Exponentials of unit root processes

Robert de Jong*
Ohio State University

Preliminary version

October 22, 2009

Abstract

This paper shows that the sum of the exponential of a unit root process converges in distribution, after rescaling by the exponential of the maximum value of a Beveridge-Nelson approximation to the unit root process. The increments of the unit root process are assumed to be linear processes, and the limit distribution that is derived depends on the structure of the linear process and its innovation distribution. This result is a new piece of econometric methodology that is applied towards deriving the limit distribution of the Dickey-Fuller test under the assumption that the data are the exponential of a unit root process, and towards the KPSS test. The limit distribution of the KPSS test applied to data that are the exponential of a unit root process turns out to be pivotal and converges at the same rate as when the data are stationary. The analytical results on the limit behavior of the KPSS test seem to pose an empirical puzzle, since the KPSS test rejects in many instances for the level series when one might expect a unit root in the logarithm of the series.

1 Introduction

This paper shows the convergence in distribution of statistics of the form

$$R_n = \exp(-m_n) \sum_{t=1}^{n} \exp(y_t) \tag{1}$$

*Department of Economics, Ohio State University, 444 Arps Hall, Columbus, OH 43210, email de-jong.8@osu.edu. I thank Magda Peligrad and Jean Bertoin for their input and discussion.

where y_t, $t = 1, \ldots, n$ is a unit root process and m_n is the maximum of a Beveridge-Nelson approximation to y_t. It works out the consequences of this new limit theory into two ways. First, the limit behavior of the Dickey-Fuller test will be derived under the assumption that the data-generating process is the exponential of a unit root process. Second, the limit distribution of the KPSS statistic when applied to the exponential of a unit root process is calculated. This limit is pivotal and allows testing the null of a unit root in logarithms against the alternative of a unit root in levels. However, these are only two possible applications of the main result of this paper, and the methodology of this paper can be potentially be applied to a range of different scenarios.

The R_n statistic was analyzed for the case of i.i.d. Δy_t by Davies and Krämer (2003), who showed the property $\sup_{n \geq 1} ER_n < \infty$ under regularity conditions. Earlier, Park and Phillips (1999) established that under conditions similar to those made in this paper,

$$n^{-1/2} \sum_{t=1}^{n} \exp(y_t - m_n) \xrightarrow{p} 0; \tag{2}$$

see also Davies and Krämer (2003, p.867). Below, I will use a Laplace transform argument and show that $\lim_{n \to \infty} E \exp(-rR_n)$ exists for all $r \geq 0$; and from that result, it is then easy to prove the convergence in distribution of R_n. The argument in the proof of this paper can be characterized as looking both backward and forward in time from the point at which the maximum of the Beveridge-Nelson approximation to the unit root process is attained and using arguments from the literature on "random walks conditioned to be positive" to deal with the differences between the maximum of the series and the value of the series at each particular point in time.

In the econometrics literature, papers such as Granger and Hallmann (1991), Ermini and Hendry (2008), and Corradi (1995) attempt to define the I(1) property (loosely defined here as the property that a series displays some type of fading memory property after differencing) in such a way that under some transformations the property is preserved. A related literature seeks to find unit root tests whose null distribution is robust to monotonic transformations; see Granger and Hallmann (1991), Burridge and Guerre (1996), Breitung and Gourieroux (1997) for work along these lines. Box-Cox approaches can be found in Franses and Koop (1998) and Kobayashi and McAleer (1999).

These papers however avoid dealing with nonlinear transformations of unit root processes directly and lack the analytical conciseness of the work of Borodin and Ibragimov (1995) and Park and Phillips (1999), which seeks to characterize the limit behavior (after rescaling) of sums of the form

$$\sum_{t=1}^{n} f(y_t). \tag{3}$$

For functions $f(\cdot)$ that are "asymptotically homogeneous" the limit behavior of the rescaled statistic can be derived because it is asymptotically equivalent to a sum of a function of $n^{-1/2}y_t$, and at that point an appeal to the functional central limit theorem can be used to derive the limit distribution. Trivial examples of such functions are the identity and the square. See Park and Phillips (1999) and Borodin and Ibragimov (1995). For integrable functions $f(\cdot)$, the convergence in distribution of $n^{-1/2}\sum_{t=1}^{n} f(y_t)$ has been derived in Borodin and Ibragimov (1995) and Park and Phillips (1999). Borodin and Ibragimov (1995) also derive a central limit theorem type result for periodic functions $f(\cdot)$. These results form the analytical basis for a follow-up literature that uses these results as an analytical basis. Two notable examples of this are Park and Phillips (2000) and Park and Phillips (2001). This paper adds the exponential function to the set of functions for which the behavior of statistics such as the one of Equation(3) is well understood.

The argument of this paper extends to statistics of the form

$$\sum_{t=1}^{n} f(y_t - m_n) \tag{4}$$

for functions $f(\cdot)$ satisfying $\sup_{x\leq 0}(|x|+1)^{2+\eta}|f(x)| < \infty$ for some $\eta > 0$. However, while the properties of the exponential can serve to weigh a statistic that depends on y_t only, this does not happen in general, and therefore this extension is not pursued here.

Section 2 of this paper states the main theorem. The case where Δy_t is a linear process and the Dickey-Fuller test when applied to the exponential of a unit root process are analyzed in Section 3. An aspect of the proof of the main theorem can be exploited to derive the limit behavior of the KPSS statistic when applied to the exponential of a unit root process; this is pursued in Section 4. Section 5 contains a small simulation study. The paper concludes with two appendices, the first one of which contains the proofs of the theorems of the paper, and the second one of which contains the tables related to Section 5.

2 Main result

Below, y_t, $t = 1, ..., n$ is assumed to satisfy $y_t = y_0 + \gamma\sum_{j=1}^{t}\varepsilon_j + c_t$, where ε_t, $t = ..., -1, 0, 1, ...$ is a mean zero i.i.d. sequence and $\gamma \neq 0$. The approximation $\gamma\sum_{j=1}^{t}\varepsilon_j$ to $y_t - y_0$ is denoted as x_t for $t = 1, ..., n$, $x_0 = 0$, and $m_n = y_0 + \max_{1\leq t\leq n} x_t$. In addition, it is assumed that $c_t = \sum_{j=-\infty}^{\infty}\alpha_j\varepsilon_{t+j}$ and $\sum_{j=-\infty}^{\infty}|\alpha_j| < \infty$. Note that because of this setup, y_0 is instantly removed from the statistic, and for all proofs we can act as though $y_0 = 0$.

