Power properties of a nonparametric test of the martingale hypothesis

Julio A. Afonso Rodríguez
Department of Applied Economics and Quantitative Methods
Universidad de La Laguna

Abstract

In this paper we study the power properties of a simple nonparametric test of the martingale hypothesis recently proposed by Park and Whang (2005), and further studied by Escanciano (2007). Particularly, we are interested in obtaining the limiting distribution of this test statistic for several possible non-martingale non-stationary processes, such as the class of stochastic unit root processes, with particular attention to the so-called local-heteroskedastic integrated and weak bilinear unit root processes introduced by McCabe and Smith (1998) and Charemza et.al. (2005) and Lifshits (2006), respectively. We also consider the class of heteroskedastic (or stochastically) integrated processes introduced by Hansen (1992) and further generalized by Harris et.al. (2002). We also conduct a simulation experiment to evaluate numerically the power behaviour of this test statistic for each of these closely related, but different, alternatives.

Key words and phrases: martingale hypothesis, near unit roots, stochastic unit roots, heteroskedastic integration

JEL Classification: C12, C14, C22

1. Introduction

One of the oldest, and more frequently analyzed, questions in economics and finance is that of the predictability of the asset prices. The formal treatment of this question has resulted in an equivalent statement known as the efficient market hypothesis which states that no systematic trading strategy that exploits the conditional mean dynamics of the observed prices can be more profitable in the long run than holding the market portfolio. Also, from the time series econometrics point of view, there exist subtle differences in the formulation of the hypothesis of non predictability when dealing with the increments of the price process. Hence, we get the connection of the martingale theory, the random walk paradigm in economic time series, and the unit root theory. In this paper we study the behavior and properties of a simple nonparametric test of the null hypothesis of a martingale process with mean independent increments against two different non-martingale nonstationary processes that could arise as plausible alternatives for describing economic and financial time series. The test statistic we examine is the one proposed by Park and Whang (2005) and the alternatives considered are two different versions of the so-called stochastic unit root (STUR) process, and the stochastically integrated (SI) process which represents a heteroskedastic integrated version of a standard integrated process.

The paper is organised as follows. Section 2 is devoted to reviewing the martingale hypothesis and its implications, and the testing procedures we study. Section 3 introduces the two types of alternatives mentioned above, their main characteristics and properties and states the main results of the paper concerning the limiting behavior of the test statistics for the null of a martingale under these nonstationary alternatives. Theoretical results are accompanied by some numerical results obtained through a small Monte Carlo experiment for assessing its finite-sample power performance.

2. The martingale hypothesis and the test statistics

The concept of martingale is closely related to that of a random walk, when making particular assumptions on the dependence structure of the increment series. Particularly, assuming a random walk with mean-independent increments, which is a stronger condition that uncorrelated increments, implies the martingale hypothesis

$$E_{t-1}[X_t] = E[X_t | I_{t-1}] = X_{t-1} \text{ a.s.} \qquad (2.1)$$

with $I_t = \{X_t, X_{t-1}, ...\}$ the information set at time *t*, and F_t be the σ-field generated by I_t Alternatively, using the first differences, $u_t = \Delta X_t = X_t - X_{t-1}$, u_t follows a martingale difference sequence (mds) when $E_{t-1}[u_t] = E[u_t | I_{t-1}] = 0$ a.s., which means that past and current information are of no use to forecasting future values of a mds. However, if u_t is a real-valued stationary time series, the martingale difference hypothesis (mdh) states that the following conditional moment restriction holds

$$E_{t-1}[u_t] = E[u_t | I_{t-1}] = \mu \text{ a.s.}, \mu \in \mathbf{R}$$

The mdh slightly generalizes the notion of mds by allowing the unconditional mean of u_t to be nonzero and unknown, and states that the best predictor, in the sense of least mean square error, of the future values given the past and current information set is just the unconditional expectation, which can be called conditional mean independence. Also, given that the characteristic property of a mds is the fact that u_t is linearly unpredictable given any linear or nonlinear transformation of the past, $w(I_{t-1})$ with

$I_t = \{u_t, u_{t-1}, ...\}$, then we get the following fundamental equivalence

$$E_{t-1}[u_t] = E[u_t | I_{t-1}] = \mu \text{ a.s., } \mu \in \mathbf{R} \quad \Leftrightarrow \quad E[(u_t - \mu) w(I_{t-1})] = 0$$

for all F_{t-1}-measurable weighting function $w(\cdot)$, such that the moment exists. As is discussed in Escanciano and Lobato (2007), this last expression is the key element in developing a great variety of tests for the mdh, depending on the choice of the information set at time t based on the series of first differences, u_{t-j} ($j \geq 0$), and the function $w(\cdot)$ to be used. However, exploiting this relationship, a general correlation test for u_t and the space spanned by X_{t-1} can be constructed as a test for the original null hypothesis of a martingale behavior. To represent the complete space spanned by X_{t-1}, of all linear and nonlinear functions of X_{t-1}, Park and Whang (2005) use the indicator function of X_{t-1}, that is $I(X_{t-1} \leq x)$ for almost all $x \in \mathbf{R}$, giving rise to the moment indicator

$$Q(x) = E[u_t \cdot I(X_{t-1} \leq x)] \tag{2.2}$$

that states the connection between the martingale hypothesis condition and the condition of no correlation, $Q(x) = 0$ for all values of x, through the expression

$$Q(x) = E[E[u_t \cdot I(X_{t-1} \leq x) | X_{t-1}]] = E[E[u_t | X_{t-1}] \cdot I(X_{t-1} \leq x)]$$

Stute (1997), and Koul and Stute (1999) have used this type of measures based on stationary sequences to develop some simple nonparametric statistics to testing for model adequacy in standard regression and time series analysis.

Alternatively, and in order to develop a general theory of testing for a martingale, Durlauf (1991) considers the periodogram-based estimate of the deviations of the spectral distribution function for the process of first differences, u_t, from its theoretical shape when the periodogram is normalized by the sample variance, that is

$$U_n(r) = \frac{\sqrt{2}}{\pi} \sum_{k=1}^{n-1} \frac{\sin(k\pi r)}{k} (n^{1/2} \hat{\rho}_n(k))$$

with $\hat{\rho}_n(k) = \hat{\gamma}_n^{-1}(0) \hat{\gamma}_n(k)$, the k-th order sample autocorrelation coefficient, with

$$\hat{\gamma}_n(k) = (1/n) \sum_{t=k+1}^{n} (u_t - \bar{u}_n)(u_{t-k} - \bar{u}_n)$$

the k-th order sample autocovariance coefficient. The correction for the sample mean extends the analysis to the case where the time series of interest is a random walk with drift in levels. Under the martingale difference hypothesis for u_t, $\hat{\rho}_n(k) \Rightarrow 0$ for all $k \geq 1$, so the analysis will focus on the autocorrelation function in which case can be shown that $U_n(r) \Rightarrow V(r) = B(r) - rB(1)$, with $V(r)$ a standard (first-level) Brownian bridge. Making use of this fundamental result, Durlauf (1991) propose several spectral shape tests (see Corollary 2.1, p.363) based on different global measures of excessive fluctuations in $U_n(r)$. Among all its proposals, the two more frequently used test statistics are those based on the Cramér-von Mises (CvM) and Kolmogorov-Smirnov (KS) measures defined as

$$CvM_n = (1/n) \sum_{k=1}^{n} U_n^2(k/n)$$

and

$$KS_n = \max_{k=1,\ldots,n} | U_n(k/n) |$$

respectively, which usually display good finite sample and asymptotic properties.

Instead of making use of some particular implications of the martingale hypothesis on some usual sample statistics, as the case of the sample serial correlation coefficients, Park and Whang (2005) consider that the random element that form the basis of the test statistics for the martingale hypothesis is defined as the following marked empirical process

$$Q_n(x) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} u_t \cdot I(X_{t-1} \le x) \tag{2.3}$$

with marks given by the random sequence $u_1, ..., u_n$, and jumps at the points $X_0, X_1, ..., X_{n-1}$, while $M_n(x) = Q_n(x\sqrt{n})$ is a stochastic process with parameter $x \in \mathbf{R}$, that takes values in $D(\mathbf{R})$, the space of cadlag functions on **R**, with $M_n(-\infty) = 0$ and $M_n(\infty) = B_n(r) = (1/\sqrt{n}) \sum_{t=1}^{n} u_t$. Given that this stochastic process can also be written as

$$M_n(x) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} u_t \cdot I\left(\frac{X_{t-1}}{\sqrt{n}} \le x \right) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} u_t \cdot I\left(\frac{X_{t-1}}{\sigma_u \sqrt{n}} \le y \right) \tag{2.4}$$

where $y = x/\sigma_u$, with σ_u^2 introduced in Assumption(b) below, it can also be defined the following modified version, scaled by $\sigma_n^2 = (1/n) \sum_{t=1}^{n} u_t^2$ to define a pivotal test statistic as

$$V_n(x) = \sigma_n^{-1} M_n(x) \tag{2.5}$$

Thus, in order to obtain useful limiting results for $V_n(x)$, Park and Whang (2005) introduce the following standard assumption on the sequence of first differences, u_t, that ensures the application of the martingale limit theory.

Assumption 2.1. *Innovation sequence*

Given the filtration F_t introduced earlier, (u_t, F_t) is a mds such that:

(a) $s_n^2 = (1/n) \sum_{t=1}^{n} E[u_t^2 \mid F_{t-1}] \to^p \sigma_u^2 > 0$

(b) $\sup_{t \ge 1} E[u_t^4 \mid F_{t-1}] < \kappa < \infty$ a.s.

