Capítulo 7

On testing for a stochastic unit root in financial time series: The case of a bilinear unit root process

Julio Afonso

J.Afonso
Universidad de la Laguna, Departamento de Economía de las Instituciones, Estadística Económica y Econometría.Facultad de Ciencias Económicas y Empresariales. Campus de Guajara. Camino La Hornera s/n 38071. San Cristóbal de la Laguna. Tenerife. Islas Canarias.
jafonsor@ull.es

M.Ramos, F.Miranda (eds.) *Optimización-Estocástica-Recursiva-Coherente-Sistémica y sus variantes (probabilidad, econometría y estadística aplicada),* Temas Selectos de Optimización-©ECORFAN-Santiago de Compostela, España, 2012.

Abstract

This paper considers a particular member of the class of stochastic (or randomized) unit root (STUR) process given by a simple bilinear process with a unit root. Under a certain reparameterization of the bilinear parameter, we use the recently proposed stochastic limit for this process to show the consistency of some commonly used nonparametric tests of the null hypothesis of stationarity against the alternative of a unit root under this form of nonstationarity, other than difference stationarity. Also, as an alternative to the existing pseudo T-ratio test for the null of a fixed unit root against a bilinear unit root, we propose a new testing procedure based on a simple modification of the KPSS test statistic that has the advantage to allow for more general forms of the deterministic component and that seems to have good size and power in finite samples to discriminate between a fixed (or linear) and a bilinear unit root. We derive the asymptotic null and alternative distributions and also we present an application to the series of log-prices of some stock market indexes with distinct time frequencies: IBEX 35, SP500, and Dow Jones Composite Average (daily), and CAC40 (weekly).

7 Introduction

In recent years there has been an active and increasing research on the generalization and extension of the concept of nonstationarity around the central case of a random walk, I(1) or fixed unit root process. Since the contributions of McCabe and Tremayne (1995), Leybourne, McCabe and Tremayne (1996), Leybourne, McCabe and Mills (1996) and Granger and Swanson (1997) introducing the so-called stochastic (or randomized) unit root processes, there has been many different contributions that consider this family of processes as a plausible alternative to the standard case of I(1) processes and as a possible explication to the rejection of the I(1) evidence in many empirical studies. In this paper we consider one particular member of this family of global nonstationary processes, that can be partially or locally stationary, that has very interesting properties and can be called the bilinear unit root (BLUR) process. This process is the nonstationary version of the diagonal bilinear process of order one, BL(1,0,1,1). We study the behaviour and properties (size and power) of several unit root and stationarity tests in the proximity of a perfect, or fixed, unit root given by a BLUR process with weak bilinear effect as has been defined and analyzed by Lifshits (2006). Given the main conclusions of this analysis, we propose a new semi-nonparametric test statistic to distinguish between a fixed and a bilinear unit root, complementing the existing test proposed by Charemza et.al. (2005) but with a very different approach.

The structure of the paper is as follows. Section 7.2 introduces the general framework to our analysis, with initial attention to the distinction between the particular cases of stationarity and nonstationarity defined in the usual way as I(0) and I(1) processes. Section 7.3 introduces the case of a nonlinear nonstationary process, the bilinear unit root (BLUR) process, which falls within the class of STUR processes, but that with suitable normalization (weak BLUR) have a closed-form limiting representation that includes, as a particular case, the fixed unit root process. This section also includes the study of the consistency and asymptotic distribution under weak BLUR of some widely used semi-nonparametric residual-based tests for the null hypothesis of stationarity against the alternative of a fixed unit root, and we will show that these test statistics have nontrivial power against this form of nonstationarity. Section 7.4 introduces a new test procedure to consistently distinguish between a fixed and a stochastic unit root when the alternative is a (weak) BLUR process. Section 7.5 present a small application of this new test procedure. Finally, all the proofs are collected in the appendixes.

7.2. A unified framework for semi-nonparametric residual-based tests for stationarity and for a fixed unit root

At a starting point we consider a generalized version of an unobserved components model for the observed time series ($Y_t, t \in N$) given by:

$$Y_t = d_t + \eta_t \quad t = 1,2,\dots,n \tag{7.2.1}$$

$$\eta_t = \alpha_t \eta_{t-1} + \varepsilon_t \tag{7.2.2}$$

Where d_t contains the deterministic trend components, while equation (7.2.2) considers a very general structure for the stochastic trend component η_t that determines the stochastic nature and memory properties (persistence) of Y_t.[43] Under stationarity of the sequence ε_t, we have that the condition for stationarity[44] (or I(0)) of Y_t is given by $\alpha_t = \alpha$ for all t = 1, ..., n, with $|\alpha| < 1$, while that if $\alpha_t = 1$ for all t = 1, ..., n, then η_t is a random walk (or I(1)) process driven by a stationary innovation process, with Y_t being nonstationary as well. There are some possible intermediate situations between I(0) and I(1) when α_t varies (deterministically or randomly) between 1 and any value $|\alpha| < 1$ through the whole sample, or even eventually when takes a value $|\alpha| > 1$ which gives the explosive case. In the next section we consider this last situation described by a stochastic process α_t, with $E[\alpha_t] = 1$ and $Var[\alpha_t] \geq 0$ which is generally called a stochastic (or randomized) unit root (STUR) process. For this reason, in what follows, we introduce the notion of a fixed unit root, as opposed to the case of a stochastic unit root process, in the standard I(1) case described above or, equivalently, when $Var[\alpha_t] = 0$. This STUR specification will be the main topic in the next two sections of the paper, where will be considered as an alternative and a generalization to both stationarity and the fixed unit root cases. These STUR processes can arise naturally in economics and in finance. Gonzalo and Lee (1998) find that a stochastic autoregressive unit root characterizing the behaviour of the consumption can arise when assuming a quadratic utility in the solution to a problem of maximization of the utility function, while that Charemza et.al. (2005) find a theoretical motivation for the particular case of a bilinear unit root process (to be introduced and defined latter) from a simple generalization of a model of speculative behaviour characterizing the formation of the dividend-adjusted logarithms of prices of shares. Also, from the econometric point of view, the stochastic process (7.2.2) can generate a wide range of interesting processes for describing the behaviour of many economic and financial time series. In particular, under very general conditions on the sequence (α_t, ε_t) we have that $E_{t-1}(\eta_t) = E_{t-1}(\alpha_t)\eta_{t-1} + E_{t-1}(\varepsilon_t) = E_{t-1}(\alpha_t)\eta_{t-1}$, and:

$$Var_{t-1}(\eta_t) = E_{t-1}(\varepsilon_t^2) + E_{t-1}[(\alpha_t - E_{t-1}(\alpha_t))^2]\eta_{t-1}^2$$

Where this conditional variance can be seen as a very general form of an ARCH-type model, that is a very widely used model for capturing some of the well known stylized facts of many economic and financial series.

---

[43] There are some other possible representations for this kind of models in this context, but they do not represent any fundamental difference from the one considered here. One possibility could be to introduce a separation between the sources of stationarity and nonstationarity, such as $Y_t = d_t + \zeta_t$, where $\zeta_t = \eta_t + u_t$ with η_t as in (2.1) and (u_t, ε_t) a stationary sequence with $Var[\varepsilon_t] = \sigma_\varepsilon^2 \geq 0$, so that $\sigma_\varepsilon^2 = 0$ corresponds to the stationarity case irrespective of the particular structure of η_t.

[44] Strictly speaking, we have that with dt 0 the sequence Yt is stationary around the deterministic trend component, dt, which can be called trend or deterministic stationarity.

To complete the initial specification of our model, we next introduce two standard assumptions concerning the structure and behaviour of the deterministic and stochastic components d_t and ε_t.

Assumption 7.2 Deterministic component: We assume that the deterministic component is given by a *p*th-order polynomial trend function, that is $d_t = \boldsymbol{\tau}'_{t,p}\boldsymbol{\beta}_p$, with $\boldsymbol{\tau}_{t,p} = (1, t, \dots, t^p)'$, $\boldsymbol{\beta}_p = (\beta_0, \beta_1, \dots, \beta_p)'$, and $p \geq 0$. This general formulation may accommodate many other possible forms of the trend function, such as incorporating a systematic (abrupt or gradual) break in the polynomial time trend, that is $d_t = \boldsymbol{\tau}'_{t,p}\boldsymbol{\beta}_{pt}(\lambda)$ with $\boldsymbol{\beta}_{pt}(\lambda) = \boldsymbol{\beta}_{1p} + h_t(\lambda)\boldsymbol{\beta}_{2p}$, where $h_t(\lambda) = I(t > [n\lambda])$ is the usual indicating function and relative break-point $\lambda \in \Lambda = (0,1)$,[45] or even models with continuous change in the mean or multiple discontinuous changes. In any case, it is assumed that there exist a diagonal, non-stochastic and non-singular weighting matrix, $\boldsymbol{\Gamma}_n$, such that $\boldsymbol{\Gamma}_n \cdot \boldsymbol{\tau}_{t,p} = \boldsymbol{\tau}_p(\tfrac{t}{n}) \rightarrow \boldsymbol{\tau}_p(r)$ uniformly over $r \in [0,1]$ as $n \rightarrow \infty$, with $\boldsymbol{\tau}_p(r)$ a continuously differentiable function on [0,1], for all $(t-1)/n < r \leq t/n$, $t = 1, \dots, n$. This assumption implies that the limiting terms in $\boldsymbol{\tau}_p(r)$ are of bounded variation. In the leading case of a *p*th-order polynomial trend function, $p \geq 0$, we have $\boldsymbol{\Gamma}_n = \mathrm{diag}(1, n^{-1}, \dots, n^{-p})$, and $\boldsymbol{\tau}_p(r) = (1, r, \dots, r^p)' \in [0,1]^{p+1}$.

Assumption 7.2 Error term: Let us consider that the zero mean error sequence ε_t satisfies either of the two following conditions:

(a) $(\varepsilon_t, t \in \mathbb{Z})$ is a stationary process with finite variance $E[\varepsilon_0^2] = \sigma^2 < \infty$ and appropriate memory restrictions that ensure a necessary invariance principle, such as:

$$n^{-1/2}\sum_{t=1}^{[nr]} \varepsilon_t \Rightarrow B(r) = \sigma_\infty W(r), \quad \sigma_\infty^2 = \sum_{k=-\infty}^{\infty} E[\varepsilon_0 \varepsilon_k] \tag{7.2.3}$$

With $W(r)$ a standard Brownian motion process, or:

(b.1) $(\varepsilon_t, t \in \mathbb{Z})$ are iid random variables with $E[\varepsilon_0^2] = \sigma^2 < \infty$, and $E[|\varepsilon_0|^m] < \infty$ for some $m > 2$.

(b.2) $(\varepsilon_t, \mathcal{F}_t, t \in \mathbb{Z})$, $\mathcal{F}_t = \sigma(\varepsilon_t, \varepsilon_{t-1}, \dots)$, is a martingale difference sequence (MDS) with $E_{t-1}[\varepsilon_t^2] = \sigma^2 < \infty$ for all *t*, $\sup_{t \in Z} E_{t-1}[|\varepsilon_t|^m] < \infty$ a.s. form some $m>2$.

Remark 7.1 Under Assumption 2.b (with either b.1 or b.2), the result (7.2.3) follows trivially from the invariance principle of McLeish (1975), with the long-run variance (LRV) $\sigma_\infty^2 = \sigma_\varepsilon^2$, while that Assumption 2.a covers many commonly used situations where it is introduced a particular set of conditions controlling both temporal dependence and heterogeneity in the innovation process.

---

[45] Observe that this compact specification is equivalent to consider $d_t = \boldsymbol{\tau}'_{t,p}\boldsymbol{\beta}_{1p} + \boldsymbol{\tau}'_{t,p}(\lambda)\boldsymbol{\beta}_{2p}$, with $\boldsymbol{\tau}'_{t,p}(\lambda) = \boldsymbol{\tau}'_{t,p} \cdot h_t(\lambda)$. Another equivalent parameterization, that introduce the separation between both regimes, is given by considering $d_t = \bar{\boldsymbol{\tau}}'_{t,p}(\lambda)\boldsymbol{\beta}_{1p} + \boldsymbol{\tau}'_{t,p}(\lambda)\boldsymbol{\alpha}_p$, with $\bar{\boldsymbol{\tau}}_{t,p}(\lambda) = \boldsymbol{\tau}_{t,p} \cdot \bar{h}_t(\lambda)$, $\bar{h}_t(\lambda) = 1 - h_t(\lambda) = I(t \leq [n\lambda])$, and $\boldsymbol{\alpha}_p = \boldsymbol{\beta}_{1p} + \boldsymbol{\beta}_{2p}$.

Two of the most common set of conditions are the linear process (LP) (cf. Phillips and Solo (1992)) driven by iid or MDS innovations, and the strong mixing (see, e.g., Phillips (1987)) with mixing coefficients of size *m*/(*m*–2), for some $m > 2$. These two conditions allow for a wide variety of possible generating mechanisms for the sequence ε_t, including all Gaussian and many other stationary finite order ARMA models under very general conditions on the underlying errors.

Also, Assumption 2b.2 determines that, with $m > 4$ and for all $k \geq 1$, then it is verified the following joint invariance principle for sample covariances:

$$\begin{pmatrix} n^{-1/2}\sum_{t=1}^{[nr]}(\varepsilon_t^2-\sigma_\varepsilon^2) \\ n^{-1/2}\sum_{t=1}^{[nr]}\varepsilon_t\varepsilon_{t+k} \end{pmatrix} \Rightarrow \begin{pmatrix} \kappa_\varepsilon W_0(r) \\ \sigma_\varepsilon^2 W_k(r) \end{pmatrix}$$

With $\kappa_\varepsilon^2 = E[(\varepsilon_t^2-\sigma_\varepsilon^2)^2]$, and $W_0(r)$, $W_k(r)$ two standard independent Brownian motion processes. Under a more general dependence setting, as in Assumption 2.a based on a LP with iid innovations, it is verified a similar result, as can be seen in, e.g., Phillips and Solo (1992) and Ibragimov and Phillips (2008).

All the test procedures that we analyse in this paper are based on OLS detrended observations, that is:

$$\hat{\eta}_{t,p} = Y_t - \boldsymbol{\tau}'_{t,p}\hat{\boldsymbol{\beta}}_{p,n} = \eta_t - \boldsymbol{\tau}'_{t,p}(\hat{\boldsymbol{\beta}}_{p,n}-\boldsymbol{\beta}_p) = \eta_t - n^{-v}\boldsymbol{\tau}'_p(\tfrac{t}{n})[n^{v}\boldsymbol{\Gamma}_n^{-1}(\hat{\boldsymbol{\beta}}_{p,n}-\boldsymbol{\beta}_p)] \quad (7.2.4)$$

With:

$$n^{v}\boldsymbol{\Gamma}_n^{-1}(\hat{\boldsymbol{\beta}}_{p,n}-\boldsymbol{\beta}_p) = \mathbf{Q}_{n,p}^{-1}n^{-(1-v)}\sum_{j=1}^{n}\boldsymbol{\tau}_p(\tfrac{j}{n})\eta_j \quad (7.2.5)$$

The suitable normalized OLS bias of the estimator of the trend parameter vector $\boldsymbol{\beta}_p$ in (7.2.1), with:

$$\mathbf{Q}_{n,p} = n^{-1}\sum_{j=1}^{n}\boldsymbol{\tau}_p(\tfrac{j}{n})\boldsymbol{\tau}'_p(\tfrac{j}{n}) \to \mathbf{Q}_p = \int_0^1 \boldsymbol{\tau}_p(s)\boldsymbol{\tau}'_p(s)ds \quad (7.2.6)$$

Where convergence follows from Assumption 1 as $n\to\infty$ with $\mathbf{Q}_p > 0$. The scaling factor *v* in the last expression of (2.4) can take the values ±1/2 depending on whether we consider the stationarity or nonstationarity (fixed unit root) case.