Throughout this paper the following assumption will be needed:

Assumption 1. *$y_t = y_0 + \gamma \sum_{j=1}^{t} \varepsilon_j + c_t$, where $\gamma \neq 0$, $c_t = \sum_{j=-\infty}^{\infty} \alpha_j \varepsilon_{t+j}$ and ε_t, $t = ..., -1, 0, 1, ...$, is i.i.d., $E\varepsilon_0 = 0$, $0 < E|\varepsilon_0|^p < \infty$ for some $p > 3$, $\sum_{j=-\infty}^{\infty} |\alpha_j| < \infty$, and the distribution of ε_t is continuous.*

The main analytical result of this paper is the following:

Theorem 1. *Assume that Assumption 1 holds. Then*

$$R_n \xrightarrow{d} \exp(\tilde{c}_0) + \sum_{t=1}^{\infty} \exp(\tilde{c}_t) \exp(\tilde{x}_t) + \sum_{t=1}^{\infty} \exp(\tilde{c}_{-t}) \exp(-\tilde{x}_t^*),$$

where $\tilde{x}_t$ and $\tilde{x}_t^*$, $t = 0, 1, ...$ are independent Markov chains, $\tilde{x}_0 = \tilde{x}_0^* = 0$, and $\tilde{c}_t = \gamma^{-1} \sum_{j=-\infty}^{\infty} \alpha_j (\tilde{x}_{t+j} - \tilde{x}_{t+j-1}) I(t+j \geq 1) + \gamma^{-1} \sum_{j=-\infty}^{\infty} \alpha_j (\tilde{x}_{t+j+1}^* - \tilde{x}_{t+j}^*) I(t+j \leq 0)$.

While the above result shows convergence in distribution to a non-defective limit, the limit distribution in general depends on the distribution of ε_t and on the α_j sequence. Therefore, the limit is not distribution-free; if the above limit distribution result were to be used for testing purposes, the limit distribution will need to be obtained through some resampling method.

It is interesting to note that the above result and its proof can be interpreted as saying that the limit distribution of R_n is determined by a finite but large number of values of x_t for which t is "close" to the index value of the maximum. To be precise, let τ_n denote the value indexing the maximum of x_t, viz., $x_{\tau_n} = \max_{1 \leq t \leq n} x_t$. Note that since the assumption that ε_t has a continuous distribution is made, τ_n is unique. It now follows from the proof that for all $M > 0$

$$\sum_{t=1}^{n} \exp(y_t - m_n) I(|t - \tau_n| \leq M) \xrightarrow{d} \exp(\tilde{c}_0) + \sum_{t=1}^{M} \exp(\tilde{c}_t) \exp(\tilde{x}_t) + \sum_{t=1}^{M} \exp(\tilde{c}_{-t}) \exp(-\tilde{x}_t^*) \quad (5)$$

while a similar result holds for the case where $|\tau_n - t| > M$. This result will be exploited in Section 4.

The strategy of the proof is to show the convergence of the Laplace transform $E\exp(-rR_n)$ of R_n. This is then followed by an argument that R_n is $O_p(1)$, which suffices for the proof that R_n converges in distribution. The proof starts with observing that for any $L \geq 1$, for an independent copy x_t^* of x_t, this Laplace transform equals an $o(1)$ term plus

$$E \sum_{k=L+1}^{n-L} I(\min_{1 \leq t \leq k-1} x_t^* > 0, \max_{1 \leq t \leq n-k} x_t \leq 0) \exp(-r \sum_{t=0}^{k-1} \exp(-x_t^* + c_{-t}) - r \sum_{t=1}^{n-k} \exp(x_t + c_t)). \quad (6)$$

By the definition of probability conditioned on an event, the above equation can then be rewritten such as to involve a "random walk conditioned to be positive" and a "random walk conditioned to be negative". The literature on random walks conditioned to be positive - i.e. results by Bertoin and Doney (1994) and by Ritter (1981) - can then be invoked at various points in the proof. As explained in Biggins (2003), probabilities involving random walks conditioned to be positive can be related, via an h-transform, to a Markov chain which is "random walk conditioned to stay positive", and random walks conditioned to be positive can be viewed as a discrete version of a Bessel-3 process. The Markov chains $\tilde{x}_t$ and $\tilde{x}_t^*$ can be characterized as obtained through such an h-transform; see Bertoin and Doney (1994) for details.

An earlier version of this paper considered the case of i.i.d. Δy_t, i.e. the case $c_t = 0$. For this case, since the dependence of c_t on both x_t and x_t^* has now disappeared, the expression of the above equation can be simplified substantially by using the independence between $\tilde{x}_t$ and $\tilde{x}_t^*$.

3 Linear processes and the Dickey-Fuller test

If we assume that $\Delta y_t = \sum_{j=0}^{\infty} \beta_j \varepsilon_{t-j}$ for $t = 1, ..., n$, then it follows from the Beveridge-Nelson decomposition (see Phillips and Solo (1992, p. 976)) that

$$y_t = y_0 + \gamma \sum_{j=1}^{t} \varepsilon_j + c_t \tag{7}$$

where $\gamma = \sum_{j=0}^{\infty} \beta_j$ and $c_t = \sum_{i=0}^{\infty} (-\tilde{\beta}_i) \varepsilon_{t-i}$, $y_0 = \sum_{i=0}^{\infty} \tilde{\beta}_i \varepsilon_{-i}$, and $\tilde{\beta}_j = \sum_{i=j+1}^{\infty} \beta_i$. Also,

$$\sum_{j=0}^{\infty} | \sum_{i=j+1}^{\infty} \beta_i | = \sum_{j=0}^{\infty} | \sum_{i=0}^{\infty} I(i \geq j+1) \beta_i | \leq \sum_{i=0}^{\infty} \sum_{j=0}^{i-1} |\beta_i| = \sum_{i=0}^{\infty} i |\beta_i|. \tag{8}$$

Therefore, it now follows that the conditions of Assumption 1 hold for linear processes under the following assumption:

Assumption 2. *Assume that $y_t = \sum_{j=1}^{t} u_j$ where $u_t = \sum_{j=0}^{\infty} \beta_j \varepsilon_{t-j}$, ε_t, $t = ..., -1, 0, 1, ...$, is i.i.d., $E\varepsilon_0 = 0$, $0 < E|\varepsilon_0|^p < \infty$ for some $p > 3$, $\sum_{i=0}^{\infty} i|\beta_i| < \infty$, $\sum_{j=0}^{\infty} \beta_j \neq 0$, and the distribution of ε_t is continuous. Then Assumption 1 holds.*

