A KPSS better than KPSS
Rank tests for short memory stationarity

Matteo M. Pelagatti\textsuperscript{a,*}, Pranab K. Sen\textsuperscript{b}

\textsuperscript{a}Department of Statistics, Università degli Studi di Milano-Bicocca
Via Bicocca degli Arcimboldi 8, 20126 Milan, Italy

\textsuperscript{b}Department of Statistics and Operations Research, University of North Carolina
318 Hanes Hall, Chapel Hill, NC 27599-3260

Abstract

We propose a rank-test of the null hypothesis of short memory stationarity possibly after linear detrending.

For the level-stationarity hypothesis, the test statistic we propose is a modified version of the popular KPSS statistic, in which ranks substitute the original observations. We prove that the rank KPSS statistic shares the same limiting distribution as the standard KPSS statistic under the null and diverges under I(1) alternatives.

For the trend-stationarity hypothesis, we apply the same rank KPSS statistic to the residual of a Theil-Sen regression on a linear trend. We derive the asymptotic distribution of the Theil-Sen estimator under short memory errors and prove that the Theil-Sen detrended rank KPSS statistic shares the same weak limit as the least-squares detrended KPSS.

We study the asymptotic relative efficiency of our test compared to the KPSS and prove that it may have unbounded efficiency gains under fat-tailed distributions compensated by very moderate efficiency losses under thin-tailed distributions. For this and other reasons discussed in the body of the article our rank KPSS test turns out to be an irresistible competitor of the KPSS for most real-world economic and financial applications.

The weak convergence results and asymptotic representations proved in this article may have an interest on their own, as they extend to ranks analogous results widely used in unit-root econometrics.

Keywords: Stationarity test, Unit roots, Robustness, Rank statistics, Theil-Sen estimator, Asymptotic efficiency

JEL: C12, C14, C22

\textsuperscript{*}Corresponding author (Tel +39-02-64485834, Fax +39-02-64485878). The first author’s research was partially financed by the Ministerial Grant PRIN 20074PFL7C.

Email addresses: matteo.pelagatti@unimib.it (Matteo M. Pelagatti), pk sen@bios.unc.edu (Pranab K. Sen)
1. Introduction

Despite its long tradition in statistics, ranks-based inference has never become popular among econometricians. Indeed, there is really no econometric textbook covering rank tests, \textit{R}-estimators or neighboring topics like \textit{U}-statistics. Few exceptions can be found in econometric journals (McCabe, 1989; Campbell and Dufour, 1995; Breitung and Gouriéroux, 1997; Hasan and Koenker, 1997; Hasan, 2001; Luger, 2003; Mukherjee, 2007), in which rank methods are mostly used for unit-root testing. We are not able to speculate on the causes of this neglect, but ranks-based techniques generally reach a good balance between efficiency and robustness, and this is certainly a property the applied econometrician should seek. In fact, in many real economic and financial applications Gaussianity is the exception, rather than the rule, and the performance of the wide-spread least-squares methods can drastically deteriorate.

In this paper we propose a rank test for the null hypothesis of short-memory stationarity (often referred to as \textit{level-stationarity}), and a rank test for the null of short-memory stationarity on a linear trend (also \textit{trend-stationarity}).

Our test for level-stationarity is based on the well-known KPSS statistic (Kwiatkowski et al., 1992) applied to the ranks of the observations. The asymptotic distribution of our rank KPSS (RKPSS) statistic under the null is the same as that of the original KPSS and, thus, any software package implementing the KPSS test implements our RKPSS test as well. The advantages of our RKPSS test over Kwiatkowski et al.’s (1992) are many: i) the existence of moments is not necessary, ii) the test statistic is invariant to monotonic transformations of the data, iii) the asymptotic relative efficiency of the RKPSS with respect to the KPSS test is moderately smaller than one under thin-tailed distribution, but may increase without bound under fat-tailed distributions. Thus, unless platykurtosis is ascertained, there is really no reason to prefer the KPSS to our RKPSS. Our test can perform better than the KPSS even under Gaussianity, whenever the data are subject to some positive autocorrelation. Notice that economic time series are usually characterized by positive autocorrelation, and financial data by leptokurtosis. Property ii) is also very valuable, since economic data are often transformed through logarithms or other strictly monotonic maps as, for instance, the Box-Cox family. Since short-memory stationarity is not altered by such transformations, it is highly desirable for the outcome of the test not to depend on their choice.

The idea for the RKPSS test was inspired by a recent work by de Jong et al. (2007), who apply the KPSS statistic to the signs of the median-adjusted observations. Unfortunately, their approach leads to a test with competitive power properties only under extremely fat-tailed distributions as the simulations in Section 5 will show. Furthermore, de Jong et al. (2007) do not provide a study of the (asymptotic) efficiency of their test and do not derive the theory for a test of trend-stationarity. Nonetheless, Bosco et al. (2010) show by simulation that detrending based on least absolute deviations leads to the expected asymptotic distribution of the test statistic.

A test of trend-stationarity could be based on the KPSS statistic computed
on the residuals of a least-squares regression of the ranks on a linear trend. This approach has the advantage of removing any monotonic trend in the data, regardless of its parametric specification. Unfortunately, there are also important drawbacks, discussed in Section 4, that made us pursue a different approach.

Our test for stationarity on a linear trend is based on the KPSS statistic applied to the residuals of a regression of the observations on a linear trend. The regression coefficients are not estimated through least-squares as in the original KPSS test, but using the Theil-Sen estimator proposed by Theil (1950) and generalized by Sen (1968). The Theil-Sen estimator is simple to compute, robust to fat-tails and guarantees that, under the null, the asymptotic distribution of the detrended RKPSS statistic is the same as that of the detrended KPSS statistic.

The weak convergence results and asymptotic representations we prove in this paper should interest a wider audience than stationarity test users. Indeed, they represent generalizations to ranks of many results commonly used in unit-root econometrics such as those listed in the widely cited paper by Phillips and Perron (1988). In particular, the asymptotic distribution of the Theil-Sen estimator under strong mixing, derived in Section 4, may be of interest to all those who seek a robust and easy way to detrend fat-tailed observations.

The plan of the paper is as follows. In Section 2 we provide the asymptotic theory for the RKPSS statistic both under the null and under the alternative of integration. In Section 3 we study the asymptotic relative efficiency of our test compared to the KPSS with particular attention to two families of distributions widely used in financial econometrics: Student’s $t$ and the Generalized Error Distribution. Section 4 derives the asymptotic distribution of the test for trend-stationarity. Section 5 presents a battery of simulations for comparing the finite-sample properties of the KPSS and RKPSS statistics, also with the IKPSS statistic of de Jong et al. (2007). Finally, in Section 7 we give some concluding remarks and directions for future research. All proofs can be found in the Appendix.

2. Rank KPSS for level stationarity

Let the observed time series be a sample path of the real random sequence $\{X_1, X_2, \ldots, X_T\}$ and let

$$R_{T,t} = \sum_{i=1}^{T} I\{X_i \leq X_t\}, \quad \text{for } t = 1, \ldots, T, \quad (1)$$

with $I_A$ indicator function of the set $A$, be the rank of $X_t$ among $\{X_1, \ldots, X_T\}$. Notice that the arithmetic mean of the rank sequence $\{R_{T,1}, \ldots, R_{T,T}\}$ is $(T + 1)/2$ and does not depend on the data.

The test statistic we propose in this paper is the KPSS applied to the ranks
of the observations. So, let $S_{T,t}$ be the sequence of demeaned partial sums:

$$S_{T,t} = \sum_{i=1}^{t} \left( \frac{R_{T,i}}{T} - \frac{T + 1}{2T} \right).$$

(2)

Notice that the KPSS statistic is invariant to scale transformations, so working with $R_{T,i}/T$ rather than $R_{T,i}$ turns out to generate the same statistic. We chose to work with the former form since this makes our partial sum process diverge at the same rate as the analogous quantity defined in Kwiatkowski et al. (1992, eq. 5).

In complete analogy with Kwiatkowski et al. (1992), define the random quantity

$$\eta_{\mu,T}^R = T^{-2} \sum_{i=1}^{T} S_{T,t}^2$$

(3)

and the RKPSS test statistic for level-stationarity as

$$\hat{\eta}_{\mu,T}^R = \eta_{\mu,T}^R / \hat{\sigma}_T^2,$$

(4)

where $\hat{\sigma}_T^2$ is a HAC estimator of the long-run variance of the process $\{R_{T,t}/T\}$:

$$\hat{\sigma}_T^2 = \frac{1}{T} \sum_{s=1}^{T} \sum_{t=1}^{T} k\left( \frac{s-t}{\gamma_T} \right) \left[ \frac{R_{T,s}}{T} - \frac{T + 1}{2T} \right] \left[ \frac{R_{T,t}}{T} - \frac{T + 1}{2T} \right],$$

(5)

with $k(\cdot)$ symmetric kernel function and $\gamma_T$ bandwidth parameter.

We state here two assumptions that will be useful in the rest of the paper.

**Assumption 1. (Short memory stationarity)**

1. $\{X_1, \ldots, X_T\}$ is a strictly stationary random sequence.
2. $\{X_1, \ldots, X_T\}$ is strong mixing with parameter $\alpha(T) = O(T^{-v}), v > 2$.
3. For all $i \in \{1, \ldots, T\}$ and $T \in \mathbb{N}$, $X_i$ has non-degenerate absolutely continuous distribution function $F(\cdot)$ defined on $\mathbb{R}$ with density $f(\cdot)$.

**Assumption 2. (Regularity of the kernel function)**

1. $k(\cdot)$ satisfies $\int_{-\infty}^{\infty} |\psi(z)| dz < \infty$, where

$$\psi(z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} k(x) \exp(-izx) dx.$$  

(6)

2. $k(\cdot)$ is continuous at all but a finite number of points, $k(x) = k(-x)$, $|k(x)| < l(x)$ where $l(x)$ is non-increasing and $\int_{0}^{\infty} |l(x)| dx < \infty$, and $k(0) = 1$.
3. $\gamma_T / \sqrt{T} \to 0$ and $\gamma_T \to \infty$ as $T \to \infty$. 

4
Define the population counterpart of Spearman’s rank autocorrelation coefficient as

\[ \rho_{i,j} = 12 \mathbb{E} \left\{ \left[ F(X_i) - 1/2 \right] \left[ F(X_j) - 1/2 \right] \right\}, \quad (7) \]

for \( 1 \leq i, j \leq T \).

The following theorem gives the asymptotic distribution of the test statistic under the null of strong mixing stationarity. Weak convergence will be denoted by \( \Rightarrow \) and is to be intended with respect to the Skorohod \( J_1 \)-topology on the space \( D[0,1] \) of càdlàg functions on the unit interval.

**Theorem 1.** Under Assumption 1,

\[ \eta^{R}_{\mu,T} \Rightarrow \sigma^2 \int_0^1 V(r)^2 \, \text{d}r, \quad (8) \]

with \( V \) standard Brownian bridge and

\[ \sigma^2 = \frac{1}{12} \left[ 1 + 2 \sum_{k=2}^{\infty} \rho_{1,k} \right]; \quad (9) \]

furthermore

\[ T^{-1/2} S_{T,t} = T^{-1/2} \left\{ \sum_{i=1}^{t} F(X_i) - t \sum_{i=1}^{T} F(X_i) \right\} + O_p(T^{-1/2}). \quad (10) \]

Under Assumptions 1 and 2,

\[ \hat{\eta}^{R}_{\mu,T} \Rightarrow \int_0^1 V(r)^2 \, \text{d}r. \quad (11) \]

The above theorem states that our RKPSS statistic has the same asymptotic distribution as the original KPSS statistic, so both tables and softwares for carrying out the KPSS test may be used for our test\(^1\).

Under the alternative that \( X_t \) is an I(1) process, we have the following result, that proves the consistency of our test under the most econometrically relevant alternative.

**Theorem 2.** Suppose there exists a strictly monotone (Borel) function \( g: \mathbb{R} \rightarrow \mathbb{R} \) such that \( T^{-1/2} g(X_{\lfloor rT \rfloor}, T) \Rightarrow \omega W(r) \), where \( \omega \) is a strictly positive real number and \( W \) is standard Brownian motion on \([0,1]\), then

\[ \frac{\eta^{R}_{\mu,T}}{T} \Rightarrow \int_0^1 \left[ \int_0^s R_0(r) \, \text{d}r \right]^2 \, \text{d}s, \quad (12) \]

\(^1\)Using the results in Oliveira and Suquet (1998), the convergence in equations (8) and (11) could be derived under the weaker assumption on the mixing coefficients \( \sum_{n=1}^{\infty} \alpha(n) < \infty \). This stronger result comes at the cost of setting the weak convergence in the \( L^2(0,1) \) topology, which restrict the functionals of the partial sums process for which the continuous mapping theorem applies. Since, this would limit the use of the asymptotic representation we derive in the theorem, we preferred the more traditional Skorohod’s setup.
with \( R_0(r) = \int_0^1 \mathbb{1}_{\{W(u) < W(r)\}} \, du - \frac{1}{2} \), and

\[
\hat{\sigma}^2_T \leq \frac{1}{T} \sum_{s=1}^{T} \sum_{t=1}^{T} k \left( \frac{s-t}{\gamma_T} \right) = O(\gamma_T). \tag{13}
\]

Since the numerator \( \eta^{R}_{\mu,T} \) of the RKPSS statistic is of order \( O_p(T) \) while the denominator is of order \( O(\gamma_T) \) with \( \gamma_T = o(T) \), the corollary below follows immediately.

**Corollary 1.** The RKPSS statistic \( \hat{\eta}^{R}_{\mu,T} \) is consistent against the hypothesis of Theorem 2.

Notice, that the alternative hypothesis we used in Theorem 2 is much weaker than the corresponding hypothesis for the KPSS statistic (Kwiatkowski et al., 1992, Section 4). Indeed, while for the KPSS test, the process \( X_t \) must be I(1), in the RKPSS case the I(1) process can be any strictly monotonic transformation of \( X_t \). Thus, all typical data transformations used in econometrics (Box-Cox, exponential, logit, etc.) change neither the value of the statistic nor its asymptotic behaviour both under the null and under the alternative.

The thesis in equation (12) of Theorem 2 suggests that the statistic \( \eta^{R}_{\mu,T}/T \) can be used to test the hypothesis \( g(X_t) \sim I(1) \) against stationarity. Indeed, \( \eta^{R}_{\mu,T}/T \) is free of nuisance parameters and converges weakly to a proper distribution under the null and to the Dirac (point mass) distribution under the alternative. However, the results of Shin and Schmidt (1992) for the KPSS statistic used as a unit root test seem to suggest a poor performance also of our statistic.

### 3. Asymptotic relative efficiency

In order to investigate the relative efficiency of our RKPSS test with respect to the original KPSS, let us consider the local alternative represented by the following data generating process:

\[
Y_t = \frac{\sigma_z}{T} \sum_{s=1}^{rT} Z_t + X_t, \quad t = 1, 2, \ldots, T, \tag{14}
\]

where \( \sigma_z \) and \( \sigma_x \) are positive real numbers, and \( Z_t \) and \( X_t \) are mutually independent stationary processes such that, for \( r \in \{0, 1\} \),

\[
T^{-1/2} \sum_{i=1}^{rT} Z_t \Rightarrow W_z(r) \quad \text{and} \quad T^{-1/2} \sum_{i=1}^{rT} X_t \Rightarrow \sigma_x W_x(r). \tag{15}
\]

Notice that, for fixed values of \( \sigma_z/T \), the process (14) is exactly the one considered by Kwiatkowski et al. (1992, eq.2-3) for testing the hypothesis \( \sigma_z = 0 \) against \( \sigma_z > 0 \). In case \( Z_t \) and \( X_t \) are i.i.d. Gaussian processes, the KPSS statistic turns out to be the (one-sided) Lagrange multiplier (LM) statistic and also
the locally best invariant (LBI) test under the group of linear transformations (see Nabeya and Tanaka, 1988, for details).

