Asymptotic null distributions of stationarity and nonstationarity tests under local-to-finite variance errors*

Nunzio Cappuccio
Department of Economics
University of Padova
35123 Padova - Italy
email: nunzio.cappuccio@unipd.it

Diego Lubian
Department of Economics
University of Verona
37129 Verona - Italy
email: diego.lubian@univr.it

This version: 8th September 2003

Abstract

The purpose of this paper is to investigate the asymptotic null distribution of stationarity and nonstationarity tests when the distribution of the error term belongs to the normal domain of attraction of a stable law in any finite sample but the error term is an i.i.d. process with finite variance as $T \uparrow \infty$. This local-to-finite variance setup is helpful to highlight the behavior of test statistics under the null hypothesis in the borderline or near borderline cases between finite and infinite variance and to assess the robustness of these test statistics to small departures from the standard finite variance context. From an empirical point of view, our analysis can be useful in settings where the (non)-existence of the (second) moments is not clear-cut, such as, for example, in the analysis of financial time series. A Monte Carlo simulation study is performed to improve our understanding of the practical implications of the limi theory we develop. The main purpose of the simulation experiment is to assess the size distortion of the unit root and stationarity tests under investigation.

Keywords: Stable distributions, unit root tests, stationarity tests, asymptotic distributions, local-to-finite variance, size distortion.

*Financial support from MIUR (COFIN 2000) is gratefully acknowledged. Paper prepared for the conference on "Linear and non-linear dynamics in time series", Bressanone, June 6-7, 2003. We wish to thank Giuseppe Cavaliere (discussant), Francesca Callegari, Marco Ferrante, Keith Knight and conference participants for comments.

1 Introduction

In the literature on testing for stationarity and unit root a crucial maintained hypothesis concerns the existence of the variance of the error term. Whenever this condition fails the standard asymptotic results are no longer valid, as reported in Hamilton (1994),. In particular, if the distribution of the i.i.d. error term belongs to the normal domain of attraction of a stable law with maximal moment exponent α (Gnedenko and Kolmogorov, 1954; Ibragimov and Linnik, 1971), the relevant asymptotic theory follows from the weak convergence results to (functionals) of α-stable Lévy processes given by Phillips (1990), who specialized the results in Resnick (1986), when the distribution of the error term belongs to the domain of attraction of a stable law.

When the Data Generating Process (hereafter DGP) is a driftless random walk, Chan and Tran (1989) for the i.i.d. case and Phillips (1990) for the dependent errors have developed the appropriate asymptotic theory for time series regressions with a unit root and infinite variance errors. For the random walk with drift Callegari et al. (2003) have shown, for i.i.d. errors, that the functional form of the asymptotic distribution of the least squares estimator and of the t-statistic depends on whether the maximal moment exponent α lies between zero and one, is equal to one or lies between one and two. The asymptotic distribution of additional unit root tests with infinite variance errors has been analyzed by Ahn et al. (2001). As for tests of the null hypothesis of stationarity, Amsler and Schmidt (1999) have studied the asymptotic distribution of the $KPSS$ test of Kwiatkowski et al. (1992) and of the modified rescaled range (MRS) test of Lo (1991).

The purpose of this paper is to investigate the robustness of standard unit root and stationarity tests to small departures from the maintained hypothesis of finite variance. To this end, we follow a "local-to" approach. This methodology, in the spirit of Pitman (1949), is standard for the study of the asymptotic power of test statistics and it entails the specification of a sequence of local alternatives which collapses to the null hypothesis as $T \uparrow \infty$. The asymptotic power of unit root tests has been investigated by, amongst others, Phillips (1987), Perron and Ng (1996) and Nabeya and Perron (1994).

However, in this paper we consider the asymptotic null distribution of selected test statistics when one of the maintained hypothesis, namely the existence of the first or the second moments of the error term, is satisfied only as $T \uparrow \infty$. In particular, we follow Amsler and Schmidt (1999) who assume that the error term of a driftless random walk belongs to the normal domain of attraction of a stable law in any finite sample but has finite variance in the limit as $T \uparrow \infty$. These local departures from the finite variance setup are helpful to highlight the behavior of unit root and stationarity tests in borderline or near borderline cases between finite and infinite variance and to assess the robustness of these statistics to small departures from the standard finite variance context[1].

We believe that this "local-to" approach has several advantages since it provides a link between the nowadays standard limiting distribution in autoregressions with integrated processes and finite variance and those obtained under the infinite variance assumption. More importantly, it allows to investigate analytically the robustness of standard asymptotic inference procedures with respect to the presence of an error term with heavy tails in finite sample. This robustness analysis may be empirically relevant in settings where the (non)-existence of the (second) moments is not clear-cut, such as, for example, in the analysis of financial time series. In fact, it is often argued that financial asset returns can be viewed as the cumulative outcome of a large number of pieces of information and individual decisions (McCulloch, 1996; Rachev and Mittnik, 2000). Since the empirical distribution of financial asset returns is usually found leptokurtic, this suggests to consider non-gaussian stable laws, as first postulated by Maldelbrot in the early 60s[2] (Mandelbrot, 1997). However the empirical evidence in favor of the stable model is not clear-cut (McCulloch, 1997). Therefore, lacking an established empirical evidence in favor or against the stable laws we believe that the local-to-finite variance approach proposed by Amsler and Schmidt (1999) can be useful for improving our understanding of the robustness of unit root and stationarity tests.

The paper is organized as follows. In the next section, after having introduced the "local-to-finite" variance approach, we present some results on the weak convergence of sample moments of a random walk process characterized by "local-to-finite" variance errors. In a Lemma we collect several convergence results on the first and second sample moments useful in our subsequent analysis. In subsection 2.1 we establish the limiting distributions of some unit root tests, whereas in subsection 2.2 we consider some tests of the null hypothesis of stationarity. Section 3 is dedicated to a MonteCarlo simulation study to assess the finite sample size distortion of stationarity and nonstationarity tests when the researcher erroneously makes use

[1] A similar approach has been used to assess the robustness of inferential procedures in cointegrating regressions when regressors are near-integrated (Elliott, 1998).

[2] See also Loretan and Phillips (1994) for an analysis on the existence of finite moments in financial time series.

of the "standard" critical values (i.e. those under finite variance) even though the error process has infinite variance as in section 2. All proofs are collected in the Appendix.

2 Asymptotic distribution under local-to-finite variance

To build intuition for the local-to-finite variance setup, let us assume that the process u_t is a weighted sum of two independent processes[3]:

$$u_t = v_{1t} + z_t v_{2t}$$

where v_{1t} is i.i.d. with zero mean and finite variance σ_1^2, v_{2t} is also i.i.d., symmetrically distributed with distribution belonging to the normal domain of attraction of a stable law with characteristic exponent α, with $0 < \alpha < 2$, denoted as $v_{2t} \in \mathcal{ND}(\alpha)$.[4] As for the weight, several specification are of interest. One possibility is to consider the Bernoulli random variable $z_t \sim \mathrm{B}(1, p)$, mutually independent on v_{1t} and v_{2t}. Intuitively, when p is made suitably small, the process u_t is, from time to time, hit by a realization from an infinite variance distribution. Loosely speaking, since the probability of "extreme" realizations is bigger when drawing from random variables whose distribution belongs to the normal domain of attraction of a stable law than for random variables whose density has finite variance, this model may be helpful in explaining outliers occurrence in time series. This intuition is made rigorous in Appendix B where we show that the tail behavior of the distribution function the process u_t is of the Pareto-Lévy form, with characteristic exponent α. However, under this specification we have infinite variance both in finite samples and asymptotically.

Here we follow an alternative specification, proposed in Amsler and Schmidt (1999), which maintain the infinite variance in finite samples but collapses to the

[3]This is the socalled innovative outlier model.

[4]Necessary and sufficient condition for v_{2t} to belong to the normal domain of attraction of a stable law with characteristic exponent α is that the tail behavior of its distribution function has asymptotically the Pareto-Lévy form,

$$\begin{aligned} \Pr(v_{2t} < h) &= (c_1 a^\alpha + \alpha_1(h)) \frac{1}{|h|^\alpha} \qquad h < 0 \\ \Pr(v_{2t} < h) &= 1 - (c_2 a^\alpha + \alpha_2(h)) \frac{1}{|h|^\alpha} \qquad h > 0 \end{aligned}$$

where c_1 and c_2 are non-negative constants such that $c_1 + c_2 = 1$, both $\alpha_1(h) \to 0$ and $\alpha_2(h) \to 0$ as $|h| \to \infty$, and the constant a is a scale parameter (Ibragimov and Linnik, 1971, pg. 92).

standard finite variance assumption asymptotically. In particular, we assume that the process u_t is generated according to the following mechanism

$$u_t = v_{1t} + \frac{\gamma}{aT^{1/\alpha - 1/2}} v_{2t}. \tag{1}$$

so that u_t exhibits infinite variance in any finite sample size but finite variance in the limit as T approaches infinity. Notice that the importance of the stable component diminishes as the sample size grows but at a slower rate as α increases. Thus, for a given γ, when α is close to 2 we need a large sample size to annihilate the stable component whereas when α is less than 1 a relatively small sample size is required.

By Donsker's theorem (Billingsley (1968)), we have $T^{-1/2} \sum_{t=1}^{[Tr]} v_{1t} \Rightarrow \sigma_1 W(r)$, where $\Rightarrow$ stands for the weak convergence of probability measures, and $W(r)$ is the standard Wiener process. Under the above assumptions on the sequence v_{2t}, from Resnick (1986) and Phillips (1990), we also have the following weak convergence result

$$\left(\frac{1}{a_T} \sum_{t=1}^{[Tr]} v_{2t}, \frac{1}{a_T^2} \sum_{t=1}^{[Tr]} v_{2t}^2 \right) \Rightarrow \left(U_\alpha(r), V(r) \right), \tag{2}$$

where the norming sequence is given by $a_T = aT^{1/\alpha}$, $U_\alpha(r)$ is the Lévy α-stable process on the space $D[0,1]$ and $V(r)$ is its quadratic variation process $V(r) = [U_\alpha, U_\alpha]_r = U_\alpha^2(r) - 2\int_0^r U_\alpha^- \mathrm{d}U_\alpha$ (see Protter, 1990, pg. 58, Phillips, 1990, eq. (11)). Here, $U_\alpha^-(r)$ stands for the left limit of the process $U_\alpha(\cdot)$ in r.