Remark 7.2 The case of no deterministic component, that is when $d_t = 0$ in (7.2.1), is also covered by this results simply by considering $\hat{\eta}_{t,p} = Y_t = \eta_t$. When $p = 0$, we have the usual demeaned observations, while that $p = 1$ represents the case of demeaned and linearly detrended observations. Because the extension to the case of higher orders of the polynomial trend function does not pose any additional restriction, we will maintain this general formulation.

Next we consider three closely-related semi-nonparametric test statistics proposed to testing the null hypothesis of stationarity, $\alpha_t = \alpha \; \forall t = 1,\ldots,n$, with $|\alpha|<1$, against the alternative of a fixed unit root, that is, $\alpha_t = 1 \; \forall t = 1,\ldots,n$.

Each of them explode the fact that under the I(1) alternative we must expect a excessive fluctuation in the residual sequence (7.2.4), given that $\hat{\eta}_{t,p} = O_p(\sqrt{n})$ in this case, to consistently distinguish between these two types of behaviour.

These test statistics are the following:

$$\hat{M}_{n,p}(q_n) = \frac{1}{n \cdot \hat{\omega}^2_{n,p}(q_n)} \sum_{t=1}^{n} \hat{V}^2_{t,p} \tag{7.2.7}$$

$$V\hat{S}_{n,p}(q_n) = \frac{1}{n \cdot \hat{\omega}^2_{n,p}(q_n)} \left\{ \sum_{t=1}^{n} \hat{V}^2_{t,p} - \frac{1}{n}\left(\sum_{t=1}^{n} \hat{V}_{t,p} \right)^2 \right\} = \hat{M}_{n,p}(q_n) - \frac{1}{\hat{\omega}^2_{n,p}(q_n)} \left(\frac{1}{n} \sum_{t=1}^{n} \hat{V}_{t,p} \right)^2 \tag{7.2.8}$$

And:

$$K\hat{S}_{n,p}(q_n) = \max_{t=1,\ldots,n} \frac{1}{\hat{\omega}_{n,p}(q_n)} \left| \hat{V}_{t,p} - \frac{t}{n} \hat{V}_{n,p} \right| \tag{7.2.9}$$

With $\hat{V}_{t,p} = n^{-1/2}\hat{S}_{t,p} = n^{-1/2} \sum_{j=1}^{t} \hat{\eta}_{j,p}$ the scaled partial sum of OLS residuals. The test statistic (7.2.7) is the widely known as KPSS statistic proposed by Kwiatkowski et.al. (1992), while (7.2.8) is the rescaled variance-ratio test statistic proposed by Giraitis et.al. (2003). This two test statistics measure a excessive fluctuation in the residual sequence through a Cramér-von Mises metric, while the CUSUM-type test statistic (7.2.9) is the one proposed by Xiao (2001) that uses the Kolmogorov-Smirnov measure of fluctuation. In these three cases, a rejection of the null hypothesis of stationarity is registered for large values of the estimated test statistic when compared with the proper critical values from its non-standard null limiting distributions. In (7.2.7)-(7.2.9) $\hat{\omega}^2_{n,p}(q_n)$ is a consistent estimator of the long-run variance of the sequence η_t under the null hypothesis of stationarity, usually a kernel nonparametric estimator of the form:

$$\hat{\omega}^2_{n,p}(q_n) = \sum_{k=-q_n}^{q_n} w(k/q_n)\hat{\gamma}_{n,p}(k) \tag{7.2.10}$$

Where $\hat{\gamma}_{n,p}(k) = n^{-1} \sum_{t=k+1}^{n} \hat{\eta}_{t,p}\hat{\eta}_{t-k,p}$ is the *k*-lag sample residual autocovariance, q_n is the bandwidth parameter that must satisfy certain upper bounding condition when depends on the sample size, and $w(\cdot/\cdot)$ is the kernel or weighting function[46]. Under the stationarity assumption, and by standard application of the weak LLN, we have that $\hat{\omega}^2_{n,p}(q_n) \rightarrow^p \sigma^2_{\eta,\infty} = \sigma^2_\infty (1-\alpha)^{-2}$ whenever the sample size-dependent bandwidth parameter q_n verify the condition $q_n = O_p(n^{1/2-a})$, with $0 < a < 1/2$, where the O_p is used to cover the cases where it is estimated from data. We also consider, for purpose of comparison of the results in the next section, a test procedure of the reverse hypothesis, that is I(1) against I(0), that uses the same idea and information that the above stationarity tests.

This is the Breitung's (2002, 2003) test for a unit root based on a semi-nonparametric variance-ratio test statistic defined as:

$$\overline{R}_{n,p} = n^{-1}\hat{R}_{n,p} \tag{7.2.11}$$

[46] For a more formal and complete treatment of the choice and combination of bandwidth and kernel functions in the context of stationarity tests see, e.g., Kurozumi (2002), Carrion-i-Silvestre and Sansó (2006), and Xiao and Lima (2007) among others.

Where the original variance-ratio test statistic, $\hat{R}_{n,p}$, is given by:

$$\hat{R}_{n,p} = \frac{1}{n\cdot\sum_{t=1}^{n}\hat{\eta}_{t,p}^{2}}\sum_{t=1}^{n}\hat{S}_{t,p}^{2} = \frac{1}{n\cdot\hat{\gamma}_{n,p}(0)}\sum_{t=1}^{n}\hat{V}_{t,p}^{2} \qquad (7.2.12)$$

Which is closely-related to the KPSS test statistic through the relation $\hat{R}_{n,p} = \hat{\lambda}_{n,p}(q_n)\cdot\hat{M}_{n}(q_n)$, where $\hat{\lambda}_{n,p}(q_n) = 1 + 2\sum_{k=1}^{q_n} w(k/q_n)\hat{\rho}_{n,p}(k)$ is the ratio of the long-run to short run variance estimators, with $\hat{\rho}_{n,p}(k) = \hat{\gamma}_{n,p}^{-1}(0)\hat{\gamma}_{n,p}(k)$ the *k*-lag sample residual autocorrelation. The test statistic $\overline{R}_{n,p}$, that has the advantage that it does not requires any correction for autocorrelation, is pivotal in the sense that its asymptotic distribution under the assumption of a fixed unit root is free of nuisance parameters (see Breitung (2002) for more details).

Given that $\overline{R}_{n,p}$ converges asymptotically to zero under stationarity, the test rejects the null of a fixed unit root for a low estimated value of the test statistic. Finally, and to complement the results for all these more traditional test procedures, we consider the recently proposed test for covariance stationary by Xiao and Lima (2007). These authors argue that this test procedure preserves the same size and power properties that of existing similar tests, while that it has higher power in the presence of a changing unconditional variance, but this was only proved through a simulation experiment. To our knowledge, under the alternative of a fixed unit root, never has been determined its consistency rate and asymptotic distribution. For that reason, and for further comparative purposes, we also consider this additional test. Their test procedure is a generalization of the CUSUM-type test by Xiao (2001) given in (7.2.9) to the case of detecting a excessive fluctuation in the first two sample moments of the detrended process. To that end, the main focus of analysis is the behaviour of the scaled partial sum functional of the bivariate process $\boldsymbol{\xi}_t = (\eta_t, \upsilon_t)'$, with $\upsilon_t = \eta_t^2 - \overline{\sigma}_{n,\eta}^2$, where $\overline{\sigma}_{n,\eta}^2 = n^{-1}\sum_{s=1}^{n}\eta_s^2$ is the sample variance of the error process which under stationarity becomes the finite population variance, that is, $\overline{\sigma}_{n,\eta}^2 = E[\eta_t^2]$. Under stationarity and Assumption 2, the process $n^{-1/2}\sum_{t=1}^{[nr]}\boldsymbol{\xi}_t$ verifies an invariance principle with weak convergence to a bivariate Brownian process with covariance matrix $\boldsymbol{\Omega}$ that is the long-run covariance matrix of $\boldsymbol{\xi}_t$, that is, $\lim_{n\to\infty}\mathrm{Var}(n^{-1/2}\sum_{t=1}^{n}\boldsymbol{\xi}_t)$. With this, an appropriate test statistic could be based on the sample version of the following generalized CUSUM-type statistic $C_n = \max_{1\le k\le n}\|\boldsymbol{\Omega}^{-1/2}n^{-1/2}\sum_{t=1}^{k}\boldsymbol{\xi}_t\|$ given by:

$$\hat{C}_{n,p}(q_n) = \max_{1\le k\le n}\|\hat{\mathbf{C}}_{n,p}(k,q_n)\| = \max_{1\le k\le n}\left\|\hat{\boldsymbol{\Omega}}_{n,p}^{-1/2}(q_n)n^{-1/2}\sum_{t=1}^{k}\hat{\boldsymbol{\xi}}_{t,p}\right\|, \qquad n^{-1/2}\sum_{t=1}^{[nr]}\hat{\boldsymbol{\xi}}_{t,p} = \begin{pmatrix} n^{-1/2}\hat{S}_{n,p}(r) \\ n^{-1/2}\hat{S}_{n,p}(r) \end{pmatrix}$$

Where $\|\cdot\|$ is any appropriate norm of vectors, $\hat{\boldsymbol{\xi}}_{t,p} = (\hat{\eta}_{t,p}, \hat{\upsilon}_{t,p})'$, $\hat{\upsilon}_{t,p} = \hat{\eta}_{t,p}^2 - \hat{\sigma}_{n,p}^2$ with $\hat{\sigma}_{n,p}^2 = n^{-1}\sum_{j=1}^{n}\hat{\eta}_{j,p}^2$, and $\hat{\boldsymbol{\Omega}}_{n,p}(q_n)$ is the kernel nonparametric estimator of $\boldsymbol{\Omega}$ given by:

$$\hat{\boldsymbol{\Omega}}_{n,p}(q_n) = \begin{pmatrix} \hat{\omega}_{n,p}^2(q_n) & \hat{\lambda}_{n,p}(q_n) \\ & \hat{\kappa}_{n,p}^2(q_n) \end{pmatrix}$$

Where $\hat{\omega}_{n,p}^2(q_n)$ is as in (7.2.10) above, with $\hat{\lambda}_{n,p}(q_n)$ and $\hat{\kappa}_{n,p}^2(q_n)$ defined similarly by using the sequences $\hat{\eta}_{t,p}\hat{\upsilon}_{t,p}$ and $\hat{\upsilon}_{t,p}$, respectively. Taking $v = -1/2$ in (7.2.4) and (7.2.5) under the fixed unit root assumption, $\alpha_t = 1$ for all $t = 1, ..., n$, we then have that:

$$\hat{B}_{n,p}(r) = n^{-1/2}\hat{\eta}_{[nr],p} = n^{-1/2}\eta_{[nr]} - [\boldsymbol{\tau}'_p(r)][n^{1/2}\boldsymbol{\Gamma}_n^{-1}(\hat{\boldsymbol{\beta}}_{p,n} - \boldsymbol{\beta}_p)]$$

$$\Rightarrow B_p(r) = B(r) - \boldsymbol{\tau}'_p(r)\mathbf{Q}_p^{-1}\int_0^1 \boldsymbol{\tau}_p(s)B(s)ds \qquad (7.2.13)$$

$$= \sigma_\infty\left(W(r) - \boldsymbol{\tau}'_p(r)\,\mathbf{Q}_p^{-1}\int_0^1 \boldsymbol{\tau}_p(s)W(s)ds\right) = \sigma_\infty W_p(r)$$

With $\hat{B}_{n,p}(r) = \hat{B}_{t,p} = n^{-1/2}\hat{\eta}_{t,p}$ for $t/n \le r < (t+1)/n$, $t = 1, ..., n-1$, which gives $\hat{V}_{[nr],p} = O_p(n)$, and $n^{-1}\hat{V}_{[nr],p} = n^{-3/2}\hat{S}_{[nr],p} = n^{-1}\sum_{t=1}^{[nr]}\hat{B}_{t,p} \Rightarrow \int_0^r B_p(s)ds$ as $n\to\infty$. Also, taking into account that $n^{-1/2}\eta_t = n^{-1/2}\eta_{t-k} + O_p(n^{-1/2})$, for any $k \ge 1$, then:

$$\hat{B}_{t-k,p} = n^{-1/2}\hat{\eta}_{t-k,p} = n^{-1/2}\eta_{t-k} - [\boldsymbol{\tau}'_p(\tfrac{t-k}{n})][n^{1/2}\boldsymbol{\Gamma}_n^{-1}(\hat{\boldsymbol{\beta}}_{p,n} - \boldsymbol{\beta}_p)]$$
$$= \hat{B}_{t,p} + (\Delta_k\boldsymbol{\tau}'_p(\tfrac{t}{n}))[n^{1/2}\boldsymbol{\Gamma}_n^{-1}(\hat{\boldsymbol{\beta}}_{p,n} - \boldsymbol{\beta}_p)] + O_p(n^{-1/2}) = \hat{B}_{t,p} + O_p(n^{-1/2}) \qquad (7.2.14)$$

Where $\Delta_k\boldsymbol{\tau}'_p(\frac{t}{n}) = 0$ for $p = 0$, and $\Delta_k\boldsymbol{\tau}'_p(\frac{t}{n}) = O(n^{-1})$ for any $p \ge 1$, so that:

$$n^{-1}\hat{\gamma}_{n,p}(k) = n^{-1}\sum_{t=k+1}^{n}\hat{B}_{t,p}\hat{B}_{t-k,p} = n^{-1}\sum_{t=1}^{n}\hat{B}_{t,p}^2 + O_p(n^{-1/2})$$
$$\Rightarrow \int_0^1 B_p(s)^2 ds = \sigma_\infty^2\int_0^1 W_p(s)^2 ds \qquad (7.2.15)$$

For any $|k| \le q_n$. Also, from (2.13), we have that:

$$n^{-1}\hat{\upsilon}_{t,p} = (n^{-1/2}\hat{\eta}_{t,p})^2 - n^{-1}\sum_{j=1}^{n}(n^{-1/2}\hat{\eta}_{j,p})^2 = \hat{B}_{t,p}^2 - n^{-1}\sum_{j=1}^{n}\hat{B}_{j,p}^2 \Rightarrow B_p^2(r) - \int_0^1 B_p^2(s)ds \qquad (7.2.16)$$

And:

$$n^{-2}\hat{S}_{n,p}(r) \Rightarrow \int_0^r\left(B_p^2(s) - \int_0^1 B_p^2(a)da\right)ds = \int_0^r B_p^2(s)ds - r\int_0^1 B_p^2(s)ds \qquad (7.2.17)$$

With this results, we can now formulate the following proposition that states the asymptotic distribution of all these test statistics under the fixed unit root assumption.