Define the partial sum processes both for the KPSS and RKPSS statistics under the local alternative (14) as

\[ S^K_{T,t} := \sum_{i=1}^{t} (Y_i - \bar{Y}), \quad S^R_{T,t} := \sum_{i=1}^{t} \left( \frac{R^u_{T,i}}{T} - \frac{T+1}{2T} \right), \]

where \( \bar{Y} := T^{-1} \sum_{i=1}^{T} Y_i \), and \( R^u_{T,i} \) is the rank of \( Y_i \) among \( \{Y_1, \ldots, Y_T\} \).

**Theorem 3.** Assume that \( Y_t \) is generated by (14)-(15), where \( X_t \) satisfies Assumption 1, then, for \( r = t/T \),

\[ \frac{1}{\sqrt{T} \sigma_x} S^K_{T,t} \Rightarrow V(r) + \frac{\sigma_x}{\sigma} K(r) \]

\[ \frac{1}{\sqrt{T} \sigma} S^R_{T,t} \Rightarrow V(r) + f_2(0) \frac{\sigma_x}{\sigma} K(r) \]

where \( V(r) \) is a standard Brownian bridge independent of

\[ K(r) := \int_{0}^{r} W_z(u) \, du - r \int_{0}^{1} W_z(u) \, du, \]

with \( f_2(0) := \mathbb{E} f(X) \) and \( \sigma \) as in (9).

Furthermore, the (Pitman) asymptotic relative efficiency (ARE) of the RKPSS test with respect to the KPSS is

\[ e_{R,K} = f_2(0) \frac{\sigma_x}{\sigma}. \]

In the i.i.d. case, we have \( \sigma^2 = 1/12, \sigma_x^2 = \text{Var}(X_t) \) and \( e_{R,K} = f_2(0) \sqrt{12 \text{Var}(X_t)}. \)

Notice that, if \( f(x) \) is the density of \( X \), then \( \sigma_x X \) has density \( f(x/\sigma_x) / \sigma_x \)

and, by a change of variable,

\[ g_2(0) = \frac{1}{\sigma_x^2} \int \left[ f \left( \frac{x}{\sigma_x} \right) \right]^2 \, dx = \frac{1}{\sigma_x^2} \int f(y)^2 \, dy = \frac{1}{\sigma_x^2} f_2(0). \]

Thus, for computing \( e_{R,K} \) we can assume without loss of generality that \( \sigma_x = 1 \). The values of \( f_2(0) \) and \( e_{R,K} \) for common probability density functions are shown in Table 1. In that table we report the results also for the Student’s \( t \) density with 5 and 3 degrees of freedom. Since this distribution is widely used in financial econometrics and robust statistics we report the formula for calculating \( f_2(0) \) relative to the Student’s \( t \) distribution with variance 1 and \( \nu \) degrees of freedom:

\[ f_2(0) = \frac{\sqrt{\pi} \Gamma \left( \frac{1}{2} + \nu \right)}{\sqrt{\nu - 2} B \left( \frac{\nu}{2}, \frac{3}{2} \right) \Gamma(1+\nu)}. \]
Table 1: ARE of the RKPSS with respect to the KPSS for a selection of symmetric distributions.

<table>
<thead>
<tr>
<th>Distribution</th>
<th>$f_2(0)$</th>
<th>$\epsilon_{R,K}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>$\frac{1}{2\sqrt{\pi}}$</td>
<td>0.977</td>
</tr>
<tr>
<td>Uniform</td>
<td>$\frac{1}{\sqrt{12}}$</td>
<td>1.000</td>
</tr>
<tr>
<td>Logistic</td>
<td>$\frac{2}{3\sqrt{\pi}}$</td>
<td>1.047</td>
</tr>
<tr>
<td>Student5</td>
<td>$\frac{4\sqrt{3\pi}}{3}$</td>
<td>1.114</td>
</tr>
<tr>
<td>Laplace</td>
<td>$\frac{1}{2\sqrt{2}}$</td>
<td>1.225</td>
</tr>
<tr>
<td>Student3</td>
<td>$\frac{\pi}{2\sqrt{3}}$</td>
<td>1.378</td>
</tr>
</tbody>
</table>

with $B$ and $\Gamma$ representing the functions Beta and Gamma. By inspecting this expression we see that the relative efficiency of the RKPSS test with respect to the KPSS approaches infinity as $\nu \downarrow 2$. Another important class of probability distributions both in statistics and financial econometrics is the generalized error distribution\(^2\) (GED) family, whose shape parameter $r \in (0, \infty]$ determines the tail thickness: for $r = 2$ the GED is a standard normal distribution, for $r < 2$ the GED is leptokurtic and for $r > 2$ the GED is platykurtic. Moreover, for $r = 1$ the GED is the standard Laplace distribution and for $r = \infty$ it is a uniform distribution on $(-\sqrt{3}, \sqrt{3})$ (cf. Nelson, 1991, for instance). For the GED($r$) we have

$$f_2(0) = 2^{-1-\frac{1}{r}} r \Gamma \left(\frac{3}{r}\right)^{1/2} \Gamma \left(\frac{1}{r}\right)^{-3/2}.$$

The ARE of the RKPSS with respect to the KPSS as a function of the shape parameter $r$ is plotted in Figure 1. The range of the ARE is $(0.934, \infty)$ and the RKPSS is more efficient than the KPSS when $r < 1.757$.

After considering the behaviour of $f_2(0)$ under frequently used classes of densities, one may ask if there is lower bound for $f_2(0)$ among all the possible choices of nonnegative functions $f$ such that $\int f(x) \, dx = 1$, $\int xf(x) \, dx = 0$ and $\int x^2f(x) \, dx = 1$. The answer is affirmative and can be found in a very old paper by Hodges and Lehmann (1956)\(^3\): such a density is

$$f(x) = \begin{cases} 
\frac{3}{20\sqrt{3}}(5-x^2), & \text{for } x^2 \leq 5 \\
0, & \text{for } x^2 > 5
\end{cases}$$

and $f_2(0) = 3\sqrt{5}/25$. Thus, indicating with $\mathcal{H}$ the class of densities with zero mean and finite variance, we have

$$\min_{f \in \mathcal{H}} \epsilon_{R,K} = \frac{6\sqrt{15}}{25} \approx 0.930.$$

\(^2\)Sometimes it is referred to as exponential power distribution or generalized normal distribution.

\(^3\)The solution can be found in the equations (1.12) and (1.13) of that paper, but notice that equation (1.13) has a couple of typos and should read $a = \sqrt{5}$, $b = 3/(20\sqrt{5})$. 

8
It is to remark that the ARE of the RKPSS test is exactly the square root of the ARE of the Wilcoxon test with respect to the $t$ test for a location parameter. This implies that its efficiency gains are larger and the efficiency losses smaller when compared to usual rank-based inference. This fact is due to the $T$-consistency of the $\eta_{\mu,T}^{R}$ statistic under I(1) alternatives as opposed to the typical $\sqrt{T}$-consistency of most statistical tests and estimators.

From the above discussion we can conclude that against an extremely mild loss of efficiency in the Gaussian case, for which the KPSS statistic enjoys some optimality properties (LBI test), the RKPSS test can have significant efficiency gains in many empirically relevant situations. The RKPSS is particularly (asymptotically relatively) efficient under heavy-tailed distributions and the loss of ARE in the least favorable case is very moderate (-7%).

If we relax the independency assumption, then the variances $\sigma^2$ and $\sigma_x^2$ are given, respectively, by (9) and

$$\sigma_x^2 = \text{Var}(X_t) \left[ 1 + 2 \sum_{i=1}^{\infty} \rho_{1,i} \right],$$

with $\rho_{i,j}$ Pearson’s correlation coefficient between $X_i$ and $X_j$. Thus, the ARE values computed above must be multiplied by the factor

$$\kappa := \sqrt{\frac{1 + 2 \sum_{i=1}^{\infty} \rho_{1,i}}{1 + 2 \sum_{i=1}^{\infty} \varrho_{1,i}}}.$$

In the Gaussian case we have

$$\rho = \frac{6}{\pi} \arcsin \left( \frac{\varrho}{2} \right), \quad (16)$$

where $\rho$ is Spearman’s and $\varrho$ Pearson’s correlation coefficient. Therefore, the signs concord, but $|\rho| \leq |\varrho|$, even though the two “rho’s” tend to be very close:
the sup $|\rho - \varrho| = 0.018$ is achieved for $\varrho = \pm 2\sqrt{1 - 9/\pi^2} \approx \pm 0.59$. If we take a Gaussian AR(1) process with autoregressive coefficient $\varrho$, the ARE (of RKPSS with respect to KPSS) increases for $\varrho > 0$ and decreases for $\varrho < 0$ if compared to the i.i.d. case. In this case the $\kappa$ ratio is given by

$$\kappa = \sqrt{\frac{(1 + \varrho)(1 - \rho)}{(1 - \varrho)(1 + \rho)}}$$

with $\rho$ given by (16). If we consider the interval $\varrho \in [-0.999, 0.999]$, then $\kappa$ is a strictly increasing function of $\varrho$ with range $[0.952, 1.050]$. Figure 2 plots $\kappa$ against $\varrho$: their relation appears to be almost linear. In the Gaussian AR(1) case, the RKPSS test is asymptotically more efficient than the KPSS for $\varrho > 0.5$. Notice that for economic time series a persistence of this kind or stronger is quite common and therefore the RKPSS statistic may be preferable even under Gaussianity. Our simulations in Section 5 cover also the case of a Gaussian AR(1) with autoregressive coefficient 0.5, in which the RKPSS and the KPSS have the same efficiency under local alternatives: the size-adjusted power of the two tests can be found in Table 5, and indeed, for small values of the signal-to-noise ratio, the power figures of the two tests are virtually indistinguishable.

We conclude this section by characterizing the distribution of the KPSS and RKPSS statistics under the local alternative (14). If, for $c \in \mathbb{R}$, we define the random variable

$$H(c) := \int_0^1 \left[V(r) + cK(r)\right]^2 \, dr,$$

the continuous mapping theorem and the consistency of the long run variance estimators under local alternatives assure that

$$\hat{\eta}_{\tau,T} \Rightarrow H(\sigma_z/\sigma_x), \quad \text{and} \quad \hat{\eta}_{\tau,T}^R \Rightarrow H(f_2(0)\sigma_z/\sigma).$$

Since $V(r) + cK(r)$ is a Gaussian process with variance increasing with $c^2$, the quantiles of $H(c)$ are also increasing functions of $|c|$. In particular, very good
approximations of the 90th, 95th and 99th percentiles are given by
\[ q_{90} = 0.347 + 0.0284c^2, \quad q_{95} = 0.461 + 0.0402c^2, \quad q_{99} = 0.745 + 0.0686c^2. \]

4. Rank KPSS for trend stationarity

In many real-world applications the econometrician is interested in testing for stationarity after a deterministic monotonic trend has been removed from the data. In principle, this case could be handled in our RKPSS framework by regressing the ranks of the observations on a linear trend and then computing the KPSS statistic on the residuals of this peculiar regression. This approach would have the advantage of making the parametric specification of the trend component unnecessary and we could reasonably expect the test statistic to converge to the same distribution as that of the KPSS for trend-stationarity. Unfortunately, this is not the case and, under the null, the asymptotic distribution of the test statistic depends on the specification of the trend and of the distribution of the observations. The following simple example helps clarifying the last statement. Consider the following data generating process:

\[ Y_t = t + X_t, \quad t = 1, \ldots, T, \]

with \( X_t \) uniformly distributed on \((-0.5, 0.5)\). The increments of \( Y_t \) are always positive, and the rank of \( Y_t \) among \( \{Y_1, \ldots, Y_T\} \) is \( t \). Thus, the residuals of the regression on a linear trend equal zero for every \( t \) and the KPSS statistic computed on the residuals is also equal to zero with probability one. On the contrary, if the slope of the linear trend in the data is zero, it is easy to prove, using the asymptotic representation in Theorem 1, that the asymptotic distribution of the KPSS statistic computed on these residuals is the same as that of standard KPSS statistic computed on linearly detrended data.

In finite samples, the event that all the residuals of the regression of the ranks on a linear trend are zero may have probability one or close to one even under the alternative hypothesis of integration,

\[ Y_t = \beta t + \sum_{i=1}^{t} X_i, \quad t = 1, \ldots, T, \]

with \( X_t \) mean-zero stationary sequence. If \( X_t \) is uniformly distributed on \((-\beta/T, \beta/T)\), then the ranks of \( Y_t \) are increasing and the residuals of their regression on a linear trend are all equal to zero (and so the KPSS statistic). Even though this case is relevant only in finite samples, it is certainly not desirable for a level-stationarity test that its power decreases with the “strength” of the deterministic trend.

---

4These quantile functions were obtained by regressing empirical percentiles of the process \( H(c) \) on \( c^2 \) for \( c = 0, 0.1, \ldots, 5 \). The empirical distribution of \( H(c) \) was obtained by generating 100,000 pairs of Gaussian series of length \( T = 1000 \). The regression \( R^2 \) was above 0.99 in all cases.
So, as in Kwiatkowski et al. (1992), we limit our attention to the null hypothesis of stationarity on a linear trend, where the data are generated by

\[ Y_t = \alpha + \beta t + X_t, \quad (18) \]

with \( X_t \) short memory stationary sequence.

Kwiatkowski et al. (1992) provide the asymptotic distribution of their statistic applied to the residuals of a regression on a linear trend estimated by means of ordinary least squares. This distribution turns out to be a functional of a second-level Brownian bridge.

Turning back to our rank KPSS test, a natural class of regression estimators to look into is that of \( R \)-estimators like those proposed by Adiche (1967), Jurečková (1971) and Jaeckel (1972). In particular, following Jurečková’s (1971) approach the \( R \)-estimator of a regression on a linear trend is defined as the \( b \) that solves

\[ \sum_{t=1}^{T} \left( t - \frac{T + 1}{2} \right) a_T(R_{T,t}(b)) = 0, \quad (19) \]

where \( a_T \) is the, so called, score function, and \( R_{T,t}(b) \) is the rank of \( Y_t - bt \) among \( \{Y_1 - b_1, \ldots, Y_T - bT\} \). Since in our RKPSS statistic we are implicitly using Wilcoxon type scores, i.e.

\[ a_T(R_{T,t}) = \frac{R_{T,t}}{T} - \frac{T + 1}{2}, \]

it is natural to use them also in (19).

Fortunately, for the case of regression on a linear trend there is an easier to compute estimator, which enjoys the same asymptotic behaviour of the \( R \)-estimator (19) with Wilcoxon scores. Indeed, Jaeckel (1972, Section 4) shows that, under this condition, the Theil-Sen estimator,

\[ \tilde{\beta}_T := \text{median} \left\{ \frac{Y_j - Y_i}{j - i}; \ 1 \leq i < j \leq T \right\}, \quad (20) \]

proposed by Theil (1950) and generalized by Sen (1968), has the same asymptotic distribution as the \( R \)-estimator with Wilcoxon scores.

Sen (1968) derives the asymptotic distribution of (20) under i.i.d. regression errors with a continuous distribution, and proves that in that case its asymptotic relative efficiency with respect to the OLS estimator (i) is equal to \( 3/\pi = 0.955 \) if the error distribution is normal, (ii) is greater than unit if the error distribution is Laplace or logistic, (iii) may be indefinitely large for distributions with heavy tails such as Student’s \( t \) with few degrees of freedom, (iv) it cannot be less than 0.864 if the error distribution is continuous.