In order to investigate the asymptotic distribution of the test statistics of interest (to be described below) when the error term is given by (1), it is convenient to obtain beforehand some convergence results concerning sample moments and partial sums of the local-to- finite variance error term . These convergence results are collected in the following Lemma whose proof can be found in the Appendix.

Lemma 2.1 *Let u_t be generated as in (1) with $v_{1t} \sim$ i.i.d.$(0, \sigma_1^2)$ and $v_{2t} \sim$ i.i.d. and $v_{2t} \in \mathcal{ND}(\alpha)$, and let $y_t = \sum_{j=1}^t u_j$, then as $T \uparrow \infty$,*

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} u_t \Rightarrow \sigma_1 W(r) + \gamma U_\alpha(r) \equiv Z_{\alpha,\gamma}(r)$$

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} (u_t - \bar{u}) \Rightarrow Z_{\alpha,\gamma}(r) - r Z_{\alpha,\gamma}(1) \equiv \tilde{Z}_{\alpha,\gamma}(r)$$

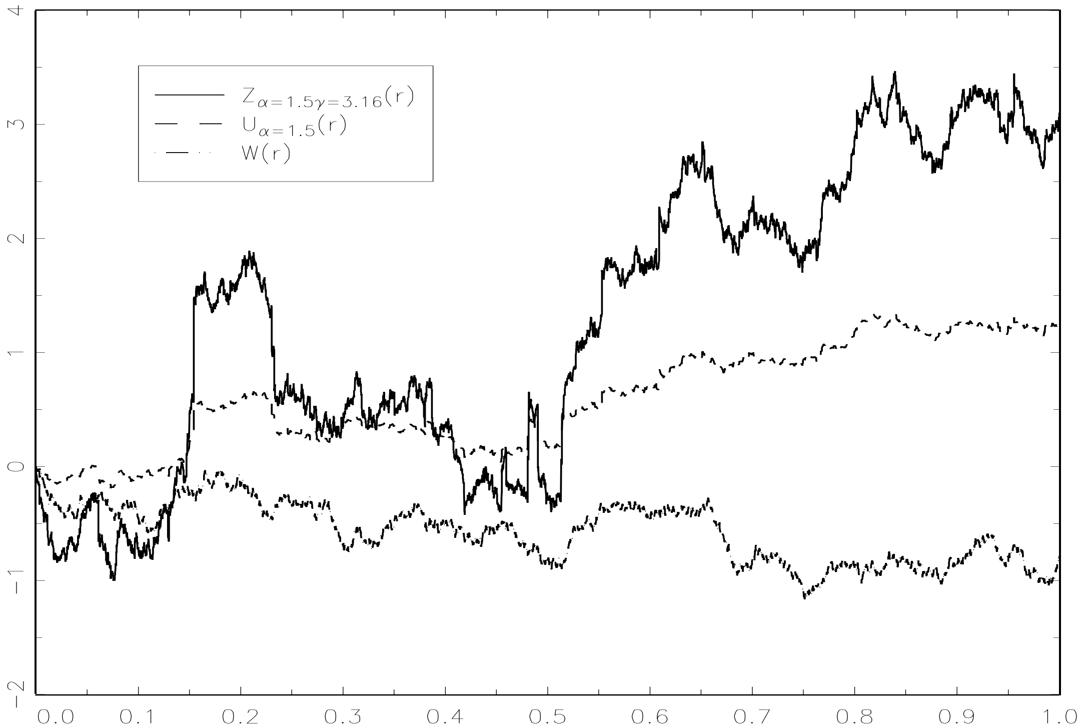

Figure 1: Sample trajectories of $W(r)$, $U_\alpha(r)$ and $Z_{\alpha,\gamma}(r)$, $\alpha = 1.5$, $\gamma = 3.16$.

$$\frac{1}{T}\sum_{t=1}^{[Tr]} u_t^2 \Rightarrow \sigma_1^2 r + \gamma^2 V(r) \equiv K_\gamma(r), \quad \frac{1}{T^{3/2}}\sum_{t=1}^{[Tr]} y_t \Rightarrow \int_0^r Z_{\alpha,\gamma}(r)$$

$$\frac{1}{T^2}\sum_{t=1}^{[Tr]} y_t^2 \Rightarrow \int_0^r Z_{\alpha,\gamma}^2(r), \quad \frac{1}{T}\sum_{t=1}^{[Tr]} y_{t-1}u_t \Rightarrow \int_0^r Z_{\alpha,\gamma}\mathrm{d}Z_{\alpha,\gamma}$$

$$\frac{1}{T^{3/2}}\sum_{t=1}^{T} tu_t \Rightarrow Z_{\alpha,\gamma}(1) - \int_0^1 Z_{\alpha,\gamma}, \quad \frac{1}{T^{5/2}}\sum_{t=1}^{T} ty_t \Rightarrow \int_0^1 rZ_{\alpha,\gamma}$$

As expected, two are the key parameters affecting these asymptotic distributions: α, the maximal moment exponent characterizing the Lévy process $U_\alpha(r)$, and γ, which provides the relative importance of the Lévy process in the limiting distributions. In short, α select the Lévy process while γ tells how important it is. The interaction between these two parameters will determine how much these limiting distributions will differ from those under standard assumptions.

To better understand the asymptotic results in the Lemma 2.1 some graphs are reported. Figure 1 presents one sample trajectory of the three processes $W(r)$, $U_\alpha(r)$ and $Z_{\alpha,\gamma}(r)$ for $\alpha = 1.5$ and $\gamma = 3.16$; it is evident the effect of the outliers in the trajectory of $U_\alpha(r)$, which is reflected in that of $Z_{\alpha,\gamma}(r)$. Replication of these trajectories permits to estimate the density of each process for a given r; these

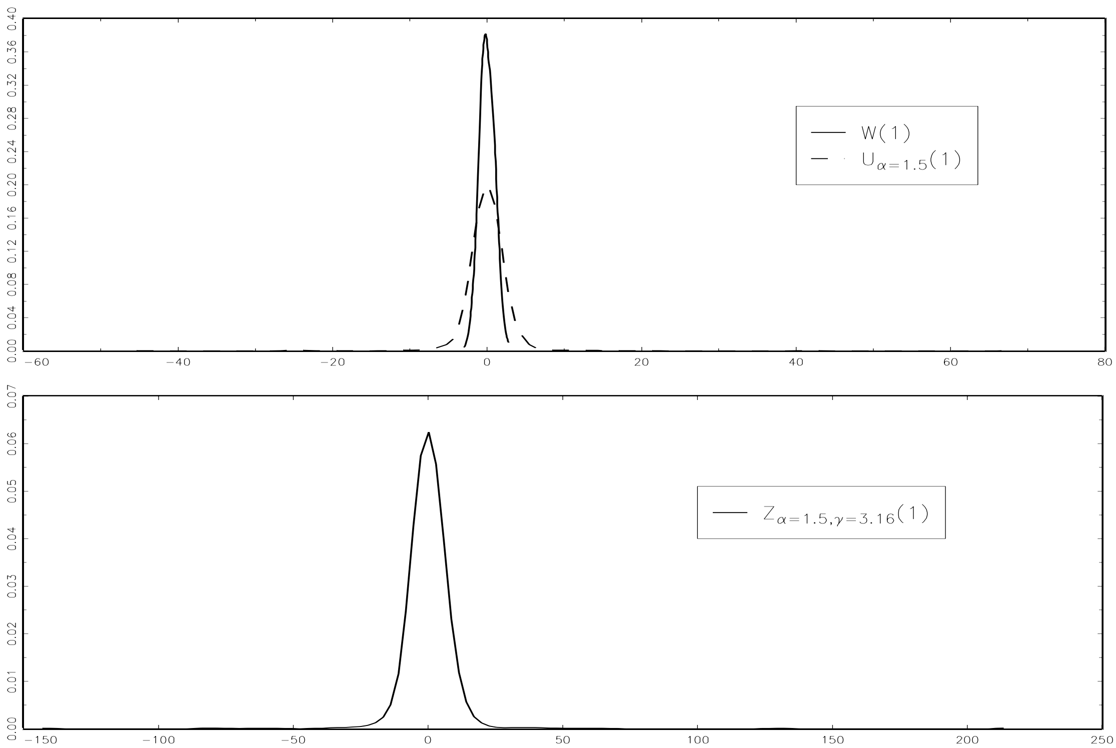

Figure 2: Estimated densities of $W(1)$, $U_\alpha(1)$ and $Z_{\alpha,\gamma}(1)$, $\alpha = 1.5$, $\gamma = 3.16$.

densities[5] are showed in Figure 2 for $r = 1$.

Figure 3 presents a nonparametric estimate of the empirical density of the limiting random variables $\int_0^1 W$, $\int_0^1 U_\alpha$ and $\int_0^1 Z_{\alpha,\gamma}$ for the $\alpha = 1.5$ and $\gamma = 3.16$ case. Finally, Figure 4 reports nonparametric estimates of the empirical density of the limiting random variable $\int_0^1 Z_{\alpha,\gamma}$ for some values of γ and $\alpha = 1.5$.

Throughout the paper we make the assumption that both v_{1t} and v_{2t} are i.i.d. error processes even though one might find it desirable to consider serially dependent errors, such as linear processes[6]. This extension is left to future research because we prefer to concentrate our efforts in the assessment of the robustness of test statistics in non-standard but neat settings and we do not want that our conclusions can be affected by some other factors such as the lag length selection in augmented Dickey-Fuller tests or the consistent estimation of the "long-run" variance.

[5]The nonparametric estimate of all densities is computed by kernel smoothing, with Epanechnikov and bandwidth as suggested in Silverman (1986).

[6]See Phillips (1990) for a treatment of unit root tests when v_{2t} is a linear process.