Under these conditions, we then have $B_n(r) = (1/\sqrt{n}) \sum_{t=1}^{n} u_t \Rightarrow B(r) = \sigma_u W(r)$, and $\sigma_n^2 \to^p \sigma_u^2$, which gives

$$M_n(x) \Rightarrow \sigma_u M(x)$$

and $V_n(x) \Rightarrow M(x)$, with

$$M(x) = \int_0^1 I(W(r) \le x) dW(r) \tag{2.6}$$

for $x \in \mathbf{R}$, with $M(-\infty) = 0$ and $M(\infty) = W(1)$. With these results, two different types of statistics from $V_n(x)$ may be constructed, namely, the KS-type statistic given by

$$S_n = \sup_{x \in \mathbf{R}} |V_n(x)| \tag{2.7}$$

and the CvM-type statistic defined as

$$T_n = (1/n) \sum_{t=1}^{n} V_n^2(X_{t-1}) \tag{2.8}$$

based on the empirical distribution measure of X_{t-1}, with limiting null distributions given by

$$S_n \Rightarrow S = \sup_{x \in \mathbf{R}} |M(x)| \tag{2.9}$$

and

$$T_n \Rightarrow T = \int_0^1 M^2(W(r))dr \qquad (2.10)$$

respectively, by standard application of the continuous mapping theorem since (2.7) and (2.8) are continuous functionals of $V_n(\cdot)$. The testing procedure based on any of these two statistics rejects the null hypothesis of a martingale for large values, with asymptotic critical values given in Table 1 of Park and Whang (2005), which are shown in the last column of the following two tables, together with the finite-sample critical values computed via direct simulation of S_n and T_n in (2.7) and (2.8) for sample sizes n = 100, 250, 500, and 1000 and 1000 independent replications.

Table 2.1. *Finite-sample and asymptotic quantiles of the null distribution of* S_n

Significance level	$n = 100$	250	500	1000	∞
0.99	0.575	0.595	0.608	0.592	0.612
0.95	0.713	0.740	0.729	0.769	0.765
0.90	0.802	0.844	0.804	0.879	0.865
0.10	2.025	2.086	2.090	2.097	2.119
0.05	2.314	2.334	2.342	2.353	2.388
0.01	2.801	2.779	2.910	3.019	2.911

Table 2.2. *Finite-sample and asymptotic quantiles of the null distribution of* T_n

Significance level	$n = 100$	250	500	1000	∞
0.99	0.052	0.062	0.055	0.060	0.055
0.95	0.090	0.101	0.099	0.105	0.101
0.90	0.133	0.138	0.132	0.161	0.145
0.10	1.635	1.643	1.559	1.687	1.650
0.05	2.089	2.119	2.163	2.165	2.165
0.01	3.073	2.814	3.331	3.600	3.328

From these results, it is quite remarkable the stability of the null distributions even for very small sample sizes, which also represents a notable advantage for the use of these test statistics in the sense that excessive random fluctuations of these measures must be mainly due to the violation of the null hypothesis.

The testing procedures proposed by Park and Whang (2005) have the attractive property of being quite simple to compute, compared to similar tests of Durlauf (1991). Durlauf's tests based on sample correlations are tests of the shape of the spectral density function, which is rectangular under the null hypothesis of no correlation. The simulation results in Park and Whang (2005) show that this martingale test is much more powerful against a wide range of non-martingale alternatives, not only the linear non-Gaussian cases that are covered by Durlauf's tests, but also the nonlinear non-Gaussian alternatives.[1]

Also, when considering the application of the Durlauf's tests to economic or financial time series, usually characterized by certain type of conditional heteroskedastic behavior, Deo (2000) shows that the null limiting distribution of $U_n(r)$ is not longer a Brownian bridge but another Gaussian process that depends on the dependence structure of the second moments of the series. The two cases considered are the two major models of conditionally heteroskedastic martingale differences, viz. the stochastic volatility (SV) and the generalized autoregressive conditionally heteroskedastic (GARCH) model. In order to preserve the usual limiting distribution, Deo (2000) propose to introduce a particular nonparametric correction of $U_n(r)$ that requires, under

[1] It is also worth to mention the work by Kumagai (2001) where, based on an earliest version of the paper (viz. Park and Whang (1999)), it is proposed a unit-root type specification test for stochastic processes generated by linear functions of nonstationary integrated processes, in order to test the random walk hypothesis in asset prices.

the null, the finiteness of eight moments for the result to hold. This technical requirement seems to be too strong and unrealistic in practice for many conditionally heteroskedastic series.

However, for the testing procedures based on the marked empirical process $Q_n(x)$, these situations do not provoke any modification of the limiting results obtained under conditional homoskedasticity, given the robustness of the invariance principles to conditionally heteroskedastic martingale difference processes.[2]

3. Power under non-martingale nonstationary alternatives

The class of first-order random coefficient autoregressive process, RCA(1), can serve as the generating mechanism of a wide range of very different behaviours of a time series, including stationarity, non-stationarity and even explosive episodes. In what follows, we assume that the data generating process (DGP) of a univariate observed time series is given by

$$X_t = d_t + \eta_t \quad t = 1, ..., n \tag{3.1}$$

where it is decomposed into d_t and η_t, the deterministic and stochastic trend components, respectively. The leading cases for d_t are the no deterministics component, $d_t = 0$, the constant case, $d_t = \alpha_0$, and the constant and linear trend case, $d_t = \alpha_0 + \alpha_1 t$. All these cases can be formulated as particular cases of the p-th order, $p \geq 0$, polynomial trend function given by $d_t = \boldsymbol{\alpha}_p' \boldsymbol{\tau}_{p,t}$, with $\boldsymbol{\alpha}_p = (\alpha_0, \alpha_1, ..., \alpha_p)'$ and $\boldsymbol{\tau}_{p,t} = (1, t, ..., t^p)'$. The case of no deterministics is simply given by $\boldsymbol{\alpha}_p = \mathbf{0}_{p+1}$. In what follows we assume this last situation, and the extension to the case of a non-martingale deterministically integrated process X_t will be considered in a further extension of the present work.[3] For the stochastic component we assume that it is given by a RCA(1) process of the form

$$\eta_t = \alpha_t \eta_{t-1} + \varepsilon_t \tag{3.2}$$

with

[2] See, e.g., Hansen (1991) for the proof of the near epoch dependence property of a GARCH(1,1) process without imposing strict stationarity, and Kim, Cho and Lee (2000) for the proof of the mixingale property of a covariance stationary GARCH(1,1) process with finite four moment. These two properties ensure the application of standard asymptotic theory as, for example, the required invariance principles. For a more recent and general treatment of this questions see Kulperger and Yu (2005) for the general case of a covariance stationary GARCH(p,q) and also the contributions by Carrasco and Chen (2002) and Hill (2009) for more general cases, including stochastic volatility models.

[3] From (2.1), we have that the marked empirical process $M_n(x)$ can be decomposed as

$$M_n(x) = (1/\sqrt{n})\sum_{t=1}^{n} \Delta d_t I(\eta_{t-1}/\sqrt{n} \leq x - d_{t-1}/\sqrt{n}) + (1/\sqrt{n})\sum_{t=1}^{n} \Delta\eta_t I(\eta_{t-1}/\sqrt{n} \leq x - d_{t-1}/\sqrt{n})$$

with $M_n(x) = (1/\sqrt{n})\sum_{t=1}^{n} \Delta\eta_t I(\frac{\eta_{t-1}}{\sqrt{n}} \leq x - \frac{\alpha_0}{\sqrt{n}})$ in the case of a constant term, $d_t = \alpha_0$, and

$$M_n(x) = \alpha_1 \sqrt{n}\left\{(1/n)\sum_{t=1}^{n} I(\tfrac{\eta_{t-1}}{\sqrt{n}} \leq x - [\tfrac{\alpha_0 - \alpha_1}{\sqrt{n}} + \alpha_1 \tfrac{t}{\sqrt{n}}])\right\} + (1/\sqrt{n})\sum_{t=1}^{n} \Delta\eta_t I(\tfrac{\eta_{t-1}}{\sqrt{n}} \leq x - [\tfrac{\alpha_0 - \alpha_1}{\sqrt{n}} + \alpha_1 \tfrac{t}{\sqrt{n}}])$$

in the case of a constant term and a linear trend, $d_t = \alpha_0 + \alpha_1 t$, with the terms $\alpha_0/\sqrt{n}$ and $(\alpha_0 - \alpha_1)/\sqrt{n}$ asymptotically negligible, but with the term $\alpha_1(\frac{t}{\sqrt{n}}) = \alpha_1\sqrt{n}(\frac{t}{n}) = O(\sqrt{n})$ in the last case. Thus, except in the first situation, the constant term case, and in large samples, it may be expected some nonnegligible effects due to these components.

$$\alpha_t = \phi + \phi_t \tag{3.3}$$

that nests a great variety of stationary and nonstationary processes with very different stochastic properties depending on the particular assumptions made on the sequences ε_t and ϕ_t. In this paper we are particularly interested in the class of the so-called stochastic unit root (STUR), or randomized unit root, processes, for which the nondegenerate random coefficient α_t has expected value one, $E[\alpha_t]=1$,[4] and that is given by a recursive combination of a random walk and a noise process. This type of processes, having a root that is not constant, but stochastic, and varying around unity, represent a quite natural extension from the perfect or fixed unit root case when $Var[\alpha_t]=0$ (which is denoted by I(1)), and its sample paths can alternate between some periods of stationarity and a mildly explosive behaviour. Since its introduction by McCabe and Tremayne (1995), Leybourne, McCabe and Tremayne (1996), Leybourne, McCabe and Mills (1996), and Granger and Swanson (1997), this type of processes has received a great attention, both at the theoretical and the empirical level of analysis, given the difficulties in distinguishing between exact and stochastic unit roots and the fact that stochastic unit roots can arise naturally in economic and financial theory.