The asymptotic behaviour of the Theil-Sen estimator for random regressors is studied by Wang (2005), while Peng et al. (2008) analyze it under more

\[ \text{median} \left\{ \frac{Y_j - Y_i}{j - i}; \ 1 \leq i < j \leq T \right\}, \quad (20) \]
general conditions on the distribution of the regression errors. But, to the best of our knowledge, there is no published result on the behaviour of the Theil-Sen estimator when the regression errors are serially dependent. The following theorem fills this gap.

**Theorem 4.** Let the linear model (18) hold with $X_t$ strictly stationary and ergodic having a continuous distribution with density $f$. Then, $\hat{\beta}_T$ is consistent for $\beta$.

Furthermore, if the regression errors $\{X_t\}$ are strong mixing with mixing coefficients $\alpha(n)$ such that $\sum_{n=1}^{\infty} \alpha(n) < \infty$, and $f_2(0) := \int f(x)^2 \, dx < \infty$, then

$$Q_T(\hat{\beta}_T - \beta) \Rightarrow N\left(0, \frac{\sigma^2}{f_2(0)^2}\right),$$

where $Q_T := \sqrt{T(T^2 - 1)/12}$, and $\sigma^2$ as in equation (9).

Finally, under the same conditions,

$$Q_T(\hat{\beta}_T - \beta) = \sqrt{\frac{12}{T}} \sum_{t=1}^{T} \left[ F(X_t) - 1/2 \right] \left[ t - \frac{T+1}{2} \right] + o_p(1).$$

Now, for $t = 1, \ldots, T$ define the regression residuals $\tilde{X}_{T,t} = Y_t - \tilde{\beta}_T t$, their ranks

$$\hat{R}_{T,t} = \sum_{i=1}^{T} I\{\tilde{X}_i \leq \tilde{X}_t\}, \quad (21)$$

and the partial sum process

$$\tilde{S}_{T,t} = \sum_{i=1}^{t} \left( \frac{\hat{R}_{T,i}}{T} - \frac{T+1}{2T} \right). \quad (22)$$

Analogously to Kwiatkowski et al. (1992), define the test statistic for trend-stationarity as

$$\eta^R_{r,T} := T^{-2} \sum_{i=t}^{T} \tilde{S}_{T,t}^2, \quad (23)$$

the kernel estimator of the long-run variance of $\hat{R}_{T,t}$ as

$$\hat{\sigma}^2_T := \frac{1}{T} \sum_{s=1}^{T} \sum_{t=1}^{T} k \left( \frac{s-t}{\gamma_T} \right) \left[ \frac{\hat{R}_{T,s}}{T} - \frac{T+1}{2T} \right] \left[ \frac{\hat{R}_{T,t}}{T} - \frac{T+1}{2T} \right], \quad (24)$$

and $\hat{\eta}^R_{r,T} := \eta^R_{r,T} / \hat{\sigma}^2_T$.

**Theorem 5.** Under model (18), Assumption 1 and $f_2(0) := \int f(x)^2 \, dx < \infty$,

$$\eta^R_{r,T} \Rightarrow \sigma \int V_2(r)^2 \, dr,$$
where \( V_2(r) \) is a second-level Brownian bridge and \( \sigma \) as in equation (9).

Under the above assumptions and Assumptions 2,

\[
\hat{\eta}^R_{\tau,T} \Rightarrow \int V_2(r)^2 \, dr.
\]

Again, the asymptotic distribution of our RKPSS statistic for trend-stationarity coincides with that of the KPSS statistic for the same null hypothesis.

5. Simulations

In this section we compare size and power of the KPSS and indicator KPSS (IKPSS) tests with the rank KPSS test that we are proposing. de Jong et al. (2007) found that their IKPSS test has good size properties regardless of the distribution, but performs better than the KPSS in term of power only for extremely fat-tailed distributions.

For our simulation experiment we reproduce the design of de Jong et al. (2007). We consider sample sizes ranging from \( T = 50 \) to \( T = 5000 \) of Student’s \( t \) distributed observations. Specifically, we generate from normal (\( t_\infty \)), \( t_5 \), \( t_3 \), \( t_2 \) and Cauchy (\( t_1 \)) random variables. Notice that \( t_2 \) has finite expectation but infinite variance, while for Cauchy neither the mean nor the variance exist. We also consider the local to finite variance (also local to finite mean) alternative in which \( x_t = x_{1,t} + T^{-1/2}x_{2,t} \), where \( x_1 \) is normal and \( x_2 \) Cauchy. Our results are based on 20,000 replications.

As in de Jong et al. (2007) most of our experiments are for the case that the data (under the null) or the innovations (under the alternative) are i.i.d. We consider few cases in which they are AR(1) with parameter \( \phi = 0.5 \).

As for the estimation of the long-run variance, we slightly modify the simulation design of de Jong et al. (2007). As de Jong et al. (2007) for i.i.d. data we implement the case of no lags (the kernel is the Dirac measure for the singleton \( \{0\} \)), that in the tables will be denoted as \( \gamma_T = \gamma_{T,0} \), and the case of Bartlett kernel\(^7\) with bandwidth \( \gamma_T = 4(T/100)^{1/4} \), which we will denote as \( \gamma_T = \gamma_{T,1/4} \). But for AR(1) data, we use a Bartlett kernel with bandwidth \( \gamma_T = 1.1447 \cdot (1.7778 \cdot T)^{1/3} \), which, according to Andrews (1991) is optimal for our AR(1) process. In the tables, we will refer to this case as \( \gamma_T = \gamma_{T,1/3} \).

We first consider the size of the test. Panel (a) of Table 2 reports the actual size of the tests for i.i.d. data and \( \gamma_T = \gamma_{T,0} \) with nominal size set to 5%. Both the RKPSS and the IKPSS have sizes very close to 0.05 for any sample size and for any distribution considered. As de Jong et al. (2007) noticed, the usual

\(^6\)The simulations have been implemented in Ox version 5 by Doornik (2007) using the internal probability package.

\(^7\)The Bartlett kernel is defined by

\[
k(x) = \begin{cases} 
1 - |x|, & \text{for } |x| \leq 1, \\
0, & \text{otherwise}.
\end{cases}
\]
KPSS test is not robust to the Cauchy and local to finite variance distributions: it rejects too seldom. Nonetheless, it is reasonably robust to other fat tailed distributions, such as $t_3$ and even $t_2$. This is surprising, since the assumption of finite variance needed to derive the asymptotic distribution of the KPSS statistic is violated by the $t_2$.

Panel (b) of the same table illustrates the actual sizes when the HAC estimator of the long-run variance is used. In the i.i.d. case (with $\gamma_T = \gamma_{T,1/4}$) the conclusions are similar to those for panel (a). For the AR(1) case (with $\gamma_T = \gamma_{T,1/3}$) there is a clear tendency to over-reject, which for the KPSS is compensated by the under-rejection in the Cauchy case. All the tests seem to approach the nominal size as $T$ grows, with the exception of the KPSS under Cauchy distributed innovations, which tends to be undersized in large samples.

In Tables 3 and 4 we report the power and the size-adjusted power of the tests for the case $\gamma_T = \gamma_{T,1/4}$. Power and size-adjusted power will differ non-trivially only in those cases for which we found size distortion in Table 2, namely the Cauchy and local to finite variance.

We parameterize the unit root alternative in the same way as in Kwiatkowski et al. (1992) and de Jong et al. (2007). That is, we generate the i.i.d. sequences $\{\varepsilon_t\}$ and $\{\eta_t\}$ from the same distribution and compute the time series $\{x_t\}$ from

$$
\mu_0 = 0, \quad \mu_t = \mu_{t-1} + \sqrt{\lambda} \eta_t, \quad x_t = \mu_t + \varepsilon_t,
$$

where $\lambda$ is the signal-to-noise ratio, that measures the relative importance of the random walk component.

The main result emerging from these simulations is that the RKPSS test behaves similarly to the KPSS under normality and similarly to the IKPSS under Cauchy, while in all the other cases it shows higher power than its competitors.

The same conclusion may be drawn by observing the figures in Table 5. In this table, we report size-adjusted powers for the three tests with HAC-estimated long-run variance. For the autoregressive cases, we generated sample paths from random-walk plus AR(1) processes (i.e. $\varepsilon_t$ is AR(1) while $\eta_t$ is i.i.d.). Again, the RKPSS test behaves very similarly to the best between the KPSS and the IKPSS in the two extreme situations (normal and Cauchy), while it achieves better power in virtually all intermediate situations.

The results of a similar simulation experiment conducted on the detrended versions of the three test statistics are reported in Tables 6, 7 and 8 (we report only size-adjusted power). The conclusions are similar to the ones drawn for the level-stationarity tests with the exception that the IKPSS is now oversized also under i.i.d. data in small samples.

---

8The size problem of the KPSS test in presence of dependence is well known and many solutions have been proposed in the literature. Most of these solutions apply to the RKPSS and IKPSS statistics as well, but their implementation goes beyond the scope of this paper, which focuses on robustness to fat-tails.
Table 2: Size, KPSS, Rank KPSS and Indicator KPSS: (a) $\gamma_T = \gamma_{T,0} = 0$, (b) $\gamma_T = \gamma_{T,1/4}$ for i.i.d. data, $\gamma_T = \gamma_{T,1/3}$ for AR(1) data.

<table>
<thead>
<tr>
<th>$T$</th>
<th>Normal</th>
<th>$t_1$</th>
<th>Normal</th>
<th>$t_1$</th>
<th>Local</th>
<th>Normal, $\phi = 5$</th>
<th>$t_1$, $\phi = 0.5$</th>
<th>CDF</th>
<th>Normal, $\phi = 5$</th>
<th>$t_1$, $\phi = 0.5$</th>
<th>CDF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>KPSS</td>
<td>KPSS</td>
<td>KPSS</td>
<td>KPSS</td>
<td>MLP</td>
<td>KPSS</td>
<td>KPSS</td>
<td>MLP</td>
<td>KPSS</td>
<td>KPSS</td>
<td>MLP</td>
</tr>
<tr>
<td>(a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>0.052</td>
<td>0.052</td>
<td>0.053</td>
<td>0.052</td>
<td>0.055</td>
<td>0.052</td>
<td>0.048</td>
<td>0.053</td>
<td>0.052</td>
<td>0.044</td>
<td>0.052</td>
</tr>
<tr>
<td>100</td>
<td>0.049</td>
<td>0.049</td>
<td>0.051</td>
<td>0.049</td>
<td>0.051</td>
<td>0.050</td>
<td>0.049</td>
<td>0.051</td>
<td>0.050</td>
<td>0.044</td>
<td>0.051</td>
</tr>
<tr>
<td>200</td>
<td>0.052</td>
<td>0.052</td>
<td>0.052</td>
<td>0.052</td>
<td>0.052</td>
<td>0.050</td>
<td>0.050</td>
<td>0.051</td>
<td>0.050</td>
<td>0.042</td>
<td>0.051</td>
</tr>
<tr>
<td>500</td>
<td>0.052</td>
<td>0.050</td>
<td>0.050</td>
<td>0.052</td>
<td>0.051</td>
<td>0.049</td>
<td>0.049</td>
<td>0.051</td>
<td>0.046</td>
<td>0.050</td>
<td>0.050</td>
</tr>
<tr>
<td>1000</td>
<td>0.051</td>
<td>0.050</td>
<td>0.052</td>
<td>0.052</td>
<td>0.051</td>
<td>0.051</td>
<td>0.051</td>
<td>0.051</td>
<td>0.046</td>
<td>0.050</td>
<td>0.050</td>
</tr>
<tr>
<td>2000</td>
<td>0.050</td>
<td>0.050</td>
<td>0.050</td>
<td>0.050</td>
<td>0.049</td>
<td>0.049</td>
<td>0.049</td>
<td>0.049</td>
<td>0.046</td>
<td>0.050</td>
<td>0.050</td>
</tr>
<tr>
<td>5000</td>
<td>0.053</td>
<td>0.053</td>
<td>0.053</td>
<td>0.050</td>
<td>0.049</td>
<td>0.049</td>
<td>0.049</td>
<td>0.049</td>
<td>0.047</td>
<td>0.048</td>
<td>0.045</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$T$</th>
<th>Normal</th>
<th>$t_1$</th>
<th>Normal</th>
<th>$t_1$</th>
<th>Local</th>
<th>Normal, $\phi = 5$</th>
<th>$t_1$, $\phi = 0.5$</th>
<th>CDF</th>
<th>Normal, $\phi = 5$</th>
<th>$t_1$, $\phi = 0.5$</th>
<th>CDF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>KPSS</td>
<td>KPSS</td>
<td>KPSS</td>
<td>KPSS</td>
<td>MLP</td>
<td>KPSS</td>
<td>KPSS</td>
<td>MLP</td>
<td>KPSS</td>
<td>KPSS</td>
<td>MLP</td>
</tr>
<tr>
<td>(b)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>0.043</td>
<td>0.045</td>
<td>0.046</td>
<td>0.042</td>
<td>0.047</td>
<td>0.046</td>
<td>0.023</td>
<td>0.045</td>
<td>0.047</td>
<td>0.078</td>
<td>0.078</td>
</tr>
<tr>
<td>100</td>
<td>0.044</td>
<td>0.045</td>
<td>0.048</td>
<td>0.045</td>
<td>0.046</td>
<td>0.045</td>
<td>0.028</td>
<td>0.047</td>
<td>0.046</td>
<td>0.077</td>
<td>0.077</td>
</tr>
<tr>
<td>200</td>
<td>0.050</td>
<td>0.050</td>
<td>0.048</td>
<td>0.046</td>
<td>0.047</td>
<td>0.048</td>
<td>0.026</td>
<td>0.048</td>
<td>0.045</td>
<td>0.073</td>
<td>0.074</td>
</tr>
<tr>
<td>500</td>
<td>0.050</td>
<td>0.049</td>
<td>0.049</td>
<td>0.047</td>
<td>0.049</td>
<td>0.050</td>
<td>0.027</td>
<td>0.050</td>
<td>0.050</td>
<td>0.069</td>
<td>0.068</td>
</tr>
<tr>
<td>1000</td>
<td>0.050</td>
<td>0.050</td>
<td>0.051</td>
<td>0.050</td>
<td>0.048</td>
<td>0.050</td>
<td>0.029</td>
<td>0.052</td>
<td>0.050</td>
<td>0.066</td>
<td>0.064</td>
</tr>
<tr>
<td>2000</td>
<td>0.048</td>
<td>0.049</td>
<td>0.050</td>
<td>0.049</td>
<td>0.048</td>
<td>0.049</td>
<td>0.027</td>
<td>0.049</td>
<td>0.047</td>
<td>0.060</td>
<td>0.060</td>
</tr>
<tr>
<td>5000</td>
<td>0.053</td>
<td>0.051</td>
<td>0.051</td>
<td>0.049</td>
<td>0.049</td>
<td>0.048</td>
<td>0.027</td>
<td>0.047</td>
<td>0.049</td>
<td>0.061</td>
<td>0.060</td>
</tr>
</tbody>
</table>
### Table 3: Power, KPSS, Rank KPSS and Indicator KPSS: $\gamma_T = \gamma_{T,0} = 0.$