For the case when no intercept is included, the Dickey-Fuller coefficient test equals $DF_1 = n(\hat{\rho}_1 - 1)$ where

$$\hat{\rho}_1 = \frac{\sum_{t=2}^{n} \exp(y_t + y_{t-1})}{\sum_{t=2}^{n} \exp(2y_{t-1})} = \frac{\sum_{t=1}^{n-1} \exp(u_{t+1} + 2c_t + 2x_t - 2m_{n-1})}{\sum_{t=1}^{n-1} \exp(2c_t + 2x_t - 2m_{n-1})}. \tag{9}$$

The Dickey-Fuller test under the assumption that it was applied to the exponential of a unit root process was considered by Davies and Krämer (2003). Under the assumption of i.i.d. Δy_t, these authors showed an $O_p(1)$ property for $\hat{\rho}_1$. Under Assumption 2, for both the numerator and the denominator, Theorem 1 holds. However, this convergence is also joint, as the entire proof of Appendix 1 could be rewritten to show convergence in distribution for

$$\sum_{t=1}^{n}(\lambda_1 \exp(c_{t1} + \gamma_1 x_t) + \lambda_2 \exp(c_{t2} + \gamma_2 x_t)) \tag{10}$$

for any $\lambda_1, \lambda_2 \in \mathbb{R}$ and linear processes u_{t1} and u_{t2}, and the joint convergence then follows from the Cramèr-Wold device.

When an intercept is included, the Dickey-Fuller coefficient test equals $DF_2 = n(\hat{\rho}_2 - 1)$, where

$$\hat{\rho}_2 = \frac{\sum_{t=2}^{n} \exp(y_t + y_{t-1}) - (n-1)^{-1} \sum_{t=2}^{n} \exp(y_t) \sum_{t=2}^{n} \exp(y_{t-1})}{\sum_{t=2}^{n} \exp(2y_{t-1}) - (n-1)^{-1}(\sum_{t=2}^{n} \exp(y_{t-1}))^2}. \tag{11}$$

After rescaling the numerator and the denominator by $\exp(-2m_{n-1})$, it is now easily seen that under Assumption 2, $\hat{\rho}_2 = \hat{\rho}_1 + o_p(1)$. Also, letting RSS_1 and RSS_2 denote the residual sum of squares from the regression of $\exp(y_t)$ on $\exp(y_{t-1})$ without and with constant, it now immediately follows from Theorem 1 that $\exp(-2m_{n-1})RSS_1 = \exp(-2m_{n-1})RSS_2 + o_p(1) \xrightarrow{d} RSS$ for some random variable RSS. A trivial bookkeeping exercise and noting that the convergence in distribution is joint now implies that the Dickey-Fuller t-tests DF_3 and DF_4, obtained respectively from the regression of $\exp(y_t)$ on $\exp(y_{t-1})$ without and with constant, satisfy $DF_3 = DF_4 + o_p(1) \xrightarrow{d} DF$ for some limit distribution DF.

To summarize, we have the following Corollary:

Corollary 1. *Under Assumption 2,* $\hat{\rho}_1 = \hat{\rho}_2 + o_p(1)$*,* $DF_1/n = DF_2/n + o_p(1)$*,* $DF_3 = DF_4 + o_p(1)$*, and* $\hat{\rho}_1$*,* $\hat{\rho}_2$*,* DF_1/n*,* DF_2/n*,* DF_3 *and* DF_4 *converge in distribution.*

For the case of i.i.d. innovations, Krämer and Davies (2003) were able to show that DF_1 and DF_2 concentrate asymptotically on negative values, implying that the Dickey-Fuller tests asymptotically reject. The above corollary can only assert that the value of the Dickey-Fuller statistics is asymptotically large; however, given the result by Davies and Krämer, it seems that we should expect to find asymptotic rejection.

Therefore, we now find that asymptotically the Dickey-Fuller coefficient tests should asymptotically reject the null of a unit root when applied to the exponential of a unit root process, while the Dickey-Fuller t-tests do not necessarily reject asymptotically. In view of the fact that limit distributions are not pivotal, using these results for inference is not straightforward.

4 The KPSS statistic

In this section, the result of Equation (5) will be exploited for deriving the limit behavior of the KPSS statistic when applied to $\exp(y_t)$. The KPSS test was introduced in Kwiatkowski, Phillips, Schmidt and Shin (1992) and is the leading test for testing the null hypothesis of stationarity against the alternative of a unit root process. In this section, the limit distribution of the KPSS test when applied to the exponential of a unit root process will be derived.

Note that by the so-called "arcsine law" (see Feller (1968, page 399, Theorem 1a)), $\tau_n/n \xrightarrow{d} \tau$ where $P(\tau \leq z) = 2\pi^{-1}\arcsin(\sqrt{z})I(0 \leq z \leq 1) + I(z > 1)$; the distribution of τ will characterize the limit distributions of the KPSS test when applied to the exponential of the unit root process. For ease of notation, let $z_{nt} = \exp(y_t - m_n)$. The KPSS statistic applied to $\exp(y_t)$ then equals

$$KPSS_1 = \frac{n^{-2}\sum_{t=1}^{n}(\sum_{j=1}^{t} z_{nj})^2}{n^{-1}\sum_{t=1}^{n}\sum_{s=1}^{n} z_{nt}z_{ns}k((t-s)/\gamma_n)}, \tag{12}$$

while the KPSS statistic applied to demeaned data, i.e. to $\exp(y_t) - n^{-1}\sum_{t=1}^{n}\exp(y_t)$, equals

$$KPSS_2 = \frac{n^{-2}\sum_{t=1}^{n}(\sum_{j=1}^{t}(z_{nj} - \bar{z}))^2}{n^{-1}\sum_{t=1}^{n}\sum_{s=1}^{n}(z_{nt} - \bar{z})(z_{ns} - \bar{z})k((t-s)/\gamma_n)}. \tag{13}$$

The (positive) γ_n sequence is the bandwidth sequence and $k(\cdot)$ is the kernel function associated with the HAC estimator that appears in the denominator.

The limit distribution of $KPSS_1$ and $KPSS_2$ is given in the following result:

Theorem 2. *Assume that Assumption 1 holds. In addition, assume that* $k(0) = 1$, $k(.)$ *is continuous on a neighborhood of 0,* $k(x) = k(-x)$ *for all* $x > 0$, $|k(\cdot)| \leq k^*$, *and* $\int_{-\infty}^{\infty}|k(z)|dz < \infty$. *Also, assume* $\lim_{n\to\infty}(\gamma_n^{-1} + \gamma_n/n) = 0$. *Then* $KPSS_1 \xrightarrow{d} \tau$ *and* $KPSS_2 \xrightarrow{d} \tau^2 - \tau + 1/3$.