An interesting property of (3.2) is that, under quite general conditions on ϕ_t and ε_t such as weak exogeneity with $Cov[\alpha_t,\eta_{t-1}]=E[\varepsilon_t\eta_{t-1}]=0$ and contemporaneously uncorrelated $Cov[\alpha_t,\varepsilon_t]=0$,[5] the process η_t exhibits conditional heteroskedasticity, with $E_{t-1}[\eta_t]=E[\eta_t|\eta_{t-1}]=\phi\eta_{t-1}$, and $Var_{t-1}[\eta_t]=Var[\eta_t|\eta_{t-1}]=Var[\phi_t]\eta_{t-1}^2+\sigma_\varepsilon^2$. This makes this type of processes plausible to capture some of the most commonly known stylized facts of many economic and financial time series. Also, with ϕ_t an iid sequence of random variables with zero mean and $Var[\phi_t]=\sigma_\phi^2$, the main stochastic properties of η_t in (3.2)-(3.3) can be determined by means of two criticality parameters associated to the random autoregressive coefficient α_t, namely $\tau=(\phi^2+\sigma_\phi^2)^{1/2}\,\mathrm{sgn}(\phi)$, where $\mathrm{sgn}(\phi)$ takes values 1 and −1 according to $\phi \geq 0$ and $\phi < 0$ respectively, and the top Lyapunov exponent defined as $\lambda = E[\log|\phi+\phi_t|]$. Note that $\tau^2=E[\alpha_t^2]=\phi^2+\sigma_\phi^2$, and $\lambda \leq (1/m)\log(E[|\phi+\phi_t|^m])$ for any $m > 0$ by Jensen's inequality, and $\lambda \leq \log(\tau^2)$ for m = 2. It follows from Nicholls and Quinn (1982)[6] that (3.2) has a strictly stationary

[4] Granger and Swanson (1997) consider a different representation for this type of processes, that can be called a exponential stochastic unit root, where $\alpha_t=\exp(\phi_t)$, and $\phi_t=\rho_0+\rho\phi_{t-1}+u_t$ is a stable Gaussian AR(1) process, $|\rho| < 1$, with $u_t \sim iidN(0,\sigma_u^2)$. In this setup, $E[\alpha_t]=1$ only if σ_u^2 if fixed at the value $\sigma_u^2=-2\rho_0(1+\rho)$, when $\alpha_0<0$ and $0<\alpha<1$, or $\alpha_0>0$ and $-1<\alpha<0$, that gives $Var[\alpha_t]=$ $\exp(-2\rho_0/(1-\rho))-1>0$. Also, taking $\mu_\phi=E[\phi_t]=\rho_0/(1-\rho)$, we have that a second order Taylor expansion of the exponential function around μ_ϕ gives $\alpha_t=\exp(\mu_\phi)(1-\mu_\phi)+\exp(\mu_\phi)\phi_t+O_p(\phi_t^2)$, so that this process cannot be written as in (3.3). If instead, simply assuming that ϕ_t is a zero mean stationary sequence, as in McCabe and Smith (1998), then we get $\alpha_t=1+\phi_t+O_p(\phi_t^2)$.

[5] McCabe and Tremayne (1995) assume that the processes α_t and ε_t are mutually dependent with contemporaneous covariance $Cov[\alpha_t,\varepsilon_t]=\psi\sqrt{Var[\alpha_t]}$ and serially independent, with $|\psi|<\sqrt{Var[\varepsilon_t]}$, so that these two processes are not contemporaneously linearly dependent.

[6] For a more recent study on the general stochastic properties of a RCA(1) process see, e.g., Aue, Horváth and Steinebach (2006), and Berkes, Horváth, and Ling (2009), and their results related to the quasi-maximum likelihood estimation of the model parameters under a variety of assumptions concerning the

solution with finite second-order stationary moment as well as almost sure convergence if and only if $\tau^2 < 1$. The solution is given by $\eta_t = \sum_{k=0}^{\infty} \pi_k \varepsilon_{t-k}$, with $\pi_0 = 1$ and $\pi_k = \Pi_{j=0}^{k-1} \alpha_{t-j} = \Pi_{j=0}^{k-1} (\phi + \phi_{t-j})$, k = 1, 2, … However, for $\tau^2 \geq 1$ the above solution for η_t continues to be a strictly stationary (and ergodic) solution to (3.2) if $\lambda < 0$, and only if $\lambda \leq 0$, but with a infinite second moment. Thus, $Var[\eta_t] = \sigma_\varepsilon^2 (\tau^{2n} - 1)(\tau^2 - 1)^{-1}$ for $\tau \neq 1$, and $Var[\eta_t] = n\sigma_\varepsilon^2$ for $\tau = 1$. Note that this variance increases exponentially for $\tau^2 > 1$, and increases linearly for $\tau^2 = 1$. Nagakura (2009) develops the relevant asymptotic theory for explosive RCA(1) models (denoted ERCA(1)), although these results seems to be of limited practical relevance given that this explosive behavior is rarely observed for actual time series data.

Recently, McCabe, Martin and Tremayne (2005) consider the general representation of some stationary and nonstationary models commonly adopted in the economics and finance literature making use of the unified framework proposed by Cramer (1961) for purely non-deterministic processes, to study their persistence properties in terms of the behaviour of traditional measures of persistence. The range of models considered extend the I(0)/I(1) paradigm including a nonstationary fractionally integrated noise model and other two nonlinear nonstationary models that do not lie into the I(d) class, the STUR model (3.2)-(3.3) and the stochastically (or heteroskedastic) integrated (SI) model, which is a multiplicative combination of a random walk and a noise process, together with an additive error term.

The so-called stochastically integrated (SI) process is a generalized version of the local level model based on unobserved components, with the addition of a third component with high, but limited, persistence that represents an intermediate case between stationarity and nonstationarity. The original idea comes from the heteroskedastic cointegrating regression equation proposed by Hansen (1992), where the error term follows a bi-integrated process that resembles the nonstationary heteroskedastic behavior of the integrated regressors. This original proposal was later extended by Harris et.al. (2002, 2003), and McCabe et.al. (2003), where McCabe et.al. (2006) give the more generalized version of this model given by

$$\eta_t = \boldsymbol{\pi}'_m \mathbf{w}_{m,t} + \varepsilon_t + \mathbf{v}'_{q,t} \mathbf{h}_{q,t} \tag{3.4}$$

with the m, and q-dimensional nonstationary vectors $\mathbf{w}_{m,t}$ and $\mathbf{h}_{q,t}$ given by

$$\mathbf{w}_{m,t} = \mathbf{w}_{m,t-1} + \boldsymbol{\upsilon}_{m,t} = \mathbf{w}_{m,0} + \sum_{j=1}^{t} \boldsymbol{\upsilon}_{m,j} \tag{3.5}$$

and

$$\mathbf{h}_{q,t} = \mathbf{h}_{q,t-1} + \boldsymbol{\xi}_{q,t} = \mathbf{h}_{q,0} + \sum_{j=1}^{t} \boldsymbol{\xi}_{q,j} \tag{3.6}$$

and where the zero mean $2q$+m+1-dimensional error vector $\boldsymbol{\zeta}_t = (\varepsilon_t, \boldsymbol{\upsilon}'_{m,t}, \mathbf{v}'_{q,t}, \boldsymbol{\xi}'_{q,t})'$ is assumed to be stationary. When $\boldsymbol{\pi}_m \neq \mathbf{0}_m$, then η_t is said to be *stochastically integrated*, SI. If, in addition, $E[\mathbf{v}'_{q,t}\mathbf{v}_{q,t}] > 0$, η_t is said to be *heteroskedastically integrated* (HI) due to the term $\mathbf{v}'_{q,t}\mathbf{h}_{q,t}$, whereas if $E[\mathbf{v}'_{q,t}\mathbf{v}_{q,t}] = 0$ (or, equivalently $Var[\boldsymbol{\xi}_{q,t}] = \mathbf{0}_{q,q}$ when $\mathbf{w}_{m,t} \neq \mathbf{h}_{q,t}$) then η_t is simply difference stationary, or I(1). So, a stochastically integrated process encompasses both ordinary and heteroskedastic integration. The

stability of the model. Hwang and Basawa (2005) also studied the consistency properties of the least squares (LS) and weighted LS estimates of the autoregressive parameter in the explosive RCA(1) model.

model considered in Harris, et.al. (2002, 2003), and in McCabe et.al. (2003) is given by (3.4)-(3.6) with $m = q$, so that $\mathbf{w}_{m,t} = \mathbf{h}_{m,t}$, and thus $\eta_t = \boldsymbol{\pi}_{m,t}\mathbf{w}_{m,t} + \varepsilon_t$, with $\boldsymbol{\pi}_{m,t} = \boldsymbol{\pi}_m + \mathbf{v}_{m,t}$. The term given by the last two components in (3.4), $\rho_t = \varepsilon_t + \mathbf{v}'_{q,t}\mathbf{h}_{q,t}$ (say), behaves like a stochastically integrated process net of its stochastic trend component in the sense that, even though the innovations $\boldsymbol{\xi}_{q,t}$ have an infinitely persistent effect on $\mathbf{h}_{q,t+s}$, their effect on the level of $\mathbf{v}'_{q,t+s}\mathbf{h}_{q,t+s}$ is only transitory. This implies that the product process $\mathbf{v}'_{q,t}\mathbf{h}_{q,t}$,[7] and thus ρ_t, is stochastically trendless, even if $\mathbf{v}_{q,t}$ is correlated with $\boldsymbol{\xi}_{q,t}$. This characterization of the nature of the dependence of ρ_t is termed the **stochastically trendless property**, and states that the behavior of the process up to time t has a negligible effect on its behavior into the infinite future. Formally, under proper assumptions on the finiteness of moments of $\mathbf{v}_{q,t}$ and $\boldsymbol{\xi}_{q,t}$ it must be show that $c_t(s) \to^p 0$ as $s\to\infty$, for fixed t, where $c_t(s)$ is defined as

$$c_t(s) = E_t[\rho_{t+s}] - E[\rho_{t+s}] = E_t[\varepsilon_{t+s} + \mathbf{v}'_{q,t+s}\mathbf{h}_{q,t}] - E[\varepsilon_{t+s} + \mathbf{v}'_{q,t+s}\mathbf{h}_{q,t}]$$

where the conditional expectation is based on the information available for all the elements of ρ_t up to time t.[8]

The discrete time Cramer's (1961) representation of an arbitrary (non-deterministic) series is given by

$$\mathbf{Z}_t = \sum_{k=0}^{\infty} \mathbf{A}_{t,t-k}\mathbf{e}_{t-k}$$

with $\mathbf{e}_t$ a zero mean non-degenerate vector white-noise process with $E[\mathbf{e}_t\mathbf{e}_t] = \mathbf{I}$, and the non-stochastic matrix coefficients $\mathbf{A}_{t,t-k}$ constructed so as to absorb any heteroskedastic structure in the disturbances. These coefficients are thus assumed to satisfy the square summability condition $\sum_{k=0}^{\infty} \| \mathbf{A}_{t,t-k} \|^2 < \infty$ for all t. In order to embed the SI model in (3.4)-(3.6) into this framework we assume that the error vector $\boldsymbol{\zeta}_t$ follows a linear process, $\boldsymbol{\zeta}_t = \mathbf{C}(L)\mathbf{e}_t$ with $\sum_{k=1}^{\infty} k^a \| \mathbf{C}_k \|^2 < \infty$, $a \geq 2$, and $\mathbf{C}_0$ having full rank, with $E[\boldsymbol{\zeta}_t\boldsymbol{\zeta}'_t] = \mathbf{C}(1)\mathbf{C}(1)' > 0$ a general (non-diagonal) covariance matrix. Then we define the random vector

$$\mathbf{G}_{t-k} = \begin{pmatrix} \boldsymbol{\upsilon}_{m,t-k} \\ \mathbf{v}'_{q,t-k}\sum_{j=1}^{t-k}\boldsymbol{\xi}_{q,j} \\ \varepsilon_{t-k} \end{pmatrix}$$

with a general covariance matrix $\boldsymbol{\Sigma}_{t-k} = E[\mathbf{G}_{t-k}\mathbf{G}'_{t-k}]$, and $\mathbf{A}_{t,t-k} = \mathbf{A}^*_{t,t-k} + \mathbf{B}_{t,t-k}$, with

---

[7] Despite of this result, in strict sense, the term $\mathbf{v}'_{q,t}\mathbf{h}_{q,t}$ is nonstationary heteroskedastic, exhibiting a linear trend in variance.