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$T$</th>
<th>Normal</th>
<th>Local</th>
<th>Cache</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$\gamma_T$</td>
<td>$\gamma_T$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$t_{\lambda}$</td>
<td>$t_{\lambda}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\gamma_T$</td>
<td>$\gamma_T$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$t_{\lambda}$</td>
<td>$t_{\lambda}$</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$T$</th>
<th>Normal</th>
<th>Local</th>
<th>Cache</th>
</tr>
</thead>
<tbody>
<tr>
<td>.001</td>
<td>50</td>
<td>0.053</td>
<td>0.051</td>
<td>0.053</td>
</tr>
<tr>
<td>.001</td>
<td>100</td>
<td>0.063</td>
<td>0.061</td>
<td>0.058</td>
</tr>
<tr>
<td>.001</td>
<td>200</td>
<td>0.097</td>
<td>0.096</td>
<td>0.082</td>
</tr>
<tr>
<td>.001</td>
<td>500</td>
<td>0.312</td>
<td>0.300</td>
<td>0.234</td>
</tr>
<tr>
<td>.001</td>
<td>1000</td>
<td>0.607</td>
<td>0.512</td>
<td>0.315</td>
</tr>
<tr>
<td>.001</td>
<td>2000</td>
<td>0.865</td>
<td>0.794</td>
<td>0.657</td>
</tr>
<tr>
<td>.001</td>
<td>5000</td>
<td>0.991</td>
<td>0.990</td>
<td>0.990</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$T$</th>
<th>Normal</th>
<th>Local</th>
<th>Cache</th>
</tr>
</thead>
<tbody>
<tr>
<td>.001</td>
<td>50</td>
<td>0.078</td>
<td>0.068</td>
<td>0.068</td>
</tr>
<tr>
<td>.001</td>
<td>100</td>
<td>0.168</td>
<td>0.131</td>
<td>0.131</td>
</tr>
<tr>
<td>.001</td>
<td>200</td>
<td>0.400</td>
<td>0.319</td>
<td>0.319</td>
</tr>
<tr>
<td>.001</td>
<td>500</td>
<td>0.753</td>
<td>0.698</td>
<td>0.698</td>
</tr>
<tr>
<td>.001</td>
<td>1000</td>
<td>0.956</td>
<td>0.917</td>
<td>0.917</td>
</tr>
<tr>
<td>.001</td>
<td>2000</td>
<td>0.997</td>
<td>0.991</td>
<td>0.991</td>
</tr>
<tr>
<td>.001</td>
<td>5000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$T$</th>
<th>Normal</th>
<th>Local</th>
<th>Cache</th>
</tr>
</thead>
<tbody>
<tr>
<td>.01</td>
<td>50</td>
<td>0.297</td>
<td>0.223</td>
<td>0.223</td>
</tr>
<tr>
<td>.01</td>
<td>100</td>
<td>0.594</td>
<td>0.492</td>
<td>0.492</td>
</tr>
<tr>
<td>.01</td>
<td>200</td>
<td>0.842</td>
<td>0.744</td>
<td>0.744</td>
</tr>
<tr>
<td>.01</td>
<td>500</td>
<td>0.988</td>
<td>0.973</td>
<td>0.973</td>
</tr>
<tr>
<td>.01</td>
<td>1000</td>
<td>0.999</td>
<td>0.998</td>
<td>0.998</td>
</tr>
<tr>
<td>.01</td>
<td>2000</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
</tr>
<tr>
<td>.01</td>
<td>5000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$T$</th>
<th>Normal</th>
<th>Local</th>
<th>Cache</th>
</tr>
</thead>
<tbody>
<tr>
<td>.1</td>
<td>50</td>
<td>0.729</td>
<td>0.631</td>
<td>0.631</td>
</tr>
<tr>
<td>.1</td>
<td>100</td>
<td>0.923</td>
<td>0.871</td>
<td>0.871</td>
</tr>
<tr>
<td>.1</td>
<td>200</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
</tr>
<tr>
<td>.1</td>
<td>500</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>.1</td>
<td>1000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>.1</td>
<td>2000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>.1</td>
<td>5000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$T$</th>
<th>Normal</th>
<th>Local</th>
<th>Cache</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50</td>
<td>0.924</td>
<td>0.860</td>
<td>0.860</td>
</tr>
<tr>
<td>1</td>
<td>100</td>
<td>0.987</td>
<td>0.967</td>
<td>0.967</td>
</tr>
<tr>
<td>1</td>
<td>200</td>
<td>0.998</td>
<td>0.997</td>
<td>0.997</td>
</tr>
<tr>
<td>1</td>
<td>500</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>1</td>
<td>1000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>1</td>
<td>2000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>1</td>
<td>5000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>
Table 4: Size-adjusted power, KPSS, Rank KPSS and Indicator KPSS: $\gamma_T = \gamma_{T,0} = 0$.

<table>
<thead>
<tr>
<th>$\lambda_T$</th>
<th>Normal</th>
<th>Local</th>
<th>50</th>
<th>100</th>
<th>200</th>
<th>500</th>
<th>1000</th>
<th>2000</th>
<th>5000</th>
<th>10000</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma_T$</td>
<td></td>
<td></td>
<td>0.0001</td>
<td>0.001</td>
<td>0.01</td>
<td>0.1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\gamma_{T,0}$</td>
<td></td>
<td>0.053</td>
<td>0.033</td>
<td>0.017</td>
<td>0.009</td>
<td>0.005</td>
<td>0.002</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>$\gamma_{T,1}$</td>
<td></td>
<td>0.053</td>
<td>0.033</td>
<td>0.017</td>
<td>0.009</td>
<td>0.005</td>
<td>0.002</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>$\gamma_{T,2}$</td>
<td></td>
<td>0.053</td>
<td>0.033</td>
<td>0.017</td>
<td>0.009</td>
<td>0.005</td>
<td>0.002</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>$\gamma_{T,3}$</td>
<td></td>
<td>0.053</td>
<td>0.033</td>
<td>0.017</td>
<td>0.009</td>
<td>0.005</td>
<td>0.002</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>$\gamma_{T,4}$</td>
<td></td>
<td>0.053</td>
<td>0.033</td>
<td>0.017</td>
<td>0.009</td>
<td>0.005</td>
<td>0.002</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>$\gamma_{T,5}$</td>
<td></td>
<td>0.053</td>
<td>0.033</td>
<td>0.017</td>
<td>0.009</td>
<td>0.005</td>
<td>0.002</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>$\gamma_{T,6}$</td>
<td></td>
<td>0.053</td>
<td>0.033</td>
<td>0.017</td>
<td>0.009</td>
<td>0.005</td>
<td>0.002</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>$\gamma_{T,7}$</td>
<td></td>
<td>0.053</td>
<td>0.033</td>
<td>0.017</td>
<td>0.009</td>
<td>0.005</td>
<td>0.002</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>$\gamma_{T,8}$</td>
<td></td>
<td>0.053</td>
<td>0.033</td>
<td>0.017</td>
<td>0.009</td>
<td>0.005</td>
<td>0.002</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>$\gamma_{T,9}$</td>
<td></td>
<td>0.053</td>
<td>0.033</td>
<td>0.017</td>
<td>0.009</td>
<td>0.005</td>
<td>0.002</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>$\gamma_{T,10}$</td>
<td></td>
<td>0.053</td>
<td>0.033</td>
<td>0.017</td>
<td>0.009</td>
<td>0.005</td>
<td>0.002</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>$\gamma_{T,11}$</td>
<td></td>
<td>0.053</td>
<td>0.033</td>
<td>0.017</td>
<td>0.009</td>
<td>0.005</td>
<td>0.002</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>$\gamma_{T,12}$</td>
<td></td>
<td>0.053</td>
<td>0.033</td>
<td>0.017</td>
<td>0.009</td>
<td>0.005</td>
<td>0.002</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>$\gamma_{T,13}$</td>
<td></td>
<td>0.053</td>
<td>0.033</td>
<td>0.017</td>
<td>0.009</td>
<td>0.005</td>
<td>0.002</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>$\gamma_{T,14}$</td>
<td></td>
<td>0.053</td>
<td>0.033</td>
<td>0.017</td>
<td>0.009</td>
<td>0.005</td>
<td>0.002</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>$\gamma_{T,15}$</td>
<td></td>
<td>0.053</td>
<td>0.033</td>
<td>0.017</td>
<td>0.009</td>
<td>0.005</td>
<td>0.002</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>$\gamma_{T,16}$</td>
<td></td>
<td>0.053</td>
<td>0.033</td>
<td>0.017</td>
<td>0.009</td>
<td>0.005</td>
<td>0.002</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>$\gamma_{T,17}$</td>
<td></td>
<td>0.053</td>
<td>0.033</td>
<td>0.017</td>
<td>0.009</td>
<td>0.005</td>
<td>0.002</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>$\gamma_{T,18}$</td>
<td></td>
<td>0.053</td>
<td>0.033</td>
<td>0.017</td>
<td>0.009</td>
<td>0.005</td>
<td>0.002</td>
<td>0.001</td>
<td>0.001</td>
</tr>
</tbody>
</table>
Table 5: Size-adjusted power, KPSS, Rank KPSS and Indicator KPSS: $\gamma_T = \gamma_{T, 1/4}$ for i.i.d. noise, $\gamma_T = \gamma_{T, 1/4}$ for AR(1) noise.

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$T$</th>
<th>Normal</th>
<th>KPSS</th>
<th>IKPSS</th>
<th>Normal</th>
<th>KPSS</th>
<th>IKPSS</th>
<th>Normal</th>
<th>KPSS</th>
<th>IKPSS</th>
<th>Normal</th>
<th>KPSS</th>
<th>IKPSS</th>
<th>Normal</th>
<th>KPSS</th>
<th>IKPSS</th>
<th>Normal</th>
<th>KPSS</th>
<th>IKPSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0001</td>
<td>50</td>
<td>0.051</td>
<td>0.050</td>
<td>0.050</td>
<td>0.052</td>
<td>0.052</td>
<td>0.052</td>
<td>0.111</td>
<td>0.214</td>
<td>0.218</td>
<td>0.074</td>
<td>0.074</td>
<td>0.069</td>
<td>0.082</td>
<td>0.077</td>
<td>0.071</td>
<td>0.142</td>
<td>0.138</td>
<td>0.134</td>
</tr>
<tr>
<td>0.001</td>
<td>50</td>
<td>0.075</td>
<td>0.073</td>
<td>0.064</td>
<td>0.089</td>
<td>0.098</td>
<td>0.093</td>
<td>0.222</td>
<td>0.444</td>
<td>0.447</td>
<td>0.078</td>
<td>0.080</td>
<td>0.073</td>
<td>0.091</td>
<td>0.087</td>
<td>0.084</td>
<td>0.199</td>
<td>0.210</td>
<td>0.208</td>
</tr>
<tr>
<td>0.01</td>
<td>50</td>
<td>0.283</td>
<td>0.284</td>
<td>0.216</td>
<td>0.314</td>
<td>0.369</td>
<td>0.328</td>
<td>0.435</td>
<td>0.724</td>
<td>0.707</td>
<td>0.148</td>
<td>0.148</td>
<td>0.127</td>
<td>0.176</td>
<td>0.179</td>
<td>0.173</td>
<td>0.318</td>
<td>0.339</td>
<td>0.339</td>
</tr>
<tr>
<td>0.1</td>
<td>50</td>
<td>0.725</td>
<td>0.715</td>
<td>0.624</td>
<td>0.722</td>
<td>0.759</td>
<td>0.700</td>
<td>0.678</td>
<td>0.882</td>
<td>0.850</td>
<td>0.420</td>
<td>0.416</td>
<td>0.370</td>
<td>0.438</td>
<td>0.432</td>
<td>0.405</td>
<td>0.497</td>
<td>0.488</td>
<td>0.473</td>
</tr>
<tr>
<td>1</td>
<td>50</td>
<td>0.921</td>
<td>0.921</td>
<td>0.896</td>
<td>0.918</td>
<td>0.939</td>
<td>0.909</td>
<td>0.814</td>
<td>0.981</td>
<td>0.973</td>
<td>0.644</td>
<td>0.637</td>
<td>0.595</td>
<td>0.643</td>
<td>0.652</td>
<td>0.623</td>
<td>0.642</td>
<td>0.697</td>
<td>0.678</td>
</tr>
<tr>
<td>0.0001</td>
<td>100</td>
<td>0.064</td>
<td>0.063</td>
<td>0.056</td>
<td>0.062</td>
<td>0.067</td>
<td>0.068</td>
<td>0.178</td>
<td>0.457</td>
<td>0.471</td>
<td>0.079</td>
<td>0.080</td>
<td>0.071</td>
<td>0.080</td>
<td>0.082</td>
<td>0.080</td>
<td>0.156</td>
<td>0.216</td>
<td>0.229</td>
</tr>
<tr>
<td>0.001</td>
<td>100</td>
<td>0.094</td>
<td>0.092</td>
<td>0.079</td>
<td>0.104</td>
<td>0.139</td>
<td>0.127</td>
<td>0.270</td>
<td>0.768</td>
<td>0.783</td>
<td>0.079</td>
<td>0.081</td>
<td>0.076</td>
<td>0.085</td>
<td>0.095</td>
<td>0.093</td>
<td>0.208</td>
<td>0.386</td>
<td>0.421</td>
</tr>
<tr>
<td>0.05</td>
<td>100</td>
<td>0.385</td>
<td>0.382</td>
<td>0.274</td>
<td>0.384</td>
<td>0.455</td>
<td>0.396</td>
<td>0.319</td>
<td>0.526</td>
<td>0.529</td>
<td>0.102</td>
<td>0.102</td>
<td>0.095</td>
<td>0.111</td>
<td>0.111</td>
<td>0.105</td>
<td>0.239</td>
<td>0.414</td>
<td>0.439</td>
</tr>
<tr>
<td>0.1</td>
<td>100</td>
<td>0.786</td>
<td>0.774</td>
<td>0.698</td>
<td>0.785</td>
<td>0.869</td>
<td>0.812</td>
<td>0.708</td>
<td>0.999</td>
<td>0.998</td>
<td>0.499</td>
<td>0.501</td>
<td>0.464</td>
<td>0.517</td>
<td>0.584</td>
<td>0.570</td>
<td>0.558</td>
<td>0.897</td>
<td>0.907</td>
</tr>
<tr>
<td>0.5</td>
<td>100</td>
<td>0.955</td>
<td>0.952</td>
<td>0.916</td>
<td>0.952</td>
<td>0.981</td>
<td>0.972</td>
<td>0.827</td>
<td>1.000</td>
<td>1.000</td>
<td>0.771</td>
<td>0.769</td>
<td>0.735</td>
<td>0.765</td>
<td>0.822</td>
<td>0.813</td>
<td>0.705</td>
<td>0.977</td>
<td>0.979</td>
</tr>
<tr>
<td>1</td>
<td>100</td>
<td>0.997</td>
<td>0.997</td>
<td>0.991</td>
<td>0.996</td>
<td>0.999</td>
<td>0.998</td>
<td>0.911</td>
<td>1.000</td>
<td>1.000</td>
<td>0.933</td>
<td>0.933</td>
<td>0.930</td>
<td>0.935</td>
<td>0.957</td>
<td>0.954</td>
<td>0.829</td>
<td>0.997</td>
<td>0.997</td>
</tr>
<tr>
<td>0.0001</td>
<td>1000</td>
<td>0.999</td>
<td>0.999</td>
<td>0.998</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>0.962</td>
<td>1.000</td>
<td>1.000</td>
<td>0.995</td>
<td>0.995</td>
<td>0.993</td>
<td>0.996</td>
<td>0.998</td>
<td>0.997</td>
<td>0.952</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>0.001</td>
<td>1000</td>
<td>0.999</td>
<td>0.999</td>
<td>0.998</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>0.962</td>
<td>1.000</td>
<td>1.000</td>
<td>0.995</td>
<td>0.995</td>
<td>0.993</td>
<td>0.996</td>
<td>0.998</td>
<td>0.997</td>
<td>0.952</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>0.05</td>
<td>1000</td>
<td>0.999</td>
<td>0.999</td>
<td>0.998</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>0.962</td>
<td>1.000</td>
<td>1.000</td>
<td>0.995</td>
<td>0.995</td>
<td>0.993</td>
<td>0.996</td>
<td>0.998</td>
<td>0.997</td>
<td>0.952</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>0.1</td>
<td>1000</td>
<td>0.999</td>
<td>0.999</td>
<td>0.998</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>0.962</td>
<td>1.000</td>
<td>1.000</td>
<td>0.995</td>
<td>0.995</td>
<td>0.993</td>
<td>0.996</td>
<td>0.998</td>
<td>0.997</td>
<td>0.952</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>0.5</td>
<td>1000</td>
<td>0.999</td>
<td>0.999</td>
<td>0.998</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>0.962</td>
<td>1.000</td>
<td>1.000</td>
<td>0.995</td>
<td>0.995</td>
<td>0.993</td>
<td>0.996</td>
<td>0.998</td>
<td>0.997</td>
<td>0.952</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>1</td>
<td>1000</td>
<td>0.999</td>
<td>0.999</td>
<td>0.998</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>0.962</td>
<td>1.000</td>
<td>1.000</td>
<td>0.995</td>
<td>0.995</td>
<td>0.993</td>
<td>0.996</td>
<td>0.998</td>
<td>0.997</td>
<td>0.952</td>
<td>1.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Note: $\gamma_T = \gamma_{T, 1/4}$ for i.i.d. noise, $\gamma_T = \gamma_{T, 1/4}$ for AR(1) noise.
Table 6: Size, detrended versions of the tests: (a) i.i.d. data and $\gamma_T = \gamma_{T,0} = 0$, (b) AR(1) data and $\gamma_T = \gamma_{T,1/3}$.