Proposition 7.2.1 Under the DGP (2.1)-(2.2), with $\alpha_t = 1$ for all t = 1, ..., n, and Assumption 2.a, we have that:

(a) $(n\cdot q_n)^{-1}\hat{\omega}_{n,p}^2(q_n) \Rightarrow K\cdot\int_0^1 B_p^2(s)\,ds = K\sigma_\infty^2\int_0^1 W_p^2(s)\,ds$

(b) $(q_n/n)\hat{M}_{n,p}(q_n) \Rightarrow M_p = \int_0^1\left(\int_0^r W_p(s)ds\right)^2 dr/K\cdot\int_0^1 W_p(s)^2 ds$

(c) $(q_n/n)\hat{VS}_{n,p}(q_n) \Rightarrow VS_p = M_p - \left(K\cdot\int_0^1 W_p(s)^2 ds\right)^{-1}\left\{\int_0^1\left(\int_0^r W_p(s)ds\right)dr\right\}^2$

(d) $(q_n/n)^{1/2}\hat{KS}_{n,p}(q_n) \Rightarrow KS_p = \left(K\cdot\int_0^1 W_p(s)^2 ds\right)^{-1/2}\sup_{0\le r\le 1}|\int_0^r W_p(s)ds - r\int_0^1 W_p(s)ds|$

(e) $\bar{R}_{n,p} \Rightarrow R_p = K\cdot M_p$

And:

(f) $\hat{\lambda}_{n,p}(q_n) = O_p(q_n n^{3/2})$, $\hat{\kappa}_{n,p}^2(q_n) = O_p(q_n n^2)$

$(n^2\cdot q_n)^{-1}\cdot\hat{\kappa}_{n,p}^2(q_n) \Rightarrow K\cdot\int_0^1\left(B_p^2(r) - \int_0^1 B_p^2(s)ds\right)^2 dr = K\sigma_\infty^4\int_0^1\left(W_p^2(r) - \int_0^1 W_p^2(s)ds\right)^2 dr$

(g) $(q_n/n)^{1/2}\hat{C}_{n,p}(q_n) = \max_{1\le k\le n}\|(n/q_n)^{-1/2}\hat{\mathbf{C}}_{n,p}(k,q_n)\| \Rightarrow C_p$

Where $K = \int_{-1}^{1} w(s)ds$ for any symmetric kernel with finite support, with the limiting distribution C_p in (g) given by:

$$C_p = \left(K \cdot \int_0^1 \left(B_p^2(s) - \int_0^1 B_p^2(a)da \right)^2 ds \right)^{-1/2} \sup_{0 \le r \le 1} \left| \int_0^r \left(B_p^2(s) - \int_0^1 B_p^2(a)da \right) ds \right| .$$

Proof. Results (a)-(e) follow from the use of equations (7.2.13)-(7.2.15) and the application of the continuous mapping theorem (CMP)[47], while that the result in (f) comes from (7.2.13) and (7.2.16) by using a similar development that in (7.2.14) and (7.2.15). Given the convergence rates in (a) and (f) for the components of the kernel estimator $\hat{\boldsymbol{\Omega}}_{n,p}(q_n)$, under the fixed unit root assumption, it is dominated by the element $\hat{\kappa}_{n,p}^2(q_n)$, which gives:

$$(n^2q_n)^{-1}\hat{\boldsymbol{\Omega}}_{n,p}(q_n) = \begin{pmatrix} O_p(n^{-1}) & O_p(n^{-1/2}) \\ & (n^2q_n)^{-1}\hat{\kappa}_{n,p}^2(q_n) \end{pmatrix}$$

With this and the use of (7.2.16) and (7.2.17) we have that:

$$\hat{\mathbf{C}}_{n,p}(r,q_n) = (nq_n^{1/2})^{-1}[(n^2q_n)^{-1}\hat{\boldsymbol{\Omega}}_{n,p}(q_n)]^{-1/2} \begin{pmatrix} n[n^{-3/2}\hat{S}_{n,p}(r)] \\ n^{3/2}[n^{-2}\hat{S}_{n,p}(r)] \end{pmatrix}$$

$$= (n/q_n)^{1/2} \left\{ [(n^2q_n)^{-1}\hat{\boldsymbol{\Omega}}_{n,p}(q_n)]^{-1/2} \begin{pmatrix} n^{-1/2}[n^{-3/2}\hat{S}_{n,p}(r)] \\ n^{-2}\hat{S}_{n,p}(r) \end{pmatrix} \right\}$$

Where the term between brackets is $O_p(1)$, with $n^{-1/2}[n^{-3/2}\hat{S}_{n,p}(r)] = O_p(n^{-1/2}) = o_p(1)$. This gives the consistency rate of $(n/q_n)^{1/2}$, which is the same as for the Xiao (2001) KS test given in (d). With this it is straightforward to show that:

$$(q_n/n)^{1/2}\hat{C}_{n,p}(q_n) = \max_{1 \le k \le n} \| (n/q_n)^{-1/2}\hat{\mathbf{C}}_{n,p}(k,q_n) \| \Rightarrow C_p$$

With C_p given above.

Remark 7.3 Note that all this limiting distributions (b)-(e) are free of nuisance parameters, under the assumption of a correct specification of the deterministic component, because the scale effect from the long-run variance σ_∞^2 in the numerator and denominator of the limits cancels. The limit distributions of the scaled stationarity tests depends on the kernel choice through the constant *K*, that takes value one in the case of the Bartlett kernel, which is the most often used in practice. Also, this results indicate the consistency rates of the stationarity tests under the alternative of a fixed unit root, which is (n/q_n) for the KPSS and VS tests, and $(n/q_n)^{1/2}$ for the KS and C tests.

[47] A more detailed proof and additional results concerning the asymptotic distribution of these test statistics and of the test for covariance stationarity by Xiao and Lima (2007) under a fixed unit root is available in Afonso-Rodríguez (2012b).

For this last testing procedure, and under the fixed unit root alternative to stationarity, the dominant term is the one corresponding to the partial sum of centered squared residuals, $n^{-2}\hat{S}_{n,p}(r)$, whose behaviour is intended to capture instabilities in the first two sample moments.

7.3 Semi-nonparametric residual-based tests for stationarity and for a fixed unit root under a weak bilinear unit root process

This section is concerned with the extension of the previous results to the case where equation (7.2.2), determining the structure of the stochastic trend component, is described as a particular member of the class of nonstationary random coefficient autoregressive (RCA) processes. Assuming that the sequence of random coefficients α_t in (7.2.2) can be decomposed as:

$$\alpha_t = \phi_0 + \phi_t \qquad (7.3.1)$$

With ϕ_0 a fixed real-valued coefficient and ϕ_t a sequence of random variables, then $\lambda = E[\log(\alpha_t)] < 0$, and $\kappa_m = E[|\alpha_t|^m] < 1$ for any $m > 0$ determine the necessary and sufficient conditions for strict stationarity and ergodicity and weak stationarity (existence of the *m*th-order noncentral moment), respectively, of the stochastic recurrence equation (2.2) (see, e.g., Nagakura (2009), and Afonso-Rodríguez (2012a)). The leading case here is when $\kappa_2 = \phi_0^2 + E[\phi_t^2] = \phi_0^2 + \sigma_\phi^2 \geq 1$ for a given random sequence ϕ_t such that $E[\phi_t] = 0$, and $\sigma_\phi^2 \geq 0$, so that η_t is not covariance stationary (while that it can preserve the strict stationary property depending on the distributional assumptions posed on the random sequence ϕ_t), so that it can partially behaves as a random walk through the full sample, being stationary for some periods, and even mildly explosive for others. In what follows we are concerning with a member of this family of nonstationary processes which is called a stochastic (or randomized) unit root (STUR) process when $\phi_0 = 1$, so that the autoregressive root is equal to 1 only on average, with the STUR process being stationary for some time, while it would be mildly explosive for some other time.[48]

The fixed unit root process is obtained as a particular case when $\sigma_\phi^2 = 0$. These arguments suggest the interest in the study of the size and power properties of unit root and stationarity tests, respectively, under nonstationarity of the STUR type. Some early studies on this topic can be found in Granger and Swanson (1997), McCabe and Smith (1998), Gonzalo and Lee (1998), and Yoon (2004), while more recently it can be cited Francq et.al. (2008) and Afonso-Rodríguez (2012b). Testing procedures for a STUR alternative to a fixed unit root process were developed by McCabe and Tremayne (1995), Leybourne, McCabe and Tremayne (1996), and Leybourne, McCabe and Mills (1996). One of the main difficulties for the analytic study of the effects of considering a STUR alternative is the lack of theoretical results about the possible application of the invariance principle to empirical process based on this kind of models.

---

[48] Granger and Swanson (1997) introduce a very different STUR process, where $\alpha_t = \exp(\phi_t)$, with ϕ_t a strictly stationary and Gaussian AR(1) process. With this assumptions, it follows that $\alpha_t = 1 + \phi_t + O_p(\phi_t^2)$, which implies that both specifications are equivalent up to the term $O_p(\phi_t^2)$. Despite the relationship between these both specifications, they present some differences so that we can call this latter case as the exponential STUR process, while that (3.1) with ϕ_0 = 1 might be called the additive STUR process.

McCabe and Smith (1998) introduce the concept of local heteroskedastic integration by considering that the random sequence ϕ_t may be factorized as $\phi_t = \omega n^{-\lambda} \upsilon_t$, $\omega \geq 0$, with υ_t a stationary sequence and $\lambda > 0$ an appropriate power factor needed to obtain an approximate representation of the random coefficients $\alpha_t = 1 + \phi_t$ which generates a local (first-order) approximation of the STUR process about the fixed unit root case. Also, by assumption, they rule out the case where the random terms ϕ_t and ε_t may be correlated, which exclude some interesting alternatives from their analysis as the first-order Markovian bilinear unit-root process proposed by Francq et.al. (2008).

Taylor and van Dijk (2002) develop an extensive set of Monte Carlo experiments to study the power of these test statistics against different forms of nonstationarity other than difference stationarity which display a greater degree of persistence. Their general conclusion is that these tests against STUR behaviour only appear to display power against processes with a higher degree of persistence than the fixed unit root process, and not against processes with lower persistence, even where those processes are also non-stationary. For that reason, as well as for having interesting economic interpretations and its empirical relevance, we consider the analysis of the stochastic bilinear (diagonal) unit root (BLUR) process[49] when α_t in (7.2.2) is given by:

$$\alpha_t = \alpha_t(\boldsymbol{\theta}) = 1 + \alpha\varepsilon_{t-1} \tag{7.3.2}$$

Where $\boldsymbol{\theta} = (1, \alpha)'$, with $\boldsymbol{\theta}_0 = (1, 0)'$ indicating the fixed unit root case. With this, the complete specification of the stochastic trend component η_t is the following:

$$\eta_t = \alpha_t(\boldsymbol{\theta})\eta_{t-1} + \varepsilon_t = (1 + \alpha\varepsilon_{t-1})\eta_{t-1} + \varepsilon_t \quad t = 1, \ldots, n \tag{7.3.3}$$

Charemza et.al. (2005) first introduce this process as being derived from a model of speculative behaviour and proved that, under Assumption 2.b, $E[\eta_t] = \alpha\sigma_\varepsilon^2(t-1)$, $E[\Delta\eta_t] = \alpha\sigma_\varepsilon^2$, and $Var[\Delta\eta_t] = (5\sigma_\varepsilon^2 + \alpha^2 E[\varepsilon_0^2])(1 + \alpha^2\sigma_\varepsilon^2)^{t-2} - 4t\alpha^2\sigma_\varepsilon^4 + 7\alpha^2\sigma_\varepsilon^4 - 4\sigma_\varepsilon^2 = O(t)$ for any value $\alpha \geq 0$, which seems to make this model not suitable as a close alternative to I(1) series. These authors were mostly concerned with testing the assumption $\alpha = 0$ against the one-sided alternative $\alpha > 0$, giving rise to the so-called α-test through an extension of the Dickey-Fuller (DF) regression and DF-type test to this framework. In section 7.4 we review this test procedure, derive some new results for this test statistic and propose a new test procedure for testing the fixed-unit root against the BLUR alternative. In order to consider a plausible alternative to the fixed-unit root case, we next consider the BLUR model with a weak bilinear effect (in short, weak BLUR process) and what can be called the weak BLUR distribution, firstly proposed by Lifshits (2006). By introducing the normalization $\alpha_n = \alpha \cdot n^{-1/2}$ of the bilinear parameter, this model can now be interpreted as a local alternative to the fixed unit root case in finite samples because that, for a fixed value of α, $\alpha_n \to 0$ as $n \to \infty$.

---

[49] Francq et.al. (2008) also consider the case of a stochastic bilinear (subdiagonal) unit-root process which allows for stationary increments and in this sense is more closely related to the fixed unit root case. Particularly, they consider the error correction form of an AR(*p*+1) process, $\Delta Y_t = \varphi Y_{t-1} + \sum_{k=1}^{p} \varphi_k \Delta Y_{t-k} + \eta_t$, where the error term η_t follows a bilinear process of order *q*, BL(*q*), of the form $\eta_t = (1 + \sum_{k=1}^{q} \phi_k \eta_{t-k})\varepsilon_t$. With a slight manipulation of these expressions, we have the following STUR form $Y_t = (1 + \varphi + \phi_1 \varepsilon_t) Y_{t-1} + \sum_{k=1}^{p} \varphi_k \Delta Y_{t-k} + \zeta_t$ in which the possible dependence is modelled parametrically through the *p* lags of the first differences. Under appropriate conditions on the coefficients ϕ_k, η_t is a centered non-correlated process (weak white noise), so that this STUR-type process becomes asymptotically indistinguishable from a fixed unit root.

The next definition, which is based on Theorem 1 by Lifshits (2006), sets the weak limit distribution of this process.

Definition 7.3.1 Let the BLUR process in (3.3) with small a bilinearity coefficient, that is with $\theta = \theta_n = (\alpha_n, 1)'$, $\alpha_n = \alpha \cdot n^{-1/2}$, $\alpha > 0$, error terms ε_t that follows Assumption 2.b, and define the scaled partial sum process of $\eta_{t,n} = \alpha_t(\theta_n)\eta_{t-1,n} + \varepsilon_t$ as $H_{n,\alpha}(r) = n^{-1/2}\eta_{t,n}$, for (t–1)/n ≤ r < (t/n), t = 1, ..., n. Then

$$H_{n,\alpha}(r) \Rightarrow H_\alpha(r) = \sigma_\varepsilon A_\alpha(r)\int_0^r \frac{1+\alpha^2\sigma_\varepsilon^2 s}{A_\alpha(s)}(dW(s) - \alpha\sigma_\varepsilon ds) + \alpha\sigma_\varepsilon^2 r \qquad (7.3.4)$$

With $A_\alpha(r) = \exp(\alpha\sigma_\varepsilon W(s) - \alpha^2\sigma_\varepsilon^2 r/2)$, and W(s) a standard Wiener process.