[8] McCabe et.al. (2003) show that, under the assumption that the sequence $\boldsymbol{\zeta}_t$ follows a linear process defined on iid innovations, the above condition is satisfied. We conjecture that this results also follows under more general assumptions on this sequence, for example, when considering that the components of the error vector $\boldsymbol{\zeta}_t$ are strong (α-) mixing when adapted to the natural filtration given by its past observations. Thus, making use of the maximal inequalities for unconditional and conditional covariances of strong mixing sequences in Roussas and Ioannides (1987) and DeMei and Lan (2013), respectively, and under additional assumptions on the existence of unconditional absolute moments of order greater than 2, then it can be shown the condition $|c_t(s)| \to 0$ *a.s.* for fixed t and $s\to\infty$.

$$\mathbf{A}^*_{t,t} = \begin{pmatrix} \boldsymbol{\pi}'_m & 1 & 0 \\ \mathbf{0}'_m & 0 & 0 \\ \mathbf{0}'_m & 0 & 0 \end{pmatrix} \boldsymbol{\Sigma}^{1/2}_t, \mathbf{A}^*_{t,t-k} = \begin{pmatrix} \boldsymbol{\pi}'_m & 0 & 0 \\ \mathbf{0}'_m & 0 & 0 \\ \mathbf{0}'_m & 0 & 0 \end{pmatrix} \boldsymbol{\Sigma}^{1/2}_{t-k}, k = 1,...,t-1$$

$$\mathbf{B}_{t,t} = \begin{pmatrix} \mathbf{0}'_m & 0 & 1 \\ \mathbf{0}'_m & 0 & 0 \\ \mathbf{0}'_m & 0 & 0 \end{pmatrix} \boldsymbol{\Sigma}^{1/2}_t$$

$\mathbf{A}^*_{t,t-k} = \mathbf{0}$ for $k \geq t$, and $\mathbf{B}_{t,t-k} = \mathbf{0}$ for $k \geq 1$, then we have that $\mathbf{Z}_t$ is given by

$$\mathbf{Z}_t = \sum_{k=0}^{\infty} \mathbf{A}_{t,t-k}(\boldsymbol{\Sigma}^{-1/2}_{t-k}\mathbf{G}_{t-k}) = \sum_{k=0}^{t-1} \mathbf{A}^*_{t,t-k}(\boldsymbol{\Sigma}^{-1/2}_{t-k}\mathbf{G}_{t-k}) + \mathbf{B}_{t,t}(\boldsymbol{\Sigma}^{-1/2}_t\mathbf{G}_t)$$

where the first component, Z_t is such that $Z_t = \boldsymbol{\pi}'_m \sum_{k=0}^{t-1} \boldsymbol{\upsilon}_{m,t-k} + \mathbf{v}'_{q,t} \sum_{j=1}^{t} \boldsymbol{\xi}_{q,j} + \varepsilon_t$, with $\eta_t = Z_t + \boldsymbol{\pi}'_m \mathbf{w}_{m,0} + \mathbf{v}'_{q,t}\mathbf{h}_{q,0}$. This gives a generalized version of the Cramer representation with the standard case given by $\mathbf{C}(L) = \mathbf{C}_0$. Making use of this general representation for a SI process based on weak white noise innovations, the results in McCabe, Martin and Tremayne (2005) imply that for $\boldsymbol{\pi}_m \neq \mathbf{0}_m$ the long-run impulse response and variance ratio measures indicate substantial persistence mimicking that of the pure random walk process, while that for $\boldsymbol{\pi}_m = \mathbf{0}_m$ the results are ambiguous. However, for a STUR process given by equations (3.2) and (3.3), with $E[\alpha_t] = 1$ and $Var[\alpha_t] = Var[\phi_t] > 0$, the evidence from both measures of persistence is unambiguous indicating a high degree of persistence, although with very different limiting behavior. Despite these differences in the persistence performance of this two different non-stationary processes, we consider the possibility of being plausible non-stationary alternatives to the null of a martingale process. Formally, we introduce this possibility in the following assumption.

Assumption 3.1. *Nonstationary non-martingale alternative processes*

We assume that, as a nonstationary non-martingale alternative, the stochastic trend component η_t of the observed process X_t in (3.1) is given either by:

(*a*) A STUR-type process, with $E[\alpha_t] = 1$ (ϕ = 1) and $Var[\alpha_t] \geq 0$ in (3.2)-(3.3), with

(*a*.1) $\phi_t = \omega n^{-\delta/2} \upsilon_t$, $\omega \geq 0$, or

(*a*.2) $\phi_t = \alpha_n \varepsilon_{t-1}$, $\alpha_n = \alpha/\sqrt{n}$, $\alpha \geq 0$

(*b*) A stochastically integrated (SI)-type process of the form

$$\eta_t = \boldsymbol{\pi}'_m \mathbf{w}_{m,t} + \varepsilon_t + \mathbf{v}'_{q,t}\mathbf{h}_{q,t}$$

with initial conditions $\mathbf{w}_{m,0}, \mathbf{h}_{q,0} = O_p(n^{1/2-\lambda}), 0 < \lambda \leq 1/2$.

The process described in Assumption 3.1(*a*.1) is the local-heteroskedastic integrated (LHI) version of a STUR process proposed by McCabe and Smith (1998), while that the process considered in Assumption 3.1(*a*.2) is the so-called weak bilinear unit root (weak BLUR) process as has been proposed by Charemza et.al. (2005) and Lifshits (2006), which is the nonstationary version of the BL(1,0,1,1) process $\eta_t = (\phi + \alpha\varepsilon_{t-1})\eta_{t-1} + \varepsilon_t$. Under iid(0, σ^2_ε) innovations ε_t, the stationary condition is given by $\phi^2 + \alpha^2\sigma^2_\varepsilon < 1$, that is clearly not satisfied with ϕ = 1. The scaling of the bilinear coefficient α by the square root of the sample size is a convenient reparameterization that allows to apply Lifshits' (2006) invariance principle.

Given that each alternative process in Assumption 3.1 is based on a different set of random error sequences, we introduce their properties in the following Assumption 3.2.

Assumption 3.2. *Linear assumption on the error sequence*

(*a*) In the case of the STUR-type nonstationary non-martingale processes given in Assumption 3.1(*a*) above, we have that $E[|\upsilon_t^k|] = \kappa_k \le \kappa^k < \infty$, $k = 0, 1, ...,$ and

$$\boldsymbol{\zeta}_t = \begin{pmatrix} \varepsilon_t \\ \upsilon_t \end{pmatrix} = \mathbf{C}(L)\mathbf{e}_t \tag{3.7}$$

with $\mathbf{C}(L) = \sum_{k=0}^{\infty} \mathbf{C}_k L^k, \mathbf{C}(1) > 0$, $\sum_{k=1}^{\infty} k^2 \|\mathbf{C}_k\|^2 < \infty$, where $\mathbf{e}_t = (e_{0,t}, e_{1,t})'$ is an iid sequence with $E[\mathbf{e}_t] = \mathbf{0}_2$, and $E[\mathbf{e}_t\mathbf{e}_t'] = \mathbf{\Sigma}_{2,2}$, so that it is verified that

$$n^{-1/2}\sum_{t=1}^{[nr]} \boldsymbol{\zeta}_t = n^{-1/2}\sum_{t=1}^{[nr]} \begin{pmatrix} \varepsilon_t \\ \upsilon_t \end{pmatrix} \Rightarrow \begin{pmatrix} B_0(r) \\ B_1(r) \end{pmatrix} = \mathbf{\Omega}^{1/2} \begin{pmatrix} W_0(r) \\ W_1(r) \end{pmatrix}, \mathbf{\Omega} = \mathbf{C}(1)\mathbf{\Sigma}_{2,2}\mathbf{C}(1)' > 0 \tag{3.8}$$

where $B_0(r) = \omega_0[\sqrt{1-\rho_{01}^2}W_0(r) + \rho_{01}W_1(r)]$, and $B_1(r) = \omega_1 W_1(r)$, with $W_0(r)$ and $W_1(r)$ two independent standard Wiener processes, and ρ_{01} the (two-sided) long-run correlation between $B_0(r)$ and $B_1(r)$.