<table>
<thead>
<tr>
<th>$T$</th>
<th>Normal</th>
<th>$t_5$</th>
<th>Cauchy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>KPSS</td>
<td>RKPSS</td>
<td>IKPSS</td>
</tr>
<tr>
<td>(a) i.i.d. data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>0.053</td>
<td>0.053</td>
<td>0.096</td>
</tr>
<tr>
<td>100</td>
<td>0.049</td>
<td>0.051</td>
<td>0.066</td>
</tr>
<tr>
<td>200</td>
<td>0.051</td>
<td>0.051</td>
<td>0.059</td>
</tr>
<tr>
<td>500</td>
<td>0.047</td>
<td>0.048</td>
<td>0.051</td>
</tr>
<tr>
<td>1000</td>
<td>0.051</td>
<td>0.050</td>
<td>0.051</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b) AR(1) with $\phi = 0.5$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>0.080</td>
<td>0.085</td>
<td>0.126</td>
</tr>
<tr>
<td>100</td>
<td>0.080</td>
<td>0.082</td>
<td>0.088</td>
</tr>
<tr>
<td>200</td>
<td>0.080</td>
<td>0.080</td>
<td>0.079</td>
</tr>
<tr>
<td>500</td>
<td>0.068</td>
<td>0.069</td>
<td>0.069</td>
</tr>
<tr>
<td>1000</td>
<td>0.068</td>
<td>0.067</td>
<td>0.065</td>
</tr>
</tbody>
</table>
Table 7: Size-adjusted power, detrended versions of the tests: i.i.d. data and $\gamma_T = \gamma_{T0} = 0$.

| $\lambda$ | $\frac{\gamma_T}{T}$ | $N_{50}$ | $T_{50}$ | $R_{50}$ | $S_{50}$ | $D_{50}$ | $O_{50}$ | $S_{25}$ | $T_{25}$ | $R_{25}$ | $O_{25}$ | $D_{25}$ | $S_{25}$ | $R_{25}$ | $S_{10}$ | $T_{10}$ | $R_{10}$ | $O_{10}$ | $D_{10}$ | $S_{10}$ | $R_{10}$ | $S_{1}$ | $T_{1}$ | $R_{1}$ | $O_{1}$ | $D_{1}$ | $S_{1}$ | $R_{1}$ | $S_{0.1}$ | $T_{0.1}$ | $R_{0.1}$ | $O_{0.1}$ | $D_{0.1}$ | $S_{0.1}$ | $R_{0.1}$ | $S_{0.01}$ | $T_{0.01}$ | $R_{0.01}$ | $O_{0.01}$ | $D_{0.01}$ | $S_{0.01}$ | $R_{0.01}$ | $S_{0}$ | $T_{0}$ | $R_{0}$ | $O_{0}$ | $D_{0}$ | $S_{0}$ | $R_{0}$ | $S_{0.001}$ | $T_{0.001}$ | $R_{0.001}$ | $O_{0.001}$ | $D_{0.001}$ | $S_{0.001}$ | $R_{0.001}$ |
|----------|---------------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0.0001   | 0.008               | 0.098  | 0.119  | 0.319  | 0.319  | 0.319  | 0.319  | 0.319  | 0.319  | 0.319  | 0.319  | 0.319  | 0.319  | 0.319  | 0.319  | 0.319  | 0.319  | 0.319  | 0.319  | 0.319  | 0.319  | 0.319  | 0.319  | 0.319  | 0.319  | 0.319  | 0.319  | 0.319  | 0.319  | 0.319  | 0.319  | 0.319  | 0.319  | 0.319  | 0.319  | 0.319  | 0.319  | 0.319  | 0.319  |
| 0.001    | 0.060               | 0.098  | 0.113  | 0.143  | 0.143  | 0.143  | 0.143  | 0.143  | 0.143  | 0.143  | 0.143  | 0.143  | 0.143  | 0.143  | 0.143  | 0.143  | 0.143  | 0.143  | 0.143  | 0.143  | 0.143  | 0.143  | 0.143  | 0.143  | 0.143  | 0.143  | 0.143  | 0.143  | 0.143  | 0.143  | 0.143  | 0.143  | 0.143  | 0.143  | 0.143  | 0.143  | 0.143  | 0.143  | 0.143  | 0.143  |
| 0.01     | 0.112               | 0.098  | 0.117  | 0.173  | 0.173  | 0.173  | 0.173  | 0.173  | 0.173  | 0.173  | 0.173  | 0.173  | 0.173  | 0.173  | 0.173  | 0.173  | 0.173  | 0.173  | 0.173  | 0.173  | 0.173  | 0.173  | 0.173  | 0.173  | 0.173  | 0.173  | 0.173  | 0.173  | 0.173  | 0.173  | 0.173  | 0.173  | 0.173  | 0.173  | 0.173  | 0.173  | 0.173  | 0.173  | 0.173  | 0.173  |
| 0.1      | 0.164               | 0.098  | 0.121  | 0.207  | 0.207  | 0.207  | 0.207  | 0.207  | 0.207  | 0.207  | 0.207  | 0.207  | 0.207  | 0.207  | 0.207  | 0.207  | 0.207  | 0.207  | 0.207  | 0.207  | 0.207  | 0.207  | 0.207  | 0.207  | 0.207  | 0.207  | 0.207  | 0.207  | 0.207  | 0.207  | 0.207  | 0.207  | 0.207  | 0.207  | 0.207  | 0.207  | 0.207  | 0.207  | 0.207  | 0.207  |
| 1        | 0.216               | 0.098  | 0.125  | 0.253  | 0.253  | 0.253  | 0.253  | 0.253  | 0.253  | 0.253  | 0.253  | 0.253  | 0.253  | 0.253  | 0.253  | 0.253  | 0.253  | 0.253  | 0.253  | 0.253  | 0.253  | 0.253  | 0.253  | 0.253  | 0.253  | 0.253  | 0.253  | 0.253  | 0.253  | 0.253  | 0.253  | 0.253  | 0.253  | 0.253  | 0.253  | 0.253  | 0.253  | 0.253  | 0.253  | 0.253  |
Table 8: Size-adjusted power, detrended versions of the tests: AR(1) data and $\gamma_T = \gamma_{T,1/3}$.

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$T$</th>
<th>Normal</th>
<th>$t_0$</th>
<th>KPSS</th>
<th>RKPSS</th>
<th>IKPSS</th>
<th>Cauchy</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0001</td>
<td>50</td>
<td>0.076</td>
<td>0.075</td>
<td>0.055</td>
<td>0.077</td>
<td>0.078</td>
<td>0.053</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.082</td>
<td>0.081</td>
<td>0.065</td>
<td>0.082</td>
<td>0.083</td>
<td>0.065</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>0.078</td>
<td>0.078</td>
<td>0.068</td>
<td>0.080</td>
<td>0.081</td>
<td>0.075</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>0.102</td>
<td>0.102</td>
<td>0.087</td>
<td>0.097</td>
<td>0.099</td>
<td>0.088</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.160</td>
<td>0.157</td>
<td>0.139</td>
<td>0.156</td>
<td>0.169</td>
<td>0.159</td>
</tr>
<tr>
<td>0.001</td>
<td>50</td>
<td>0.077</td>
<td>0.076</td>
<td>0.053</td>
<td>0.079</td>
<td>0.080</td>
<td>0.056</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.091</td>
<td>0.090</td>
<td>0.071</td>
<td>0.092</td>
<td>0.092</td>
<td>0.075</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>0.115</td>
<td>0.112</td>
<td>0.095</td>
<td>0.116</td>
<td>0.122</td>
<td>0.106</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>0.296</td>
<td>0.293</td>
<td>0.250</td>
<td>0.296</td>
<td>0.315</td>
<td>0.279</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.601</td>
<td>0.590</td>
<td>0.550</td>
<td>0.606</td>
<td>0.636</td>
<td>0.607</td>
</tr>
<tr>
<td>0.01</td>
<td>50</td>
<td>0.098</td>
<td>0.099</td>
<td>0.068</td>
<td>0.097</td>
<td>0.098</td>
<td>0.065</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.162</td>
<td>0.159</td>
<td>0.121</td>
<td>0.166</td>
<td>0.172</td>
<td>0.140</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>0.353</td>
<td>0.349</td>
<td>0.297</td>
<td>0.362</td>
<td>0.382</td>
<td>0.344</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>0.784</td>
<td>0.780</td>
<td>0.731</td>
<td>0.781</td>
<td>0.794</td>
<td>0.761</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.956</td>
<td>0.955</td>
<td>0.942</td>
<td>0.953</td>
<td>0.958</td>
<td>0.949</td>
</tr>
<tr>
<td>0.1</td>
<td>50</td>
<td>0.224</td>
<td>0.221</td>
<td>0.138</td>
<td>0.226</td>
<td>0.230</td>
<td>0.148</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.474</td>
<td>0.466</td>
<td>0.384</td>
<td>0.475</td>
<td>0.477</td>
<td>0.402</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>0.752</td>
<td>0.744</td>
<td>0.691</td>
<td>0.756</td>
<td>0.762</td>
<td>0.714</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>0.959</td>
<td>0.957</td>
<td>0.938</td>
<td>0.955</td>
<td>0.956</td>
<td>0.936</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.993</td>
<td>0.992</td>
<td>0.988</td>
<td>0.993</td>
<td>0.993</td>
<td>0.988</td>
</tr>
<tr>
<td>1</td>
<td>50</td>
<td>0.425</td>
<td>0.415</td>
<td>0.277</td>
<td>0.441</td>
<td>0.431</td>
<td>0.283</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.692</td>
<td>0.679</td>
<td>0.583</td>
<td>0.687</td>
<td>0.680</td>
<td>0.586</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>0.862</td>
<td>0.852</td>
<td>0.799</td>
<td>0.865</td>
<td>0.860</td>
<td>0.808</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>0.973</td>
<td>0.970</td>
<td>0.951</td>
<td>0.973</td>
<td>0.969</td>
<td>0.948</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.996</td>
<td>0.995</td>
<td>0.989</td>
<td>0.995</td>
<td>0.995</td>
<td>0.989</td>
</tr>
</tbody>
</table>
6. Empirical example

In order to explore the different behaviour of the three tests (KPSS, IKPSS and RKPSS) on real data, we consider daily electricity prices that typically show very heavy tails in their distributions. In particular, we take daily averages (excluding weekends) of the 24 hourly prices registered at three important European power exchanges: Powernext (France), EEX (Germany) and APX (Netherlands). The time span we consider is January 1st, 2004 through June 25th, 2009, which amounts to 1431 weekday observations.

As robustly shown by Bosco et al. (2010) the three series are non-stationary. Indeed, all three tests we are considering reject the hypothesis of short-memory stationarity at any usual level and reasonable bandwidth (up to \( \gamma_T = 12(T/100)^{1/3} \) for a Bartlett kernel, cf. Andrews 1991, equation 5.3 for the choice of the rate of divergence of \( \gamma_T \)). Even though some (non-robust) early works on electricity prices tended to consider them stationary, the non-stationarity of electricity prices is not surprising, in fact, in the long run, electricity prices are heavily dependent on oil and gas prices, and the logarithm of these is generally well approximated by an integrated process.

An interesting feature to test on European electricity prices is whether their pairwise ratios are stationary, or stated alternatively, if pairs of log-price time series are cointegrated with cointegrating space spanned by \((1, -1)\). This condition, termed by De Vany and Walls (1999) as strong integration of electricity markets, is important for at least two aspects: firstly, since the European Union introduced reforms to favour the creation of one European market for electricity, finding strong integration would in some way certify its success in this direction; secondly, since the degree of development of purely financial electricity derivatives markets may differ substantially in each country, strong integration would allow investors in countries with less-developed financial markets to hedge with instruments exchanged in other countries.

We tested the hypothesis \( \log p_i - \log p_j \sim I(0) \), where \( i \) and \( j \) represent any two markets in our sample, using the KPSS, IKPSS and RKPSS tests. Figure 3 shows the time series (left panels) and the test statistics as functions of the bandwidth parameter ranging for 0 to \( \gamma_T = 12(T/100)^{1/3} \) in a Bartlett kernel (right panels). The horizontal lines in the right panels denote the three critical values common to the three tests. While there is little doubt about the rejection of short-memory stationarity for the pairs Netherlands-France and Netherlands-Germany, for Germany-France the KPSS statistic is somewhat ambiguous. For bandwidths up to 10, it rejects the null at any usual level. For bandwidths in the range 10-22, it rejects at a 5% level, but not at 1%. For bandwidths larger than 22 it rejects only at a 10% level. The IKPSS and RKPSS statistics always reject at 5%, while at 1% they reject with bandwidths up to 20.

If we relay on Andrews’s (1991) data-driven bandwidth selection method, we obtain the results in Table 9, which confirm the rejection of the null for the IKPSS and RKPSS statistics at any usual level, while the KPSS statistic \( p \)-value lays in the 5%–1% limbo. Thus, at 1% level the conclusions of the robust tests and of the KPSS differ. This may be due either to the fact that, under fat
Figure 3: Log price ratio series and test statistics as functions of the bandwidth parameter $\gamma_T$.

Table 9: Stationarity tests with bandwidth parameter $\gamma_T$ chosen as in Andrews (1991). Critical values: 10%, 0.347; 5%, 0.463; 1%, 0.739.

