu}$	$\widehat{\beta}$	$\widehat{\rho}$	$\widehat{\alpha}_1$	$\widehat{\alpha}_2$	$\widehat{\alpha}_3$	R^2	AIC
− 0.03 (0.02)	0.0006 (0.0002)	0.32 (0.06)	0.00036 (0.0004)	−0.00049 (0.0004)	0.0016 (0.0003)	*0.16*	−0.35

Table 9: *TARSUR* regime roots and conditional probabilities for U.S. Stock Prices

Z_{t-d}	$\widehat{\rho}_1 = \widehat{\gamma} + \widehat{\rho}_2$	$\widehat{\rho}_2 = 1 - \widehat{\gamma}\widehat{p}(r)$	p_{22}	p_{12}
$\Delta rgnp_t$	0.98	1.05	0.41	0.12

Table 10: *TARSUR* model for Interest Rates

$\Delta Y_t =$	**ΔBUK**	**ΔBJP**	**ΔBWG**
$Z_{t-d} =$	ΔBUS_t	ΔBUS_{t-1}	ΔBUS_{t-1}
$\widehat{\mu}_1$	0.014 (0.023)	0.037 (0.018)	0.005 (0.008)
$\widehat{\mu}_2$	−0.034 (0.052)	−0.015 (0.005)	−0.009 (0.011)
$\widehat{\beta}_1$	0.000005 (0.00002)	0.00003 (0.000008)	0.000009 (0.000006)
$\widehat{\beta}_2$	0.000004 (0.00001)	− 0.000005 (0.000005)	0.000001 (0.000009)
$\widehat{\gamma}$	− 0.0098 (0.008)	− 0.016 (0.005)	−0.005 (0.008)
$\widehat{r}$	0.023	−0.037	−0.005
$\widehat{d}$	0	1	1
$\widehat{p(r)}$	0.317	0.231	0.573
W_T	1.39	9.78	1.73
R^2	*0.033*	*0.031*	*0.067*

Note: standard errors in brackets.

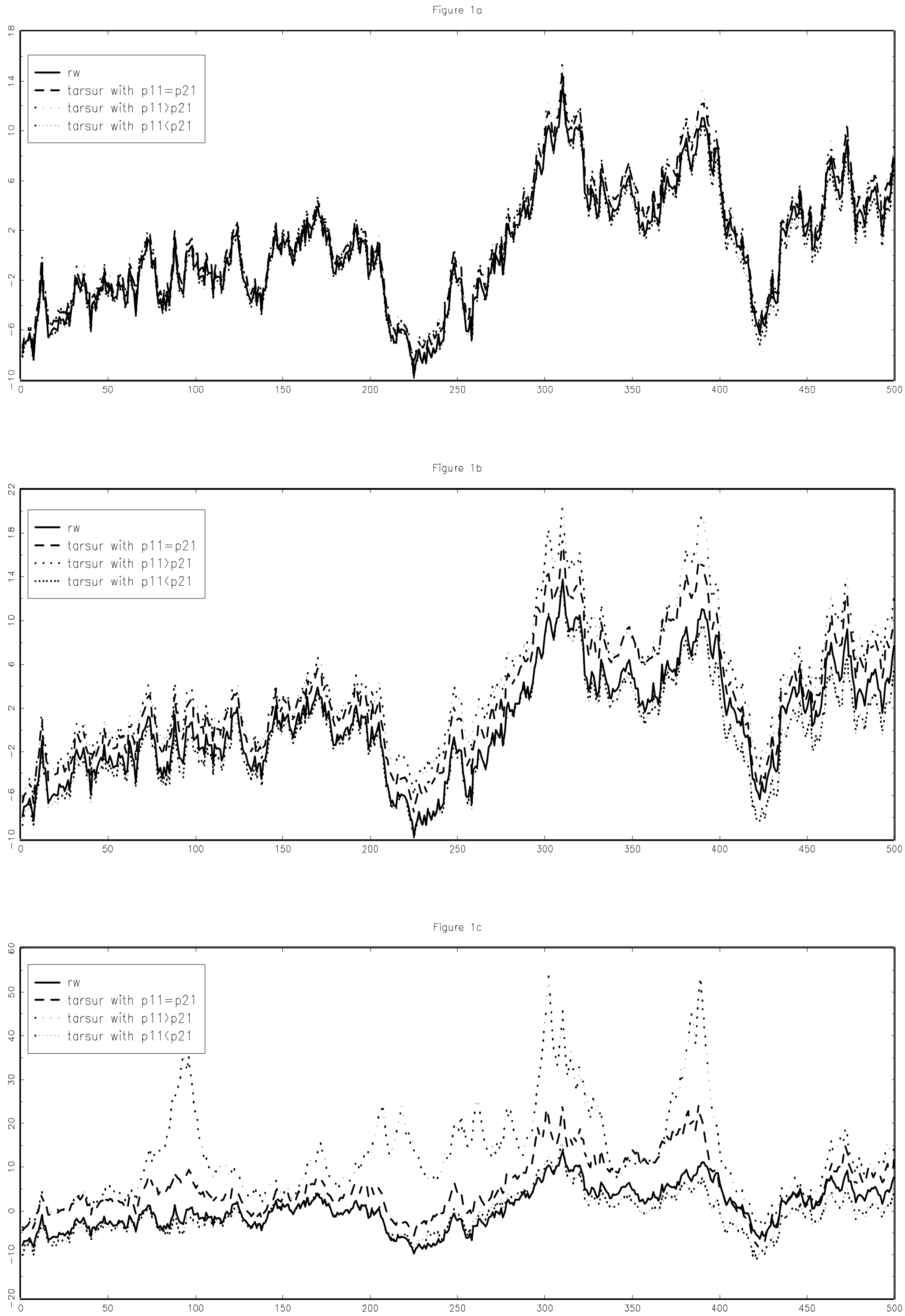

Figure 1: Random Walk versus different TARSUR series. Each figure differs by $V(\delta_t)$.