(*b*) In the case of the SI-type nonstationary non-martingale process given in Assumption 3.1(*b*) above, with $\boldsymbol{\zeta}_t = (\varepsilon_t, \boldsymbol{\upsilon}_{m,t}', \mathbf{v}_{q,t}', \boldsymbol{\xi}_{q,t}')'$ a zero mean 2*q*+*m*+1 dimensional error vector, we have that $\boldsymbol{\zeta}_t = \mathbf{C}(L)\mathbf{e}_t$ with $\mathbf{C}(L) = \sum_{k=0}^{\infty} \mathbf{C}_k L^k$, $\mathbf{C}(1) > 0$, $\sum_{k=1}^{\infty} k^2 \|\mathbf{C}_k\|^2 < \infty$, and $\mathbf{e}_t$ an iid sequence with $E[\mathbf{e}_t] = \mathbf{0}$, and $E[\mathbf{e}_t\mathbf{e}_t'] = \mathbf{\Sigma}$. Then, it is verified that

$$n^{-1/2}\sum_{t=1}^{[nr]} \boldsymbol{\zeta}_t \Rightarrow \mathbf{B}(r) = \mathbf{\Omega}^{1/2}\mathbf{W}(r) \tag{3.9}$$

with $\mathbf{W}(r)$ a standard vector Wiener process, and $\mathbf{\Omega} = \mathbf{C}(1)\mathbf{\Sigma}\mathbf{C}(1)' > 0$ the long-run covariance matrix of $\mathbf{B}(r) = (B_0(r), \mathbf{B}_m(r)', \mathbf{B}_{q1}(r)', \mathbf{B}_{q2}(r)')'$.

This assumption is quite standard in time series econometrics when dealing with nonstationary processes, and allows to introduce a certain controlled degree of serial dependence through the linear process framework. Due to the fact that the STUR-type process given by the weak BLUR specification in Assumption 3.1(*a*.1) above only depends on the sequence ε_t, we use the results in (3.7)-(3.8) only concerning to the stochastic distributional limit of the scaled partial sum process of ε_t, that is, $n^{-1/2}\sum_{t=1}^{[nr]} \varepsilon_t \Rightarrow B_0(r) = \omega_0 W_0(r)$. Also, for any particular form of the stochastic trend component η_t in (3.2) or (3.4) above, we define the normalized partial sum processes

$$H_n(r) = (1/\sqrt{n})\eta_t \quad (t/n) \le r < (t+1)/n \quad 0 \le t \le n-1 \tag{3.10}$$

and $H_n(\frac{t-h}{n}) = (1/\sqrt{n})\eta_{t-h}$, $h \le t \le n$, $0 \le h \le n-1$, and as in Section 2, let $u_t = \eta_t - \eta_{t-1}$ be the sequence of first differences given by

$$u_t = \phi_t \eta_{t-1} + \varepsilon_t \tag{3.11}$$

for a STUR-type process, and

$$\begin{aligned} u_t &= \boldsymbol{\pi}_m' \boldsymbol{\upsilon}_{m,t} + \Delta\varepsilon_t + (\mathbf{v}_{q,t}' - \mathbf{v}_{q,t-1}')\mathbf{h}_{q,t-1} + \mathbf{v}_{q,t}'\boldsymbol{\xi}_{q,t} \\ &= \boldsymbol{\pi}_m' \boldsymbol{\upsilon}_{m,t} + z_t + \mathbf{z}_{q,t}'\mathbf{h}_{q,t-1} + \mathbf{v}_{q,t}'\boldsymbol{\xi}_{q,t} \end{aligned} \tag{3.12}$$

for the SI alternative, with $z_t = \Delta\varepsilon_t$, and $\mathbf{z}_{q,t} = \mathbf{v}_{q,t} - \mathbf{v}_{q,t-1}$. We are now in position to formulate the following results concerning the distributional limits of the stochastic trend component η_t and of the process of first differences, u_t, for each type of nonstationary nonlinear alternative considered in Assumption 3.1, as well as for the sample variance σ_n^2.

Proposition 3.1. *Under Assumption 3.2 on the error terms:*

(*a*) *When the stochastic trend component* η_t *is generated under the LHI alternative, with* $\delta = 1$, $\omega \geq 0$, *and* $\lambda_{10} = \sum_{k=1}^{\infty} E[\varepsilon_t \upsilon_{t-k}]$, *then:*

$$(a.1)\ H_n(r) \Rightarrow H_\omega(r) = B_0(r) + \omega\left\{\int_0^r B_1(s)dB_0(s) + r\cdot\lambda_{10}\right\} \tag{3.13}$$

$$(a.2)\ u_t = \omega H_n(\tfrac{t-1}{n})\upsilon_t + \varepsilon_t = O_p(1)$$

$$(a.3)\ \sigma_n^2 \Rightarrow Y_\omega = \sigma_0^2 + \omega^2\sigma_1^2 \int_0^1 H_\omega(s)^2 ds + 2\omega\gamma_{01}\int_0^1 H_\omega(s)ds \tag{3.14}$$

with $\sigma_0^2 = E[\varepsilon_t^2]$, $\sigma_1^2 = E[\upsilon_t^2]$, *and* $\gamma_{01} = E[\varepsilon_t \upsilon_t]$;

(*b*) *When the stochastic trend component* η_t *is generated under the weak bilinear integrated alternative, then:*

$$(b.1)\ H_n(r) \Rightarrow H_\alpha(r) = \omega_0 A_\alpha(r)\int_0^r \frac{1+\alpha^2\omega_0^2 s}{A_\alpha(s)}(dW_0(s) - \alpha\omega_0 ds) + \alpha\omega_0^2 r \tag{3.15}$$

$$(b.2)\ u_t = \alpha H_n(\tfrac{t-1}{n})\varepsilon_{t-1} + \varepsilon_t = O_p(1)$$

$$(b.3)\ \sigma_n^2 \Rightarrow Y_\alpha = \sigma_0^2 + \alpha^2\sigma_0^2 \int_0^1 H_\alpha(s)^2 ds + 2\alpha\gamma_0(1)\int_0^1 H_\alpha(s)ds \tag{3.16}$$

with $A_\alpha(r) = \exp(\alpha B_0(s) - \alpha^2\omega_0^2(\tfrac{r}{2}))$, $\omega_0^2 = \sigma_e^2 C(1)^2$, $\sigma_0^2 = E[\varepsilon_t^2]$, $\gamma_0(1) = E[\varepsilon_t\varepsilon_{t-1}]$;

(*c*) *When the stochastic trend component* η_t *is generated by a stochastically integrated sequence, then:*

$$(c.1)\ H_n(r) \Rightarrow H_{m,q}(r) = C_m(r) + C_q(r) \tag{3.17}$$

$$(c.2)\ u_t = \boldsymbol{\pi}_m'\boldsymbol{\upsilon}_{m,t} + z_t + \mathbf{z}_{q,t}'\sqrt{n}\left\{(\mathbf{h}_{q,0}/\sqrt{n}) + (1/\sqrt{n})\sum_{k=1}^{t-1}\boldsymbol{\xi}_{q,k}\right\} + \mathbf{v}_{q,t}'\boldsymbol{\xi}_{q,t} = O_p(\sqrt{n})$$

$$(c.3)\ (1/n)\sum_{t=1}^{[nr]} u_t \Rightarrow J_q(r) + r\cdot\Lambda_{q,1}$$

$$(c.4)\ \sigma_n^2 = O_p(n),\ (1/n)\sigma_n^2 \Rightarrow Y_q = \int_0^1 \mathbf{B}_{q2}(s)'\boldsymbol{\Gamma}_{q,q}\mathbf{B}_{q2}(s)ds \tag{3.18}$$

where $C_m(r) = \boldsymbol{\pi}_m'\mathbf{B}_m(r)$ *and* $C_q(r)$ *are the weak limits of* $\boldsymbol{\pi}_m'[(1/\sqrt{n})\mathbf{w}_{m,[nr]}]$ *and* $n^{-1/2}\sum_{t=1}^{[nr]}(\mathbf{v}_{q,[nr]}'\boldsymbol{\xi}_{q,t} - E[\mathbf{v}_{q,[nr]}'\boldsymbol{\xi}_{q,t}])$, *respectively, and with*

$$J_q(r) = \int_0^r \mathbf{B}_{q2}(s)'d\mathbf{Z}_q(s),\ n^{-1/2}\sum_{t=1}^{[nr]}\mathbf{z}_{q,t} \Rightarrow \mathbf{Z}_q(r),\ \Lambda_{q,1} = \sum_{j=1}^{\infty} E[\mathbf{z}_{q,t}'\boldsymbol{\xi}_{q,t-j}]$$

$$\boldsymbol{\Gamma}_{q,q} = E[\mathbf{z}_{q,t}\mathbf{z}_{q,t}'] = 2\boldsymbol{\gamma}_{q,q}(0) - (\boldsymbol{\gamma}_{q,q}(1) + \boldsymbol{\gamma}_{q,q}'(1)),\ \boldsymbol{\gamma}_{q,q}(i) = E[\mathbf{v}_{q,t}\mathbf{v}_{q,t-i}'], i = 0,1.$$

Proof. Part (a) mainly follows from Theorem 1 in McCabe and Smith (1998). See Appendix A for specific details in our case. Part (b) follows from Lifshits (2006) in the case of an iid error sequence ε_t, and from Afonso-Rodríguez (2012) under the linear assumption for the innovations in the weak BLUR process. For the proof of part (c) see Appendix B.

Remark 3.1. First note that, under a STUR-type LHI alternative with serially correlated sequences $(\varepsilon_t, \upsilon_t)$, the limiting process $H_\omega(r)$ can be decomposed as

$$H_{\omega}(r) = \omega_0 \left\{ \sqrt{1-\rho_{01}^2} W_0(r) + \omega\omega_1 \sqrt{1-\rho_{01}^2} \int_0^r W_1(s)dW_0(s) \right\}$$
$$+ \omega_0\rho_{01} \left\{ W_1(r) + \omega\omega_1 \int_0^r W_1(s)dW_1(s) + r\cdot\omega \frac{\omega_1\lambda_{10}}{\omega_0\rho_{01}} \right\}$$

where $\int_0^r W_1(s)dW_1(s) = (1/2)(W_1^2(r) - r)$, and $E[H_{\omega}(r)] = r\cdot\omega[\omega_0\omega_1^2](\lambda_{10}/\omega_0\omega_1)$, with $(\lambda_{10}/\omega_0\omega_1)$ the one-sided long-run correlation coefficient between ε_t and υ_t. In this case, under serial correlated error terms ε_t and υ_t this limiting process induces an additional bias component though the drift term $E[H_{\omega}(r)]$.