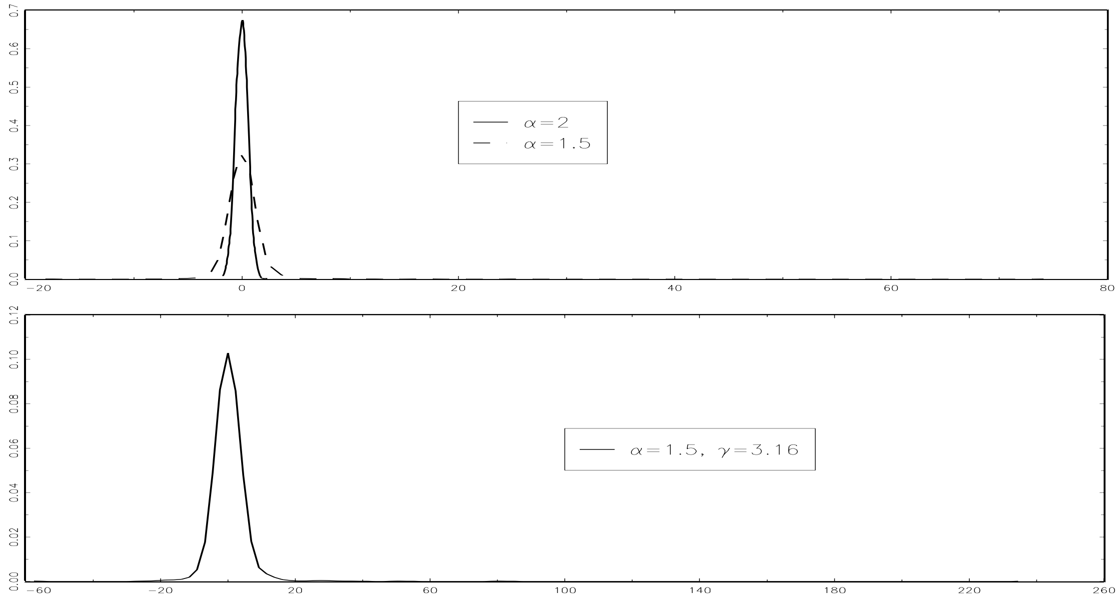

Figure 3: Nonparametric estimate of the empirical density of $\int_0^1 W$, $\int_0^1 U_\alpha$ and $\int_0^1 Z_{\alpha,\gamma}$, $\alpha = 1.5$, $\gamma = 3.16$.

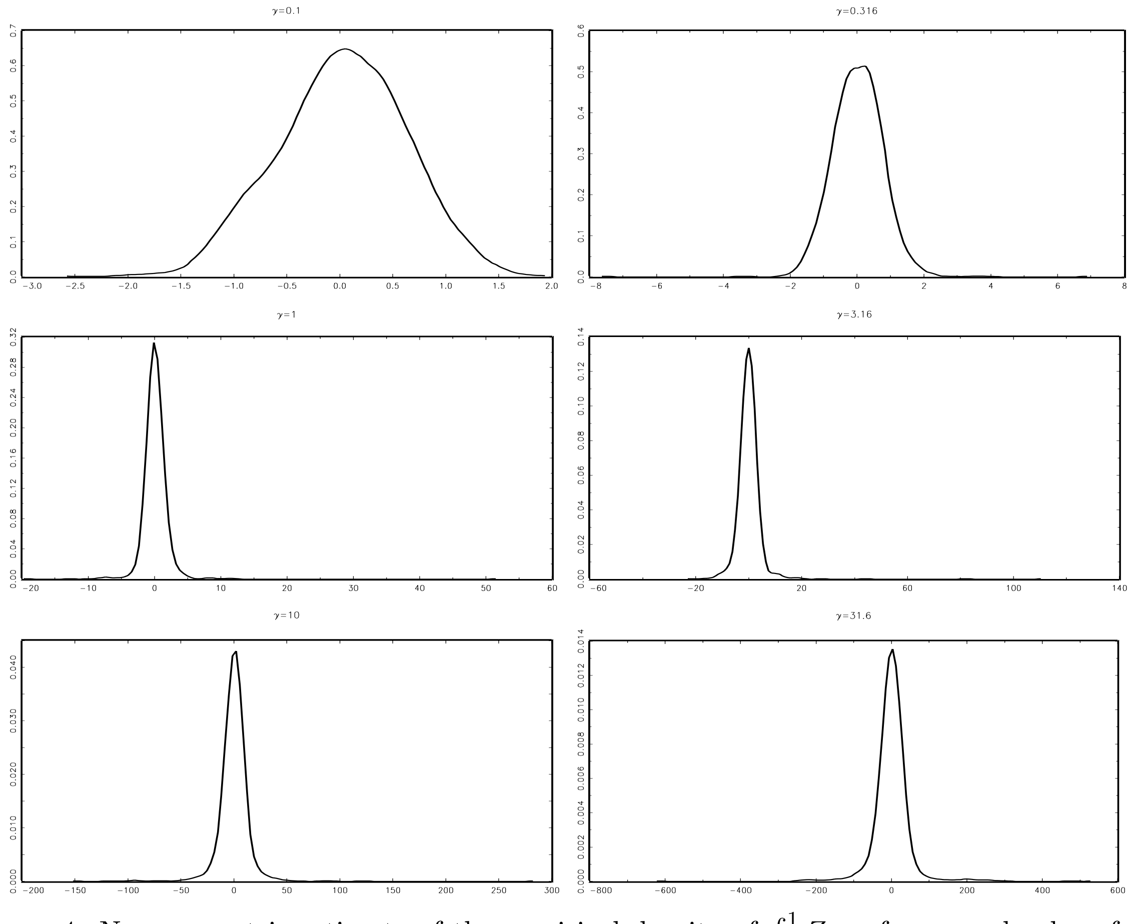

Figure 4: Nonparametric estimate of the empirical density of $\int_0^1 Z_{\alpha,\gamma}$, for several values for γ and $\alpha = 1.5$.

2.1 Unit root tests

In order to investigate unit root tests, we assume that $\{y_t\}$ is generated according to the mechanism

$$y_t = \rho y_{t-1} + u_t, \qquad t = 1, \ldots, T \tag{3}$$

with $\rho = 1$ and that the initial condition y_0 is any random variable.

We consider several well-known test statistics for testing the null hypothesis $H_{DS} : \rho = 1$ in (3) against the alternative hypothesis $|\rho| < 1$. First, we study two nowadays standard test statistics proposed by Dickey and Fuller (1976), the $T(\hat{\rho}-1)$ and the t-ratio statistics. We also consider the Lagrange Multiplier test (hereafter LM) proposed by Ahn (1993), and the well-known Durbin-Watson (DW) test. The interest in the DW test stems from the optimality properties of the test statistic in the first-order autoregressive model with i.i.d. Gaussian errors. In fact, Sargan and Bhargava (1983) and Bhargava (1986) show that the DW test statistics can be used for constructing uniformly most powerful tests of the null hypothesis of a random walk against stationary alternatives in a driftless and with drift DGP, respectively. Thus, besides $T(\hat{\rho} - 1)$, our test statistics for the null of a unit root are given by

$$t_{\hat{\rho}} = \left(\sum_{t=2}^{T} y_{t-1}^2 \right)^{1/2} (\hat{\rho} - 1)/s \tag{4}$$

$$LM = \frac{\left(\sum_{t=2}^{T} (y_t - y_{t-1}) y_{t-1} \right)^2}{\bar{s}^2 \sum_{t=2}^{T} y_{t-1}^2} \tag{5}$$

$$DW = \frac{\sum_{t=2}^{T} (y_t - y_{t-1})^2}{\sum_{t=2}^{T} y_{t-1}^2} \tag{6}$$

where $\hat{\rho}$ is the OLS estimator of ρ given by $\hat{\rho} = \left(\sum_{t=2}^{T} y_{t-1}^2 \right)^{-1} \sum_{t=2}^{T} y_t y_{t-1}$, $s^2 = T^{-1} \sum_{t=2}^{T} (y_t - \hat{\rho} y_{t-1})^2$ and $\bar{s}^2 = \sum_{t=2}^{T} (y_t - y_{t-1})^2 / T$.

The limiting behavior of the above test statistics under the local-to-finite variance setup is summarized in the following theorem whose proof is omitted, but available upon request, since it follows directly by Lemma 2.1 and repeated application of the continuous mapping theorem

THEOREM 2.2 *When y_t is generated according to (1) and (3), under the null hy-*

pothesis $H_{DS}:\rho=1$ *and as* $T\uparrow\infty$*, we have*

$$T(\hat{\rho}-1) \;\Rightarrow\; \frac{\int_0^1 Z_{\alpha,\gamma}\mathrm{d}Z_{\alpha,\gamma}}{\int_0^1 Z_{\alpha,\gamma}^2} \tag{7}$$

$$t(\hat{\rho}) \;\Rightarrow\; \frac{\int_0^1 Z_{\alpha,\gamma}\mathrm{d}Z_{\alpha,\gamma}}{\left(K_\gamma(1)\gamma^2V(1)\int_0^1 Z_{\alpha,\gamma}^2\right)^{1/2}} \tag{8}$$

$$LM \;\Rightarrow\; \frac{\left(\int_0^1 Z_{\alpha,\gamma}\mathrm{d}Z_{\alpha,\gamma}\right)^2}{K_\gamma(1)\int_0^1 Z_{\alpha,\gamma}^2} \tag{9}$$

$$TDW \;\Rightarrow\; \frac{K_\gamma(1)}{\int_0^1 Z_{\alpha,\gamma}^2} \tag{10}$$

It is noticeable that, notwithstanding asymptotically the process u_t has finite variance, the limiting distribution of the unit root test statistics is a (complicated) function of both the Wiener process $W(r)$ and the Lévy α-stable process $U_\alpha(r)$. In contrast with the asymptotic distributions available in the infinite variance case (see Ahn et al., 2001), here they depend not only on the maximal moment exponent α but also on the nuisance parameters σ_1^2 and γ. The role played by $U_\alpha(r)$ in shaping the asymptotic distribution of the test statistics depends on the magnitude of the weight γ. Of course, the limit distribution of the test statistics collapse to the standard ones as $\gamma\to 0$.

2.2 Stationarity tests

Following Kwiatkowski et al. (1992) let us assume that the observable time series y_t is generated according to

$$y_t = d_t + r_t + u_t \qquad t=1,\ldots,T \tag{11}$$

$$r_t = r_{t-1}+\eta_t \tag{12}$$

where $d_t=\delta' x_t$ depends on the unknown coefficients δ of the (known) deterministic components, typically a constant and a linear time trend, u_t is generated as in (1) and $\eta_t\sim$ i.i.d.$(0,\sigma_\eta^2)$. In DGP (11)-(12) the null hypothesis of stationarity is specified as $H_S:\sigma_\eta^2=0$. For simplicity, we restrict ourselves to the level stationarity case, i.e. we set $x_t=1$.