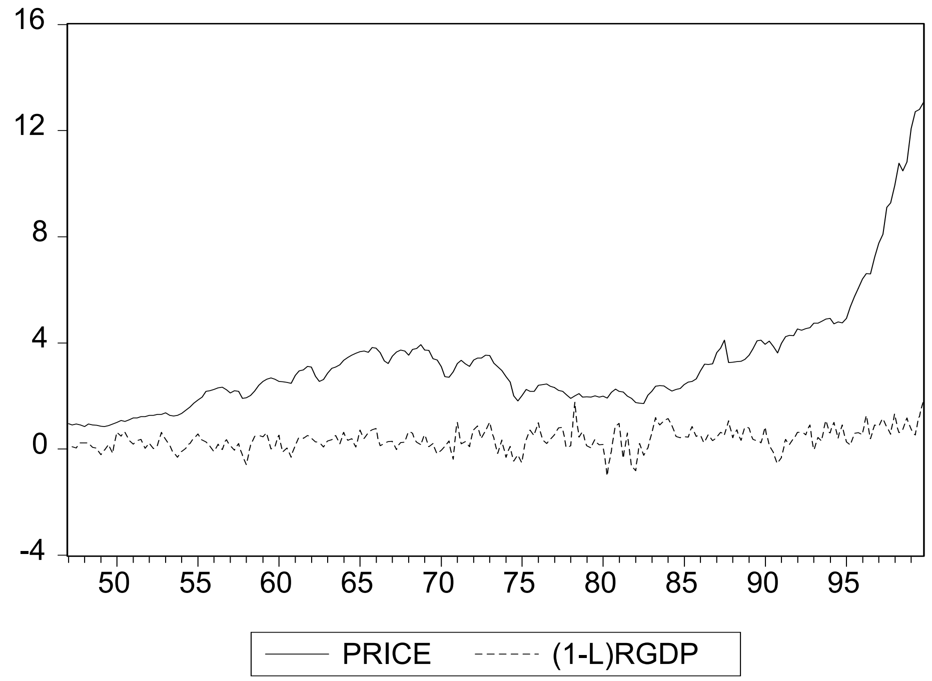

Figure 2: US stock prices and $(1-L)$ real GDP, 1947:1-1999:4.

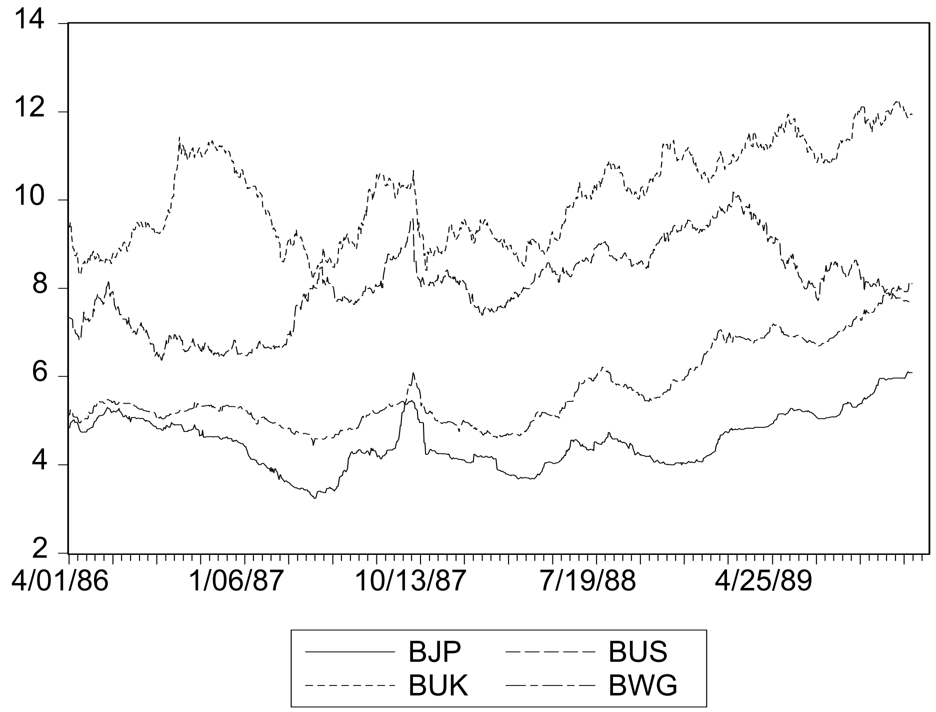

Figure 3: Interest rates from Japan, UK, US and West Germany.