By using $\alpha = \alpha_n$ as in the above definition and the explicit form of $Var[\Delta\eta_t]$ in Charemza et.al. (2005) it is immediate to show that this STUR process is asymptotically weakly stationary in first differences. Also, given that this fact is explicitly considered in the derivation of the limit diffusion process $H_\alpha(r)$ in (7.3.4), we have that $H_0(r) = B(r) = \sigma_\varepsilon W(r)$ for α = 0, $E[H_\alpha(r)] = \alpha\sigma_\varepsilon^2 r$ (so that the weak BLUR process introduce a location displacement to the right of the limit distribution compared to that of the Brownian motion), and $E[H_\alpha^2(r)] = 5\alpha^{-2}[\exp(\alpha^2\sigma_\varepsilon^2 r) - 1] - 4\alpha^2\sigma_\varepsilon^4 r$ (see Liftshits (2006), p.4544). With these two moments it can be checked that $Var[H_\alpha(r)] > 0$ only for values of α below the upper limit of $1/(\sigma_\varepsilon\sqrt{r})$ for any 0 < r ≤ 1. Also, the first derivative of $Var[H_\alpha(r)]$ with respect to α is negative for values of α below the upper limit of $1/(\sigma_\varepsilon\sqrt{r})$ for any 0 < r ≤ 1, so that it is a strictly decreasing function in α for values α ∈ (0, $1/(\sigma_\varepsilon\sqrt{r})$) which determines that this distribution becomes flat around the mean value. Thus, from (7.3.4) it is evident that this limiting distribution is a function of two parameters, α and σ_ε^2, where the error variance σ_ε^2 plays an important role not only in the scale of the distribution but also in the extent of the drift term $\alpha\sigma_\varepsilon^2 r$. Also, by using a second-order Taylor series expansion of $H_\alpha(r)$ around α = 0, it is possible to determine more precisely how this process depends on the Brownian motion $B(r) = \sigma_\varepsilon W(r)$ as a function of the bilinear parameter α:

$$H_\alpha(r) = B(r) + \alpha((1/2)B(r)^2) - \frac{\alpha^2\sigma_\varepsilon^2}{2}\int_0^r B(s)ds + O_p(\alpha^3)$$

Finally, by application of the concept of summability defined by Gonzalo and Pitarakis (2006) and Berenguer-Rico (2011)[50], it is immediate to appreciate that this weak BLUR process is an asymptotic plausible alternative to the fixed unit root case, I(1) = *S*(1), because both processes are of the same order of summability. In order to make possible the comparison of the next results with what were obtained in section 2, we formulate a proposition that establish the invariance of the weak BLUR limit distribution when the error term from the bilinear equation is weakly dependent and, in particular, when follows a linear process.

[50] The concept of summability is a generalization of the order of integration of a stochastic process which is defined as follows. A stochastic process X_t with positive variance is said to be summable of order δ, denoted as S(δ), if $n^{-(1/2+\delta)}L(n)\sum_{t=1}^{[nr]}(X_t - m_t) = O_p(1)$ as $n \to \infty$, where δ is the minimum real number that makes this scaled partial sum process bounded in probability, with m_t a deterministic sequence, and L(n) a slowly-varying function.

Proposition 7.3.1 Let $\varepsilon_t = C(L)u_t$, with $C(L) = \sum_{j=0}^{\infty} c_j L^j$, and $\sum_{j=1}^{\infty} j c_j^2 < \infty$, where the error process u_t is given either by (a) $u_t \sim iid(0, \sigma_u^2)$, $\sigma_u^2 = E[u_0^2] < \infty$, with $E[|u_t|^m] < \infty$, or (b) (u_t, F_t) is a MDS with respect to the information set $F_t = \sigma(u_k, k \le t)$, $E[u_t^2 | F_{t-1}] = \sigma_u^2 < \infty$, and $\sup_{t\in\mathbb{Z}} E[|u_t|^m | F_{t-1}] < \infty$ a.s., for some m>2. Then, given the weak BLUR(1,1) process with $\boldsymbol{\theta} = \boldsymbol{\theta}_n = (1, \alpha_n)'$, $\alpha_n = n^{-1/2}\alpha$, and $\alpha \geq 0$ fixed, and if $m \geq 4$, then:

$$H_{n,\alpha}(r) = n^{-1/2} H_{t,n}(r) \Rightarrow \sigma_\infty Q_\alpha(r) \tag{7.3.5}$$

With:

$$Q_\alpha(r) = A_\alpha(r) \int_0^r \frac{1 + \alpha^2 \sigma_\infty^2 s}{A_\alpha(s)} \left(dW(s) - \alpha \sigma_\infty ds \right) \tag{7.3.6}$$

And $A_\alpha(r) = \exp(\alpha \sigma_\infty W(s) - \alpha^2 \sigma_\infty^2 r/2)$, where $\sigma_\infty^2 = \sigma_u^2 C(1)^2$, and $W(r)$ is a standard Wiener processes.

Proof. See Appendix A.

With these results, next proposition trivially states the limiting distributions of the residual-based tests for stationarity against a fixed-unit root, and for the reverse hypothesis, that we introduce in section 7.2 under the alternative of a weak BLUR process by simply replacing all the functionals defined in terms of the Brownian motion process by the same expressions but in terms of the diffusion process $H_\alpha(r) = \sigma_\infty Q_\alpha(r)$.

Proposition 7.3.2 Under the DGP (2.1)-(2.2), with the weak BLUR alternative given in Definition **7.**3.1 and under the Proposition **7.**3.2, we have that:

(a) $(n \cdot q_n)^{-1} \hat{\omega}^2_{\alpha,p,n}(q_n) \Rightarrow K \int_0^1 H^2_{\alpha,p}(s) ds = K\sigma_\infty^2 \int_0^1 Q^2_{\alpha,p}(s) ds$

(b) $(q_n/n)\hat{M}_{n,p}(q_n) \Rightarrow M_p(\alpha)$

(c) $(q_n/n)\hat{VS}_{n,p}(q_n) \Rightarrow VS_p(\alpha)$

(d) $(q_n/n)^{1/2}\hat{KS}_{n,p}(q_n) \Rightarrow KS_p(\alpha)$

And:

(e) $\bar{R}_{n,p} \Rightarrow R_p(\alpha) = K M_p(\alpha)$

Where $M_p(\alpha)$, $VS_p(\alpha)$, and $KS_p(\alpha)$ are as in Proposition 2.1 with $W_p(r)$ replaced by:

$$Q_{\alpha,p}(r) = Q_\alpha(r) - \boldsymbol{\tau}_p'(r)\left(\int_0^1 \boldsymbol{\tau}_p(s)\boldsymbol{\tau}_p'(s) ds \right)^{-1} \int_0^1 \boldsymbol{\tau}_p(s) Q_\alpha(s) ds \tag{7.3.7}$$

Where $Q_\alpha(r)$ is given in (3.6). For the components of the kernel estimator of the long-run covariance matrix in the Xiao and Lima (2007) test statistic we have the same divergence rates that under the fixed unit root alternative, with the limiting distribution given now by $C_p(\alpha)$ as in:

Proposition 2.1 with $B_p(r)^2 = \sigma_\infty^2 W_p(r)^2$ replaced by $H_{\alpha,p}(r)^2 = \sigma_\infty^2 Q_{\alpha,p}(r)^2$.

Remark 7.7.4 First of all it is clear from the definition of $Q_\alpha(r)$ that for $\alpha = 0$ we get the results stated in Proposition 7.2.1 under the fixed unit root assumption. Second, from the definition of $Q_\alpha(r)$ in (7.3.6) it is evident that all these limit distributions depend both on α, the pure bilinear effect, and the variance of the error process, σ_∞^2 (or σ_ε^2 with weak white noise innovations), so that in strict sense are not free of nuisance parameters. Also, given that all these results are obtained with the same divergence rates of the original tests statistics that under the fixed unit root assumption it is not of application the usual consistency concept against this form of nonstationarity other that difference stationarity. In this case the behaviour is mainly determined by the value of the bilinear parameter α.

The next result states approximately the scale shift and, more importantly, the displacement to the right of $M_p(\alpha)$, as a function of the pure bilinear effect α, which determines the approximate power of the KPSS test and the source of size distortion for the variance-ratio test $\bar{R}_{n,p}$ under the weak BLUR alternative.

Proposition 3.3 Given the limit distribution of the KPSS test under the weak BLUR alternative in result (b) of Proposition 7.3.2, and up to the order of magnitude α^2, we have that:

$$M_p(\alpha) = K^{-1}\frac{M_{1,p}(\alpha)}{M_{2,p}(\alpha)} = M_p + 2\alpha M_{2,p}^{-1} \times \left\{ K^{-1}\int_0^1\left(\int_0^r B_p(s)ds\right)\left(\int_0^r B_{2,p}(s)ds\right)dr - M_p \cdot \int_0^1 B_p(r)B_{2,p}(r)dr \right\} + O_p(\alpha^2) \qquad (7.3.8)$$

With:

$$B_{2,p}(r) = B(r)^2 - \boldsymbol{\tau}_p'(r)\left(\int_0^1 \boldsymbol{\tau}_p(s)\boldsymbol{\tau}_p'(s)ds\right)^{-1}\int_0^1 \boldsymbol{\tau}_p(s)B(s)^2 ds$$

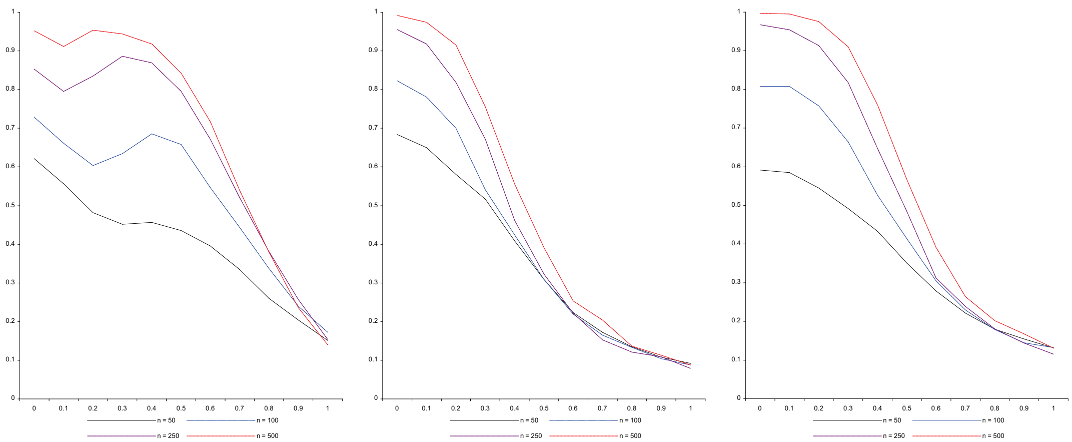
Proof. See Appendix B.

Remark 7.5 Given that, under Gaussianity, $E[B_p(r)B_{2,p}(r)] = 0$, then for any given value of $\alpha > 0$, the leading term in (7.3.8) is given by the first term between brackets that determines a displacement to the right of this limit distribution.

This result allows to establish that there we must expect a reduction in the empirical power of the KPSS test under this alternative as long as α increase, and also an increase of the rejection frequencies of the null of a fixed-unit root with the Breitung's variance-ratio test statistic given that we must expect to find more frequently small values of this test statistic.We have performed a small Monte Carlo study to numerically illustrate the performance of both test procedures under the alternative of a weak BLUR process for different sample sizes and values of α, keeping constant the variance of the error term. To save space we present some figures representing the power performance of the KPSS test in the three standard cases of analysis in terms of the structure of the deterministic component (no trend, OLS demeaned ($p = 0$), and OLS demeaned and detrended ($p = 1$) data), with the Bartlett kernel for the computation of the LRV with deterministic bandwidth parameter $q_n = [c\cdot(n/100)^{1/4}]$, with $c = 4$, and iid Gaussian noise ε_t with variance $\sigma_\varepsilon^2 = 1$.

The power results, see Figure 1 below, are based on 10000 independent realizations, sample size-adjusted quantiles with sample sizes n = 50, 100, 250, and 500 and values of α_n = 0, 0.1, ..., 0.9, 1.

Figure 7.1 Power performance of the KPSS tests under a weak BLUR alternative



(*a*) *No trend component* (*b*) *OLS demeaned data* (*c*) *OLS demeaned and detrended data*

As was discussed below, increasing the value of the bilinear parameter for a fixed sample size creates a spurious appearance of stationary when using this testing procedures due to the joint effect of the displacement and flattening of the distribution $H_\alpha(r)$.

In Afonso-Rodríguez (2012a) we develop a similar study but with the DF tests (normalized bias of OLS estimator and OLS T-test) and we found that the size profile of both test procedures can be seriously distorted under a weak BLUR alternative, with a greater impact that in the above cases of both BLUR parameters, α and σ_ε^2, even for very small values of α. Thus, there could be situations where a standard test for a fixed-unit root could wrongly reject in favour of stationarity due to the effects induced by this kind of nonstationary processes. This findings can be considered as the main reason to look for a testing procedure that allows to discriminate between both types of nonstationarity, that is, fixed-unit root against weak bilinear nonstationarity.

7.4 A new semi-nonparametric residual-based test for detecting weak bilinear nonstationarity

The first attempt to develop a testing procedure for discriminating between a fixed-unit root process and a bilinear unit root, with η_t given as in (7.3.3), was made by Charemza et.al. (2005). Their procedure is inspired by the DF T-test statistic based on the OLS fitting of an auxiliary regression related to the AR(1) model. In this case these authors propose to compute the pseudo T-test statistic based on the OLS fitting of an auxiliary regression which is a feasible version of the following:

$$\Delta Y_t = \alpha\varepsilon_{t-1}Y_{t-1} + \Delta d_t + \xi_t \qquad (7.4.1)$$

Where, from (7.2.1) and Assumption 1, we have that $\Delta d_t = \boldsymbol{\tau}'_{t,p-1}\boldsymbol{\phi}_{p-1}$, with $\boldsymbol{\phi}_{p-1}$ a new set of trend parameters formed as a linear combination of the components of $\boldsymbol{\beta}_p$, and with the error term given by $\xi_t = \varepsilon_t - \alpha\varepsilon_{t-1}d_{t-1}$. Using ε_t from ξ_t we have $\varepsilon_{t-1} = \Delta Y_{t-1} - \alpha\varepsilon_{t-2}\eta_{t-2} - \Delta d_{t-1}$, so that we can write(7.4.1) as:

$$\Delta Y_t = \alpha Z_{t-1} + \boldsymbol{\tau}'_{t,p-1}\boldsymbol{\phi}_{p-1} + \zeta_t \qquad (7.4.2)$$

Where $Z_{t-1} = Y_{t-1}\Delta Y_{t-1}$ is the nonstationary regressor, and $\zeta_t = \varepsilon_t$ under the null hypothesis α = 0. Thus, the propose testing procedure is based on the use the standard OLS-based pseudo T-ratio test statistic, $\hat{T}_{n,p-1}$, for testing $\alpha = 0$ against the one-sided alternative $\alpha > 0$. The next two propositions, that we state here without proof but that can be requested from the author, establish the limit distribution of $\hat{T}_{n,p-1}$ in the case of no deterministic component in the DGP (that is, when d_t = 0 in (2.1)) both under the null of a fixed-unit root and under the alternative of a weak BLUR nonstationary process.