The first set of stationarity tests look at some measure of "magnitude" of the cumulated sums of the residual series obtained by demeaning or detrending the observable time series. If the observable process is stationary it does have finite mean, finite variance and it cannot grow without bounds. On the other hand, a unit root process has trending variance so that its fluctuations are much larger than those of a stationary process. This suggests that test statistics based on some measure of the fluctuations in the time series might be useful in deciding between stationarity and nonstationarity by rejecting the null hypothesis of stationarity whenever the time series fluctuates too much bewildering.

Let e_t be the residuals of a regression of the observable time series y_t on a constant, namely $e_t = y_t - \bar{y}$, define the cumulative process $S_t = \sum_{j=1}^{t} e_j$ and the estimated variance $\hat{\sigma}_e^2 = T^{-1} \sum_{t=1}^{T} e_t^2$, we consider the tests

$$KPSS = \frac{1}{\hat{\sigma}_e^2} \frac{1}{T^2} \sum_{t=1}^{T} S_t^2 \tag{13}$$

$$MRS = \frac{1}{\sqrt{T}\hat{\sigma}_e} \left(\max_t S_t - \min_t S_t \right) \tag{14}$$

$$KS = \max_{k=1,\dots,T} \frac{k}{\hat{\sigma}_e T^{1/2}} \left| \frac{S_k}{k} - \frac{S_T}{T} \right| \tag{15}$$

where the $KPSS$ is due to Kwiatkowski et al. (1992), the Modified Range Statistic, MRS, has been proposed by Lo (1991) and the KS test, a Kolmogorv-Smirnov test, is in Xiao (2001).

A different strategy to test the null of stationarity is based on the Lagrange Multiplier principle as proposed by Choi (1994), Choi and Ahn (1999) and Choi and Yu (1997). These authors consider the following DGP

$$y_t = d_t + r_t \tag{16}$$

and the null hypothesis is $r_t \sim I(0)$. Defining the cumulated sum $C_t = \sum_{i=1}^{t} y_i$, they show that the above null hypothesis is equivalent to $\beta = 1$ and $y_t \sim I(0)$ in the model

$$C_t = \beta C_{t-1} + y_t$$

Writing the log-likelihood for this DGP under Gaussian error, they obtain LM tests, according to how the estimator of the information matrix is chosen

$$LM_1 = \frac{1}{\hat{\omega}^4} \left(\frac{1}{T} \sum_{t=1}^{T} Q_{t-1} \Delta Q_t \right)^2 \tag{17}$$

$$LM_2 = \frac{1}{\hat{\omega}^2} \frac{\left(\sum_{t=1}^{T} Q_{t-1} \Delta Q_t \right)^2}{\sum_{t=1}^{T} Q_{t-1}^2} \tag{18}$$

where Q_t are the residuals of a regression of C_t on the trend variable t, $\hat{\omega}^2 = T^{-1}\sum_{t=1}^{T}\Delta Q_t^2$ and Δ is the first-difference operator.

We also consider the Sargan-Barghava-Durbin-Hausman ($SBDH$) test given by

$$SBDH = \frac{1}{\hat{\omega}^2}\frac{1}{T^2}\sum_{t=1}^{T} Q_t^2 \tag{19}$$

This statistic is clearly a Durbin-Watson test on the residuals Q_t and can be interpreted as a stationarity test for the observable y_t when the rejection region is the right tail of the distribution (Stock, 1994).

The asymptotic distributions of the above test statistics are summarized in the following Theorem (see the Appendix for the proof).

THEOREM 2.3 *Let y_t be generated as in (11)-(12) and u_t be as in (1), then under the null hypothesis $\sigma_\eta^2 = 0$ and as $T \uparrow \infty$,*

$$KPSS \Rightarrow \frac{\int_0^1 \tilde{Z}_{\alpha,\gamma}^2}{K_\gamma(1)} \tag{20}$$

$$MRS \Rightarrow \frac{1}{K_\gamma(1)^{1/2}}\left(\sup_r \tilde{Z}_{\alpha,\gamma}(r) - \inf_r \tilde{Z}_{\alpha,\gamma}(r)\right) \tag{21}$$

$$KS \Rightarrow \frac{1}{\sqrt{K_\gamma(1)}}\sup_r |\tilde{Z}_{\alpha,\gamma}(r)| \tag{22}$$

$$LM_1 \Rightarrow \frac{1}{K_\gamma(1)}\left(\int_0^1 \bar{Z}_{\alpha,\gamma}\mathrm{d}\bar{Z}_{\alpha,\gamma}\right)^2 \tag{23}$$

$$LM_2 \Rightarrow \frac{1}{\sqrt{K_\gamma(1)}}\frac{\left(\int_0^1 \bar{Z}_{\alpha,\gamma}\mathrm{d}\bar{Z}_{\alpha,\gamma}\right)^2}{\int_0^1 Z_{\alpha,\gamma}^2 - 3\left(\int_0^1 rZ_{\alpha,\gamma}\right)^2} \tag{24}$$

$$SBDH \Rightarrow \frac{1}{\sqrt{K_\gamma(1)}}\left(\int_0^1 Z_{\alpha,\gamma}^2 - 3\left(\int_0^1 rZ_{\alpha,\gamma}\right)^2\right) \tag{25}$$

where $\bar{Z}_{\alpha,\gamma} = Z_{\alpha,\gamma} - 3r\int_0^1 rZ_{\alpha,\gamma}\mathrm{d}r$ and $\mathrm{d}\bar{Z}_{\alpha,\gamma} = \mathrm{d}Z_{\alpha,\gamma} - 3\int_0^1 rZ_{\alpha,\gamma}\mathrm{d}r$.

As for the unit root test, the limiting distribution of stationarity test is a complicated function of the compound process $Z_{\alpha,\gamma} = \sigma_1 W(r) + \gamma U_\alpha(r)$ and of its quadratic variation $K_\gamma(1)$. These asymptotic distribution are also affected by the maximal moment exponent and the weight γ. Once again the relative importance of the Wiener and stable components depends on the size of weight attached to the infinite variance component.

3 Finite sample size

In order to improve our understanding of the practical implications of Theorems 2.2 and 2.3, we carry out a MonteCarlo experiment whose main purpose is to investigate the size distortion of the unit root and stationarity test under local-to-finite variance. In the experiment we set $a = \sigma_1^2 = 1$, $T = \{100, 1000, 10000\}$ and $\alpha = \{1.5, 1, 0.5\}$. Moreover, as in Amsler and Schmidt (1999), we consider the following values for γ

$$\gamma = \{0.1, 0.316, 1, 3.16, 10, 31.6\}$$

where $3.16 \approx \sqrt{10}$. The MonteCarlo experiment has been carried out using Gauss 5 and the number of replications N has been set to 20,000. We have 18 parameter combinations for each sample size and three different sample sizes amounting to a total of 54 experiments. For each experiment and in each replication, we simulate a driftless random walk with the error term generated as in (1), calculate each test statistic and store their values. Thus, we end up with a sample of 20,000 values of each test statistic for each of the 54 experiments.

The effective size of the tests, when the nominal size is fixed at 5% and the critical values for the finite variance case are used[7], is reported in tables 1 and 2 for the unit root and stationarity tests, respectively.

With regard to unit root tests we have that as expected, for a given sample size, the effective size worsens as the stable component becomes more important, i.e. as γ increases. In general, these tests have effective size smaller than the reference 5% nominal size. This leads to fewer rejections than admissible under the probability of type I error chosen which in turns makes them conservative tests. A noticeable exception is the behavior of the DW test for any value of γ, $\alpha = 0.5$ and $T = 100$ and for the smallest value of γ when $T = 1000$. On the other hand and for all tests considered, for a given value of γ, the effective size is closer to the nominal size as α increases. The DW test is little sensitive both to increases in α and in γ displaying a rather constant behavior across all simulation experiments. This good performance of the DW test is accordance with previous simulation experiments in standard settings, see Stock (1990). The $T(\hat{\rho} - 1)$, $t(\hat{\rho})$, and LM tests have close patterns of effective size with little differences in relative performance, even though the $T(\hat{\rho} - 1)$ test has somehow less severe size distortion.

Analogously, the effective size of stationarity tests at the 5% nominal level, for a

[7] The 5% critical values have been obtained for all tests by simulation of the case $\gamma = 0$, $T = 100,000$ and 20,000 replications .

Table 1: Effective size (in %) with 5% nominal size for nonstationarity tests under (1)

		γ						γ					
		0.1	0.316	1	3.16	10	31.6	0.1	0.316	1	3.16	10	31.6
T	α	$T(\hat{\rho}-1)$ Left tail						$t(\hat{\rho})$ Left tail					
	0.5	4.38	3.76	3.23	2.69	2.32	2.02	4.64	3.92	3.34	2.58	2.21	1.88
100	1.0	4.96	4.53	4.14	3.15	2.92	3.31	5.44	4.98	4.37	3.25	2.98	3.34
	1.5	5.10	4.86	4.50	4.06	3.94	4.14	5.50	5.19	4.79	4.15	4.13	4.38
	0.5	4.37	3.68	3.14	2.73	2.21	2.14	4.39	3.76	3.06	2.60	2.02	1.93
1000	1.0	5.03	4.78	3.88	3.02	3.28	3.38	5.04	4.86	3.91	2.98	3.16	3.23
	1.5	5.26	4.92	4.61	4.09	4.04	4.02	5.39	5.00	4.76	4.11	4.06	4.04
	0.5	4.38	3.80	3.16	2.54	2.36	1.93	4.37	3.74	3.13	2.41	2.17	1.80
10000	1.0	4.92	4.77	3.72	3.41	3.22	2.94	4.98	4.75	3.69	3.31	3.11	2.77
	1.5	5.09	4.97	4.44	4.41	4.38	3.83	5.13	5.03	4.44	4.38	4.35	3.84
		LM Right tail						DW Right tail					
	0.5	4.25	3.68	3.02	2.26	1.93	1.59	5.41	5.29	5.32	5.26	5.35	5.34
100	1.0	4.78	4.52	3.81	2.92	2.52	2.86	5.29	5.26	5.10	4.64	4.62	5.05
	1.5	5.04	4.81	4.46	3.80	3.67	3.76	5.37	5.16	5.00	4.64	4.79	4.99
	0.5	4.40	3.55	3.01	2.38	1.88	1.71	5.26	5.04	4.86	5.02	5.04	5.00
1000	1.0	4.90	4.71	3.75	2.86	2.81	2.94	5.37	5.32	4.84	4.56	4.73	4.60
	1.5	5.25	4.89	4.50	3.89	3.78	3.84	5.51	5.05	4.94	4.75	4.56	4.81
	0.5	4.18	3.70	2.99	2.28	2.08	1.60	5.24	5.24	5.04	4.87	5.03	4.65
10000	1.0	4.86	4.74	3.52	3.12	2.83	2.61	5.19	5.31	4.71	4.75	4.72	4.56
	1.5	5.14	4.96	4.36	4.20	4.11	3.62	5.36	5.33	5.01	5.08	5.03	4.71