Remark 3.2. All the above results in Proposition 3.1 determine that the pivotal test statistic $V_n(x)$ defined in (2.5),

$$V_n(x) = \sigma_n^{-1} M_n(x) = \frac{1}{\sigma_n \sqrt{n}} \sum_{t=1}^{n} u_t \cdot I\left(\frac{X_{t-1}}{\sqrt{n}} \le x \right)$$

has a well defined limiting distribution under any of the alternatives particularly considered in this paper, but with a very different behavior in each case. This implies that the testing procedure is inconsistent in the usual sense, that is, the finite-sample power performance only depends on the parameters defining each nonstationary alternative and does not depend on the sample size. The following Proposition 3.2 states the distributional limit of the test statistic $V_n(x)$.

Proposition 3.2. *Under Assumption 3.2 on the error terms and taking the results in Proposition 3.1, we have for each of the non-martingale nonstationary alternatives considered that*

$$V_n(x) \Rightarrow VH(x) = Y^{-1/2} MH(x)$$

with $M_n(x) \Rightarrow MH(x) = \int_0^1 I(H(r) \le x)dH(r)$, *and* $\sigma_n^2 \Rightarrow Y$, *where H(r) is given by the corresponding weak limit* $H_{\omega}(r)$, $H_{\alpha}(r)$, *or* $H_{m,q}(r)$ *in (3.13),(3.15),and (3.17), and Y is given by the weak limit* Y_{ω}, Y_{α} *or* Y_q *defined in (3.14), (3.16), and (3.18).*

Proof. This limiting distribution is readily derived from the results in Proposition 3.1 and the continuous mapping theorem, since the statistic $V_n(x)$ is a continuous functional of $H_n(r)$. Also, simply replacing $M(x)$ in (2.6) by the above limiting result $VH(x)$ in (2.9) and (2.10) we get the weak limit of the test statistics S_n and T_n defined in (2.7) and (2.8) under each of the non-martingale nonstationary alternative considered here.

Remark 3.3. The main implication of these results is that the asymptotic behavior of the test statistic under each alternative is determined by the combination of the weak limit of its two components, the empirical marked process $M_n(x)$ in the numerator, and the sample variance σ_n^2 in denominator, which are plagued of many nuisance parameters. Particularly relevant is the result for the two types of STUR processes considered, where for the element in the denominator we have that $Y_{\omega} = O_p(\omega^2)$ and $Y_{\alpha} = O_p(\alpha^2)$, so that $VH_{\omega}(x) = O_p(\omega^{-1})$ and $VH_{\alpha}(x) = O_p(\alpha^{-1})$, respectively, determining a reduction in the estimated value of the test statistic when increasing the value of ω and α in each case. Thus, when considering the test statistic based on the scaled version of $M_n(x)$, we might expect a serious problem of under-rejection of the null hypothesis, and thus an incorrect detection of the martingale behavior.

In order to illustrate these findings numerically, we perform a small Monte Carlo experiment to evaluate the power performance of this testing procedure for each of the three alternatives considered in this section. The results are shown in Tables 3.1-3.3 in Appendix C, where the power based on the test statistics S_n and T_n is computed by making use of the finite sample critical values for sample sizes n = 100, 250 and 500. As indicated in Remark 3.3, for the two version of the STUR-type alternatives, we display the results for the empirically marked process $M_n(x)$ (that is, for the unnormalized test statistic $V_n(x)$).

References

Afonso-Rodríguez, J.A. (2012). On testing for a stochastic unit root in financial time series: The case of a bilinear unit root process. In: Tópicos Selectos de Optimización, Volume 1, M.R. Fernando Miranda (Ed.), ECORFAN, Mexico.

Aue, A., L. Horváth, J. Steinebach (2006). Estimation in random coefficient autoregressive models. Journal of Time Series Analysis, 27(1), 61-76.

Berkes, I., L. Horváth, S. Ling (2009). Estimation in nonstationary random coefficient autoregressive models. Journal of Time Series Analysis, 30(4), 395-416.

Carrasco, M., X. Chen (2002). Mixing and moment properties of various GARCH and stochastic volatility models. Econometric Theory, 18(1), 17-39.

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

Cramer, H. (1961). On some classes of nonstationary stochastic processes. Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability. University of Los Angeles Press: Berkeley and Los Angeles, 57-77.

DeMei, Y., L. Lan (2013). Some conditional results for conditionally strong mixing sequences of random variables. Science China Mathematics, 56(4), 845-859.

Deo, R.S. (2000). Spectral tests of the martingale hypothesis under conditional heteroscedasticity. Journal of Econometrics, 99(2), 291-315.

Durlauf, S.N. (1991). Spectral based testing of the martingale hypothesis. Journal of Econometrics, 50(3), 355-376.

Escanciano, J.C. (2007). Weak convergence of non-stationary multivariate marked processes with applications to martingale testing. Journal of Multivariate Analysis, 98(7), 1321-1336.

Escanciano, J.C., I.N. Lobato (2007). Testing the martingale hypothesis. Working paper.

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

Hansen, B.E. (1991). GARCH(1,1) processes are near epoch dependent. Economics Letters, 36(2), 181-186.

Hansen, B.E. (1992). Heteroskedastic cointegration. Journal of Econometrics, 54(1-3), 139-158.

Harris, D., B.P.M. McCabe, S. Leybourne (2002). Stochastic cointegration: estimation and inference. Journal of Econometrics, 111(2), (363-384).

Harris, D., B.P.M. McCabe, S. Leybourne (2003). Some limit theory for autocovariances whose order depends on sample size. Econometric Theory, 19(5), (829-864).

Hill, J.B. (2009). On functional central limit theorems for dependent, heterogeneous arrays with applications to tail index and tail dependence estimation. Journal of Statistical Planning and Inference, 139(6), 2091-2110.
Hwang, S.Y., I.V. Basawa (2005). Explosive random-coefficient AR(1) processes and related asymptotics for leasts-squares estimation. Journal of Time Series Analysis, 26(6), 807-824.
Kim, S., S. Cho, S. Lee (2000). On the cusum test for parameter changes in GARCH(1,1) models. Communications in Statistics – Theory and Methods, 29(2), 445-462.
Koul, H.L., W. Stute (1999). Nonparametric model checks for time series. The Annals of Statistics, 27(1), 204-236.
Kulperger, R., H. Yu (2005). High moment partial sum processes of residuals in GARCH models and their applications. The Annals of Statistics, 33(5), 2395-2422.
Kumagai, T. (2001). Properties and applications of a martingale hypothesis test. Ph.D. Thesis in Economics, Massachusetts Institute of Technology, Dept. of Economics.
Leybourne, S., B.P.M. McCabe, and 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, and 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., and A.R. Tremayne (1995). Testing a time series for difference stationarity. The Annals of Statistics, 23(3), 1015-1028.
McCabe, B.P.M., G.M. Martin, A.R. Tremayne (2005). Assessing persistence in discrete nonstationary time-series models. Journal of Time Series Analysis, 26(2), 305-317.
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., S. Leybourne, D. Harris (2003). Testing for stochastic cointegration and evidence for present value models. EconWPA Working Paper in Econometrics 0311009.
McCabe, B.P.M., S. Leybourne, D. Harris (2006). A residual-based test for stochastic cointegration. Econometric Theory, 22(3), (429-456).
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 and Probability Letters, 79(24), 2476-2483.
Nicholls, D.F., B.G. Quinn (1982). Random coefficient autoregressive models: An Introduction. Springer, New York.
Park, J.Y., Y-J. Whang (1999). Testing for the martingale hypothesis. Cowles Foundation for Research in Economics working paper.
Park, J.Y., Y-J. Whang (2005). A test of the martingale hypothesis. Studies in Nonlinear Dynamics and Econometrics, 9(2), Article 2.
Roussas, G.G., D. Ioannides (1987). Moment inequalities for mixing sequences of random variables. Stochastic Analysis and Applications, 5(1), 60-120.
Stute, W. (1997). Nonparametric model checks for regression. The Annals of Statistics, 25(2), 613-641.

Appendixes

A. Proof of Proposition 3.1(a) *Asymptotics under local-heteroskedastic integration*
By backward substitution in (3.2) we have that

$$\eta_t = \alpha_t \eta_{t-1} + \varepsilon_t = \eta_0 Z_{1,t} + \varepsilon_t + \sum_{k=1}^{t-1} \varepsilon_k Z_{k+1,t}$$

with

$$Z_{k,t} = \prod_{j=k+1}^{t} \alpha_j = \prod_{j=k+1}^{t} (1+\phi_j),\ k = 1, ..., t-1.$$

As indicated in McCabe and Smith (1998), from a notational perspective it is more convenient to deal with forward summations than with backward summations, so the order of the subscripts on α_j and ε_k are reversed, so that η_t can also be written as

$$\eta_t = \eta_0 Z_{1,t} + \varepsilon_1 + \sum_{k=2}^{t} \varepsilon_k Z_{1,k-1}$$

where

$$Z_{1,k} = \prod_{j=1}^{k} \alpha_j = \prod_{j=1}^{k} (1+\phi_j),\ k = 1, 2, ..., t$$