Note that the above theorem implies that the KPSS statistic when applied to the exponential of a unit root process converges in distribution without any type of rescaling, as $KPSS_2$ does when applied to weakly dependent data and $KPSS_1$ when applied to mean zero weakly dependent data. When applied to a unit root process, the KPSS statistic is known to diverge at rate n; see see Kwiatkowski, Phillips, Schmidt and Shin (1992) for details. Therefore, the KPSS test should have nontrivial asymptotic power against the alternative of a unit root in levels when used as a test for the null of a random walk in the logarithm of the series.

The distribution of $\tau^2 - \tau + 1/3$ can easily be calculated explicitly. Some elementary math gives $P(\tau^2 - \tau + 1/3 \leq z) = 2\pi^{-1} \arcsin(\sqrt{4z - 1/3}) I(z \in [1/12, 1/3]) + I(z > 1/3)$. Table 1 contains critical values for τ and $\tau^2 - \tau + 1/3$. These critical values were obtained analytically (by inverting both distribution functions).

Table 1: Critical values for τ and $\tau^2 - \tau + 1/3$.

Upper tail critical values	0.10	0.05	0.01
τ	0.97553	0.99384	0.99975
$\tau^2 - \tau + 1/3$	0.32722	0.33179	0.33327

Under the assumption of stationarity of the data, Kwiatkowski, Phillips, Schmidt and Shin (1992) give the 0.10, 0.05 and 0.01 upper tail critical values for $KPSS_2$ as 0.347, 0.463 and 0.739 respectively. Since these critical values all exceed 1/3 - the upper bound for $\tau^2 - \tau + 1/3$ - this means that conventional rejection of the null of stationarity using $KPSS_2$ also implies the rejection of the hypothesis that the series has a unit root in its logarithm.

Theorem 2 gives rise to an apparent empirical puzzle. This is because the KPSS statistic appears to reject in many instances for the level series when one might expect a unit root in the logarithm of the series. For example, Kwiatkowski, Phillips, Schmidt and Shin (1992) obtain values for $KPSS_2$ for the Nelson-Plosser data. For their series “real GNP”, “nominal GNP”, “real per capita GNP”, “industrial production”, “employment”, “consumer prices”, “GNP deflator”, “wages”, “real wages”, “money”, “velocity”, and “stock prices”, they obtain values for $KPSS_2$ that are far inside the rejection region. Through Theorem 2, this can now be interpreted as evidence against the presence of a unit root in the logarithm of these series. This seems to pose an empirical puzzle, as series with a stable growth rate are often assumed to have a unit root in their logarithms.

Finally, it should be noted that these authors obtained relatively small values for $KPSS_2$ for “interest rate” and “unemployment rate”. Given the result of Theorem 2, it follows that these values are consistent not only with stationarity of these series, but also with the presence of a unit root in their logarithms.

5 Simulations

Theorem 1 is easily illustrated with a simple simulation for the case where $y_t = x_t$ and $\Delta x_t = \varepsilon_t$, $x_0 = 0$. While Theorem 1 holds for any value of the scaling parameter $E\varepsilon_t^2$, we should expect that the approximation will be poor for relatively low values of $E\varepsilon_t^2$. In that case after all, $x_t - m_n$ will be relatively small as well, and

$$\sum_{t=1}^{n} \exp(x_t - m_n) \approx \sum_{t=1}^{n} \exp(0) = n \tag{14}$$

and a large value for n will be needed in such a situation in order to achieve a good approximation to the limit distribution.

Similarly, for large values of $E\varepsilon_t^2$, $|x_t - m_n|$ will be relatively large for all values of t except those for which $x_t = m_n$. Essentially, the random walk will jump towards its maximum and away from it with a "large" jump, implying that $\exp(x_t - m_n) \approx 0$ for all t for which $x_t \neq m_n$. This also may be problematic in terms of the quality of the asymptotic approximation for moderate values of n, as the statistic will be close to 1 in that case if we assume ε_t to be continuously distributed.

A Fortran simulation program (available from the author upon request) was used to generate simulation results for

$$Q_n(c) = \sum_{t=1}^{n} \exp(c(x_t - \tilde{m}_n)) \tag{15}$$

for various values of n and c and various distributions for i.i.d. ε_t that had a variance of 1. Instead of m_n, $\tilde{m}_n = \max(0, m_n)$ was used since theorems such as those in Spitzer (1956) are capable of showing that small sample moments of $\tilde{m}_n$ are well-behaved; however, $P(m_n \neq \tilde{m}_n) = O(n^{-1/2})$, so any small sample effect is minor. For the distribution of ε_t, I used a standard normal, a uniform$[-\sqrt{3}, \sqrt{3}]$, and a Rademacher distribution (i.e., $P(\varepsilon_t = -1) = P(\varepsilon_t = 1) = 0.5$), all of which have a variance of 1. Note that the Rademacher distribution is obviously not continuous and falls outside of the scope of the formal results of the previous chapter.

In order to observe the convergence in distribution from the simulation, I used the values 50, 100, 500, 1000, 5000 and 10,000 in all situations, and added simulations for $n = 50,000$ and $n = 100,000$ for the case $c = 1$, as convergence seemed to be slow for that case. The values $c = 1$, 2, 5 and 10 were used. Everywhere, $1,000,000$ replications were used to obtain the results, except for the simulations conducted for $n = 50,000$ and $n = 100,000$, where $100,000$ replications were used.

Simulation results can be found in the tables of Appendix 2. A striking feature is that the quantiles seem to be increasing with n. In the simulations, it can be observed that some quantiles end up below 1. This is possible because $\tilde{m}_n$ equals 0 in the case where the entire random walk is negative, and ε_t will not take the value $\tilde{m}_n$ for any t in that situation.

However, since under my assumptions the random walk is oscillating, the probability of the entire random walk being negative vanishes asymptotically.

For $c = 1$, the convergence in distribution appears to be relatively slow. For higher values of c such as 5 and 10, the convergence in distribution appears to be rapid, with a good approximation to the limit distribution being reached for $n = 100$ to $n = 500$. For $c = 10$, the convergence is rapid, but the distribution of the $Q_n(10)$ statistic degenerates into excessive closeness to 1.

References

Bertoin, J. and R. A. Doney (1994), On conditioning a random walk to stay nonnegative, *Annals of Probability* 4, 2152-2167.