<table>
<thead>
<tr>
<th></th>
<th>KPSS $\gamma_T$</th>
<th>IKPSS $\hat{\eta}_T$</th>
<th>RKPSS $\hat{\gamma}_T$</th>
<th>IKPSS $\hat{\eta}_T$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Germany - France</td>
<td>12 0.697</td>
<td>5 1.540</td>
<td>9 1.218</td>
<td></td>
</tr>
<tr>
<td>Netherlands - France</td>
<td>17 1.550</td>
<td>17 2.914</td>
<td>20 2.047</td>
<td></td>
</tr>
<tr>
<td>Netherlands - Germany</td>
<td>10 1.173</td>
<td>7 1.496</td>
<td>9 1.597</td>
<td></td>
</tr>
</tbody>
</table>

...tailed processes, the KPSS test is undersized or because of its lack of power, in particular when the signal to noise ratio is low (cf. Table 3). The conjecture that the right decision is that based on the RKPSS and IKPSS tests is confirmed by the application of the tests to the extended time-span January 1st, 2002 - June 25th, 2009. In this case all three tests reject the null of short-memory stationarity for any choice of bandwidth up to $\gamma_T = 12(T/100)^{1/3}$.

Now, the visual inspection of the left panels of Figure 3 suggests a mean-reversing behaviour of the three series, and this may seem contradictory with respect to the results of the tests. But, even though we did not consider the stationary long memory alternative in this paper, as Lee and Schmidt (1996) prove for the KPSS and de Jong et al. (2007) conjecture for the IKPSS, we expect the RKPSS to have power against long memory alternatives. Indeed, mixing is a necessary condition for proving the first thesis of Theorem 1.
Figure 4: Sample ACF of log \((p^\text{DE}/p^\text{FR})\) with population ACF of two ARFIMA models fitted to the data.

4 allows the comparison of the sample autocorrelation function (ACF) of the log-ratio of German and French electricity prices with those of an ARMA(2, 2) and an ARFIMA(2, 0.22, 2) process fitted to the data\(^9\). There is no doubt that the ACF implied by the ARFIMA model is a much better approximation to the sample ACF than the one implied by the ARMA model. The same conclusion can be drawn for the other two time series.

7. Concluding remarks

In this article we have shown how ranks-based inference can be fruitfully applied to econometric issues, despite the scarce attention econometricians have traditionally devoted to this important field of nonparametric statistics.

The level-stationarity test we propose is based on the widely used KPSS statistic applied to the ranks of the original data. Our approach presents a long list of advantages over the standard KPSS test and almost no drawbacks. Namely, the existence of any moment is not necessary, the test is invariant to monotonic transformations of the data, the efficiency gains for typical economic time series may be relevant while efficiency losses in the other cases are moderate. Moreover, the asymptotic distribution of our test statistic is the same as that of the KPSS and, thus, the numerous software packages implementing the latter may be used for our RKPSS without any additional effort.

For testing trend-stationarity we apply the KPSS statistic to the data detrended using the Theil-Sen estimator. Despite its computational simplicity, the Theil-Sen estimator shares the same asymptotic properties of an \(R\)-estimator with Wilcoxon scores, that is, robustness to fat-tailed distributions and moderate efficiency losses under normality or thin-tailed densities. Again, the asymptotic distribution of our test statistic is the same as that of Kwiatkowski et al.

\(^9\)The models were estimated using the Ox package ARFIMA by Doornik and Ooms (2006) (see also Doornik and Ooms, 2004) with the maximum likelihood option.
(1992) after least-squares based linear detrending.

For the aforementioned reasons, we envisage no reasons why our RKPSS test shouldn’t substitute or, at least, accompany the classical KPSS test in economic and financial applications. Furthermore, the weak convergence results and asymptotic representations proved in this article may have an interest on their own, as they extend to ranks analogous results routinely used in econometrics.

Future research related to this article should extend the RKPSS statistic to the multivariate case as in Nyblom and Harvey (2000) and Bosco et al. (2010) and derive robust cointegration tests. Finally, generalizing our results to the case of general score functions, thus including our test as well as de Jong et al.’s (2007) as special cases, would be of great theoretical interest.

Appendix A. Proofs

Proof of Theorem 1.

Let

\[ F_T(x) = \frac{1}{T} \sum_{i=1}^{T} I\{X_i \leq x\} \quad \text{and} \quad F_t(x) = \frac{1}{t} \sum_{i=1}^{t} I\{X_i \leq x\} \]

be the empirical distribution functions of, respectively, \( \{X_1, \ldots, X_T\} \) and \( \{X_1, \ldots, X_t\} \).

Weak convergence of \( \eta_{\mu,T}^R \)
Notice that

\[ T^{-1/2}S_{T,t} = tT^{-1/2} \left[ \frac{1}{t} \sum_{i=1}^{t} \left( \frac{R_{T,i}}{T} \right)^{T} \right] \]

\[ = tT^{-1/2} \int \left( F_{T}(x) - \frac{T + 1}{2T} \right) dF_{i}(x) \]

\[ = tT^{-1/2} \int \left( F_{T}(x) - \frac{T + 1}{2T} \right) \left[ d(F_{i}(x) - F_{T}(x)) + dF_{T}(x) \right] \]

\[ = tT^{-1/2} \int \left( F_{T}(x) - \frac{T + 1}{2T} \right) d(F_{i}(x) - F_{T}(x)) \]

\[ = -tT^{-1/2} \int (F_{i}(x) - F_{T}(x)) dF_{T}(x) \]

\[ = -tT^{-1/2} \int (F_{i}(x) - F_{T}(x)) [dF(x) + d(F_{T}(x) - F(x))] \]

\[ = -tT^{-1/2} \int (F_{i}(x) - F_{T}(x)) dF(x) + \]

\[ - tT^{-1/2} \int (F_{i}(x) - F_{T}(x)) d(F_{i}(x) - F(x)) \]

\[ = -tT^{-1/2} \int (F_{i}(x) - F(x)) dF(x) + \]

\[ + tT^{-1/2} \int (F_{T}(x) - F(x)) dF(x) + \]

\[ - tT^{-1/2} \int (F_{T}(x) - F(x)) d(F_{T}(x) - F(x)) \]

\[ = -\frac{t}{\sqrt{T}} \int (F_{i}(x) - F(x)) dF(x) + \]

\[ + \frac{t}{T} \int \sqrt{T}(F_{T}(x) - F(x)) dF(x) + \]

\[ - \frac{1}{\sqrt{T}} \int \frac{\sqrt{T}}{\sqrt{T}} \left( F_{i}(x) - F_{T}(x) \right) d\left[ \sqrt{T}(F_{T}(x) - F(x)) \right] \quad \text{(A.3)} \]

\[ = A_{T} + B_{T} + R_{T}, \quad \text{(A.4)} \]

where \( A_{T} \), \( B_{T} \) and \( R_{T} \) are just names for the lines, respectively, (A.1), (A.2) and (A.3) above.

Let us now consider a sequence of stochastic processes \( \{W^{*}_{T}\} = \{W^{*}_{T}(x,r); x \in \mathbb{R}, r \in [0,1]\} \), such that

\[ W^{*}_{T}(x,r) = W^{*}_{T}(x,\lfloor Tr \rfloor/T) \]

and

\[ W^{*}_{T}(x,t/T) = tT^{-1/2}(F_{i}(x) - F(x)) \quad 0 \leq t \leq T. \quad \text{(A.5)} \]

Let \( \{W^{*}\} = \{W^{*}_{T}(x,r); x \in \mathbb{R}, r \in [0,1]\} \) be a Gaussian (Kiefer) process with
\( E W^*(x, r) = 0, \forall x, r, \) and

\[
E [W^*(x, r)W^*(y, s)] = \min(r, s)\zeta(x, y) \tag{A.6}
\]

with

\[
\zeta(x, y) = E \left[ \mathbb{I}_{\{X_1 \leq x\}} \mathbb{I}_{\{X_2 \leq y\}} - F(x)F(y) \right] + \\
\sum_{k=2}^{\infty} E \left[ \mathbb{I}_{\{X_1 \leq x\}} \mathbb{I}_{\{X_k \leq y\}} + \mathbb{I}_{\{X_k \leq x\}} \mathbb{I}_{\{X_1 \leq y\}} - 2F(x)F(y) \right]. \tag{A.7}
\]

Now we need the following result: under Assumption 1 \( \zeta(x, y) < \infty, \forall x, y \in \mathbb{R} \)

and \( W_T^* \Rightarrow W^* \).

This was proven by Sen (1974, Theorem 2.2) under \( \phi \)-mixing, generalized by Yoshihara (1975, Theorem 2) to strong mixing with \( \alpha(T) = O(T^v) \) and \( v > 5/2 \), and further generalized by Shao (1986) to the same setup with \( v > 2 \).

Let us look back at equation (A.4) and note that

\[
A_T = -\int W_T^*(x, t/T) dF(x) =: -W_T^{0*}(t/T), \quad \text{say} \tag{A.8}
\]

\[
B_T = \frac{t}{T} \int W_T^*(x, 1) dF(x) = \frac{t}{T} W_T^{0*}(1). \tag{A.9}
\]

Also

\[
R_T = -T^{-1/2} \int \left[ W_T^*(x, t/T) - \frac{t}{T} W_T^*(x, 1) \right] dW_T^*(1, 1) = O_p(T^{-1/2}). \tag{A.10}
\]

Finally, by (A.6) and (A.7), we note that \( W_T^{0*} \) is an integrated stochastic process converging in distribution to the Gaussian process \( \{W^{0*}(r), 0 \leq r \leq 1\} \)

with

\[
E [W^{0*}(r)W^{0*}(s)] = \min(r, s)\sigma^2 \tag{A.11}
\]

and

\[
\sigma^2 = \int \int \zeta(x, y) dF(x) dF(y). \tag{A.12}
\]

Thus,

\[
\sigma^{-1}W_T^{0*} \Rightarrow W \tag{A.13}
\]

where \( W \) is a standard Brownian motion on \([0, 1]\).

Therefore letting \( \xi_T(r) = T^{-1/2} S_{\lfloor Tr \rfloor} \), we have

\[
\xi_T(r) = -\left[ W_T^{0*}(r) - r W_T^{0*}(1) \right] + O_p(T^{-1/2}) \tag{A.14}
\]

for every \( r \in [0, 1] \), so that

\[
\xi_T(r) \Rightarrow \sigma V(r) \tag{A.15}
\]
where $V$ is a standard Brownian bridge. By applying the continuous mapping theorem, the first statement of Theorem 1 follows:

$$
\eta^R_{\mu,T} \Rightarrow \sigma^2 \int_0^1 V(r)^2 \, dr.
$$

Now we have to prove that the variance in the equation (9) of the theorem coincides with the one defined above in equation (A.12). First, notice that $\zeta(x,y)$ may be restated as

$$
\zeta(x,y) = [F(x \wedge y) - F(x)F(y)] + 2 \sum_{k=2}^{\infty} [F_{1,k}(x,y) - F(x)F(y)],
$$

(A.16)

where $F_{1,k}$ denotes the joint distribution function of the pair $(X_1, X_k)$. We shall prove the identity of the following two expressions:

$$
\int \int [F(x \wedge y) - F(x)F(y)] \, dF(x) \, dF(y) + 2 \sum_{k=2}^{\infty} \int \int [F_{1,k}(x,y) - F(x)F(y)] \, dF(x) \, dF(y),
$$

(A.17)

and

$$
\mathbb{E} \left[ F(X_1) - \frac{1}{2} \right]^2 + \lim_{T \to \infty} 2 \sum_{k=2}^{T} \mathbb{E} \left[ F(X_1) - \frac{1}{2} \right] \left[ F(X_k) - \frac{1}{2} \right].
$$

(A.18)

It is straightforward to show using uniform random variables arguments that the first summand in the sums (A.17) and (A.18) equals 1/12.

Next we shall show that, for all $k$

$$
\int \int [F_{1,k}(x,y) - F(x)F(y)] \, dF(x) \, dF(y) = \mathbb{E} \left[ F(X_1) - \frac{1}{2} \right] \left[ F(X_k) - \frac{1}{2} \right],
$$

which, by computing the products and writing expectations as integrals, may be rewritten as

$$
\int \int F_{1,k}(x,y) \, dF(x) \, dF(y) - \frac{1}{4} = \int \int F(x)F(y) \, dF_{1,k}(x,y) - \frac{1}{4}.
$$

Now, by setting $u = F(x)$, $v = F(y)$, the integrals on both sides of the equal sign can be rewritten, respectively, as

$$
\int_0^1 \int_0^1 C_{1,k}(u,v) \, du \, dv \quad \text{and} \quad \int_0^1 \int_0^1 uv \, dC_{1,k}(u,v),
$$

where $C_{1,k}(u,v) = F_{1,k}(F^{-1}(u), F^{-1}(v))$ is the copula function relative to the bivariate distribution $F_{1,k}$, with both marginals equal to $F$. The existence and
uniqueness of $C_{1,k}(u,v)$ are guaranteed by Sklar’s theorem and the continuity of $F$ (Nelsen, 1998, Theorem 2.3.3 and Corollary 2.3.7).

From Nelsen (1998, Theorem 5.1.6) we have

$$
\frac{1}{12} \rho_{1,k} = \int_0^1 \int_0^1 C_{1,k}(u,v) \, du \, dv - \frac{1}{4} = \int_0^1 \int_0^1 uv \, dC_{1,k}(u,v) - \frac{1}{4},
$$

where $\rho_{1,k}$ is the population version of Spearman’s rank correlation coefficient. Thus, the long run variance may be written also as

$$
\sigma^2 = \frac{1}{12} \left[ 1 + 2 \sum_{k=2}^{\infty} \rho_{1,k} \right].
$$

Asymptotic representation of $\eta_{R,T}^R$

Let’s get back to the decomposition in equation (A.4): $T^{-1/2}S_{T,t} = A_T + B_T + R_T$. Since

$$
\int F_t(x) \, dF(x) = E \left[ \frac{1}{T} \sum_{i=1}^{T} I_{\{X_i \leq x\}} \left| X_1, \ldots, X_T \right| \right] = 1 - \frac{1}{T} \sum_{i=1}^{T} F(X_i)
$$

and

$$
\int F_T(x) \, dF(x) = E \left[ \frac{1}{T} \sum_{i=1}^{T} I_{\{X_i \leq x\}} \left| X_1, \ldots, X_T \right| \right] = 1 - \frac{1}{T} \sum_{i=1}^{T} F(X_i),
$$

we have

$$
T^{-1/2}S_{T,t} = T^{-1/2} \left\{ \sum_{i=1}^{T} F(X_i) - \frac{T}{T} \sum_{i=1}^{T} F(X_i) \right\} + O_p(T^{-1/2}).
$$

Weak convergence of $\hat{\eta}_{R,T}^R$

In order to prove the second statement of the theorem, it suffices to show that $\hat{\sigma}_T^2 \overset{p}{\rightarrow} \sigma^2$.

We shall prove that as $T$ grows

$$
\hat{\sigma}_T^2 := \frac{1}{T} \sum_{s=1}^{T} \sum_{t=1}^{T} k \left( \frac{s-t}{\gamma T} \right) \left[ F_T(X_s) - \frac{T+1}{2T} \right] \left[ F_T(X_t) - \frac{T+1}{2T} \right] \tag{A.19}
$$

converges in probability to the same limit as

$$
\hat{\sigma}_T^2 := \frac{1}{T} \sum_{s=1}^{T} \sum_{t=1}^{T} k \left( \frac{s-t}{\gamma T} \right) \left[ F(X_s) - \frac{1}{2} \right] \left[ F(X_t) - \frac{1}{2} \right], \tag{A.20}
$$

that is, $|\hat{\sigma}_T^2 - \hat{\sigma}_T^2| \overset{p}{\rightarrow} 0$. Indeed, expression (A.20) denotes a kernel estimator for the long run variance of the sequence of strong mixing random variables.
\{ F(X_1), \ldots, F(X_T) \}, that in view of the results of de Jong and Davidson (2000) is consistent for \( \sigma^2 \).