Let k–1 = [na], with $a \in$ [0,1]. Then $Z_{1,k-1} = Z_{1,[na]}$ can be written as

$$Z_{1,[na]} = \prod_{j=1}^{[na]} (1+\phi_j) = \exp\left\{\sum_{j=1}^{[na]} \log(1+\phi_j)\right\}$$
$$= \exp\left\{\sum_{j=1}^{[na]} \phi_j\right\} \cdot \exp\left\{\sum_{k=0}^{\infty} \frac{(-1)^{k+1}}{k+2} \sum_{j=1}^{[na]} \phi_j^{k+2}\right\}$$

where

$$\sum_{j=1}^{[na]} \phi_j = \omega n^{(1-\delta)/2} \left\{n^{-1/2} \sum_{j=1}^{[na]} \upsilon_j\right\}$$

and

$$\sum_{j=1}^{[na]} \phi_j^{k+2} = \omega^2 n^{1-\delta} (\omega^k n^{-\delta k/2}) \left\{n^{-1} \sum_{j=1}^{[na]} \upsilon_j^{k+2}\right\}$$

making use of the LHI specification in Assumption 3.1(a.1). Then we have

$$E\left|\sum_{k=0}^{\infty} \frac{(-1)^{k+1}}{k+2} \sum_{j=1}^{[na]} \phi_j^{k+2}\right| \le \omega^2 n^{1-\delta} \sum_{k=0}^{\infty} (\omega n^{-\delta/2})^k \left\{n^{-1} \sum_{j=1}^{[na]} E[|\, \upsilon_j^{k+2}\, |]\right\}$$
$$\le a\omega^2 \kappa^2 n^{1-\delta} \sum_{k=0}^{\infty} (\kappa\omega n^{-\delta/2})^k$$

where the element on the second hand of the above inequality is $O(1)$ for any $\delta \ge 1$ whenever $\kappa\omega = o(n^{\delta/2})$, which gives

$$Z_{1,[na]} = \exp\left\{\omega n^{(1-\delta)/2} \left\{n^{-1/2} \sum_{j=1}^{[na]} \upsilon_j\right\}\right\} (1+O_p(n^{1-\delta}))$$
$$= \left(1+\omega n^{(1-\delta)/2} \left\{n^{-1/2} \sum_{j=1}^{[na]} \upsilon_j\right\} + O_p(n^{1-\delta})\right)(1+O_p(n^{1-\delta}))$$

that is

$$Z_{1,[na]} = 1 + \omega n^{(1-\delta)/2}\left\{ n^{-1/2}\sum_{j=1}^{[na]} \upsilon_j \right\} + O_p(n^{1-\delta})$$

Then, substituting $Z_{1,k-1}$ in η_t we have

$$\eta_t = \eta_0\left(1 + \omega n^{(1-\delta)/2}\left\{ n^{-1/2}\sum_{j=1}^{t} \upsilon_j \right\} + O_p(n^{1-\delta})\right)$$

$$+\varepsilon_1 + \sum_{k=2}^{t}\varepsilon_k\left(1 + \omega n^{(1-\delta)/2}\left\{ n^{-1/2}\sum_{j=1}^{k-1} \upsilon_j \right\} + O_p(n^{1-\delta})\right)$$

that is

$$\eta_t = \eta_0(1 + O_p(n^{1-\delta})) + \sqrt{n}\left\{ n^{-1/2}\sum_{k=1}^{t}\varepsilon_k \right\} + \omega n^{(1-\delta)/2}\sum_{k=2}^{t}\varepsilon_k\left\{ n^{-1/2}\sum_{j=1}^{k-1}\upsilon_j \right\}$$

$$+\left\{ n^{-1/2}\sum_{k=1}^{t}\varepsilon_k - n^{-1/2}\varepsilon_1 \right\} O_p(n^{3/2-\delta})$$

Thus, scaling by $n^{-1/2}$ and taking $\delta = 1$ we get

$$n^{-1/2}\eta_{[nr]} = (n^{-1/2}\eta_0)(1 + O_p(1)) + \left\{ n^{-1/2}\sum_{k=1}^{[nr]}\varepsilon_k \right\} + \omega n^{-1/2}\sum_{k=2}^{[nr]}\varepsilon_k\left\{ n^{-1/2}\sum_{j=1}^{k-1}\upsilon_j \right\} + O_p(1)$$

with

$$H_n(r) = n^{-1/2}\eta_{[nr]} \Rightarrow B_0(r) + \omega\left\{ \int_0^r B_1(s)dB_0(s) + r\cdot\lambda_{10} \right\}$$

where $\lambda_{10} = \sum_{k=1}^{\infty} E[\varepsilon_t\upsilon_{t-k}]$ is the limit in probability of $n^{-1}\sum_{k=2}^{[nr]}\sum_{j=1}^{k-1} E[\upsilon_j\varepsilon_k]$. Given that the sequence of first differences, $u_t = \eta_t - \eta_{t-1}$, is given by $u_t = \omega\upsilon_t n^{-\delta/2}\eta_{t-1} + \varepsilon_t$, then

$$\sigma_n^2 = (1/n)\sum_{t=1}^{n} u_t^2 = (1/n)\sum_{t=1}^{n}\varepsilon_t^2 + \sigma_1^2\omega^2(1/n)\sum_{t=1}^{n}(n^{-\delta/2}\eta_{t-1})^2$$

$$+(2\gamma_{01}\omega)\left\{ n^{-1}\sum_{t=1}^{n}(n^{-\delta/2}\eta_{t-1}) \right\}$$

$$+(\omega^2/\sqrt{n})\left\{ n^{-1/2}\sum_{t=1}^{n}(n^{-\delta/2}\eta_{t-1})^2(\upsilon_t^2 - \sigma_1^2) \right\}$$

$$+(2\omega/\sqrt{n})\left\{ n^{-1/2}\sum_{t=1}^{n}(n^{-\delta/2}\eta_{t-1})(\varepsilon_t\upsilon_t - \gamma_{01}) \right\}$$

Clearly, when $\delta = 1$ we get

$$\sigma_n^2 \Rightarrow \sigma_0^2 + \sigma_1^2\omega^2\int_0^1 H_\omega(s)^2 ds + 2\gamma_{01}\omega\int_0^1 H_\omega(s)ds$$

given that the two last terms between brackets are both $O_p(1)$.

■

B. Proof of Proposition 3.1(c) *Asymptotics under stochastic integration*

First of all, given the specification of the stochastically integrated process as in Assumption 3.1(b), we have that

$$n^{-1/2}\eta_{[nr]} = \boldsymbol{\pi}'_m(n^{-1/2}\mathbf{w}_{m,[nr]}) + n^{-1/2}\varepsilon_{[nr]} + \mathbf{v}'_{q,[nr]}(n^{-1/2}\mathbf{h}_{q,[nr]})$$

where for the first component we have the standard result

$$\boldsymbol{\pi}'_m(n^{-1/2}\mathbf{w}_{m,[nr]}) \Rightarrow C_m(r) = \boldsymbol{\pi}'_m\mathbf{B}_m(r)$$

where $E[C_m(r)]=0$, and $E[C_m(r)^2]=\sigma_m^2\cdot r$, with $\sigma_m^2=\boldsymbol{\pi}_m'\boldsymbol{\Omega}_{m,m}\boldsymbol{\pi}_m$. Second, for the last component we have that

$$\begin{aligned}\mathbf{v}_{q,[nr]}'(n^{-1/2}\mathbf{h}_{q,[nr]}) &= \mathbf{v}_{q,[nr]}'(n^{-1/2}\mathbf{h}_{q,0})+n^{-1/2}\sum_{j=1}^{[nr]}\mathbf{v}_{q,[nr]}'\boldsymbol{\xi}_{q,j}\\ &= \mathbf{v}_{q,[nr]}'(n^{-1/2}\mathbf{h}_{q,0})+\sqrt{n}\left\{(1/n)\sum_{j=1}^{[nr]}E[\mathbf{v}_{q,[nr]}'\boldsymbol{\xi}_{q,j}]\right\}\\ &\quad+n^{-1/2}\sum_{j=1}^{[nr]}(\mathbf{v}_{q,[nr]}'\boldsymbol{\xi}_{q,j}-E[\mathbf{v}_{q,[nr]}'\boldsymbol{\xi}_{q,j}])\end{aligned}$$

where $\mathbf{v}_{q,[nr]}'(n^{-1/2}\mathbf{h}_{q,0})=O_p(n^{-\delta})=o_p(1)$, and

$$E[\mathbf{v}_{q,[nr]}'\boldsymbol{\xi}_{q,j}]=Tr(E[\boldsymbol{\xi}_{q,j}\mathbf{v}_{q,[nr]}'])\to Tr(E[\boldsymbol{\xi}_{q,j}\mathbf{v}_{q,\infty}'])=0\text{, as } n\to\infty$$

which gives

$$n^{-1/2}\sum_{j=1}^{[nr]}(\mathbf{v}_{q,[nr]}'\boldsymbol{\xi}_{q,j}-E[\mathbf{v}_{q,[nr]}'\boldsymbol{\xi}_{q,j}])\Rightarrow C_q(r)$$

and thus $n^{-1/2}\eta_{[nr]}\Rightarrow H_{m,q}(r)=C_m(r)+C_q(r)$. For the process of first differences, that is given by

$$u_t=\boldsymbol{\pi}_m'\boldsymbol{\upsilon}_{m,t}+z_t+\mathbf{z}_{q,t}'\mathbf{h}_{q,t-1}+\mathbf{v}_{q,t}'\boldsymbol{\xi}_{q,t}$$

we have that the normalized partial sums can be written as

$$\begin{aligned}n^{-1/2}\sum_{t=1}^{[nr]}u_t &= \boldsymbol{\pi}_m'n^{-1/2}\sum_{t=1}^{[nr]}\boldsymbol{\upsilon}_{m,t}+n^{-1/2}\sum_{t=1}^{[nr]}z_t+n^{-1/2}\sum_{t=1}^{[nr]}\mathbf{h}_{q,t-1}'\mathbf{z}_{q,t}+n^{-1/2}\sum_{t=1}^{[nr]}\mathbf{v}_{q,t}'\boldsymbol{\xi}_{q,t}\\ &= \boldsymbol{\pi}_m'n^{-1/2}\sum_{t=1}^{[nr]}\boldsymbol{\upsilon}_{m,t}+\sqrt{n}\left\{n^{-1}\sum_{t=1}^{[nr]}\mathbf{h}_{q,t-1}'\mathbf{z}_{q,t}\right\}\\ &\quad+\sqrt{n}\left(\frac{[nr]}{n}\right)\left\{[nr]^{-1}\sum_{t=1}^{[nr]}\mathbf{v}_{q,t}'\boldsymbol{\xi}_{q,t}\right\}+n^{-1/2}\varepsilon_{[nr]}\end{aligned}$$

where the last two terms between brackets are both $O_p(1)$ so that by standard application of the weak convergence results to stochastic integrals we have

$$\begin{aligned}n^{-1}\sum_{t=1}^{[nr]}u_t &= n^{-1}\sum_{t=1}^{[nr]}\mathbf{h}_{q,t-1}'\mathbf{z}_{q,t}+\frac{[nr]}{n}\left\{[nr]^{-1}\sum_{t=1}^{[nr]}\mathbf{v}_{q,t}'\boldsymbol{\xi}_{q,t}\right\}+O_p(n^{-1/2})\\ &\Rightarrow J_q(r)+r\cdot\Lambda_{q,1}\end{aligned}$$

where

$$J_q(r)=\int_0^r\mathbf{B}_{q2}(s)'d\mathbf{Z}_q(s)\text{, } n^{-1/2}\textstyle\sum_{t=1}^{[nr]}\mathbf{z}_{q,t}\Rightarrow\mathbf{Z}_q(r)\text{, and } \Lambda_{q,1}=\sum_{j=1}^{\infty}E[\mathbf{z}_{q,t}'\boldsymbol{\xi}_{q,t-j}].$$