Proposition 7.4.1 Let the series Y_t be generated by (7.2.1)-(7.2.2) and (7.3.2), with $d_t = 0$. Under the null hypothesis of a fixed unit root process, with $\boldsymbol{\theta} = \boldsymbol{\theta}_0 = (1,0)'$, $\alpha_t(\boldsymbol{\theta}_0) = 1$ for all t = 1, ..., n, and the Assumption 2(b) on the error process, then as $n \to \infty$:

$$\hat{T}_{n,p-1} \Rightarrow \int_0^1 W(s)dW^*(s)\left(\int_0^1 W(s)^2 ds\right)^{-1/2} =^d N(0,1) \qquad (7.4.3)$$

Where $W(s)$, and $W^*(s)$ are two independent standard Wiener processes.

This result extends to the case of OLS detrending with a general order $p \geq 0$ of the polynomial trend function the one obtained by Charemza et.al. (2005) only for the case of raw observations and for the inclusion of a constant term (p = 0). The derivation of a similar result for the case where the DGP contains a deterministic component is somewhat more complex due to the nonlinearity of the OLS estimator of α in (7.4.2) with respect to the observed series Y_t and its components. This is one of the major drawbacks of this approach, together with the possible specification and treatment of more complex deterministic terms in the auxiliary regression and the effect of considering weakly dependent error terms.

Proposition 7.4.2. Let the series Y_t be generated by (7.2.1)-(7.2.3) and (7.3.2), with $d_t = 0$. Under the sequence of local alternatives to the null hypothesis given by the weak BLUR process as in Definition 7.3.1, with $\boldsymbol{\theta} = \boldsymbol{\theta}_n = (1, \alpha_n)'$, $\alpha_t(\boldsymbol{\theta}_n) = 1 + \alpha_n \varepsilon_{t-1}$, $\alpha_n = \alpha n^{1/2}$, and $\alpha > 0$ fixed, and the Assumption 2(b) on the error process with m > 4, then as $n \to \infty$:

(a) $$\sqrt{n}\hat{\alpha}_n - \alpha = \sqrt{n}(\hat{\alpha}_n - \alpha_n) \Rightarrow -\alpha^3 \frac{\int_0^1 H_\alpha^4(s)ds}{\int_0^1 (H_\alpha^2(s) + \alpha^2 H_\alpha^4(s))ds}$$

(b) $\hat{T}_{n,p-1} = O_p(n^{1/2})$

And:

(c) $$n^{-1/2}\hat{T}_{n,p-1} \Rightarrow \frac{\alpha \int_0^1 H_\alpha(s)^2 ds}{\sqrt{1+\alpha^2 G_\alpha} \cdot \sqrt{\int_0^1 (H_\alpha(s)^2 + \alpha^2 H_\alpha(s)^4)ds}} \qquad (7.4.4)$$

With:

$$G_\alpha = \int_0^1 H_\alpha(s)^2 ds + 3\left(\int_0^1 H_\alpha(s)^2 ds\right)^2 \left(\int_0^1 (H_\alpha(s)^2 + \alpha^2 H_\alpha(s)^4)ds\right)^{-1}$$

And $H_\alpha(s)$ the limit diffusion process given in (7.3.4).

Remark 7.6 This result extends the proof of a similar proposition in Lifshits (2006) to the case of OLS detrended data, and also is a modification of his results, where the term $(1+\alpha\tilde{G}_d)^{1/2}$ in the denominator of the right hand side of (c) does not appears. Also, the result in 7.4.2(a) is very important given that it establish the root-*n* consistency of $\hat{\alpha}_n$ as an estimator of the scaled bilinear parameter $\alpha_n = \alpha n^{1/2}$ under the assumption that this effect is present in the observed process.

Despite the merits and relatively good performance of this test procedure, given some of its limitations and difficulties[51], we propose a new alternative testing procedure that is based on comparing the different degree and amplitude of the fluctuations for the first difference of a fixed-unit root process (when $\alpha = 0$), which is stationary, and of a weak BLUR process (when $\alpha > 0$), that is stationary only asymptotically. With this idea we present two versions of this semi-nonparametric residual based test, with the advantage that both are asymptotically equivalent and the first one is based on the same OLS residual sequence as the stationarity tests considered in section 7.7.2. Both test statistics are of KPSS-type in the sense that they are built as the test statistic $\hat{M}_{n,p}(q_n)$ in (7.7.2.7) but with a different set of residuals. The first one, $\tilde{D}_{n,p}(q_n)$, is defined as:

$$\tilde{D}_{n,p}(q_n) = \frac{1}{n\cdot\tilde{v}_n^2(q_n)}\sum_{t=1}^{n}\left(n^{-1/2}\sum_{j=1}^{t}\tilde{\eta}_{j,p}\right)^2 = \frac{1}{n\cdot\tilde{v}_n^2(q_n)}\sum_{t=1}^{n}\left(n^{-1/2}\hat{\eta}_{t,p}\right)^2 \qquad (7.4.5)$$

Where $\tilde{\eta}_{t,p} = \Delta\hat{\eta}_{t,p}$ is the first difference of the OLS residuals given in (7.2.4) from the estimation of (7.2.1) under Assumption 1 on the deterministic component, while the second is given by:

$$\hat{D}_{n,p}(q_n) = \frac{1}{n\cdot\hat{v}_n^2(q_n)}\sum_{t=1}^{n}\left(n^{-1/2}\sum_{j=1}^{t}\hat{\eta}_{j,p-1}\right)^2 \qquad (7.4.6)$$

Based on the sequence of OLS residuals from the fitting of (7.2.1) in first differences, that is $\Delta Y_t = \tau'_{t,p-1}\hat{\phi}_{p-1,n} + \hat{\eta}_{t,p-1}$ $t = 1, ..., n$. In (7.4.5) and (7.4.6), the scaling factors $\tilde{v}_n^2(q_n)$ and $\hat{v}_n^2(q_n)$ are kernel estimators of the LRV as in (7.2.10) based on the corresponding residual sequences. In Appendix C it is proved that both estimators have the same limit because the two residual sequences are asymptotically equivalent. However, despite this, there is a major difference between $\tilde{D}_{n,p}(q_n)$ and $\hat{D}_{n,p}(q_n)$ due to the term in their numerators. The following proposition states the stochastic limit distribution of each test statistics under the general assumption of a weak BLUR process generating the stochastic trend component η_t.

Proposition 7.4.3 Given the DGP (7.2.1)-(7.2.2),(7.3.2), the sequence of local alternatives to the fixed-unit root case given by the weak BLUR process as in Definition 3.1, with $\theta = \theta_n = (1,\alpha_n)'$, $\alpha_t(\theta_n) = 1 + \alpha_n\varepsilon_{t-1}$, $\alpha_n = \alpha n^{1/2}$, and $\alpha > 0$ fixed, and the Assumption 2(a) on the error process, then as $n\to\infty$:

---

[51] The difficulties cited above are mainly related with the use of quantiles of a null limit distribution (the standard normal) that could not be appropriate in any case because the DGP may not correspond to the specified auxiliary regression used to compute the T-test statistic $\hat{T}_{n,p-1}$, along with the fact that the auxiliary regression (7.4.2) is only an approximation to the true DGP (7.4.1).

(a) $\tilde{v}_n^2(q_n), \hat{v}_n^2(q_n) \Rightarrow \omega_\alpha^2 = \sigma^2 + 2\sigma^2\alpha \int_0^1 [H_\alpha^2(s) + H_\alpha(s) + \tfrac{1}{2}] ds$ (7.4.7)

(b) $\tilde{D}_{n,p}(q_n) \Rightarrow \tilde{D}_\alpha = \omega_\alpha^{-2} \int_0^1 H_{\alpha,p}(s)^2 ds$ (7.4.8)

(c) $\hat{D}_{n,p}(q_n) \Rightarrow \hat{D}_\alpha = \omega_\alpha^{-2} \int_0^1 B_{\alpha,p-1}(s)^2 ds$ (7.4.9)

With $H_{\alpha,p}(r)$ given in Proposition 3.2 and $B_{\alpha,p-1}(r)$, a (p−1)th-level $H_\alpha(r)$ process, defined in Appendix C.

Proof. See Appendix C.

It is easy to check that under the fixed-unit root assumption, that is $\alpha = 0$, both limit distributions (7.4.8) and (7.4.9) are free of nuisance parameters and are simply given by

$$\tilde{D}_0 = \int_0^1 Q_{0,p}(s)^2 ds = \int_0^1 W_p(s)^2 ds \qquad (7.4.10)$$

And:

$$\hat{D}_0 = \int_0^1 V_{p-1}(s)^2 ds \qquad (7.4.11)$$

With $W_p(r)$ given in (2.13) and $V_{p-1}(r)$ a $(p-1)th$-level Brownian bridge defined as:

$$V_{p-1}(r) = W(r) - \int_0^r \boldsymbol{\tau}'_{p-1}(s)ds\left(\int_0^1 \boldsymbol{\tau}_{p-1}(s)\boldsymbol{\tau}'_{p-1}(s)ds\right)^{-1}\int_0^1 \boldsymbol{\tau}_{p-1}(s)dW(s) \qquad (7.4.12)$$

Given that for any value $\alpha > 0$, the limit distributions $\tilde{D}_\alpha$ and $\hat{D}_\alpha$ are dominated by the effect of α in the denominator with respect to (7.4.10) and (7.4.11), we compute the lower quantiles of these distributions as the critical values and the rejection of the null hypothesis $\alpha = 0$ using (7.4.5) or (7.4.6) will be for small values of the corresponding test statistic. The following table presents these critical values for the first version of the test statistic, $\tilde{D}_{n,p}(q_n)$ defined in (7.4.5), computed via Monte Carlo simulation, for different sample sizes, iid Gaussian white noise errors and with 10000 independent realizations.

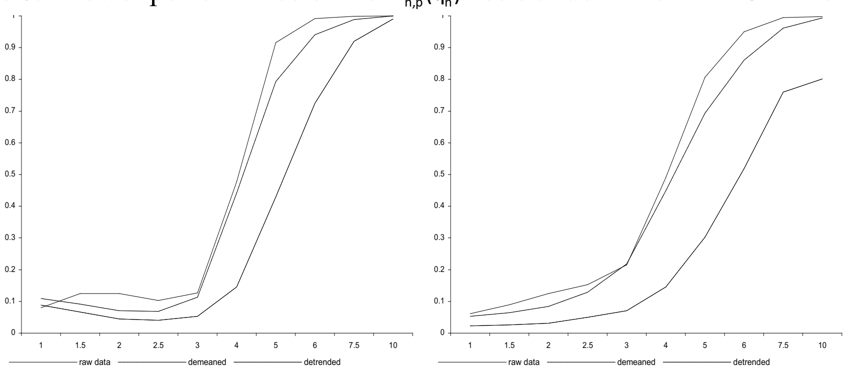
Table 7.1 Finite sample simulated critical values for $\tilde{D}_{n,p}(q_n)$

			No deterministic component	Demeaned data ($p = 0$)	Detrended data ($p = 1$)
Significance level	0.01	$n = 50$	0.03356	0.02415	0.01840
	0.025		0.04254	0.02858	0.02071
	0.05		0.05223	0.03304	0.02335
	0.1		0.06793	0.03981	0.02699
	0.01	$n = 100$	0.03548	0.02352	0.01755
	0.025		0.04621	0.02938	0.02080
	0.05		0.05883	0.03469	0.02373
	0.1		0.08047	0.04256	0.02781
	0.01	$n = 250$	0.03497	0.02427	0.01735
	0.025		0.04392	0.02937	0.01987
	0.05		0.05748	0.03455	0.02263
	0.1		0.07834	0.04213	0.02654
	0.01	$n = 500$	0.03558	0.02343	0.01625
	0.025		0.04396	0.02819	0.01905
	0.05		0.05525	0.03365	0.02225
	0.1		0.07609	0.04159	0.02630
	0.01	$n = 1000$	0.03624	0.02332	0.01672
	0.025		0.04638	0.02885	0.01988
	0.05		0.05738	0.03440	0.02252
	0.1		0.07690	0.04212	0.02646

As can be seen in this table, there are not significant differences in these values for very small sample sizes with respect to large ones.

Also, despite of the differences among (7.4.10) and (7.4.11), the simulated lower quantiles for the null distribution of the second version of the test statistic, $\hat{D}_{n,p}(q_n)$ in (7.4.6), take almost the same values, so that in practice we can use the values in Table 1 for both tests. We have performed a Monte Carlo experiment to evaluate the power behaviour of these tests, based on 10000 independent realizations, for different sample sizes and values of the scaled bilinear parameter $\alpha_n = \alpha/n$. The following Figure 7.2 presents a small part of these results, which is the power behaviour of the $\tilde{D}_{n,p}(q_n)$ test for a sample of size $n = 250$, with $q_n = [c\cdot(n/100)^{1/4}]$, $c = 0$ and 4, the Bartlett kernel, and values of $\alpha = (1, 1.5, 2, 2.5, 3, 4, 5, 6, 7.5, 10)$ (which corresponds to values of α_n from 0.063 to 0.632).

Figure 7.2 Power performance of the $\tilde{D}_{n,p}(q_n)$ tests under a weak BLUR alternative



(a) $c = 0$ (b) $c = 4$

The results from this simulation experiment show that there is a significant increase in power for a given sample size for increasing values of α, and with the sample size for each value of α (consistency), and that the power increase with the use of the nonparametric correction for persistence through the kernel estimator of the long-run variance ($c > 0$).

In order to determine the behaviour of these new test statistics against other forms of a stochastic unit root process we consider a generalized version of the alternative given by the local heteroskedastic integrated (LHI) process introduced in McCabe and Smith (1998).