given α decreases as γ increases and, for a given γ increases with α. These findings are not surprising since both γ and α govern the behavior of the test statistics under the sequence of local-to finite variances. Considering the 5% nominal size the $KPSS$ and $SBDH$ have effective size closer to the nominal level. The LM_1 test displays effective sizes larger than the nominal ones for the sample size $T = 100$, being a liberal test in this case. The effective size decreases as the sample size grows and it adjusts around the true nominal level as $T = 10,000$. The MRS test statistics has the worst behavior with smallest effective size amongst all test statistics. These findings are in accordance with the simulation results in Amsler and Schmidt (1999). The KS test exhibits a slightly better behavior which is clearly dominated by the remaining test statistics. The rather bad behavior of the KS and MRS test statistics can be rationalized by noticing that these tests consider the maximum or the difference between the maximum and the minimum and that these values may be the most sensitive to the presence of the outliers induced by the infinite variance error terms. The performance of the LM_2 test is in between the $KPSS$, LM_1 and $SBDH$, on the one hand, and the MRS and KS, on the other hand, even tough it is not negatively affected by small sample sizes as the LM_1 test.

So far, by looking at the size distortion we have compared the distributions under finite variance and the local-to-finite variance at the 5% percentile for the $T(\widehat{\rho}-1)$ and $t(\widehat{\rho})$ statistics and at the 95% percentile for all other tests. A better understanding of the differences between these distributions can be achieved using graphical methods. In particular, following Davidson and MacKinnon (1998) we use the P-value discrepancy plots which are built as follows. For each of the $j = 1, 2, \ldots, 20,000$, realizations of the test statistics, we compute its P-value, say p_j (using the distribution under the finite variance case). Then, we estimate the empirical distribution function of the P-values, at m points, as

$$\hat{F}(r_i) = \frac{1}{N}\sum_{j=1}^{N} I(p_j \leq r_i)$$

where $I(\cdot)$ is the indicator function and

$$r_i = .001, .002, \ldots, .010, .015, \ldots, .990, .991, \ldots, .999 \qquad (m = 215).$$

When plotted against r_i, $\hat{F}(r_i)$ should be close to the 45^o line. This is the so called P-value plot. Instead, we plot $\hat{F}(r_i) - r_i$ against r_i, which should result in a horizontal line with zero intercept, which is just the deviation of the actual size from the nominal size obtaining the so-called P-value discrepancy plots.

Table 3 reports P-value discrepancies for all test statistics, $\alpha = 1.5$, two values of γ (medium and large) and for three different nominal levels, 1%, 5% and 10 per cent. The conservative behavior of all test statistics testified by negative discrepancies is evident, apart from the DW test and the LM_1 test for small and medium sample sizes. This feature tends to be more pronounced as the nominal size increases and/or γ increases. Two facts are noteworthy: first, in most cases the P-value discrepancies are small in magnitude, usually much less than 1%, and second, the MRS and KS (and the LM_1 for small sample sizes) have large P-value discrepancies casting consistent doubts on their robustness to the kind of local departures from finite variance we are considering.

In Figure 5 and 6 we graph these P-values discrepancies for our battery of unit root (top panel) and stationarity (bottom panel) tests for a sample size of $T = 1000$ and $\alpha = 1.5$, which may be the most relevant case when working with financial time series, and for $\gamma = 31.6$ and $\gamma = 1$, a large and a moderately small weight on the infinite variance component, respectively. From the top panels of Figure 5 and 6, it is clear how the DW test outperforms other unit root tests both at the 5% nominal size and at all other significance level the researcher might choose. It is also clear that the LM test has the worst discrepancy while the $T(\hat{\rho} - 1)$ and $t(\hat{\rho})$ moves very closely. Of course, one should also notice the negative effect of increasing γ on the P-values discrepancies.

As for the bottom panels of figure 5 and 6, first one notices the striking bad behavior of the KS and MRS. A look at the behavior of the LM_2 test is instructive of the kind of information one is able obtain from P-values discrepancy plots. In fact, from both figures we notice that the graph of the LM_2 test is very close to those of the $KPSS$, $SBDH$, and LM_1 tests at the 5% nominal level. However, the P-value discrepancy of the LM_2 test differs remarkably from those of the above mentioned tests as the nominal size increases. Thus, the LM_2 tests has a tendency to under-reject at all nominal sizes while $KPSS$, $SBDH$, and LM_1 do not display such a behavior.

4 Conclusions

In this paper we have investigated the null distribution of several stationarity and nonstationarity tests when the maintained hypotheses of finite variance is almost satisfied. Considering the local-to-finite variance approach suggested by Amsler and Schmidt (1999) we establish the limiting null distributions of the test statistics and remark that they depend on the maximal moment exponent and on the weight attached to the stable component. Simulation results on the empirical size of the test statistics indicate clearly that some test are more sensitive than others to a local departure from the maintained hypothesis of finite variance and allow us to rank the test statistics according to their empirical size distortion. Our simulation results suggest that using the DW statistics when testing for a unit root and the $KPSS$ or the $SBDH$ statistics when the null is the stationarity one is not likely to induce significant size distortions. Therefore, when one is uncertain about the presence of a stable error term or does not want to rely too much in the estimated

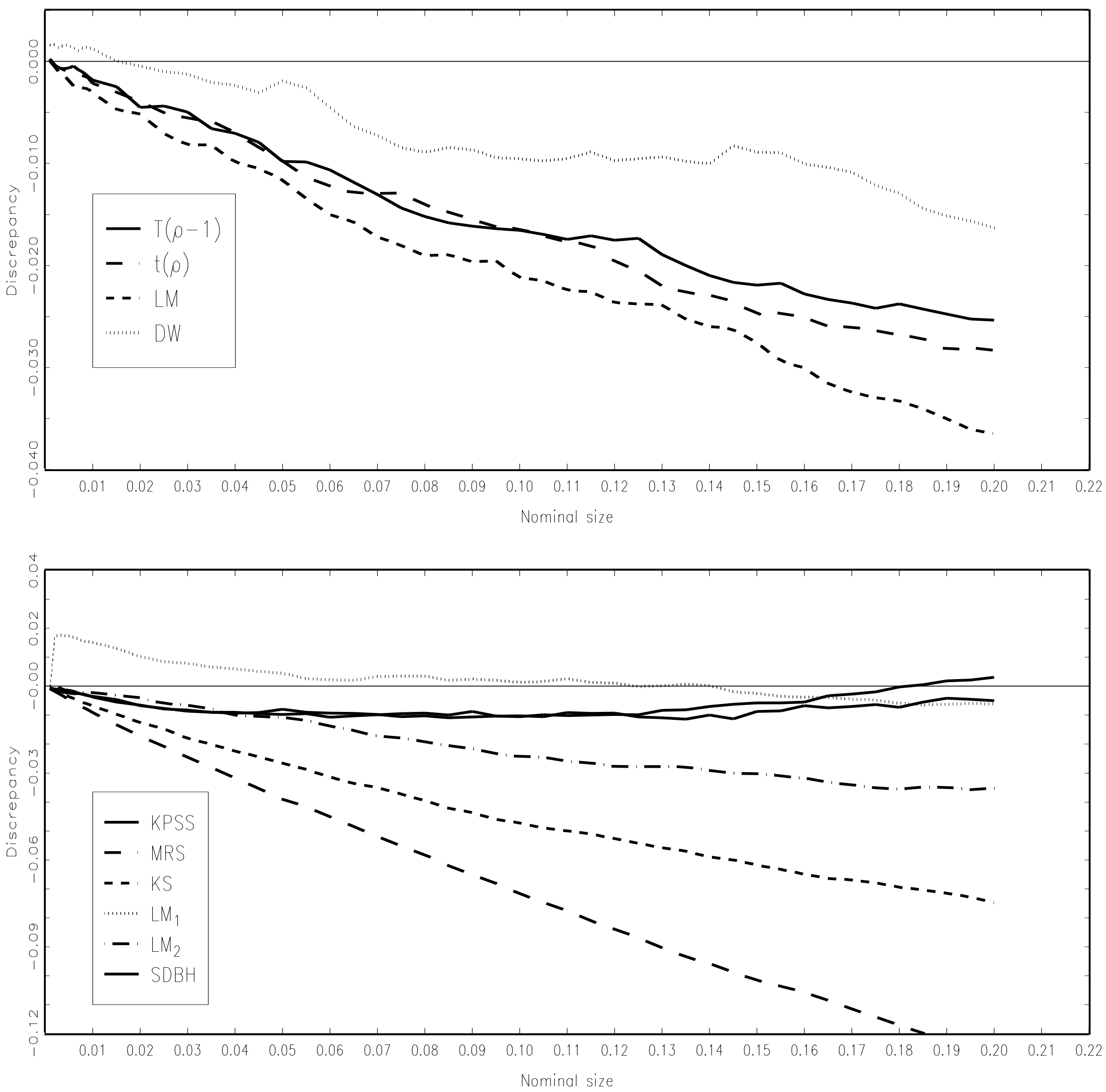

Figure 5: Discrepancy between effective and nominal size for all tests under local to finite variance, unit root tests in the upper window and stationarity tests in the lower window, $T = 1000$, $\alpha = 1.5$, $\gamma = 31.6$

value of the maximal moment exponent, a reasonable strategy could be to use the standard critical values under finite variance.