Now, define \( F_T(r) = F_T(r) - F(r) \), which by weak convergence results for empirical processes, under the mixing condition in Assumption 1, is \( O_p(T^{-1/2}) \).

Of course, it makes no difference in the limit if in equation (A.19) we substitute \((n+1)/(2n)\) with \(1/2\). Using the identity \( F_T = F + F_T \), we can write

\[
\frac{1}{T} \sum_{s=1}^{T} \sum_{t=1}^{T} k \left( \frac{s-t}{\gamma_T} \right) [F_T(X_s) - 1/2] [F_T(X_t) - 1/2] =
\]

\[
\frac{1}{T} \sum_{s=1}^{T} \sum_{t=1}^{T} k \left( \frac{s-t}{\gamma_T} \right) [F(X_s) - 1/2] [F(X_t) - 1/2] +
\]

\[
+ \frac{1}{T} \sum_{s=1}^{T} \sum_{t=1}^{T} k \left( \frac{s-t}{\gamma_T} \right) F_T(X_s) [F(X_t) - 1/2] +
\]

\[
+ \frac{1}{T} \sum_{s=1}^{T} \sum_{t=1}^{T} k \left( \frac{s-t}{\gamma_T} \right) F_T(X_s) F_T(X_t). \]

So, for completing the proof it suffices to show that the limit in probability of the expressions in the last three lines above is zero. Let’s start from the last line:

\[
\left| \frac{1}{T} \sum_{s=1}^{T} \sum_{t=1}^{T} k \left( \frac{s-t}{\gamma_T} \right) F_T(X_s) F_T(X_t) \right| 
\]

\[
\leq \frac{1}{T} \left[ \sup_x |F_T(x)| \right]^2 \sum_{h=-T+1}^{T} \sum_{h=1}^{T-h} k(h/\gamma_T) = O_p(\gamma_T/T), \]

since \([\sup_x |F_T(x)|]^2 = O_p(T^{-1})\). The second and the third last lines are clearly of the same order, so let’s just consider one of them:

\[
\left| \frac{1}{T} \sum_{s=1}^{T} \sum_{t=1}^{T} k \left( \frac{s-t}{\gamma_T} \right) F_T(X_s) [F(X_t) - 1/2] \right| 
\]

\[
\leq \sup_x |F_T(x)| \sum_{h=-T+1}^{T+1} k(h/\gamma_T) \frac{1}{T} \sum_{t=1}^{T-h} |F(X_t) - 1/2| = O_p(\gamma_T/\sqrt{T}), \]

recalling that \(\sup_x |F_T(x)| = O_p(T^{-1/2})\). Since by assumption \(\gamma_T/\sqrt{T} \to 0\) as \(T\) diverges, this completes the proof.

□
Proof of Theorem 2. Let us consider the limit behaviour of ranks under the hypotheses of the theorem: take \( r \in [0, 1] \)

\[
T^{-1} R_{\lfloor Tr \rfloor, T} = T^{-1} \sum_{t=1}^{T} \mathbb{I}\{X_t \leq X_{\lfloor Tr \rfloor}\}
\]

\[
= T^{-1} \sum_{t=1}^{T} \mathbb{I}\{T^{-1/2}g(X_t) \leq T^{-1/2}g(X_{\lfloor Tr \rfloor})\}
\]

\[
\Rightarrow R(r) := \int_{0}^{1} \mathbb{I}\{W(s) \leq W(r)\} \, ds.
\]

The result of last line is a straightforward application of a weakened version of the continuous mapping theorem to functions with discontinuities of null measure with respect to the limit measure (Billingsley, 1968, Theorems 5.1 and 5.2). This result can be found also in Breitung and Gouriéroux (1997), who also derive some properties of the process \( R(r) \), that can be seen as the occupation time of the set \((\infty, W(r)]\) by the Brownian motion \( W(\cdot) \).

Repeated applications of the continuous mapping theorem yield the first result in the theorem. Indeed, recall that \( S_{T,t} := T^{-1} \sum_{i=1}^{t} [R_{T,i} - (T + 1)/2] \) and let \( R_0(r) := R(r) - 1/2 \), then

\[
\frac{S_{T,r}}{T} \Rightarrow \int_{0}^{r} R_0(s) \, ds
\]

\[
\frac{1}{T} \sum_{t=1}^{T} \left[ \frac{S_{T,t}}{T} \right]^2 \Rightarrow \int_{0}^{1} \left[ \int_{0}^{r} R_0(s) \, ds \right]^2 \, dr. \tag{A.21}
\]

The second thesis in the theorem follows straightforwardly from the fact that \( 0 < R_{T,t}/T \leq 1 \) and the cross-products in the HAC estimator (5) are all equal or smaller than 1.

\[\square\]

Proof of Theorem 3. Setting \( r = t/T \), we obtain (using, for instance, Phillips and Perron, 1988, Sec. 2)

\[
\frac{1}{\sqrt{T}\sigma_x} S_{T,t}^K = \frac{1}{\sqrt{T}\sigma_x} \left\{ \sum_{i=1}^{t} X_i - \frac{t}{T} \sum_{i=1}^{T} X_i \right\} +
\]

\[
\frac{\sigma_z}{T^{3/2}\sigma_x} \left\{ \sum_{i=1}^{t} \sum_{s=1}^{i} Z_i - \frac{t}{T} \sum_{i=1}^{T} \sum_{s=1}^{i} Z_i \right\}
\]

\[
\Rightarrow V(r) + \frac{\sigma_z}{\sigma_x} K(r)
\]

with \( V(r) := W_x(r) - rW_x(1) \), \( K(r) := \int_{0}^{r} W_z(u) \, du - r \int_{0}^{1} W_z(u) \, du \), and with \( W_x \) and \( W_z \) independent standard Wiener processes.
As for $S_{R,T}^R$, first notice that

$$F \left( X_i + \frac{\sigma_x}{T} \sum_{s=1}^{i} Z_s \right) = F(X_i) + f(X_i) \frac{\sigma_x}{T} \sum_{s=1}^{i} Z_s + o \left( \frac{\sigma_x}{T} \sum_{s=1}^{i} Z_s \right)$$

$$= F(X_i) + f(X_i) \frac{\sigma_x}{T} \sum_{s=1}^{i} Z_s + o_p \left( T^{-1/2} \right).$$

Then, using the asymptotic representation of Theorem 1

$$\frac{1}{\sqrt{T} \sigma} S_{R,T} = \frac{1}{\sqrt{T} \sigma} \left\{ \sum_{i=1}^{T} F(Y_i) - \frac{t}{T} \sum_{i=1}^{T} F(Y_i) \right\} + o_p(T^{-1/2})$$

$$= \frac{1}{\sqrt{T} \sigma} \left\{ \sum_{i=1}^{T} F(X_i) - \frac{t}{T} \sum_{i=1}^{T} F(X_i) \right\} +$$

$$+ \frac{\sigma_x}{T^{3/2} \sigma} \left\{ \sum_{i=1}^{T} f(X_i) \sum_{s=1}^{i} Z_s - \frac{t}{T} \sum_{i=1}^{T} f(X_i) \sum_{s=1}^{i} Z_s \right\} + o_p(T^{-1/2})$$

$$\Rightarrow V(r) + f_2(0) \frac{\sigma_x}{\sigma} K(r),$$

where, again, the processes $V$ and $K$ are independent and $f_2(0) = \mathbb{E} f(X_t) = \int f(x)^2 \, dx$.

For $c > 0$ the process $V(r) + cK(r)$ is Gaussian and has variance that grows with $c^2$ for all $r \in (0,1)$. Then the percentiles of $H(c) = \int [V(r) + cK(r)]^2 \, dr$ are strictly increasing functions of $|c|$ as well. This implies that the asymptotic power function of a test statistic that under the alternative has distribution $H(b)$ dominates another test statistic with distribution $H(a)$ with $a < b$ under the alternative.

Now, if we name $T_R$ the sample size for the RKPSS statistic and $T_K$ the sample size for the KPSS statistic, then the two tests achieve the same asymptotic power for $T_R = T_K f_2(0) \sigma_x/\sigma$, so the asymptotic relative efficiency of the RKPSS test with respect to the KPSS, in Pitman’s sense, is given by

$$e_{R,K} = f_2(0) \frac{\sigma_x}{\sigma}. $$

**Proof of Theorem 4.** First of all note that

$$\hat{\beta}_T = \text{median} \left\{ \frac{Y_j - Y_i}{j - i} : 1 \leq i < j \leq T \right\}$$

$$= \text{median} \left\{ \frac{\alpha + \beta j + X_j - \alpha - \beta i - X_i}{j - i} : 1 \leq i < j \leq T \right\}$$

$$= \text{median} \left\{ \beta + \frac{X_j - X_i}{j - i} : 1 \leq i < j \leq T \right\}$$

$$= \beta + \text{median} \left\{ \frac{X_j - X_i}{j - i} : 1 \leq i < j \leq T \right\}$$
so there is no loss of generality in assuming $\beta = 0$. The estimator $\hat{\beta}_T$ is a zero of the function

$$M_T(b) := \sum_{i<j} \text{sign} \left( \frac{X_j - X_i}{j - i} - b \right)$$

which is a non-increasing step function with range $[-T(T - 1)/2, T(T - 1)/2]$ and increments equal to $-2$ at all discontinuity points.

**Consistency**

In order to prove consistency, take $\varepsilon > 0$ and consider

$$\Pr\{\hat{\beta} > \varepsilon\} \leq \Pr\{M_T(\varepsilon) \geq 0\}$$

$$\Pr\{\hat{\beta} < -\varepsilon\} \leq \Pr\{M_T(-\varepsilon) \leq 0\}.$$

Notice that the stationarity of $X_t$ implies the symmetry of the distribution of $X_j - X_i$ around zero, and, thus, of the distribution of $M_T$. As a consequence,

$$\Pr\{M_T(\varepsilon) \geq 0\} = \Pr\{M_T(-\varepsilon) \leq 0\}.$$

Define

$$N_T(b) := \sum_{i<j} \text{sign} (X_j - X_i - b)$$

and notice that for $\varepsilon > 0$, $N_T(\varepsilon) > M_T(\varepsilon)$ almost surely.

Since $(T^{-1/2}N_T(\varepsilon))$ is a $U$-statistic with kernel $g_\varepsilon(x, y) = \text{sign}(y - x - \varepsilon)$, if $X_t$ is stationary and ergodic, by the (strong) law of large number for $U$-statistics (Aaronson et al., 1996) we have

$$\left(\frac{T}{2}\right)^{-1} N_T \xrightarrow{a.s.} \theta_\varepsilon := \int_{\mathbb{R} \times \mathbb{R}} g_\varepsilon(x, y) \, dF(x) \, dF(y) = 1 - 2F_2(\varepsilon),$$

with $F_2(\varepsilon) = \int_{\mathbb{R}} F(y + z) \, dF(y)$, symmetric about zero. It follows: $\varepsilon > 0 \iff \theta_\varepsilon < 0$.

Now,

$$\Pr\{M_T(\varepsilon) \geq 0\} = \Pr \left\{ \left(\frac{T}{2}\right)^{-1} M_T(\varepsilon) \geq 0 \right\}$$

$$\leq \Pr \left\{ \left(\frac{T}{2}\right)^{-1} N_T(\varepsilon) \geq 0 \right\}$$

$$= \Pr \left\{ \left(\frac{T}{2}\right)^{-1} N_T(\varepsilon) - \theta_\varepsilon \geq -\theta_\varepsilon \right\} \rightarrow 0 \quad \text{for} \ T \rightarrow \infty.$$
Asymptotic normality
The asymptotic distribution of $\hat{\beta}_T$ can be obtained using the following asymptotic representation (proved below)

$$\left(\frac{T}{2}\right)^{-1}M_T(b_T) = \theta_T(b_T) + \frac{8}{T^2} \sum_{t=1}^{T} (F(X_t) - 1/2)q_t + o_p(T^{-1/2}), \quad (A.22)$$

where $b_T$ is a sequence such that $b_T \cdot T \to 0$ as $T$ diverges, $q_t := t - (T + 1)/2$ and

$$\theta_T(b_T) := \left(\frac{T}{2}\right)^{-1}\mathbb{E} M_T(b_T) = \left(\frac{T}{2}\right)^{-1} \sum_{1 \leq i < j \leq T} \left[1 - 2F_2((j - i)b_T)\right]. \quad (A.23)$$

Set $Q_T^2 := \sum_{t=1}^{T} q_t^2 = T(T^2 - 1)/12$, $Q_T$ equal to its positive square root and notice that for all $b \in \mathbb{R}$

$$\theta_T(b/Q_T) = \left(\frac{T}{2}\right)^{-1} \sum_{1 \leq i < j \leq T} \left(-2bf_2(0)\frac{j - i}{Q_T} + o(T^{-1/2})\right)$$

$$= -4b\left(\frac{T}{2}\right)^{-1} f_2(0)/Q_T + o(T^{-1/2})$$

$$= -4b\left(\frac{T}{2}\right)^{-1} f_2(0)/\sqrt{3T} + o(T^{-1/2})$$

where $f_2(0) := \int f^2(x) \, dx$ is the derivative of $F_2$ in zero and

$$\sum_{1 \leq j < i \leq T} (j - i) = 2Q_T^2.$$

Now, suppose for the moment that $T(T - 1)/2$ is odd, so that $\hat{\beta}_T$ is uniquely identified, then the following identities among sets hold for all $b \in \mathbb{R}$

$$\{Q_T \hat{\beta}_T > b\} = \left\{\left(\frac{T}{2}\right)^{-1}M_T(b/Q_T) > 0\right\}$$

$$= \left\{\frac{8}{T^{3/2}} \sum_{t=1}^{T} [F(X_t) - 1/2]q_t + o_p(1) > -T^{1/2}\theta_T(b/Q_T)\right\}$$

$$= \left\{\frac{8}{T^{3/2}} \sum_{t=1}^{T} [F(X_t) - 1/2]q_t + o_p(1) > \frac{4bf_2(0)}{\sqrt{3}}\right\}$$

$$= \left\{\frac{\sqrt{12}}{T^{3/2}f_2(0)} \sum_{t=1}^{T} [F(X_t) - 1/2]q_t + o_p(1) > b\right\},$$

and we can conclude that

$$Q_T \hat{\beta}_T = \frac{\sqrt{12}}{T^{3/2}f_2(0)} \sum_{t=1}^{T} [F(X_t) - 1/2]q_t + o_p(1). \quad (A.24)$$
If on the contrary $T(T - 1)/2$ is even, $\tilde{\beta}_T$ could be arbitrarily chosen in the open interval $(\beta_T^-, \beta_T^+)$, with

$$\beta_T^- := \sup \{ b : M_T(b) > 0 \}, \quad \text{and} \quad \beta_T^+ := \inf \{ b : M_T(b) < 0 \};$$
even though generally $\tilde{\beta}_T := (\beta_T^- + \beta_T^+)/2$. Now,

$$\{ Q_T \beta_T^- > b \} = \left\{ \left( \frac{T}{2} \right)^{-1} M(b/Q_T) > 0 \right\}$$

$$= \left\{ \frac{\sqrt{T^2}}{T^{3/2} f_2(0)} \sum_{i=1}^{T} \left[ F(X_i) - 1/2 |q_i + o_p(1)| > b \right] \right\},$$

and

$$\{ Q_T \beta_T^+ > b \} = \left\{ \left( \frac{T}{2} \right)^{-1} M(b/Q_T) \geq 0 \right\}$$

$$= \left\{ \frac{\sqrt{T^2}}{T^{3/2} f_2(0)} \sum_{i=1}^{T} \left[ F(X_i) - 1/2 |q_i + o_p(1)| \geq b \right] \right\},$$

but since $T^{-3/2} \sum_{i=1}^{T} [F(X_i) - 1/2] q_i$ is an absolutely continuous random variable for all $T \in \{1, 2, \ldots, \infty\}$, then $\beta_T^+ - \beta_T^-$ is $o_p(1)$ and (A.24) holds.