Under the LHI alternative, it is assumed that $\alpha_t = 1 + \phi_t$, with $\phi_t = \omega n^{-\lambda} \upsilon_t$ for any $\omega \geq 0$ with υ_t a stationary sequence such that it is verified the joint weak convergence $(B_n(r), B_{n,\upsilon}(r))' = n^{-1/2} \sum_{t=1}^{[nr]} (\varepsilon_t, \upsilon_t)' \Rightarrow (B(r), B_\upsilon(r))'$, $B(r) = \sigma_\infty W(r)$ and $B_\upsilon(r) = \sigma_{\upsilon,\infty}[\rho_{\varepsilon\upsilon} W(r) + (1-\rho_{\varepsilon\upsilon}^2)^{1/2} W_\upsilon(r)]$, with $\rho_{\varepsilon\upsilon}$ the long-run correlation between $B(r)$ and $B_\upsilon(r)$ and $(W(r), W_\upsilon(r))$ two standard independent Brownian motion processes. Taking into account that the equation (2.2) for the stochastic trend component can be expressed by backsubstitution as $\eta_t = \eta_0 M_{t,0} + \varepsilon_t + \sum_{k=1}^{t-1} \varepsilon_k M_{t,k}$, with $M_{t,k} = \prod_{i=k+1}^{t} \alpha_i$, $k = 0, 1, \ldots, t-1$, and $M_{t,t} = 1$, we can reverse the order of the subscripts on ϕ_i and ε_k without altering the result in view of the stationarity assumption concerning these error terms, which gives the more convenient representation $\eta_t = \eta_0 M_{t,0} + \varepsilon_1 + \sum_{k=2}^{t} \varepsilon_k M_{k-1,0}$ based on forward summations.

Now, using the same development as in Theorem 1 in McCabe and Smith (1998) we have that:

$$M_{[na],0} = \exp\left(\sum_{i=1}^{[na]} \log(1+\phi\beta_i)\right) = 1 + \omega n^{1/2-\lambda} B_{n,\upsilon}(a) + O_p(n^{1-2\lambda}) \quad (7.4.13)$$

For $[na] = k-1$, $k = 2, \ldots, t+1$, which gives:

$$\begin{aligned} n^{-1/2}\eta_{[nr]} &= B_n(r) + \omega n^{1/2-\lambda}\int_0^r B_{n,\upsilon}(s)dB_n(s) + O_p(n^{1-2\lambda})B_n(r) \\ &\quad + n^{-1/2}\eta_0(1 + \omega n^{1/2-\lambda} B_{n,\upsilon}(r) + O_p(n^{1-2\lambda})) \\ &= B_n(r) + \omega\int_0^r B_{n,\upsilon}(s)dB_n(s) + O_p(n^{-1/2+\gamma}) \end{aligned} \quad (7.4.14)$$

Where the last equality follows from taking $\lambda = 1/2$ and under the assumption that the initial value is of order $\eta_0 = O_p(n^{\gamma}), 0 \le \gamma < 1/2$. Then, by application of the CMT we have that

$$\begin{aligned} n^{-1/2}\eta_{[nr]} &\Rightarrow B(r) + \omega\int_0^r B_{\upsilon}(s)dB(s) \\ &= \sigma_{\infty}\left\{W(r) + \omega\sigma_{\upsilon,\infty}\rho_{\varepsilon\upsilon}\int_0^r W(s)dW(s)\right\} + \omega\sigma_{\infty}\sigma_{\upsilon,\infty}(1-\rho_{\varepsilon\upsilon}^2)^{1/2}\int_0^r W_{\upsilon}(s)dW(s) \end{aligned} \quad (7.4.15)$$

Where this limiting result depends on the magnitude of ω, the two long-run variances $\sigma_{\infty}, \sigma_{\upsilon,\infty}$ and the correlation coefficient $|\rho_{\varepsilon\upsilon}| \le 1$. With this we get an asymptotically comparable result to the limiting distribution $H_{\alpha}(r)$ for a weak bilinear unit root process and in this sense it seems an interesting exercise to evaluate the power of our tests against this alternative. The next Table 2 presents the results of a simulation experiment based on 5000 independent replications with $(\varepsilon_t, \upsilon_t)' \sim iidN(\mathbf{0}_2, \mathbf{I}_2)$ for different sample sizes, deterministic components, and values of the scale parameter ω.

Table 7.2.1 Finite sample power of the new tests statistics for a weak BLUR(1) process under a local heteroskedastic STUR(1) alternative

			ω = 1.0	2.5	5.0	10.0
No deterministics	n = 100	c = 0	0.0552	0.0740	0.1302	0.3142
		c = 4	0.0528	0.0722	0.0886	0.1894
		c = 12	0.0400	0.0412	0.0544	0.0950
	n = 250	c = 0	0.0620	0.0822	0.1334	0.3162
		c = 4	0.0618	0.0696	0.0964	0.1984
		c = 12	0.0432	0.0482	0.0582	0.1032
	n = 500	c = 0	0.0640	0.0766	0.1356	0.2958
		c = 4	0.0646	0.0752	0.1052	0.1930
		c = 12	0.0466	0.0550	0.0698	0.1066
Demeaned	n = 100	c = 0	0.0340	0.0496	0.0832	0.1994
		c = 4	0.0174	0.0212	0.0300	0.0578
		c = 12	0.0014	0.0014	0.0016	0.0030
	n = 250	c = 0	0.0526	0.0708	0.1086	0.2516
		c = 4	0.0412	0.0426	0.0632	0.1280
		c = 12	0.0160	0.0186	0.0238	0.0366
	n = 500	c = 0	0.0620	0.0728	0.1198	0.2714
		c = 4	0.0504	0.0630	0.0874	0.1596
		c = 12	0.0328	0.0382	0.0468	0.0788
Demeaned and detrended	n = 100	c = 0	0.0058	0.0074	0.0168	0.0326
		c = 4	0.0000	0.0002	0.0006	0.0024
		c = 12	0.0000	0.0000	0.0002	0.0010
	n = 250	c = 0	0.0312	0.0434	0.0714	0.1400
		c = 4	0.0084	0.0100	0.0144	0.0292
		c = 12	0.0002	0.0002	0.0006	0.0008
	n = 500	c = 0	0.0480	0.0566	0.0970	0.2020
		c = 4	0.0290	0.0374	0.0556	0.0894
		c = 12	0.0074	0.0112	0.0126	0.0176

In general we can appreciate that our test statistics has no significant power against this STUR alternative, except for relatively high values of the scale parameter ω.

7.5. Empirical illustration

In order to illustrate the application of this new testing procedure for detecting a bilinear unit root as an alternative to the fixed-unit root process, we consider the analysis of four major stock market indices: daily prices of the IBEX35, SP500 and DJCA stock market indexes, and weekly prices of the CAC40 index. To our knowledge, the T-test for a bilinear unit root has been used in practice in very few papers, Charemza et.al. (2005), Hristova (2005), and Tabak (2007). In all the cases, the authors recommend the use of stock log-price series obtained from return series corrected by GARCH-type effects.

We don't formally explore the consequences of omitting this type of correction on the size and power properties of any of this test statistics, but we conduct the empirical analysis on the original log-price series[52]. In order to incorporate some kind of weak dependence in the observed process it is possible to consider an extended version of the auxiliary regression in (4.2) given by:

$$\Delta Y_t = \alpha Z_{t-1} + \boldsymbol{\tau}'_{t,p-1}\boldsymbol{\phi}_{p-1} + \sum_{k=1}^{q}\varphi_k \Delta Y_{t-k} + \zeta_t$$

This model directly results from the assumption that η_t follows a BL(q+1,0,1,1) model, q ≥ 1, with a fixed unit root component, that is $(1-\sum_{k=1}^{q+1}\alpha_k L^k)\eta_t = \alpha\varepsilon_{t-1}\eta_{t-1} + \varepsilon_t$, where the autoregressive polynomial can be decomposed as $1-\sum_{k=1}^{q+1}\alpha_k L^k = (1-\sum_{k=1}^{q+1}\alpha_k L) + (1-L)(1+\sum_{k=1}^{q}(\sum_{j=k+1}^{q+1}\alpha_j)L^k)$, which gives $\Delta\eta_t = \alpha\varepsilon_{t-1}\eta_{t-1} + \sum_{k=1}^{q}\varphi_k\Delta\eta_{t-k} + \varepsilon_t$ under the autoregressive unit root restriction $\sum_{k=1}^{q+1}\alpha_k = 1$, and with $\varphi_k = -\sum_{j=k+1}^{q+1}\alpha_j$. A similar result is obtained under the assumption of a BL(1,q+1,1,1) model, that is $\eta_t = (\alpha_1 + \alpha\varepsilon_{t-1})\eta_{t-1} + \sum_{k=1}^{q+1}c_k\varepsilon_{t-k} + \varepsilon_t$, and also under the BL(1,0,1,1) parameterization with the error term ε_t following a linear process, $\varepsilon_t = C(L)u_t$. In any case, under the null of a fixed unit root, $\alpha = 0$, the augmentation terms $\Delta\eta_{t-k}$ (and ΔY_{t-k}) are all stationary, so that their presence should not interfere with the asymptotic results. In practice, to avoid a collinearity effect with the nonstationary regressor $Z_{t-1} \triangleq Y_{t-1}\Delta Y_{t-1}$, it is precise to run the OLS fitting of this augmented auxiliary regression without the first lag of ΔY_{t-1}, that is $\Delta Y_t = \alpha Z_{t-1} + \boldsymbol{\tau}'_{t,p-1}\boldsymbol{\phi}_{p-1} + \sum_{k=1}^{q}\varphi_k\Delta Y_{t-k-1} + \zeta_t$. The following Table 7.3 presents the details of this analysis for each series.

Table 7.3 Results of tests for a bilinear unit root

2.A. Daily closing log-prices of IBEX35 (01.09.1995-17.09.1999) $n = 1015$		Deterministic component (polinomial trend) None	$p = 0$	$p = 1$	$p = 2$
$\hat{T}_{n,p}(q)$	$q = 0$	3.6036	3.4561	3.4403	3.3851
	1		3.7683		
	2		3.6659		
	3		3.6642		
	4		3.6691		
	5		3.6917		
$\tilde{D}_{n,p}(q_n)$	$c = 0$	2.6648	0.0707	0.1932	
	$c = 12$	2.0004	0.0585	0.1615	

[52] For the price series of the IBEX35 stock market index we choose the same sample period as in Charemza et.al. (2005) to compare the results of both test procedures but without the correction for GARCH effects. The results obtained are very similar to the former ones, thus indicating that this type of potential effect could have little influence on the estimated value of the T-test statistics.

Table 7.3.1 Results of tests for a bilinear unit root (continuation)

2.B. Daily closing log-prices of SP500 (03.01.2000-26.01.2012) $n = 3036$		Deterministic component (polinomial trend)			
		None	$p = 0$	$p = 1$	$p = 2$
$\hat{T}_{n,p}(q)$	$q = 0$	−4.8153	−4.8155	−4.8257	−4.8267
	1		−5.1279		
	2		−5.0582		
	3		−5.0639		
	4		−5.1023		
	5		−5.0936		
$\tilde{D}_{n,p}(q_n)$	$c = 0$	0.1229	0.1229	0.0928	
	$c = 12$	0.1758	0.1757	0.1333	
2.C. Daily closing log-prices of DJCA (03.01.2000-26.01.2012) $n = 3048$		**Deterministic component (polinomial trend)**			
		None	$p = 0$	$p = 1$	$p = 2$
$\hat{T}_{n,p}(q)$	$q = 0$	−3.6586	−3.6634	−3.6678	−3.6682
	1		−3.8145		
	2		−3.7421		
	3		−3.7666		
	4		−3.7930		
	5		−3.7765		
$\tilde{D}_{n,p}(q_n)$	$c = 0$	0.0731	0.0772	0.0807	
	$c = 12$	0.1002	0.1061	0.1109	
2.D. Weekly closing log-prices of CAC40 (04.01.2002-27.01.2012) $n = 526$		**Deterministic component (polinomial trend)**			
		None	$p = 0$	$p = 1$	$p = 2$
$\hat{T}_{n,p}(q)$	$q = 0$	−2.4586	−2.4673	−2.4675	−2.5188
	1		−2.4058		
	2		−2.3757		
	3		−2.4304		
	4		−2.3661		
	5		−2.4482		
$\tilde{D}_{n,p}(q_n)$	$c = 0$	0.1229	0.1262	0.3763	
	$c = 12$	0.1369	0.1404	0.4019	

Note. For the T-test of Charemza et.al. (2005), the case $q = 0$ corresponds to the standard OLS T-ratio test statistic computed from (4.2), with $q > 1$ indicating the OLS T-ratio test statistic computed from (4.2) with the augmentation of q lags of ΔY_{t-1}. For the computation of the kernel LRV in the new test statistic, we consider the deterministic bandwidth parameter given by $q_n = [c\cdot(n/100)^{1/4}]$, with $c = 0$ and $c = 12$.

The analysis of this results indicate that there is no clear evidence about the existence of a bilinear unit root in this series, except for the IBEX35 and DJCA log-price series where for a 10% significance level both test procedures agree when considering the specification of the deterministic component with a constant term.

Further analysis must be done in both cases on the effects of misspecification of the deterministic component such as, for example, the existence of structural breaks that seems to be a prominent feature of some of this series for long time periods.

7.6 Conclusions

In this paper we have set out to investigate the effects that a particular member of the class of stochastic (or randomized) unit root (STUR) processes have on some commonly used tests for the null hypothesis of stationarity against the alternative of a fixed (or linear) unit root. This STUR process, the weak bilinear unit root process (weak BLUR), is given by a restricted and reparameterized version of a simple first order diagonal bilinear model, the BL(1,0,1,1) model, that seems to be useful for characterizing the behaviour of some financial time series and for replicating some of their stylized facts, such as periods of high persistence and conditional heteroskedasticity. This type of nonstationary processes was initially considered by Charemza et.al. (2005) and was subsequently studied by Lifshits (2006), who established an invariance principle that is the basis for our developments.

With this results, we found that these test procedures are consistent, at the usual rates, against this alternative but for high values of the bilinear parameter can lead to wrongly identify stationarity as a consequence of the way in which this limit distribution depends on the value of this key parameter. As an alternative to an existing parametric testing procedure proposed by Charemza et.al. (2005), we propose two alternative and equivalent test statistics based on a simple modification of the KPSS test that have the main advantage of allowing for a very general specification and treatment of the deterministic component. We derive their limit distributions under the null of a fixed unit root and under the alternative of a weak bilinear effect and found through a simulation experiment that have acceptable power in finite samples.