Biggins, J.D. (2003), Random walk conditioned to stay positive, *Journal of the London Mathematical Society* 67, 259-272.

Borodin, A. N. and I.A. Ibragimov (1995), Limit theorems for functionals of random walks, *Proceedings of the Steklov Institute of Mathematics 195.*

Breitung, J. and C. Gourieroux (1997), Rank tests for unit roots, *Journal of Econometrics* 81, 7-27.

Burridge, P. and E. Guerre (1996), The limit distribution of level crossings of a random walk, and a simple unit root test, *Econometric Theory* 12, 705-723.

Corradi, V. (1995), Nonlinear transformation of integrated time series: a reconsideration?, *Journal of Time Series Analysis* 16, 539 - 549.

Davies, P.L. and W. Krämer (2003), The Dickey-Fuller test for exponential random walks, *Econometric Theory* 19, 865-877.

Ermini, L. and D.F. Hendry (2008), Log income vs. linear income: An application of the encompassing principle, *Oxford Bulletin of Economics and Statistics* 70, 807-827.

Feller, W. (1968), *An introduction to probability theory and its application*, Volume 2. New York: Wiley.

Franses, P.H. and G. Koop (1998), On the sensitivity of unit root inference to nonlinear data transformations, *Economics Letters* 59, 7-15.

Granger, C.W.J. and J. Hallmann (1991), Nonlinear transformations of integrated time series, *Journal of Time Series Analysis* 12, 207-224.

Kobayashi, M. and M. McAleer (1999), Tests of linear and logarithmic transformations for integrated processes, *Journal of the American Statistical Association* 94, 860-868.

Kozlov, M. (1976), On the asymptotic behavior of the probability of non-extinction for critical branching processes in a random environment, *Theory of Probability and its Applications* 21, 709-804.

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

Park, J.Y. and P.C.B. Phillips (1999), Asymptotics for nonlinear transformations of integrated time series, *Econometric Theory* 15, 269-298.

Park, J.Y. and P.C.B. Phillips (2000), Nonstationary binary choice, *Econometrica* 68, 1249-1280.

Park, J.Y. and P.C.B. Phillips (2001), Nonlinear regressions with integrated time series, *Econometrica* 69, 117-161.

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

Ritter, G.A. (1981), Growth of random walks conditioned to stay positive, *Annals of Probability* 9, 699-704.

Spitzer, F.L. (1956), A combinatorial lemma and its application to probability theory, *Transactions of the American Mathematical Society,* 82, 323-339 .

Spitzer, F.L. (1960), A Tauberian theorem and its probability interpretation, *Transactions of the American Mathematical Society* 94, 150-179.

Appendix 1: Proofs

Note that everywhere in the proof, I set $\gamma = 1$, since there is no loss of generality in doing so.

Proof of Theorem 1

The proof uses the following result from the literature. The first lemma is contained in the proof of Theorem 1 of Krämer and Davies (2003). Ritter (1981) attributes a similar result to Kozlov (1976) in his Lemma 2. Note that the second result of the lemma follows from a simple application of the first result to $-x_t$.

Lemma 1. *Assume that* $E\varepsilon_0 = 0$ *and* $E|\varepsilon_0|^3 < \infty$. *Then for all* $x > 0$,

$$P(\max_{1\leq t\leq n} x_t \leq x) \leq n^{-1/2}(a + bx)$$

and

$$P(\min_{1\leq t\leq n} x_t \geq -x) \leq n^{-1/2}(a + bx).$$

Define $x_t^* = \sum_{j=1}^t \varepsilon_{-j+1}$ for $t = 1, ..., n$ and $x_0^* = 0$, and note that x_t^* constitutes a random walk that is independent of x_t. In addition, let $B_{1k} = \{\min_{1\leq t\leq k} x_t^* > 0\}$ and $B_{2k} = \{\max_{1\leq t\leq k} x_t \leq 0\}$.

The following preliminary result will be needed later on:

Lemma 2. *Assume that Assumption 1 holds. Then there exist constants* c^1 *and* c^2 *such that*

$$\sup_{i\in\mathbb{Z},k\geq 2} E(|\varepsilon_i|^{p-1}|B_{1k}) \leq c^1 E|\varepsilon_0|^p + c^2 E|\varepsilon_0|^{p-1}$$

and

$$\sup_{i\in\mathbb{Z},k\geq 2} E(|\varepsilon_i|^{p-1}|B_{2k}) \leq c^1 E|\varepsilon_0|^p + c^2 E|\varepsilon_0|^{p-1}.$$

Proof. I will show this for the case of conditioning on B_{1k} and note that the proof for the case of conditioning on B_{2k} is completely analogous. Note that for all $i \in \mathbb{Z}$,

$$E(|\varepsilon_i|^{p-1}|B_{1k})$$

$$= E(|\varepsilon_i|^{p-1}| \min_{1\leq t\leq k} \sum_{j=1}^{t} \varepsilon_{-j+1} > 0)$$

$$= P(B_{1k})^{-1} E(|\varepsilon_i|^{p-1} I(\min_{1\leq t\leq k} \sum_{j=1}^{t} \varepsilon_{-j+1} > 0))$$

$$\leq P(B_{1k})^{-1} E(|\varepsilon_i|^{p-1} P(\min_{1\leq t\leq k} \sum_{j=1,j\neq i}^{t} \varepsilon_{-j+1} > -|\varepsilon_i||\varepsilon_i)).$$

Now $\sum_{j=1,j\neq i}^{t}\varepsilon_{-j+1}$ is again a random walk for which Lemma 1 holds. Therefore,

$$P(B_{1k})^{-1}E(|\varepsilon_i|^{p-1}P(\min_{1\le t\le k}\sum_{j=1,j\neq i}^{t}\varepsilon_{-j+1}>-|\varepsilon_i||\varepsilon_i))$$

$$\le P(B_{1k})^{-1}E(|\varepsilon_i|^{p-1}(a+b|\varepsilon_i|))(k-1)^{-1/2}$$
$$\le c^1E|\varepsilon_0|^p+c^2E|\varepsilon_0|^{p-1},$$

where it was used that $\inf_{k\geq 2}k^{1/2}P(B_{1k})>0$ because $P(B_{1k})>0$ for all $k\geq 2$ and $k^{1/2}P(B_{1k})\to c$ as $k\to\infty$ by Lemma 1 of Ritter (1981) (who attributes this result to Spitzer (1960)). □