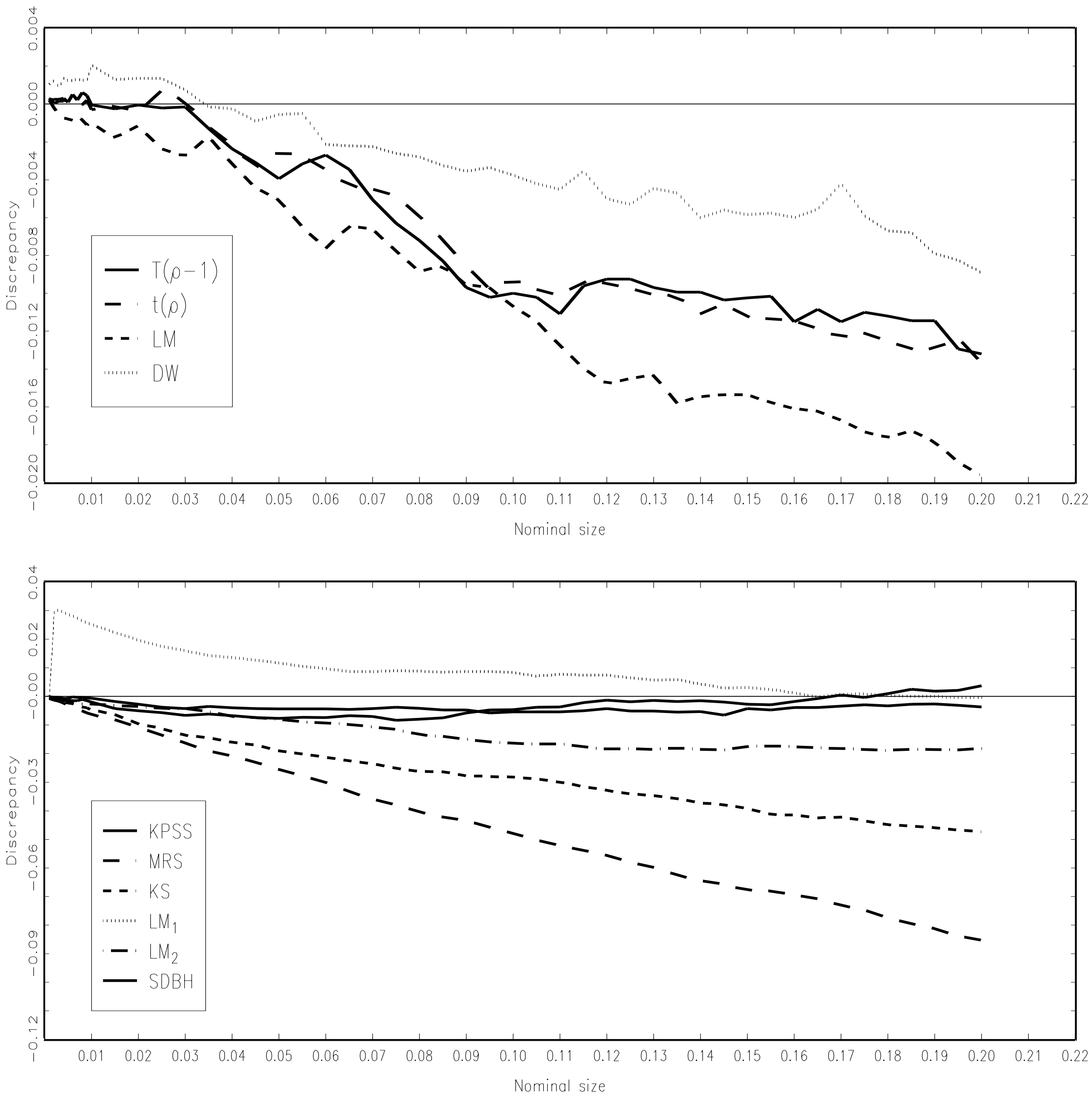

Figure 6: Discrepancy between effective and nominal size for all tests under local to finite variance, unit root tests in the upper window and stationarity tests in the lower window, $T = 1000$, $\alpha = 1.5$, $\gamma = 1$

A Appendix

A.1 Proof of Lemma 2.1

We begin by establishing the weak convergence of $T^{-1/2}y_{[Tr]}$. We have

$$\begin{aligned}\frac{1}{\sqrt{T}}y_{[Tr]} &= \frac{1}{\sqrt{T}}\sum_{t=1}^{[Tr]}u_t \\ &= \frac{1}{\sqrt{T}}\sum_{t=1}^{[Tr]}v_{1t} + \frac{\gamma}{aT^{1/\alpha}}\sum_{t=1}^{[Tr]}v_{2t} \\ &\Rightarrow \sigma_v W(r) + \gamma U_\alpha(r) \equiv Z_{\alpha,\gamma}(r)\end{aligned}$$

Next, we turn to the second convergence result

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{[Tr]}u_t^2 &= \frac{1}{T}\sum_{t=1}^{[Tr]}v_{1t}^2+\frac{\gamma^2}{aT^{2/\alpha}}\sum_{t=1}^{[Tr]}v_{2t}^2+\frac{2\gamma}{aT^{1/\alpha+1/2}}\sum_{t=1}^{[Tr]}v_{1t}v_{2t}\\ &\Rightarrow \sigma_v^2 r+\gamma^2V(r)\end{aligned}$$

since the first term of the second line converges in probability to $\sigma_v^2 r$, and thus in distribution, the second term converges in distribution to $\gamma^2V(r)$ and the last term converges in probability to zero, as $T\uparrow\infty$. The convergence of the last term to zero follows from the fact that the tail behavior of the product, say $\varepsilon_t = v_{1t}v_{2t}$, of independent variates belongs to the normal domain of attraction of the variate with the smallest maximal moment exponent, see Phillips (1990, Appendix A). Since $\varepsilon_t\in\mathcal{ND}(\alpha)$ it follows that $(aT^{1/\alpha})^{-1}\sum_{t=1}^{[Tr]}\varepsilon_t$ converges in distribution to a Lévy α-stable process while $T^{-1/2}$ converges to zero, as $T\uparrow\infty$. Thus, the product tends to zero as $T\uparrow\infty$.

The third and fourth convergence results follow by direct application of the continuous mapping theorem. The fifth one can be obtained as follows. After simple manipulations, we have that

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{[Tr]}y_{t-1}u_t &= \frac{1}{T}\sum_{t=1}^{[Tr]}x_{t-1}v_{1t}+\frac{1}{T}\sum_{t=1}^{[Tr]}z_{t-1}v_{1t}+\frac{\gamma}{T^{1/\alpha+1/2}}\sum_{t=1}^{[Tr]}x_{t-1}v_{2t}+\frac{\gamma^2}{T^{2/\alpha}}\sum_{t=1}^{[Tr]}z_{t-1}v_{2t}\\ &\Rightarrow \sigma_v^2\int_0^r W(s)\mathrm{d}W(s)+\gamma\sigma_v\int_0^r U_\alpha^-(s)\mathrm{d}W(s)+\\ &\qquad \gamma\sigma_v\int_0^r W(s)\mathrm{d}U_\alpha(s)+\gamma^2\int_0^r U_\alpha^-(s)\mathrm{d}U_\alpha(s)\\ &\equiv \int_0^r Z_{\alpha,\gamma}(s)\mathrm{d}Z_{\alpha,\gamma}(s)\end{aligned}$$

where $x_t=\sum_{s=1}^t v_{1s}$ and $z_t=\dfrac{\gamma}{aT^{1/\alpha-1/2}}\sum_{s=1}^t v_{2s}$. The weak convergence of the first terms follows from the weak convergence to stochastic integrals for sample covariances of i.i.d. processes, convergence of the second term follows from Hansen (1992), while the weak convergence of the third and fourth terms follows from Caner (1997). Finally, rearranging terms gives the result in the text.

A.2 Proof of Theorem 2.3

The limiting behavior of the $KPSS$, MRS, and KS test are obtained from Lemma 2.1, namely,

$$\frac{S_{[Tr]}}{\sqrt{T}}=\frac{1}{\sqrt{T}}\sum_{t=1}^{[Tr]}(u_t-\bar{u})\Rightarrow Z_{\alpha,\gamma}(r)-rZ_{\alpha,\gamma}(1)\equiv\tilde{Z}_{\alpha,\gamma}(r)$$

and application of the continuous mapping theorem. Amsler and Schmidt (1999) provide a proof for the $KPSS$ and MRS tests whereas for the KS test we have

$$\begin{aligned} KS &= \max_{k=1,\dots,T} \frac{k}{\hat{\sigma}_e T^{1/2}} \left| \frac{S_k}{k} - \frac{S_T}{T} \right| \\ &\Rightarrow \sup_r |Z_{\alpha,\gamma}(r) - rZ_{\alpha,\gamma}(1)| \equiv \sup_r \left| \tilde{Z}_{\alpha,\gamma}(r) \right| \end{aligned}$$

As for the LM and $SBDH$ tests we proceed as follows. First, we establish the limiting distribution of the OLS estimator of a regression of C_t on a time trend under our DGP (11)-(12) and (1), then we derive the asymptotic behavior of the test statistics. Letting $\hat{\mu} = \sum_{t=1}^T tC_t / \sum_{t=1}^T t^2$ it is easy to obtain

$$\sqrt{T}(\hat{\mu} - \mu) \Rightarrow 3 \int_0^1 rZ_{\alpha,\gamma} \mathrm{d}r \tag{26}$$

Then, given the definition of Q_t as the residuals of the above regression we have

$$\begin{aligned} Q_t &= C_t - \hat{\mu}t \\ &= \sum_{j=1}^t u_j - (\hat{\mu} - \mu)t, \qquad \text{so that} \\ Q_{t-1} &= \sum_{j=1}^{t-1} u_j - (\hat{\mu} - \mu)t + (\hat{\mu} - \mu), \qquad \text{and} \\ \Delta Q_t &= u_t - (\hat{\mu} - \mu) \end{aligned}$$