Appendix A. Proof of Proposition 7.3.1

Given the polynomial decomposition $C(L) = C(1) - (1 - L)\tilde{C}(L)$, with $\tilde{C}(L) = \sum_{j=0}^{\infty} \tilde{c}_j L^j$, $\tilde{c}_j = \sum_{i=j+1}^{\infty} c_i$, and $\sum_{j=0}^{\infty} \tilde{c}_j^2 < \infty$, then ε_t can be represented as

$$\varepsilon_t = C(1)u_t - (1 - L)\tilde{C}(L)u_t = v_t - (\tilde{u}_t - \tilde{u}_{t-1}) \quad \text{(A.1)}$$

With $v_t = C(1)u_t$, and $\tilde{u}_t = \tilde{C}(L)u_t$. Then, for the linear process $\varepsilon_t = C(L)u_t$, this decomposition yields directly the martingale approximation to the partial sum process of a stationary time series, and thus it is verified that:

$$n^{-1/2}\sum_{t=1}^{[nr]} \varepsilon_t \Rightarrow \sigma_\infty W(r) = B(r) \quad \text{(A.2)}$$

From the distributional assumptions on u_t we have that v_t is either iid (Case (a)), or a MDS (Case (b)), with zero mean and variance $\omega_v^2 = E[v_t^2] = C(1)^2\sigma_u^2$, with $\omega_v^2 = \sigma_\infty^2$. Also, it is verified that $E[\tilde{u}_t^2] < \infty$, which implies that $\tilde{u}_t = O_p(1)$. Using (A.1), we have that $\alpha_t(\boldsymbol{\theta})$ can be decomposed as $\alpha_t(\boldsymbol{\theta}) = \alpha_0 + \alpha\varepsilon_{t-1} = \alpha_0 + \alpha[v_{t-1} - (\tilde{u}_{t-1} - \tilde{u}_{t-2})]$, that is:

$$\alpha_t(\boldsymbol{\theta}) = \alpha_0 + \alpha v_{t-1} - \alpha(\tilde{u}_{t-1} - \tilde{u}_{t-2}) = \alpha_{v,t}(\boldsymbol{\theta}) - \alpha(\tilde{u}_{t-1} - \tilde{u}_{t-2})$$

So that we can write:

$$\eta_t = \alpha_t(\boldsymbol{\theta})\eta_{t-1} + \varepsilon_t = [\alpha_{v,t}(\boldsymbol{\theta}) - \alpha(\tilde{u}_{t-1} - \tilde{u}_{t-2})]\eta_{t-1} + v_t - (\tilde{u}_t - \tilde{u}_{t-1})$$
$$= \alpha_{v,t}(\boldsymbol{\theta})\eta_{t-1} + v_t - \alpha(\tilde{u}_{t-1} - \tilde{u}_{t-2})\eta_{t-1} - (\tilde{u}_t - \tilde{u}_{t-1})$$

Under $\boldsymbol{\theta} = \boldsymbol{\theta}_n = (1, \alpha_n)'$, we then have:

$$\eta_{t,n} = \alpha_{v,t}(\boldsymbol{\theta}_n)\eta_{t-1,n} + v_t - \alpha_n(\tilde{u}_{t-1} - \tilde{u}_{t-2})\eta_{t-1,n} - (\tilde{u}_t - \tilde{u}_{t-1}) \quad \text{(A.3)}$$

Where $\alpha_{v,t}(\boldsymbol{\theta}_n) = 1 + \alpha_n v_{t-1}$. Now, as in Lifshits (2006), we consider the auxiliary sequence $Y_{t,n}$ defined as $Y_{t,n} = (1 + \alpha_n v_t)\eta_{t,n} = \alpha_{v,t+1}(\boldsymbol{\theta}_n)\eta_{t,n}$ which, from (A.3), can also be written as $Y_{t,n} = \eta_{t+1,n} - v_{t+1} + \alpha_n(\tilde{u}_t - \tilde{u}_{t-1})\eta_{t,n} + (\tilde{u}_{t+1} - \tilde{u}_t)$. Now, using:

$$\eta_{t,n} = Y_{t-1,n} + v_t - \alpha_n(\tilde{u}_{t-1} - \tilde{u}_{t-2})\eta_{t-1,n} - (\tilde{u}_t - \tilde{u}_{t-1})$$

We have:

$$Y_{t,n} - Y_{t-1,n} = \alpha_{v,t+1}(\boldsymbol{\theta}_n)\eta_{t,n} - Y_{t-1,n}$$
$$= \alpha_{v,t+1}(\boldsymbol{\theta}_n)[Y_{t-1,n} + v_t - \alpha_n(\tilde{u}_{t-1} - \tilde{u}_{t-2})\eta_{t-1,n} - (\tilde{u}_t - \tilde{u}_{t-1})] - Y_{t-1,n}$$
$$= \alpha_{v,t+1}(\boldsymbol{\theta}_n)(Y_{t-1,n} + v_t) - Y_{t-1,n}$$
$$- \alpha_{v,t+1}(\boldsymbol{\theta}_n)[\alpha_n(\tilde{u}_{t-1} - \tilde{u}_{t-2})\eta_{t-1,n} + (\tilde{u}_t - \tilde{u}_{t-1})]$$

Where the first term is given by:

$$\alpha_{v,t+1}(\boldsymbol{\theta}_n)(Y_{t-1,n} + v_t) - Y_{t-1,n} = (1 + \alpha_n v_t)(Y_{t-1,n} + v_t) - Y_{t-1,n} = v_t + \alpha_n v_t (Y_{t-1,n} + v_t)$$
$$= (1 + \alpha_n Y_{t-1,n})v_t + \alpha_n[(v_t^2 - \omega_v^2) + \omega_v^2]$$

Then, scaling by $n^{-1/2}$ gives:

$$\frac{Y_{t,n}}{\sqrt{n}} - \frac{Y_{t-1,n}}{\sqrt{n}} = \left(1 + \alpha\frac{Y_{t-1,n}}{\sqrt{n}}\right)\frac{v_t}{\sqrt{n}} + \frac{\alpha}{\sqrt{n}}\frac{(v_t^2 - \omega_v^2)}{\sqrt{n}} + \frac{\alpha\omega_v^2}{n}$$
$$-\alpha_{v,t+1}(\boldsymbol{\theta}_n)\left\{\alpha\frac{(\tilde{u}_{t-1} - \tilde{u}_{t-2})}{\sqrt{n}}\frac{\eta_{t-1,n}}{\sqrt{n}} + \frac{(\tilde{u}_t - \tilde{u}_{t-1})}{\sqrt{n}}\right\}$$

Under the additional condition $m \geq 4$, then $n^{-1/2}\alpha[n^{-1/2}(v_t^2 - \omega_v^2)] = n^{-1/2}\alpha O_p(\eta^{-1/2})$. Also, given that:

$$\alpha_{v,t}(\boldsymbol{\theta}_n) = 1 + \alpha_n v_{t-1} = 1 + \alpha(n^{-1/2}v_{t-1}) = 1 + \alpha O_p(n^{-1/2})$$

And:

$$[n^{-1/2}(\tilde{u}_{t-1} - \tilde{u}_{t-2})][n^{-1/2}\eta_{t-1,n}] = [n^{-1/2}(\mu_{t-1} - \mu_{t-2})]H_\alpha\left(\frac{t-1}{n}\right) = O_p(n^{-1/2})O_p(1) = O_p(n^{-1/2})$$

Then we can write:

$$\frac{Y_{t,n}}{\sqrt{n}} - \frac{Y_{t-1,n}}{\sqrt{n}} = \left(1 + \alpha\frac{Y_{t-1,n}}{\sqrt{n}}\right)\frac{v_t}{\sqrt{n}} + \frac{\alpha\omega_v^2}{n} + O_p(n^{-1/2})$$

Which gives the same solution to $H_\alpha(r)$ that under the iid assumption but with the short-run variance σ_ε^2 replaced by σ_∞^2, that is:

$$H_\alpha(r) = \sigma_\infty A_\alpha(r)\int_0^r \frac{1 + \alpha^2\sigma_\infty^2 s}{A_\alpha(s)}(dW(s) - \alpha\sigma_\infty ds) + \alpha\sigma_\infty^2 r$$

With $A_\alpha(r) = \exp(\alpha\sigma_\infty W(s) - \alpha^2\sigma_\infty^2 r/2)$.

Appendix B. Proof of Proposition 7.3.3.

Given the asymptotic distribution of the KPSS tests under the weak BLUR alternative:

$$M_{p}(\alpha) = K^{-1}\int_0^1\left(\int_0^r H_{\alpha,p}(s)ds\right)^2 dr \Big/ \int_0^1 H_{\alpha,p}(r)^2 dr = K^{-1}\frac{M_{1,p}(\alpha)}{M_{2,p}(\alpha)}$$

We consider the first-order Taylor series expansion in α, that is,

$$M_p(\alpha) = M_p + \alpha \cdot \left.\frac{\partial M_p(\alpha)}{\partial \alpha}\right|_{\alpha=0} + O_p(\alpha^2)$$

With the aim of get evidence about the behaviour of the power profile as a function of α. Given that:

$$\frac{\partial M_p(\alpha)}{\partial \alpha} = M_{2,p}(\alpha)^{-1}\left\{K^{-1}\frac{\partial M_{1,p}(\alpha)}{\partial \alpha} - M_p(\alpha)\frac{\partial M_{2,p}(\alpha)}{\partial \alpha}\right\}$$

$$\frac{\partial}{\partial \alpha}M_{1,p}(\alpha) = 2\int_0^1\left(\int_0^r H_{\alpha,p}(s)ds\right)\left(\int_0^r \frac{\partial H_{\alpha,p}(s)}{\partial \alpha}ds\right)dr$$

And:

$$\frac{\partial}{\partial \alpha}M_{2,p}(\alpha) = 2\int_0^1 H_{\alpha,p}(r)\frac{\partial H_{\alpha,p}(r)}{\partial \alpha}dr$$

Where:

$$\begin{aligned}\frac{\partial H_{\alpha,p}(r)}{\partial \alpha} &= \frac{\partial H_{\alpha}(r)}{\partial \alpha} - \boldsymbol{\tau}_p'(r)\left(\int_0^1 \boldsymbol{\tau}_p(s)\boldsymbol{\tau}_p'(s)ds\right)^{-1}\int_0^1 \boldsymbol{\tau}_p(s)\frac{\partial H_{\alpha}(s)}{\partial \alpha}ds \\ &= B(r)^2 - \boldsymbol{\tau}_p'(r)\left(\int_0^1 \boldsymbol{\tau}_p(s)\boldsymbol{\tau}_p'(s)ds\right)^{-1}\int_0^1 \boldsymbol{\tau}_p(s)B(s)^2ds + O_p(\alpha) \\ &= B_{2,p}(r) + O_p(\alpha)\end{aligned}$$

Then:

$$\left.\frac{\partial M_p(\alpha)}{\partial \alpha}\right|_{\alpha=0} = 2M_{2,p}^{-1}\left\{K^{-1}\int_0^1\left(\int_0^r B_p(s)ds\right)\left(\int_0^r B_{2,p}(s)ds\right)dr - M_p \cdot \int_0^1 B_p(r)B_{2,p}(r)dr\right\}$$

Which finally gives:

$$\begin{aligned}M_p(\alpha) &= M_p + 2\alpha M_{2,p}^{-1} \\ &\times\left\{K^{-1}\int_0^1\left(\int_0^r B_p(s)ds\right)\left(\int_0^r B_{2,p}(s)ds\right)dr - M_p \cdot \int_0^1 B_p(r)B_{2,p}(r)dr\right\} + O_p(\alpha^2)\end{aligned}$$

To this order of magnitude (i.e., where terms in α^2 and higher powers are ignored), the second expression above determine the power distortion of the KPSS tests for small α under the weak BLUR alternative. Given that the expected value of the elements of the two integrals between brackets is not zero, there is both a scale shift and a displacement of the distribution to the left.

Appendix C. Proof of Proposition 7.4.3

From the OLS residual sequence in (2.4), we have that its first difference is given by:

$$\tilde{\eta}_{t,p} = \Delta\hat{\eta}_{t,p} = \tilde{\eta}_t - n^{-v}\Delta\boldsymbol{\tau}_p'(\tfrac{t}{n})\cdot[n^v\boldsymbol{\Gamma}_n^{-1}(\hat{\boldsymbol{\beta}}_{p,n} - \boldsymbol{\beta}_p)] \tag{E.1}$$

Where $\tilde{\eta}_t = \Delta\eta_t$. Thus, the scaled partial sum process of this first differenced residuals is given by:

$$n^{-1/2}\sum_{t=1}^{[nr]}\tilde{\eta}_{t,p} = \hat{H}_{\alpha,[nr],p}(r) - n^{-(v+1/2)}\boldsymbol{\tau}_p'(\tfrac{[nr]}{n})[n^v\boldsymbol{\Gamma}_{p,n}^{-1}(\hat{\boldsymbol{\beta}}_{p,n} - \boldsymbol{\beta}_p)] \tag{E.2}$$

With $n^{-1/2}\sum_{t=1}^{[nr]}\tilde{\eta}_{t,p} \Rightarrow Q_{\alpha,p}(r)$, defined in (3.7), under the weak BLUR assumption (that is, with $v = -1/2$ and $\boldsymbol{\theta}_n = (1, \alpha_n)'$). Also, in (E.1), with $v = -1/2$ we have that $\Delta\boldsymbol{\tau}_p'(\tfrac{t}{n}) = O(n^{-1})$, so that $\tilde{\eta}_{t,p} = \tilde{\eta}_t + O_p(n^{-1/2})$ and thus:

$$\tilde{\gamma}_n(h) = n^{-1}\sum_{t=h+1}^{n}\tilde{\eta}_{t,p}\tilde{\eta}_{t-h,p} = n^{-1}\sum_{t=h+1}^{n}\tilde{\eta}_t\tilde{\eta}_{t-h} + O_p(n^{-1/2}) \tag{E.3}$$

Where $\tilde{\eta}_t = (\alpha_t(\boldsymbol{\theta}_n) - 1)\eta_{t-1} + \varepsilon_t$. As an alternative to the use of first differences of OLS residuals from the original auxiliary regression in levels, (E.1), we can make use of the auxiliary regression based on first differences, that is:

$$\Delta Y_t = \boldsymbol{\tau}_{t,p-1}'\boldsymbol{\phi}_{p-1} + \tilde{\eta}_t \tag{E.4}$$

With OLS residuals given by:

$$\hat{\eta}_{t,p-1} = \Delta Y_t - \boldsymbol{\tau}_{t,p-1}'\hat{\boldsymbol{\phi}}_{p-1,n} = \tilde{\eta}_t - n^{-v}\boldsymbol{\tau}_{p-1}'(\tfrac{t}{n})\cdot[n^v\boldsymbol{\Gamma}_{p-1,n}^{-1}(\hat{\boldsymbol{\phi}}_{p-1,n} - \boldsymbol{\phi}_{p-1})] \tag{E.5}$$

With:

$$n^v\boldsymbol{\Gamma}_{p-1,n}^{-1}(\hat{\boldsymbol{\phi}}_{p-1,n} - \boldsymbol{\phi}_{p-1}) = \left(\frac{1}{n}\sum_{j=1}^{n}\boldsymbol{\tau}_{p-1}(\tfrac{j}{n})\boldsymbol{\tau}_{p-1}'(\tfrac{j}{n})\right)^{-1} n^{-(1-v)}\sum_{j=1}^{n}\boldsymbol{\tau}_{p-1}(\tfrac{j}{n})\tilde{\eta}_j \tag{E.6}$$

Taking now $v = 1/2$, and under the assumption of a weak bilinear unit root process, we have the following distribution limit for (E.6):

$$n^{-1/2}\boldsymbol{\Gamma}_{p-1,n}^{-1}(\hat{\boldsymbol{\phi}}_{p-1,n} - \boldsymbol{\phi}_{p-1}) \Rightarrow \left(\int_0^1\boldsymbol{\tau}_{p-1}(s)\boldsymbol{\tau}_{p-1}'(s)ds\right)^{-1}\int_0^1\boldsymbol{\tau}_{p-1}(s)dH_\alpha(s) \tag{E.7}$$