Finally, relaxing the harmless hypothesis $\beta = 0$ and exploiting the asymptotic representation (A.24), we obtain (cf. Phillips and Perron, 1988, Section 2)

$$Q_T (\tilde{\beta}_T - \beta) \Rightarrow \frac{\sqrt{T^2} \sigma}{f_2(0)} \left[ \frac{1}{2} W(1) - \int_0^1 W(u) \, du \right] \sim N \left( 0, \frac{\sigma^2}{f_2(0)} \right),$$

(A.25)

where $\sigma^2$ has been defined in equations (7) and (9).

Derivation of the asymptotic representation (A.22)

Following standard $U$-statistics theory, let

$$h_{Ti}(x; b) := \frac{1}{T-1} \sum_{j \neq i} \mathbb{E} \text{sign}(X_j - x - b(j - i))$$

$$= \frac{1}{T-1} \sum_{j \neq i} \left[ 1 - 2F(x + b(j - i)) \right] (\mathbb{I}_{(j>i)} - \mathbb{I}_{(j<i)}),$$

and

$$H_T(b) := \frac{1}{T} \sum_{i=1}^{T} h_{Ti}(X_i; b)$$

$$= \frac{1}{T(T-1)} \sum_{i=1}^{T} \sum_{j \neq i} \left[ 1 - 2F(X_i + b(j - i)) \right] (\mathbb{I}_{(j>i)} - \mathbb{I}_{(j<i)}).$$

\text{10}The $j \neq i$ under the sum sign is to be read as $j \in \{1, \ldots, i-1, i+1, \ldots, T\}$. 

36
By the Hoeffding decomposition of \( U \)-statistics (with bounded kernel) we have
\[
\left( \frac{T}{2} \right)^{-1} M_T(b) = \theta_T(b) + 2(H_T(b) - \theta_T(b)) + O_p(T^{-1}).
\]

It remains to show that, for any sequence \( b_T \) such that \( b_T T \to 0 \),
\[
H_T(b_T) - \theta_T(b_T) - H_T(0) = o_p(T^{-1/2}). \tag{A.26}
\]

In fact, after straightforward manipulations of the expression defining \( H_T(0) \), we obtain
\[
H_T(0) = 4 T \left( \frac{T}{2} - 1 \right) \sum_{i=1}^T \left[ 1 - 2 F_2(b_T(j - i)) \right] \left( I_{(j_i)} - I_{(j < i)} \right),
\]
so we can rewrite the lhs of (A.26) as
\[
S_T(b_T(j - i)) := \frac{1}{T(T - 1)} \sum_{i=1}^T \sum_{j \neq i} \left[ 1 - 2 F(X_i + b_T(j - i)) \right] + \left[ 1 - 2 F(X_i) \right] \left( I_{(j > i)} - I_{(j < i)} \right)
\]
\[
= \frac{2}{T(T - 1)} \sum_{i=1}^T \sum_{j \neq i} \left( F(X_i) - F(X_i + b_T(j - i)) + F_2(0) - F_2(b_T(j - i)) \right) \left( I_{(j > i)} - I_{(j < i)} \right). \tag{A.27}
\]

Let
\[
Z_i(b_T(j - i)) := F(X_i) - F(X_i + b_T(j - i)) - F_2(0) + F_2(b_T(j - i))
\]
and notice that \( \mathbb{E} Z_i(x) = 0 \) for any real \( x \), as \( \mathbb{E} F(X_i) = F_2(0) \) and \( \mathbb{E} F(X_i + x) = F_2(x) \). It follows that \( \mathbb{E} S_T(x) = 0 \). Now, it remains to show that
\[ T \text{Var} \left[ S_T(b_T(j - i)) \right] \to 0, \]
but since \( F \) is continuous and nondecreasing \( \text{Var} \left[ S_T(b_T(j - i)) \right] \leq \text{Var} \left[ S_T(b_T) \right] \), where \( S_T(b_T) \) can be rewritten as
\[
S_T(b_T) = \frac{2}{T(T - 1)} \sum_{i=1}^{T-1} Z_i(b_T)(T - i).
\]
It follows that
\[
\text{Var } [S_T(b_tT)] = \frac{4}{T^2(T-1)^2} \sum_{i=1}^{T-1} \sum_{j=1}^{T-1} Z_i(b_T T) Z_j(b_T T) (T-i)(T-j)
\]
\[
= \frac{2(2T-1)}{T(T-1)} \mathbb{E} Z_1(b_T T)^2 \left\{ 1 + \sum_{i=1}^{T-1} \sum_{j \neq 1} \rho_{i,j} w_i w_j \right\},
\]

with
\[
w_i := (T-i)^2 \mathbb{E} \left[ F(X_1) - F(X_1 + b_T T) \right] \quad \text{and} \quad \rho_{i,j} := \text{Corr } [Z_i(b_T T), Z_j(b_T T)].
\]

The strong-mixing condition assures that the sum of correlations in curly brackets converges as \( T \to \infty \), thus, the order of \( T \text{Var } [S_T(b_T T)] \) is the same as that of
\[
\mathbb{E} Z_1(b_T T)^2 \leq \mathbb{E} [F(X_1) - F(X_1 + b_T T)]^2 \leq \mathbb{E} |F(X_1) - F(X_1 + b_T T)|
\]
where the last inequality follows from the fact that the random variable \( F(X_1) - F(X_1 + b_T T) \) takes value in the unit interval. Furthermore
\[
\mathbb{E} |F(X_1) - F(X_1 + b_T T)| = \begin{cases} F_2(0) - F_2(b_T T), & \text{for } b_T \leq 0, \\ F_2(b_T T) - F_2(0), & \text{for } b_T > 0; \end{cases}
\]
but since \( F_2 \) is symmetric around zero
\[
\mathbb{E} |F(X_1) - F(X_1 + b_T T)| = F_2(|b_T| T) - F_2(0) = f_2(0)|b_T| T + O(b_T^2 T^2).
\]
So, when \( f_2(0) < \infty \), \( T \text{Var } [S_T(b_T T)] = O(b_T T) \), and this proves equation (A.26).

**Proof of Theorem 5.** First of all, notice that
\[
Y_t = \alpha + \beta t + X_t = \left( \alpha + \beta \frac{T+1}{2} \right) + \beta \left( t - \frac{T+1}{2} \right) + X_t
\]
\[
= \alpha_T + \beta \left( t - \frac{T+1}{2} \right) + X_t,
\]
but as ranks are translation-invariant \( \alpha_T := \alpha + \beta (T+1)/2 \) can be set to zero without loss of generality. Let
\[
q_t := \left( t - \frac{T+1}{2} \right)
\]
and \( Q_T^2 := \sum_{t=1}^{T} q_t^2 = T(T^2-1)/12 \). Moreover, let \( \lambda := Q_T (b - \beta) \) with \( b \in \mathbb{R} \), then
\[
Y_t(\lambda) := Y_t - bq_t = Y_t - \left( \beta + \frac{\lambda}{Q_T} \right) q_t = X_t - \frac{\lambda}{Q_T} q_t,
\]

38
and suppose that $|\lambda| \leq C < \infty$ holds. Let

$$R_{T,t}(\lambda) := \sum_{s=1}^{T} I\{Y_s(\lambda) \leq Y_t(\lambda)\}$$

$$S_{T,t}(\lambda) := \frac{1}{T} \sum_{s=1}^{t} \left[ R_{T,s}(\lambda) - \frac{T+1}{2} \right],$$

and, for $x \in \mathbb{R}$,

$$F_{T}(x; \lambda) := \frac{1}{T} \sum_{t=1}^{T} I\{Y_t(\lambda) \leq x\}.$$ 

Then $T^{-1} R_{T,t}(\lambda) = F_{T}(Y_{t}(\lambda); \lambda)$,

$$\frac{1}{T} S_{T,t}(\lambda) = \frac{1}{T} \sum_{s=1}^{t} \left[ F_{T}(Y_{s}(\lambda); \lambda) - \frac{T+1}{2T} \right]$$

and

$$\frac{1}{T} [S_{T,t}(\lambda) - S_{T,t}(0)] = \frac{1}{T} \sum_{s=1}^{t} \left[ F_{T}(Y_{s}(\lambda); \lambda) - F_{T}(Y_{s}(0); 0) \right].$$

Let

$$\bar{F}_{T}(x; \lambda) := \mathbb{E} F_{T}(x; \lambda) = \frac{1}{T} \sum_{t=1}^{T} \Pr \{ Y_t(\lambda) \leq x \}$$

$$= \frac{1}{T} \sum_{t=1}^{T} \Pr \left\{ X_t \leq x + \frac{\lambda}{Q_T} q_t \right\}$$

$$= \frac{1}{T} \sum_{t=1}^{T} F \left( x + \frac{\lambda}{Q_T} q_t \right)$$

$$= \frac{1}{T} \sum_{t=1}^{T} \left[ F(x) + \frac{\lambda f(x)}{Q_T} q_t + \frac{\lambda^2 f'(x)}{2 Q_T^2} q_t^2 + o(T^{-1}) \right]$$

$$= F(x) + \frac{\lambda^2}{2T} f'(x) + o(T^{-1})$$

$$= F(x) + o(T^{-1}).$$

Consider the norm defined by

$$m_T := \sup_{|\lambda| \leq C} \sup_x T^{1/2} \left| F_{T}(x; \lambda) - F_{T}(x; 0) - \bar{F}_{T}(x; \lambda) + \bar{F}_{T}(x; 0) \right|$$

$$= \sup_{|\lambda| \leq C} T^{1/2} \left| F_{T}(x; \lambda) - F_{T}(x; 0) \right| + O(T^{-1/2}).$$
By the Bahadur representation of sample quantiles, as extended to linear models (Jurečková and Sen, 1996, Ch.4-6) and strong mixing processes (Yoshihara, 1995), we obtain that $m_T = o(1)$ almost surely as $T \to \infty$. Therefore,

$$
\frac{1}{T} [S_{T,t}(\lambda) - S_{T,t}(0)] = 
$$

$$
\frac{1}{T} \sum_{s=1}^t \left\{ [F_T(Y_s(\lambda); \lambda) - F_T(Y_s(\lambda); 0)] + [F_T(Y_s(\lambda); 0) - F_T(Y_s(0); 0)] \right\} = 
$$

$$
\frac{1}{T} \sum_{s=1}^t \left\{ F_T(Y_s(\lambda); \lambda) - F_T(Y_s(0); 0) \right\} + o(T^{-1/2}) \quad \text{a.s.}
$$

Finally, by the same Bahadur representation,

$$
\sup_{\lambda} \sup_{y \in \mathbb{R}} |F_T(Y_s(\lambda); 0) - F(Y_s(\lambda)) - F_T(Y_s(0); 0) + F(Y_s(0))| = o(T^{-1/2}) \quad \text{a.s.}
$$

and

$$
F_T(Y_s(\lambda); 0) - F_T(Y_s(0); 0) = F(Y_s(\lambda)) - F(Y_s(0)) + o(T^{-1/2}) \quad \text{a.s.}
$$

$$
= -f(X_s) \lambda q_s/Q_T + o(T^{-1/2}) \quad \text{a.s., (A.28)}
$$

as $Y_s(0) = X_s$.

As a result,

$$
\frac{1}{T} [S_{T,t}(\lambda) - S_{T,t}(0)] = -\frac{\lambda}{T} \sum_{s=1}^t f(X_s)q_s/Q_T + o(T^{-1/2}) \quad \text{a.s.}
$$

Now the CLT and the functional CLT apply to $f(X_s)q_s/Q_T$. Let

$$
U_{T,t} = \frac{1}{TQ_T} \sum_{s=1}^t f(X_s)q_s.
$$

Then, letting $f_2(0) := \mathbb{E} f(X_s) = \int f^2(x) \, dx$, we have

$$
\mathbb{E} U_{T,t} = \frac{f_2(0)}{TQ_T} \sum_{s=1}^t q_s = -f_2(0) \frac{t(T-t)}{2TQ_T} = -f_2(0) \frac{t(T-t)}{2T^2} \sqrt{\frac{12}{T}} + o(T^{-1/2}).
$$

Therefore, as $T \to \infty$

$$
T^{-1/2} [S_{T,t}(\lambda) - S_{T,t}(0)] - \lambda \sqrt{12} f_2(0) \frac{t(T-t)}{2T^2} = o(1) \quad \text{a.s.}
$$
and
\[ \frac{1}{T} S_{T,t}^2(\lambda) = \left[ \frac{1}{\sqrt{T}} S_{T,t}(0) + \lambda \gamma \frac{t(T-t)}{2T^2} \right]^2 + o(1) \text{ a.s.} \]

with \( \gamma := \sqrt{12f_2(0)} \).

Now, for \( r := t/T \) and \( t = 0, 1, 2, \ldots, T \), as already seen in the proof of Theorem 1
\[ \frac{1}{\sigma \sqrt{T}} S_{T,t}(0) \Rightarrow W(r) - r W(1), \quad (A.29) \]
where \( W \) is a standard Brownian motion on \([0, 1]\) and \( \sigma^2 \) as in equations (7) and (9).

Exploiting the asymptotic representation of the Theil-Sen estimator (A.22) we obtain (using Sec.2 of Phillips and Perron, 1988, for instance)
\[ Z_T := Q_T(\hat{\beta}_T - \beta) = \frac{\sqrt{T\sigma}}{f_2(0)} \sum_{t=1}^{T} \left[ F(X_t) - 1/2 \right] q_t + o_p(1) \]
\[ \Rightarrow W(r) - r W(1) + 6 r(1-r) \left[ \frac{1}{2} W(1) - \int_0^1 W(u) \, du \right] =: V_2(r), \]
where, because of the representation (10), the Wiener process \( W \) is the same of equation (A.29). Therefore,
\[ \frac{1}{\sigma \sqrt{T}} S_{T,t}(Z_T) = \frac{1}{\sigma \sqrt{T}} S_{T,t}(0) + Z_T \gamma \frac{t(T-t)}{2T^2} + o(1) \text{ a.s.} \]
\[ \Rightarrow W(r) - r W(1) + 6 r(1-r) \left[ \frac{1}{2} W(1) - \int_0^1 W(u) \, du \right] =: V_2(r), \]
which is the, so called, second-level Brownian bridge (cf. Kwiatkowski et al., 1992, eq.16). A simple application of the continuous mapping theorem proves the first thesis of the theorem.

For proving the second statement of the theorem, we need to assess the consistency of the kernel estimator of \( \sigma^2 \). Using equation (A.28) and the consistency of the Theil-Sen estimator, we have the following representation of the residuals’ ranks
\[ \frac{R_{T,t}}{T} = F_{T}(Y_t(Z_T)) = F_{T}(Y_t(0)) + O_p(T^{-1/2}) = F(X_t) + O_p(T^{-1/2}). \]
It follows that, under Assumption 2, we can apply de Jong and Davidson (2000) and obtain
\[ \hat{\sigma}_T^2 = \frac{1}{T} \sum_{s=1}^{T} \sum_{t=1}^{T} k \left( \frac{s-t}{\gamma T} \right) [F(X_s) - 1/2][F(X_t) - 1/2] + O_p(\gamma T/T) \quad \sigma^2. \]

\[ \square \]
References