Upon substitution of these expressions in (17)-(19) we can derive the asymptotic distributions of Theorem 2.3. Thus, letting $x_t = \sum_{j=1}^t u_j$, we have

$$\begin{aligned} \frac{1}{T}\sum_{t=1}^T Q_{t-1}\Delta Q_t &= \frac{1}{T}\sum_{t=1}^T (x_{t-1} - t(\hat{\mu} - \mu) + \quad (\hat{\mu} - \mu))(u_t - (\hat{\mu} - \mu)) \\ &= \frac{1}{T}\sum_{t=1}^T x_{t-1}u_t - \sqrt{T}(\hat{\mu} - \mu)\frac{1}{T^{3/2}}\sum_{t=1}^T tu_t + \\ &\quad +\sqrt{T}(\hat{\mu} - \mu)\frac{1}{T^{3/2}}\sum_{t=1}^T u_t + \sqrt{T}(\hat{\mu} - \mu)\frac{1}{T^{3/2}}\sum_{t=1}^T x_{t-1} + \\ &\quad +T(\hat{\mu} - \mu)^2\frac{1}{T^2}\sum_{t=1}^T t + T(\hat{\mu} - \mu)^2\frac{1}{T^2}\sum_{t=1}^T 1 \\ &\Rightarrow \int_0^1 Z_{\alpha,\gamma}\mathrm{d}Z_{\alpha,\gamma} - 3\int_0^1 \mathrm{d}Z_{\alpha,\gamma}\int_0^1 rZ_{\alpha,\gamma} + \int_0^1 r\left(3\int_0^1 rZ_{\alpha,\gamma}\right)^2 \\ &\equiv \int_0^1 \bar{Z}_{\alpha,\gamma}\mathrm{d}\bar{Z}_{\alpha,\gamma} \end{aligned}$$

where $\bar{Z}_{\alpha,\gamma}$ and $\mathrm{d}\bar{Z}_{\alpha,\gamma}$ are defined in Theorem 2.3,

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T}\Delta Q_t^2 &= \frac{1}{T}\sum_{t=1}^{T}\big(u_t + (\hat{\mu}-\mu)^2 - 2(\hat{\mu}-\mu)u_t\big) \\
&= \frac{1}{T}\sum_{t=1}^{T}u_t + T(\hat{\mu}-\mu)^2\frac{1}{T^2}\sum_{t=1}^{T}1 - 2\sqrt{T}(\hat{\mu}-\mu)\frac{1}{T^{3/2}}\sum_{t=1}^{T}u_t \\
&\Rightarrow \sigma_1^2 + \gamma V(1) \equiv K_\gamma(1)
\end{aligned}$$

and

$$\begin{aligned}
\frac{1}{T^2}\sum_{t=1}^{T}Q_{t-1}^2 &= \frac{1}{T}\sum_{t=1}^{T}x_{t-1}^2 + T(\hat{\mu}-\mu)^2\frac{1}{T^3}\sum_{t=1}^{T}t^2 + T(\hat{\mu}-\mu)^2\frac{1}{T^3}\sum_{t=1}^{T}1 - \\
&\quad -2\sqrt{T}(\hat{\mu}-\mu)\frac{1}{T^{5/2}}\sum_{t=1}^{T}tx_{t-1} + 2\sqrt{T}(\hat{\mu}-\mu)^2\frac{1}{T^{5/2}}\sum_{t=1}^{T}x_{t-1} - \\
&\quad -2\sqrt{T}(\hat{\mu}-\mu)\frac{1}{T^{5/2}}\sum_{t=1}^{T}t \\
&\Rightarrow \int_0^1 Z_{\alpha,\gamma}^2 - \left(3\int_0^1 rZ_{\alpha,\gamma}\right)^2
\end{aligned}$$

which, together with the continuous mapping theorem yield Theorem 2.3.

B Tail behavior of $u_t = v_{1t} + z_t v_{2t}$

In this section, we provide a Lemma establishing the tail behavior of the process $u_t = v_{1t} + z_t v_{2t}$ introduced in section 2.

LEMMA B.1 *Let v_{1t} be an i.i.d. process with zero mean and finite variance σ_1^2 and distribution function $F_{1t}(\cdot)$, let v_{2t} belong to the normal domain of attraction of a stable law with characteristic exponent α with $0 < \alpha < 2$ with distribution function $F_{2t}(\cdot)$, independent on v_{1t} and let $z_t \sim \mathrm{B}(1,p)$. Let $u_t = v_{1t} + z_t v_{2t}$, then the distribution function $F_u(\cdot)$ of u_t belongs to the normal domain of attraction of a stable law with characteristic exponent α.*

PROOF. The proof is in two steps. First, we show that the random variable $z_t v_{2t}$ belongs to the normal domain of attraction of a stable law with characteristic exponent α, secondly, we show that u_t also belongs to the normal domain of attraction of a stable law with the same characteristic exponent α.

Letting $d_t = z_t v_{2t}$ we have

$$d_t = \begin{cases} 0 & \text{if} \quad z_t = 0 \\ v_{2t} & \text{if} \quad z_t = 1 \end{cases}$$

for each t. Letting $F_d(\cdot)$ and $f_d(\cdot)$ be the distribution function and the density function of d_t, respectively, we factorize the marginal density of d_t as

$$\begin{aligned} f_d(h) &= f_{d|u}(h|z_t = 0)\text{prob}(z_t = 0) + f_{d|u}(h|z_t = 1)\text{prob}(z_t = 1) \\ &= (1-p)\delta_0(h) + pf_u(h) \end{aligned}$$

where $\delta_0(\cdot)$ is a p.d.f. that assigns probability one to the value zero and $f_2(\cdot)$ is the density of v_{2t}. Integrating the p.d.f., we obtain the distribution function of d_t as

$$\begin{aligned} F_d(h) &= (1-p)\int_{-\infty}^{h} \delta_0(s)\text{d}s + p\int_{-\infty}^{h} f_u(s)\text{d}s \\ &= (1-p)I_{(h\geq 0)}(h) + pF_u(h) \end{aligned}$$

where

$$I_{(h\geq 0)}(h) = \begin{cases} 0 & \text{if} \quad h < 0 \\ 1 & \text{if} \quad h \geq 0 \end{cases}$$

It is immediate to see that

$$F_d(h) = \begin{cases} pF_u(h) & \text{if} \quad h < 0 \\ 1 - p + pF_u(h) & \text{if} \quad h \geq 0 \end{cases}$$

which, taking into account the tail behavior of $f_2(\cdot)$, can be written as

$$F_d(h) = \begin{cases} [\tilde{c}_1 a^\alpha + \tilde{\alpha}_1(h)]\dfrac{1}{|h|^\alpha} & \text{if} \quad h < 0 \\ 1 - [\tilde{c}_2 a^\alpha + \tilde{\alpha}_2(h)]\dfrac{1}{|h|^\alpha} & \text{if} \quad h \geq 0 \end{cases}$$

where $\tilde{c}_1 = pc_1$, $\tilde{c}_2 = pc_2$, $\tilde{\alpha}_1 = p\alpha_1$, and $\tilde{\alpha}_2 = p\alpha_2$. Since $\tilde{c}_1 > 0$, $\tilde{c}_2 > 0$ with $\tilde{c}_1 + \tilde{c}_2 > 0$, and $\lim_{h\to-\infty} \widetilde{\alpha_1(h)} = p\lim_{h\to-\infty} \alpha_1(h) = 0$ and $\lim_{h\to\infty} \widetilde{\alpha_2(h)} = p\lim_{h\to-\infty} \alpha_2(h) = 0$, the distribution function $F_d(\cdot)$ belongs to the normal domain of attraction of a stable law with characteristic exponent α by Theorem 2.6.7 of Ibragimov and Linnik (1971).

Next, we turn to the compound process u_t. Let us define the normed sums

$$S_{d,T} = \frac{d_1 + d_2 + \cdots + d_T}{B_{d,T}} - A_{d,T}$$

and

$$S_{u,T} = \frac{u_1 + u_2 + \cdots + u_T}{B_{d,T}} - A_{d,T}$$

where the norming factors $B_{d,T}$ and $A_{d,T}$ are those required for the convergence of $\sum_{t=1}^{T} d_t$ to a non-degenerate random variable. Since $F_d(\cdot)$ belongs to the normal domain of attraction of stable law with characteristic exponent α and the norming factor $B_{d,T}$ is given by $aT^{1/\alpha}$, it follows that $S_{\epsilon,T}$ may be rearranged as

$$S_{\epsilon,T} = \frac{1}{aT^{1/\alpha-1/2}} \frac{v_1 + v_2 + \cdots + v_T}{T^{1/2}} + \frac{d_1 + d_2 + \cdots + d_T}{aT^{1/\alpha}} - A_{d,T}$$

Now, $T^{-1/2} \sum_{t=1}^{T} v_t \to_d N(0, \sigma_v^2)$ by a CLT for i.i.d. sequences and $aT^{1/2}/T^{1/\alpha} \to 0$ because of $1/\alpha > 1/2$ for $\alpha \in (0, 2)$, the first term converges in distribution to zero, and hence in probability. Therefore, $S_{u,T}$ is asymptotically equivalent to $S_{d,T}$. It follows that the distribution function $F_u(\cdot)$ belongs to the normal domain of attraction of a stable law with the same characteristic exponent α of $F_d(\cdot)$, which, in turn, is the same characteristic exponent of the distribution function $F_u(\cdot)$.

References

Ahn, S.K. (1993), 'Some tests for unit roots in autoregressive-integrated-moving average models with deterministic trends', *Biometrika* **80**, 855–868.

Ahn, S.K., S.B. Fotopoulos and L. He (2001), 'Unit root tests with infinite variance errors', *Econometric Reviews* **20**, 461–483.

Amsler, C. and P. Schmidt (1999), Tests of short memory with thick tailed errors. paper presented at the Cowles Foundation Econometrics Conference "New Developments in Time Series Econometrics".

Bhargava, A. (1986), 'On the theory of testing for unit roots in observed time series', *Review of Economic Studies* **53**, 369–384.

Billingsley, P. (1968), *Convergence of Probability Measures*, Wiley.

Callegari, F., N. Cappuccio and D. Lubian (2003), 'Asymptotic inference in time series regressions with a unit root and infinite variance errors', *Journal of Statistical Planning and Inference* **116**, 277–303.

Caner, M. (1997), 'Weak convergence to a matrix stochastic integral with stable processes', *Econometric Theory* **13**, 506–528.

Chan, N.H. and L.T. Tran (1989), 'On the first order autoregressive process with infinite variance', *Econometric Theory* **5**, 354–362.

Choi, I. (1994), 'Resiual based tests for the null of stationarity with applications to u.s. macroeconomic time series', *Econometric Theory* **10**, 720–746.

Choi, I. and B. Yu (1997), 'A general framework for testing $i(m)$ against $i(m+k)$', *Journal of Economic Theory and Econometrics* **3**, 103–138.

Choi, I. and B.C. Ahn (1999), 'Testing the null of stationarity for multiple time series', *Journal of Econometrics* **88**, 41–77.