So that the scaled partial sum of these OLS residuals is given by:

$$n^{-1/2}\sum_{t=1}^{[nr]}\hat{\eta}_{t,p-1} = \hat{H}_{\alpha,n}(r) - n^{-(v+1/2)}\sum_{t=1}^{[nr]}\boldsymbol{\tau}_{p-1}'(\tfrac{t}{n})[n^v\boldsymbol{\Gamma}_{p-1,n}^{-1}(\hat{\boldsymbol{\phi}}_{p-1,n} - \boldsymbol{\phi}_{p-1})] \tag{E.8}$$

With weak limit given by $n^{-1/2}\sum_{t=1}^{[nr]}\hat{\eta}_{t,p-1} \Rightarrow B_{\alpha,p-1}(r)$, where $B_{\alpha,p-1}(r)$ is a (p–1)th-level $H_\alpha(r)$ process given by:

$$B_{\alpha,p-1}(r) = H_\alpha(r) - \int_0^r\boldsymbol{\tau}_{p-1}'(s)ds\left(\int_0^1\boldsymbol{\tau}_{p-1}(s)\boldsymbol{\tau}_{p-1}'(s)ds\right)^{-1}\int_0^1\boldsymbol{\tau}_{p-1}(s)dH_\alpha(s) \tag{E.9}$$

Also, from the fact that $\boldsymbol{\tau}'_{p-1}(\frac{t}{n}) = O(1)$, then $\hat{\eta}_{t,p-1} = \tilde{\eta}_t + O_p(n^{-1/2})$ as $n\to\infty$, and:

$$\hat{\gamma}_n(h) = n^{-1}\sum_{t=h+1}^{n}\hat{\eta}_{t,p-1}\hat{\eta}_{t-h,p-1} = n^{-1}\sum_{t=h+1}^{n}\tilde{\eta}_t\tilde{\eta}_{t-h} + O_p(n^{-1/2}) \tag{E.10}$$

As in (E.3), which gives:

$$\begin{aligned}\frac{1}{n}\sum_{t=h+1}^{n}\tilde{\eta}_{t,n}\tilde{\eta}_{t-h,n} &= \frac{1}{n}\sum_{t=h+1}^{n}\left(\alpha H_{n,\alpha}(\tfrac{t-1}{n})\varepsilon_{t-1} + \varepsilon_t\right)\left(\alpha H_{n,\alpha}(\tfrac{t-h-1}{n})\varepsilon_{t-h-1} + \varepsilon_{t-h}\right)\\ &= \alpha^2\frac{1}{\sqrt{n}}\left\{\frac{1}{\sqrt{n}}\sum_{t=h+1}^{n}H_{n,\alpha}(\tfrac{t-1}{n})H_{n,\alpha}(\tfrac{t-h-1}{n})\varepsilon_{t-1}\varepsilon_{t-h-1}\right\}\\ &\quad+\alpha\frac{1}{n}\sum_{t=h+1}^{n}\left(H_{n,\alpha}(\tfrac{t-1}{n})\varepsilon_{t-1}\varepsilon_{t-h} + H_{n,\alpha}(\tfrac{t-h-1}{n})\varepsilon_{t-h-1}\varepsilon_t\right) + \frac{1}{n}\sum_{t=h+1}^{n}\varepsilon_t\varepsilon_{t-h}\end{aligned} \tag{E.11}$$

From the recursive relation $H_{n,\alpha}(\frac{t-1}{n}) = H_{n,\alpha}(\frac{t}{n}) + O_p(n^{-1/2})$, we then have that $H_{n,\alpha}(\frac{t-h-1}{n}) = H_{n,\alpha}(\frac{t-1}{n}) + O_p(n^{-1/2})$ for any $h \geq 1$, so that the term between brackets can be written as:

$$\frac{1}{\sqrt{n}}\sum_{t=h+1}^{n}H_{n,\alpha}(\tfrac{t-1}{n})H_{n,\alpha}(\tfrac{t-h-1}{n})\varepsilon_{t-1}\varepsilon_{t-h-1} = \frac{1}{\sqrt{n}}\sum_{t=h+1}^{n}H^2_{n,\alpha}(\tfrac{t-1}{n})\varepsilon_{t-1}\varepsilon_{t-h-1} + O_p(n^{-1/2}) \tag{E.12}$$

Which is $O_p(1)$, so that the first summand term in (E.11) is $O_p(n^{-1/2})$. Also, for the second summand term we have that:

$$\alpha\frac{1}{n}\sum_{t=h+1}^{n}H_{n,\alpha}(\tfrac{t-h-1}{n})\varepsilon_{t-h-1}\varepsilon_t = \alpha\frac{1}{\sqrt{n}}\left\{\frac{1}{\sqrt{n}}\sum_{t=h+1}^{n}H_{n,\alpha}(\tfrac{t-h-1}{n})\varepsilon_{t-h-1}\varepsilon_t\right\} \tag{E.13}$$

Where the stochastic limit of the term between brackets is finite and thus it is again of order $O_p(n^{-1/2})$. The last term in (E.11) is the usual h-lag order sample covariance of the error sequence ε_t, which by standard application of the WLLN under stationarity and weak dependence gives $n^{-1}\sum_{t=h+1}^{n}\varepsilon_t\varepsilon_{t-h} \to^p \gamma_\varepsilon(h) = E[\varepsilon_t\varepsilon_{t-h}]$. Finally, for the first term in the second summand of (E.11) we have the following two possible situations. When $h > 1$, it is of application the same argument as before in (E.13), that is:

$$\alpha\frac{1}{n}\sum_{t=h+1}^{n}H_{n,\alpha}(\tfrac{t-1}{n})\varepsilon_{t-1}\varepsilon_{t-h} = \alpha\frac{1}{\sqrt{n}}\left\{\frac{1}{\sqrt{n}}\sum_{t=h+1}^{n}H_{n,\alpha}(\tfrac{t-1}{n})\varepsilon_{t-1}\varepsilon_{t-h}\right\} = O_p(n^{-1/2}) \tag{E.14}$$

While that for $h = 1$, then it can be written as:

$$\begin{aligned}\alpha\frac{1}{n}\sum_{t=h+1}^{n}H_{n,\alpha}(\tfrac{t-1}{n})\varepsilon^2_{t-1} &= \alpha\frac{1}{n}\sum_{t=h+1}^{n}H_{n,\alpha}(\tfrac{t-1}{n})[(\varepsilon^2_{t-1} - \sigma^2_\varepsilon) + \sigma^2_\varepsilon]\\ &= \alpha\frac{1}{\sqrt{n}}\left\{\frac{1}{\sqrt{n}}\sum_{t=h+1}^{n}H_{n,\alpha}(\tfrac{t-1}{n})(\varepsilon^2_{t-1} - \sigma^2_\varepsilon)\right\} + \alpha\sigma^2_\varepsilon\frac{1}{n}\sum_{t=h+1}^{n}H_{n,\alpha}(\tfrac{t-1}{n})\end{aligned}$$

By the stationarity of the sequence $(\varepsilon^2_{t-h} - \sigma^2_\varepsilon)$, $h \geq 0$, and under the condition of existence of the fourth moment $E[\varepsilon^4_t] < \infty$, then:

$$\frac{1}{\sqrt{n}}\sum_{t=h+1}^{[nr]}(\varepsilon^2_t - \sigma^2_\varepsilon) \Rightarrow B_2(r) = \kappa_2\cdot W(r)$$

With $\kappa_\varepsilon^2 = E[(\varepsilon_t^2 - \sigma_\varepsilon^2)^2]$. With this, the term between brackets is $O_p(1)$ and thus we have:

$$\alpha \frac{1}{n}\sum_{t=h+1}^{n} H_{n,\alpha}(\tfrac{t-1}{n})\varepsilon_{t-1}^2 = \alpha\sigma_\varepsilon^2 \frac{1}{n}\sum_{t=h+1}^{n} H_{n,\alpha}(\tfrac{t-1}{n}) + O_p(n^{-1/2}) \Rightarrow \alpha\sigma_\varepsilon^2 \int_0^1 H_\alpha(s)ds$$

Then, putting together all these results we have $\hat{\gamma}_n(h), \tilde{\gamma}_n(h) \to^p \gamma(h)$, $h > 1$, and:

$$\hat{\gamma}_n(h), \tilde{\gamma}_n(h) \Rightarrow \gamma(h) + \alpha\sigma_\varepsilon^2 \int_0^1 H_\alpha(s)ds \quad h = 1 \tag{E.15}$$

With $\gamma(h) = 0$ for all $h \geq 1$ under serially uncorrelated error terms. For the particular case where $h = 0$ we have that:

$$\begin{aligned}\frac{1}{n}\sum_{t=1}^{n}\tilde{\eta}_{t,n}^2 &= \frac{1}{n}\sum_{t=1}^{n}\left(\alpha H_{n,\alpha}(\tfrac{t-1}{n})\varepsilon_{t-1} + \varepsilon_t\right)^2 \\ &= \alpha^2 \frac{1}{\sqrt{n}}\left\{\frac{1}{\sqrt{n}}\sum_{t=1}^{n} H_{n,\alpha}^2(\tfrac{t-1}{n})(\varepsilon_{t-1}^2 - \sigma_\varepsilon^2)\right\} + \alpha^2\sigma_\varepsilon^2 \frac{1}{n}\sum_{t=1}^{n} H_{n,\alpha}^2(\tfrac{t-1}{n}) \\ &+ \frac{1}{n}\sum_{t=1}^{n}\varepsilon_t^2 + 2\alpha \frac{1}{\sqrt{n}}\left\{\frac{1}{\sqrt{n}}\sum_{t=1}^{n} H_{n,\alpha}(\tfrac{t-1}{n})\varepsilon_t\varepsilon_{t-1}\right\}\end{aligned}$$

Which gives:

$$\hat{\gamma}_n(0), \tilde{\gamma}_n(0) \Rightarrow \sigma_\varepsilon^2\left(1 + \alpha^2\int_0^1 H_\alpha(s)^2 ds\right) \tag{E.16}$$

Using some of the above results concerning the probability order of magnitude of the terms between brackets.

References

Afonso-Rodríguez, J.A. (2012a). On the analysis of joint confirmation of a unit root under a weak bilinear unit root process. Mimeo. University of La Laguna.

Afonso-Rodríguez, J.A. (2012b). On the behaviour of stationarity tests under a stochastic unit root (STUR) alternative. Mimeo. University of La Laguna.

Berenguer-Rico, V. (2011). Co-summability. From linear to non-linear co-integration. PH-D Thesis, Universidad Carlos III de Madrid.

Breitung, J. (2002). Nonparametric tests for unit roots and cointegration. Journal of Econometrics, 108(2), 343-363.

Breitung, J., A.M.R. Taylor (2003). Corrigendum to "Nonparametric tests for unit roots and cointegration". Journal of Econometrics, 117(2), 401-404.

Carrion-i-Silvestre, J.L., A. Sansó (2006). A guide to the computation of stationarity tests. Empirical Economics, 31, 433-448.

Charemza, W.W., M. Lifshifts, S. Makarova (2005). Conditional testing for unit-root bilinearity in financial time series: some theoretical and empirical results. Journal of Economic Dynamics & Control, 29(1-2), 63-96.

Francq, C., S. Makarova, J.M. Zakoian (2008). A class of stochastic unit-root bilinear processes: mixing properties and unit-root tests. Journal of Econometrics, 142(1), 312-326.

Giraitis, L.P., R.L. Kokoszka, G. Teyssiere (2003). Rescaled variance and related tests for long memory in volatility and levels. Journal of Econometrics, 112(2), 265-294.

Gonzalo, J., J.Y. Pitarakis (2006). Threshold effects in cointegrating relationships. Oxford Bulletin of Economics and Statistics, 68(1), 813-833.

Gonzalo,J., T.H. Lee (1998). Pitfalls in testing for long run relationships. Journal of Econometrics, 86(1), 129-154.

Granger, C.W.J., N.R. Swanson (1997). An introduction to stochastic unit-root processes. Journal of Econometrics, 80(1), 35-62.

Hristova, D. (2005). Maximum likelihood estimation of a unit root bilinear model with an application to prices. Studies in Nonlinear Dynamics and Econometrics, 9(1), 1-13.

Ibragimov, R., P.C.B. Phillips (2008). Regression asymptotics using martingale convergence methods. Econometric Theory, 24(4), 888-947.

Kurozumi, E. (2002). Testing for stationarity with a break. Journal of Econometrics, 108(1), 63-99.

Kwiatkowski, D., P.C.B. Phillips, P. Schmidt, Y. Shin (1992). Testing the null hypothesis of stationarity against the alternative of unit root. Journal of Econometrics, 54(1-3), 159-178.

Leybourne, S., B.P.M. McCabe, A.R. Tremayne (1996). Can economic time series be differenced to stationarity?. Journal of Business and Economic Statistics, 14(4), 435-446.

Leybourne, S., B.P.M. McCabe, T.C. Mills (1996). Randomized unit root processes for modelling and forecasting financial time series. Journal of Forecasting, 15(3), 253-270.

Lifshits, M.A. (2006). Invariance principle in a bilinear model with weak nonlinearity. Journal of Mathematical Sciences, 137(1) (4541-4545).

McCabe, B.P.M., A.R. Tremayne (1995). Testing a time series for difference stationarity. The Annals of Statistics, 23(3), 1015-1028.

McCabe, B.P.M., R.J. Smith (1998). The power of some tests for difference stationarity under local heteroscedastic integration. Journal of the American Statistical Association, 93(442), 751-761.

McCabe,B.P.M., R.J. Smith (1998). The power of some tests for difference stationarity under local heteroscedastic integration. Journal of the American Statistical Association, 93(442),751-761.

McLeish, D.L. (1975). Invariance principles for dependent variables. Probability Theory and Related Fields, 32(3), 165-178.

Nagakura, D. (2009). Asymptotic theory for explosive random coefficient autoregressive models and inconsistency of a unit root test against a stochastic unit root process. Statistics & Probability Letters, 79(24), 2476-2483.

Phillips, P.C.B. (1987). Time series regression with a unit root. Econometrica, 55(2), 277-301.

Phillips, P.C.B., V. Solo (1992). Asymptotics for linear processes. The Annals of Statistics, 20(2), 971-1001.

Tabak, B.M. (2007). Testing for unit root bilinearity in the Brazilian stock market. Physica A, 385, 261-269.

Xiao, Z. (2001). Testing the null hypothesis of stationarity against an autoregressive unit root alternative. Journal of Time Series Analysis, 22(1), 87-103.

Xiao, Z., L.R. Lima (2007). Testing covariance stationarity. Econometric Reviews, 26(6), 643-667.

Yoon, G. (2004). A re-evaluation of the performance of the unit root and cointegration tests under STUR: no more pitfalls. Department of Economics, Pusan National University, Pusan, S. Korea, preprint.