Define $c_{t+}=\sum_{j=-\infty}^{\infty}\alpha_j\varepsilon_{t+j}I(t+j\geq 1)$, and $c_{t-}=\sum_{-\infty}^{\infty}\alpha_j\varepsilon_{t+j}I(t+j\leq 0)$. Lemma 2 can be exploited to find the following upper bound for conditional expectations of c_{t+} and c_{t-}:

Lemma 3. *Under Assumption 1,*

$$u_{p-1}^{+}=\sup_{t\in\mathbb{Z},k\geq 2}E(|c_{t+}|^{p-1}|B_{2k})<\infty$$

and

$$u_{p-1}^{-}=\sup_{t\in\mathbb{Z},k\geq 2}E(|c_{t-}|^{p-1}|B_{1k})<\infty$$

Proof. By Jensen's inequality and Lemma 2,

$$E(|c_{t+}|^{p-1}|B_{2k})=E(|\sum_{j=1-t}^{\infty}\alpha_j\varepsilon_{t+j}|^{p-1}|B_{2k})\le(\sum_{j=-\infty}^{\infty}|\alpha_j|)^{p-2}\sum_{j=-\infty}^{\infty}|\alpha_j|E(|\varepsilon_{t+j}|^{p-1}|B_{2k})$$

$$\le(\sum_{j=-\infty}^{\infty}|\alpha_j|)^{p-1}(c^1E|\varepsilon_0|^p+c^2E|\varepsilon_0|^{p-1})<\infty.$$

Since the same bound can be obtained for $E(|c_{t-}|^{p-1}|B_{1k})$, the result now follows. □

The proof uses a minor modification of a result by Bertoin and Doney (1994) involving random walks conditioned to be positive:

Lemma 4. *Assume that Assumption 1 holds. If $f(\cdot)$ is a continuous and bounded function of $(x_1, ..x_L, x_1^*, ..., x_L^*)$, then there exist independent Markov chains $\tilde{x}_t$, $t = 1, ...$ and $\tilde{x}_t^*$, $t = 1, ...$ such that*

$$\lim_{k\to\infty}\lim_{m\to\infty} E(f(x_1, ..x_L, x_1^*, ..., x_L^*)|B_{1k}, B_{2m}) = \lim_{m\to\infty}\lim_{k\to\infty} E(f(x_1, ..x_L, x_1^*, ..., x_L^*)|B_{1k}, B_{2m})$$

$$= Ef(\tilde{x}_1, ..., \tilde{x}_L, \tilde{x}_1^*, ..., \tilde{x}_L^*).$$

Proof. This result is based on Theorem 1 of Bertoin and Doney (1994), which states that there exists Markov chains $\tilde{x}_t$ and $\tilde{x}_t^*$ such that for bounded Borel measurable functions $g(\cdot)$,

$$\lim_{m\to\infty} E(g(x_1, ..., x_L)|B_{2m}) = Eg(\tilde{x}_1, ..., \tilde{x}_L)$$

and

$$\lim_{k\to\infty} E(g(x_1^*, ..., x_L^*)|B_{1k}) = Eg(\tilde{x}_1^*, ..., \tilde{x}_L^*).$$

By boundedness of $f(\cdot)$,

$$E(f(x_1, ..x_L, x_1^*, ..., x_L^*)I(|\max_{1\le i\le L}|x_i| > K)|B_{1k}, B_{2m})$$

$$\le CE(I(\max_{1\le i\le L}|x_i| > K)|B_{1k}, B_{2m}) = CE(I(\max_{1\le i\le L}|x_i| > K)|B_{2m}) \to CEI(\max_{1\le i\le L}|\tilde{x}_i| > K),$$

and the last expression vanishes as $K \to \infty$. Since a similar argument holds for x_t^*, we can restrict attention to

$$E(f(x_1, ..x_L, x_1^*, ..., x_L^*)I(\max_{1\le i\le L}|x_i^*| \le K)I(\max_{1\le i\le L}|x_i| \le K)|B_{1k}, B_{2m}).$$

Let $s^j : j = 1, ..., M_\varepsilon$ denote the centers of M_ε balls of radius ε that cover $[-K, K]^L$. Define

$$f^\varepsilon(x_1, ..x_L, x_1^*, ..., x_L^*) = \sum_{j=1}^{M_\varepsilon}\sum_{i=1}^{M_\varepsilon} f(\theta_i, \theta_j)I((x_1, ..x_L) \in B(s^i, \varepsilon))I((x_1^*, ..x_L^*) \in B(s^j, \varepsilon))$$

and note that by continuity of $f(\cdot, \cdot)$,

$$\lim_{\varepsilon\downarrow 0}\sup_{(s_1,s_1^*)'\in[-K,K]^{2L},(s_2^*,s_2^*)'\in[-K,K]^{2L}:|(s_1,s_1^*)'-(s_2,s_2^*)'|\le\varepsilon} |f^\varepsilon(s_1, s_1^*) - f(s_2, s_2^*)| = 0.$$

Also,

$$E(f^{\varepsilon}(x_1,..x_L,x_1^*,...,x_L^*)|B_{1k},B_{2m})$$

$$=\sum_{j=1}^{M_\varepsilon}\sum_{i=1}^{M_\varepsilon} f(\theta_i,\theta_j)E(I((x_1,..x_L)\in B(s^i,\varepsilon))|B_{2m})E(I((x_1^*,..x_L^*)\in B(s^j,\varepsilon))|B_{1k})$$

$$\to \sum_{j=1}^{M_\varepsilon}\sum_{i=1}^{M_\varepsilon} f(\theta_i,\theta_j)EI((\tilde{x}_1,..\tilde{x}_L)\in B(s^i,\varepsilon))EI((\tilde{x}_1^*,..\tilde{x}_L^*)\in B(s^j,\varepsilon))$$

$$=Ef^{\varepsilon}(\tilde{x}_1,..\tilde{x}_L,\tilde{x}_1^*,..\tilde{x}_L^*)$$

as $k,m\to\infty$ (in either order) by Theorem 1 of Bertoin and Doney (1994) as quoted at the start of this proof. Therefore,

$$|E(f(x_1,..x_L,x_1^*,...,x_L^*)|B_{1k},B_{2m})-Ef(\tilde{x}_1,..\tilde{x}_L,\tilde{x}_1^*,...,\tilde{x}_L^*)|$$

$$\le 2\sup_{(s_1,s_1^*)'\in[-K,K]^{2L},(s_2^*,s_2^*)'\in[-K,K]^{2L}:|(s_1,s_1^*)'-(s_2,s_2^*)'|\le\varepsilon}|f^{\varepsilon}(s_1,s_1^*)-f(s_2,s_2^*)|$$