Davidson, R. and J. G. MacKinnon (1998), 'Graphical methods for investigating the size and power of test statistics', *The Manchester School* **66**.

Dickey, D.A. and W.A. Fuller (1976), 'Distribution of the estimators for autoregressive time series with a unit root', *Journal of the American Statistical Association* **74**, 427–431.

Elliott, G. (1998), 'On the robustness of cointegration methods when regressors almost have unit roots', *Econometrica* **66**, 149–158.

Gnedenko, B.V. and A.N. Kolmogorov (1954), *Limit distributions for sums of independent random variables*, Addison-Wesley Publishing Company.

Hamilton, J.D. (1994), *Time Series Analysis*, Princeton University Press.

Hansen, B.E. (1992), 'Convergence to stochastic integrals for dependent heterogeneous processes', *Econometric Theory* **8**, 489–500.

Ibragimov, I.A. and Y.V. Linnik (1971), *Independent and stationary sequences of random variables*, Wolters-Nordhoff Publishing Groningen.

Kwiatkowski, D., P.C.B. Phillips, P. Schmidt and Y. Shin (1992), 'Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?', *Journal of Econometrics* **54**, 159–178.

Lo, A. (1991), 'Long-term memory in stock market prices', *Econometrica* **59**, 1279–1313.

Loretan, M. and P.C.B. Phillips (1994), 'Testing the covariance stationarity of heavy-tailed time series', *Journal of Empirical Finance* **1**, 211–248.

Mandelbrot, B.B. (1997), *Fractals and Scaling in Finance*, Springer-Verlag, New York.

McCulloch, J. H. (1996), Financial applications of stable distributions, *in* G. S.Maddala and C. R.Rao, eds, 'Handbook of Statistics', Vol. 14, Elsevier, Amsterdam, pp. 393–425.

McCulloch, J. H. (1997), 'Measuring tail thickness to estimate the stable index α: a critique', *Journal of Business & Economic Statistics* **15**, 74–81.

Nabeya, S. and P. Perron (1994), 'Local asymptotic distribution related to the $ar(1)$ model with dependent errors', *Journal of Econometrics* **62**, 229–264.

Perron, P and S. Ng (1996), 'Useful modifications to some unit root tests with dependent errors and their local asymptotic properties', *Review of Economic Studies* **63**, 435–563.

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

Phillips, P.C.B. (1990), 'Time series regression with a unit root and infinite variance errors', *Econometric Theory* **6**, 44–62.

Pitman, E. J. G. (1949), *Notes on non-parametric statistical inference*, mimeo, Columbia University, New York.

Protter, P. (1990), *Stochastic Integration and Differential Equations*, Springer and Verlag.

Rachev, S. and S. Mittnik (2000), *Stable Paretian Models in Finance*, John Wiley & Sons, New York.

Resnick, S.I. (1986), 'Point processes regular variation and weak convergence', *Advances in Applied Probability* **18**, 66–138.

Sargan, J.D. and A. Bhargava (1983), 'Testing for residuals from least squares regression for being generated by the gaussian random walk', *Econometrica* **51**, 153–174.

Silverman, B. W. (1986), *Density estimation for statistics and data analysis*, Chapman & Hall.

Stock, J.H. (1990), A class of tests for integration and cointegration. mimeo, Kennedy School of Government.

Stock, J.H. (1994), 'Deciding between $I(1)$ and $I(0)$', *Journal of Econometrics* **63**, 105–131.

Xiao, Z. (2001), 'Testing the null hypothesis of stationarity against an autoregressive unit root', *Journal of Time Series Analysis* **22**, 87–105.

Table 2: Effective size (in %) with 5% nominal size for stationarity tests under (1)

		γ						γ					
		0.1	0.316	1	3.16	10	31.6	0.1	0.316	1	3.16	10	31.6
		0.1	0.316	1	3.16	10	31.6	0.1	0.316	1	3.16	10	31.6
T	α	$KPSS$						MRS					
	0.5	4.12	3.45	2.80	2.34	1.71	1.34	1.74	1.32	0.86	0.50	0.14	0.01
100	1.0	4.69	4.56	3.66	3.12	2.94	2.76	2.09	1.76	0.88	0.33	0.09	0.05
	1.5	5.05	4.76	4.40	4.10	4.04	4.13	2.29	2.11	1.25	0.77	0.65	0.63
	0.5	4.39	3.47	3.03	2.05	1.59	1.24	3.12	2.36	1.61	0.77	0.23	0.07
1000	1.0	4.83	4.25	3.80	2.99	2.67	2.90	3.79	2.97	1.60	0.48	0.18	0.14
	1.5	4.76	4.77	4.23	3.96	3.93	4.20	3.91	3.72	2.45	1.32	1.15	1.09
	0.5	4.28	3.56	2.95	2.10	1.58	1.37	3.51	3.06	1.91	0.80	0.25	0.03
10000	1.0	4.65	4.44	3.61	3.04	3.02	2.74	4.48	3.60	1.87	0.58	0.20	0.18
	1.5	4.97	4.61	4.31	3.93	4.02	4.10	4.87	4.20	2.71	1.47	1.27	1.23
		KS						LM_1					
	0.5	2.74	2.20	1.48	0.94	0.41	0.19	12.85	11.77	10.51	8.05	7.04	5.95
100	1.0	3.18	2.83	1.89	0.96	0.64	0.56	15.15	13.57	11.63	8.77	8.09	7.85
	1.5	3.43	3.32	2.51	1.99	1.88	1.81	15.20	14.78	13.43	11.49	11.05	10.98
	0.5	3.75	2.71	2.04	0.97	0.49	0.19	6.18	5.69	5.17	4.71	4.22	3.72
1000	1.0	4.37	3.52	2.46	1.12	0.89	0.81	6.75	6.46	5.85	4.74	4.65	4.47
	1.5	4.54	4.15	3.10	2.37	2.17	2.35	7.35	6.78	6.17	5.67	5.57	5.45
	0.5	3.78	3.16	2.23	1.12	0.55	0.19	4.84	4.53	4.13	4.11	3.87	3.46
10000	1.0	4.45	3.88	2.44	1.29	0.77	0.78	5.19	4.88	4.64	4.31	4.02	4.32
	1.5	4.84	4.22	3.20	2.31	2.35	2.27	5.34	5.45	4.65	4.58	4.68	4.55
		LM_2						$SBDH$					
	0.5	3.89	3.30	2.79	2.26	2.00	1.96	4.98	4.44	4.06	3.48	3.04	2.80
100	1.0	4.53	3.96	3.15	2.43	2.56	2.69	5.15	5.15	4.78	4.23	3.86	4.10
	1.5	4.42	4.23	3.81	3.36	3.43	3.33	5.05	5.48	4.81	4.70	4.72	4.92
	0.5	4.13	3.55	2.65	2.18	1.69	1.82	4.19	4.13	3.42	2.73	2.18	1.97
1000	1.0	4.76	4.41	3.67	2.83	2.69	2.92	4.93	4.43	4.20	3.55	3.19	3.17
	1.5	5.01	4.61	4.20	3.97	3.83	3.93	4.96	5.04	4.56	4.47	4.30	4.02
	0.5	4.04	3.69	3.01	2.20	1.71	1.54	4.36	3.80	3.41	2.75	2.26	1.95
10000	1.0	4.85	4.63	3.59	2.92	2.48	2.65	4.97	4.57	3.99	3.53	3.30	3.23
	1.5	5.01	4.92	4.68	3.88	3.77	3.82	5.21	4.92	4.40	4.50	4.25	4.18

Table 3: Discrepancy (in %) between effective and nominal size for all tests under local to finite variance, $\alpha = 1.5$.

Nominal Size	$T(\hat{\rho}-1)$	$t(\hat{\rho})$	LM	DW	$KPSS$	MRS	KS	LM_1	LM_2	$SBDH$
					$T = 100$, $\gamma = 1$					
0.01	-0.155	-0.030	-0.275	0.070	-0.355	-0.830	-0.705	10.840	-0.400	0.040
0.05	-0.505	-0.205	-0.535	0.000	-0.605	-3.745	-2.495	8.435	-1.190	-0.190
0.10	-1.230	-0.850	-1.300	-0.415	-0.570	-6.825	-4.300	6.395	-2.175	0.010
					$T = 1000$, $\gamma = 1$					
0.01	-0.005	-0.030	-0.100	0.205	-0.195	-0.620	-0.465	2.520	-0.255	-0.060
0.05	-0.395	-0.245	-0.505	-0.055	-0.765	-2.545	-1.905	1.175	-0.800	-0.440
0.10	-1.000	-0.920	-1.070	-0.375	-0.465	-4.790	-2.820	0.840	-1.640	-0.535
					$T = 10000$, $\gamma = 1$					
0.01	-0.215	-0.255	-0.265	0.015	-0.365	-0.555	-0.525	0.210	-0.120	-0.270
0.05	-0.565	-0.575	-0.635	0.005	-0.690	-2.295	-1.800	-0.345	-0.315	-0.605
0.10	-0.770	-0.770	-1.075	-0.325	-0.860	-4.100	-2.755	-0.500	-1.175	-1.015
					$T = 100$, $\gamma = 31.6$					
0.01	-0.200	-0.100	-0.285	0.170	-0.340	-0.940	-0.810	8.235	-0.365	-0.045
0.05	-0.860	-0.620	-1.240	-0.010	-0.870	-4.365	-3.185	5.975	-1.670	-0.085
0.10	-1.480	-1.325	-2.100	-0.640	-1.000	-8.26	-5.445	4.520	-2.860	-0.195
					$T = 1000$, $\gamma = 31.6$					
0.01	-0.185	-0.215	-0.305	0.125	-0.375	-0.915	-0.680	1.515	-0.215	-0.345
0.05	-0.980	-0.970	-1.160	-0.190	-0.805	-3.910	-2.650	0.455	-1.070	-0.980
0.10	-1.655	-1.640	-2.120	-0.955	-1.055	-7.165	-4.720	0.135	-2.435	-1.030
					$T = 10000$, $\gamma = 31.6$					
0.01	-0.325	-0.355	-0.420	-0.035	-0.435	-0.905	-0.770	-0.145	-0.115	-0.325
0.05	-1.170	-1.160	-1.385	-0.290	-0.905	-3.770	-2.730	-0.450	-1.185	-0.820
0.10	-2.095	-2.195	-2.410	-1.010	-0.985	-6.795	-4.405	-0.570	-2.210	-0.735