Finally, the sample variance of the sequence of first differences can be decomposed as

$$\begin{aligned}\sigma_n^2=(1/n)\sum_{t=1}^n u_t^2 &= \boldsymbol{\pi}_m'\left\{(1/n)\sum_{t=1}^n\boldsymbol{\upsilon}_{m,t}\boldsymbol{\upsilon}_{m,t}'\right\}\boldsymbol{\pi}_m+(1/n)\sum_{t=1}^n z_t^2+(1/n)\sum_{t=1}^n(\mathbf{v}_{q,t}'\boldsymbol{\xi}_{q,t})^2\\ &\quad+(1/n)\sum_{t=1}^n\mathbf{h}_{q,t-1}'\mathbf{z}_{q,t}\mathbf{z}_{q,t}'\mathbf{h}_{q,t-1}\\ &\quad+2\boldsymbol{\pi}_m'(1/n)\sum_{t=1}^n\boldsymbol{\upsilon}_{m,t}(z_t+\mathbf{z}_{q,t}'\mathbf{h}_{q,t-1}+\mathbf{v}_{q,t}'\boldsymbol{\xi}_{q,t})\\ &\quad+2(1/n)\sum_{t=1}^n z_t(\mathbf{h}_{q,t-1}'\mathbf{z}_{q,t}+\mathbf{v}_{q,t}'\boldsymbol{\xi}_{q,t})+2(1/n)\sum_{t=1}^n\mathbf{h}_{q,t-1}'\mathbf{z}_{q,t}\boldsymbol{\xi}_{q,t}'\mathbf{v}_{q,t}\end{aligned}$$

where we can write

$$(1/n)\sum_{t=1}^{n}\mathbf{h}'_{q,t-1}\mathbf{z}_{q,t}\mathbf{z}'_{q,t}\mathbf{h}_{q,t-1} = (1/n)\sum_{t=1}^{n}\mathbf{h}'_{q,t-1}(E[\mathbf{z}_{q,t}\mathbf{z}'_{q,t}] + (\mathbf{z}_{q,t}\mathbf{z}'_{q,t} - E[\mathbf{z}_{q,t}\mathbf{z}'_{q,t}]))\mathbf{h}_{q,t-1}$$

$$= n\left\{(1/n)\sum_{t=1}^{n}(n^{-1/2}\mathbf{h}'_{q,t-1})E[\mathbf{z}_{q,t}\mathbf{z}'_{q,t}](n^{-1/2}\mathbf{h}_{q,t-1})\right\}$$

$$+ (1/n)\sum_{t=1}^{n}\mathbf{h}'_{q,t-1}(\mathbf{z}_{q,t}\mathbf{z}'_{q,t} - E[\mathbf{z}_{q,t}\mathbf{z}'_{q,t}])\mathbf{h}_{q,t-1}$$

$$(1/n)\sum_{t=1}^{n}\boldsymbol{\upsilon}_{m,t}\mathbf{z}'_{q,t}\mathbf{h}_{q,t-1} = \sqrt{n}E[\boldsymbol{\upsilon}_{m,t}\mathbf{z}'_{q,t}]\left\{(1/n)\sum_{t=1}^{n}(n^{-1/2}\mathbf{h}_{q,t-1})\right\} + (1/n)\sum_{t=1}^{n}(\boldsymbol{\upsilon}_{m,t}\mathbf{z}'_{q,t} - E[\boldsymbol{\upsilon}_{m,t}\mathbf{z}'_{q,t}])\mathbf{h}_{q,t-1},$$

$$(1/n)\sum_{t=1}^{n}\mathbf{h}'_{q,t-1}\mathbf{z}_{q,t}z_t = \sqrt{n}\left\{(1/n)\sum_{t=1}^{n}(n^{-1/2}\mathbf{h}'_{q,t-1})\right\}E[\mathbf{z}_{q,t}z_t] + (1/n)\sum_{t=1}^{n}\mathbf{h}'_{q,t-1}(\mathbf{z}_{q,t}z_t - E[\mathbf{z}_{q,t}z_t]),$$

and

$$(1/n)\sum_{t=1}^{n}\mathbf{h}'_{q,t-1}\mathbf{z}_{q,t}\boldsymbol{\xi}'_{q,t}\mathbf{v}_{q,t} = \sqrt{n}\left\{(1/n)\sum_{t=1}^{n}(n^{-1/2}\mathbf{h}'_{q,t-1})\right\}E[\mathbf{z}_{q,t}\boldsymbol{\xi}'_{q,t}\mathbf{v}_{q,t}] + (1/n)\sum_{t=1}^{n}\mathbf{h}'_{q,t-1}(\mathbf{z}_{q,t}\boldsymbol{\xi}'_{q,t}\mathbf{v}_{q,t} - E[\mathbf{z}_{q,t}\boldsymbol{\xi}'_{q,t}\mathbf{v}_{q,t}])$$

All the above terms given between brackets are finite in probability, that is $O_p(1)$, which gives

$$(1/n)\sum_{t=1}^{n}(n^{-1/2}\mathbf{h}'_{q,t-1})E[\mathbf{z}_{q,t}\mathbf{z}'_{q,t}](n^{-1/2}\mathbf{h}_{q,t-1}) \Rightarrow Y_q = \int_0^1 \mathbf{B}_{q2}(s)'\boldsymbol{\Gamma}_{q,q}\mathbf{B}_{q2}(s)ds$$

with $\boldsymbol{\Gamma}_{q,q} = E[\mathbf{z}_{q,t}\mathbf{z}'_{q,t}] = 2\boldsymbol{\gamma}_{q,q}(0) - (\boldsymbol{\gamma}_{q,q}(1) + \boldsymbol{\gamma}'_{q,q}(1))$, where $\boldsymbol{\gamma}_{q,q}(i) = E[\mathbf{v}_{q,t}\mathbf{v}'_{q,t-i}], i = 0,1$. This result clearly means that $\sigma_n^2 = O_p(n)$, so that $(1/n)\sigma_n^2 \Rightarrow Y_q$.

■

C. Finite-sample power results

Table 3.1. *Finite-sample adjusted empirical power at 5% nominal level of the test statistics S_n and T_n without scaling by σ_n^2. The case of the LHI process*

	n = 100		n = 250		n = 500	
	S_n	T_n	S_n	T_n	S_n	T_n
ω = 1.00	0.149	0.117	0.162	0.126	0.142	0.112
1.25	0.165	0.110	0.175	0.128	0.176	0.129
1.50	0.207	0.164	0.228	0.161	0.243	0.177
1.75	0.232	0.161	0.261	0.170	0.242	0.175
2.00	0.294	0.211	0.275	0.197	0.276	0.190
2.25	0.322	0.219	0.327	0.213	0.303	0.213
2.50	0.320	0.212	0.342	0.230	0.354	0.226
2.75	0.395	0.241	0.393	0.233	0.369	0.227
3.00	0.398	0.244	0.416	0.247	0.410	0.264
4.00	0.527	0.313	0.506	0.321	0.567	0.343
5.00	0.647	0.515	0.628	0.392	0.673	0.414

Note. Results based on 1.000 independent replications, with $(\varepsilon_t, \upsilon_t)' \sim iidN(\mathbf{0}_2, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma} = diag(\sigma_\varepsilon^2, \sigma_\upsilon^2) = \mathbf{I}_{2,2}$.

Table 3.2. *Finite-sample adjusted empirical power at 5% nominal level of the test statistics S_n and T_n without scaling by σ_n^2. The case of the weak BLUR(1,1) process*

	n = 100		n = 250		n = 500	
	S_n	T_n	S_n	T_n	S_n	T_n
α = 1.00	0.251	0.199	0.267	0.200	0.267	0.208
1.25	0.329	0.273	0.349	0.274	0.290	0.227
1.50	0.372	0.289	0.390	0.308	0.398	0.290
1.75	0.399	0.317	0.454	0.331	0.480	0.363
2.00	0.517	0.391	0.502	0.384	0.507	0.404
2.25	0.574	0.430	0.556	0.433	0.581	0.435
2.50	0.592	0.455	0.647	0.496	0.619	0.440
2.75	0.673	0.489	0.669	0.509	0.714	0.536
3.00	0.746	0.613	0.719	0.536	0.708	0.550
4.00	0.860	0.752	0.849	0.733	0.868	0.748
5.00	0.947	0.915	0.950	0.926	0.969	0.923

Note. Results based on 1.000 independent replications, with $\varepsilon_t \sim iidN(0,\sigma_\varepsilon^2), \sigma_\varepsilon^2 = 1$.

Table 3.3. *Finite-sample adjusted empirical power at 5% nominal level of the test statistics S_n and T_n. The case of the SI process*

	$\pi_1 = 0$		$\pi_1 = 1$	
(a) *Version without scaling by* σ_n^2	S_n	T_n	S_n	T_n
n = 100	1.000	1.000	0.996	0.994
250	1.000	1.000	1.000	1.000
500	1.000	1.000	1.000	1.000
	$\pi_1 = 0$		$\pi_1 = 1$	
(b) *Version with scaling by* σ_n^2	S_n	T_n	S_n	T_n
n = 100	0.799	0.991	0.357	0.661
250	1.000	1.000	0.940	0.959
500	1.000	1.000	0.992	0.997

Note. Results based on 1.000 independent replications, in the case $m = q = 1$, with $w_{1,t} \neq h_{1,t}$, and $\boldsymbol{\zeta}_t = (\varepsilon_t, \upsilon_{1,t}, \nu_{1,t}, \xi_{1,t})' \sim iidN(\mathbf{0}_4, \mathbf{I}_{4,4})$.