$$+E(|f(x_1,..x_L,x_1^*,...,x_L^*)|(I(\max_{1\le i\le L}|x_i|>K)+I(\max_{1\le i\le L}|x_i^*|>K))|B_{1k},B_{2m})$$

$$+E|f(\tilde{x}_1,..\tilde{x}_L,\tilde{x}_1^*,...,\tilde{x}_L^*)|(I(\max_{1\le i\le L}|\tilde{x}_i|>K)+I(\max_{1\le i\le L}|\tilde{x}_i^*|>K))$$

$$+|E(f^{\varepsilon}(x_1,..x_L,x_1^*,...,x_L^*)|B_{1k},B_{2m})-Ef^{\varepsilon}(\tilde{x}_1,..\tilde{x}_L,\tilde{x}_1^*,...,\tilde{x}_L^*)|,$$

and by letting k and m approach infinity first, then letting ε approach 0 and finally letting K approach infinity, the lemma now follows. □

Using the following result, it can be shown that the smallest values for τ_n and the largest ones are asymptotically irrelevant:

Lemma 5. *Assume that Assumption 1 holds. Then for any sequence of random variables Z_{nk} such that $P(Z_{nk}\in[0,1])=1$ and for any integer $L\ge 1$,*

$$E\sum_{k=1}^{n} I(\tau_n=k)Z_{nk}=o(1)+\sum_{k=L+1}^{n-L} I(\tau_n=k)Z_{nk}.$$

Proof. For n large enough, the absolute value of the difference between both expressions equals

$$E(\sum_{k=1}^{L} I(\tau_n = k) + \sum_{k=n-L+1}^{n} I(\tau_n = k))Z_{nk}$$

which is bounded by

$$P(1 \leq \tau_n \leq L) + P(n - L + 1 \leq \tau_n \leq n).$$

By the arcsine law (see e.g. Feller (1968, page 399, Theorem 1a), both probabilities converge to 0. Note that the definition of Feller's K_n is slightly different from that of τ_n, with K_n taking the value 0 if the entire random walk is negative; however, $P(K_n \neq \tau_n) = P(K_n = 0) \to 0$. □

Define $x_t^* = \sum_{j=1}^{t} \varepsilon_{-j+1}$ for $t = 1, ..., n$ and $x_0^* = 0$, and note that x_t^* constitutes a random walk that is independent of x_t. In addition, define for $k \geq 1$ $B_{1k} = \{\min_{1 \leq t \leq k} x_t^* > 0\}$ and $B_{2k} = \{\max_{1 \leq t \leq k} x_t \leq 0\}$.

Lemma 6. *Assume that Assumption 1 holds. Then*

$$\sum_{k=L+1}^{n-L} EI(\tau_n = k) \exp(-r \sum_{t=1}^{n} \exp(y_t - x_k))$$

$$= E \sum_{k=L+1}^{n-L} I(B_{1,k-1}, B_{2,n-k}) \exp(-r \sum_{t=0}^{k-1} \exp(-x_t^* + c_{-t})) \exp(-r \sum_{t=1}^{n-k} \exp(x_t + c_t)).$$

Proof. It is noted in Spitzer (1956) and Feller (1968, p. 573) that the event $\{\tau_n = k\}$, for $k = 2, ..., n - 1$ is identical to the event

$$\{\max_{1 \leq t \leq k-1}(x_t - x_k) < 0, \max_{k+1 \leq t \leq n}(x_t - x_k) \leq 0\}.$$

Using this observation, the last expression from the above lemma can be simplified by noting that for any $L \geq 1$ and $n \geq 2L + 1$

$$\sum_{k=L+1}^{n-L} EI(\tau_n = k) \exp(-r \sum_{t=1}^{n} \exp(y_t - x_k))$$

$$= \sum_{k=L+1}^{n-L} EI(\max_{1\le t\le k-1}(x_t - x_k) < 0, \max_{k+1\le t\le n}(x_t - x_k) \le 0)$$

$$\times \exp(-r\exp(c_k))\exp(-r\sum_{t=1}^{k-1}\exp(x_t - x_k + c_t))\exp(-r\sum_{t=k+1}^{n}\exp(x_t - x_k + c_t))$$

$$= \sum_{k=L+1}^{n-L} EI(\min_{1\le t\le k-1}(\sum_{j=t+1}^{k}\varepsilon_j) > 0, \max_{k+1\le t\le n}(\sum_{j=k+1}^{t}\varepsilon_j) \le 0)$$

$$\times \exp(-r\exp(c_k))\exp(-r\sum_{t=1}^{k-1}\exp(-\sum_{j=t+1}^{k}\varepsilon_j + c_t))\exp(-r\sum_{t=k+1}^{n}\exp(\sum_{j=k+1}^{t}\varepsilon_j + c_t)).$$

By strict stationarity of (ε_t, c_t), if we replace inside the above expectation $\{(\varepsilon_t, c_t), t = ..., -1, 0, 1, ...\}$ by $\{(\varepsilon_{t-k}, c_{t-k}), t = ..., -1, 0, 1, ...\}$, the expectation remains unchanged; this observation then gives us the identical expression

$$\sum_{k=L+1}^{n-L} EI(\min_{1\le t\le k-1}(\sum_{j=t+1}^{k}\varepsilon_{j-k}) > 0, \max_{k+1\le t\le n}(\sum_{j=k+1}^{t}\varepsilon_{j-k}) \le 0)$$

$$\times \exp(-r\exp(c_0))\exp(-r\sum_{t=1}^{k-1}\exp(-\sum_{j=t+1}^{k}\varepsilon_{j-k}+c_{t-k}))\exp(-r\sum_{t=k+1}^{n}\exp(\sum_{j=k+1}^{t}\varepsilon_{j-k}+c_{t-k}))$$

$$= \sum_{k=L+1}^{n-L} EI(\min_{1\le t\le k-1}(\sum_{j=t+1-k}^{0}\varepsilon_j) > 0, \max_{k+1\le t\le n}(\sum_{j=1}^{t-k}\varepsilon_j) \le 0)$$

$$\times \exp(-r\exp(c_0))\exp(-r\sum_{t=1}^{k-1}\exp(-\sum_{j=t+1-k}^{0}\varepsilon_j+c_{t-k}))\exp(-r\sum_{t=k+1}^{n}\exp(\sum_{j=1}^{t-k}\varepsilon_j+c_{t-k}))$$

$$= E\sum_{k=L+1}^{n-L} I(\min_{1\le t\le k-1} x_t^* > 0, \max_{1\le t\le n-k} x_t \le 0)$$

$$\times \exp(-r\exp(c_0))\exp(-r\sum_{t=1}^{k-1}\exp(-x_{k-t}^* + c_{t-k}))\exp(-r\sum_{t=k+1}^{n}\exp(x_{t-k} + c_{t-k}))$$

$$= E\sum_{k=L+1}^{n-L} I(B_{1,k-1}, B_{2,n-k})\exp(-r\sum_{t=0}^{k-1}\exp(-x_t^* + c_{-t}))\exp(-r\sum_{t=1}^{n-k}\exp(x_t + c_t)),$$

which concludes the proof. □

This last expression, which in some sense is normalized around the index value at which the maximum is attained, is simpler to manipulate because x_t^* and x_t are two independent random walks with i.i.d. innovations. The first result of this type, which limits the growth of the c_t sequence, that I need is the following:

Lemma 7. *Assume that Assumption 1 holds. For any $A > 0$ and $k \geq 1$, let $B_{3knA} = \{\max_{-k \leq j \leq n-k} |c_j|(|j|+1)^{-\phi} \leq A\}$ for some $\phi \in (1/(p-1), 1/2)$. Then there exists a function $f_1 : (0,\infty) \to [0,\infty)$ such that for any array of random variables Z_{nk} such that $P(Z_{nk} \in [0,1]) = 1$ and for any $A > 0$ and integer $L \geq 1$,*

$$0 \leq \limsup_{n\to\infty}(E\sum_{k=L+1}^{n-L} I(B_{1,k-1}, B_{2,n-k})Z_{nk} - E\sum_{k=L+1}^{n-L} I(B_{1,k-1}, B_{2,n-k}, B_{3knA})Z_{nk}) \leq f_1(A)$$

and $\lim_{A\to\infty} f_1(A) = 0$.

Proof. The lower bound is obvious, and the upper bound follows from noting that for any $L \geq 1$, using the definition of conditional expectation, $c_t = c_{t+} + c_{t-}$, subadditivity, the Markov inequality and Lemma 3,

$$\limsup_{n\to\infty}(E\sum_{k=L+1}^{n-L} I(B_{1,k-1}, B_{2,n-k})Z_{nk} - E\sum_{k=L+1}^{n-L} I(B_{1,k-1}, B_{2,n-k}, B_{3knA})Z_{nk})$$

$$\leq \limsup_{n\to\infty} E\sum_{k=L+1}^{n-L} I(B_{1,k-1}, B_{2,n-k})I(\max_{-k\leq j\leq n-k} |c_j|(|j|+1)^{-\phi} > A)$$

$$= \limsup_{n\to\infty} \sum_{k=L+1}^{n-L} P(B_{1,k-1}, B_{2,n-k})P(\max_{-k\leq j\leq n-k} |c_j|(|j|+1)^{-\phi} > A|B_{1,k-1}, B_{2,n-k})$$

$$\leq \limsup_{n\to\infty} \sum_{k=L+1}^{n-L} P(B_{1,k-1}, B_{2,n-k})P(\max_{-k\leq j\leq n-k} |c_{j-}|(|j|+1)^{-\phi} > A/2|B_{1,k-1})$$

$$+\limsup_{n\to\infty} \sum_{k=L+1}^{n-L} P(B_{1,k-1}, B_{2,n-k})P(\max_{-k\leq j\leq n-k} |c_{j+}|(|j|+1)^{-\phi} > A/2|B_{2,n-k})$$

$$\leq \limsup_{n\to\infty} E \sum_{k=L+1}^{n-L} P(B_{1,k-1}, B_{2,n-k}) u_{p-1}^{+} (\sum_{j=-\infty}^{\infty} (|j|+1)^{-(p-1)\phi}) 2^{p-1} A^{1-p}$$

$$+ \limsup_{n\to\infty} E \sum_{k=L+1}^{n-L} P(B_{1,k-1}, B_{2,n-k}) u_{p-1}^{-} (\sum_{j=-\infty}^{\infty} (|j|+1)^{-(p-1)\phi}) 2^{p-1} A^{1-p}$$

$$\leq 2^{p-1} A^{1-p} (u_{p-1}^{+} + u_{p-1}^{-}) (\sum_{j=-\infty}^{\infty} (|j|+1)^{-(p-1)\phi}) = f_1(A),$$

and obviously $f_1(A) \to 0$ as $A \to \infty$. □

A second result is the key to the proof and shows the asymptotic irrelevance to the statistic of values of x_t^* and x_t that are relatively close to zero:

Lemma 8. *Assume that Assumption 1 holds. Let* $B_{4kn\varepsilon} = \{\min_{1\leq t\leq k}(x_t^* - \varepsilon t^{\alpha}) \geq 0\} \cap \{\max_{1\leq t\leq n-k}(x_t + \varepsilon t^{\alpha}) \leq 0\}$ *for some* $\alpha \in (\phi, 1/2)$. *Then there exists a function* $f_2 : (0,\infty) \to [0,\infty)$ *such that for any array of random variables* Z_{nk} *such that* $P(Z_{nk} \in [0,1]) = 1$ *and for any* $\varepsilon > 0$ *and integer* $L \geq 1$,

$$0 \leq \limsup_{n\to\infty} (E \sum_{k=L+1}^{n-L} I(B_{1,k-1}, B_{2,n-k}, B_{3knA}) Z_{nk} - E \sum_{k=L+1}^{n-L} I(B_{1,k-1}, B_{2,n-k}, B_{3knA}, B_{4kn\varepsilon}) Z_{nk}) \leq f_2(\varepsilon)$$

and $\lim_{\varepsilon\downarrow\infty} f_2(\varepsilon) = 0$.

Proof. The lower bound is again obvious, and using the identity $\sum_{k=1}^{n} P(B_{1,k-1}, B_{2,n-k}) = 1$,

$$\limsup_{n\to\infty} (E \sum_{k=L+1}^{n-L} I(B_{1,k-1}, B_{2,n-k}, B_{3kA}, B_{4k\varepsilon}) Z_{nk} - E \sum_{k=L+1}^{n-L} I(B_{1,k-1}, B_{2,n-k}, B_{3kA}) Z_{nk})$$

$$\leq \limsup_{n\to\infty} \sum_{k=L+1}^{n-L} P(B_{1,k-1}, B_{2,n-k})$$

$$\times (P(\min_{1\leq t\leq k-1}(x_t^* - \varepsilon t^{\alpha}) < 0 | B_{1,k-1}, B_{2,n-k}) + P(\max_{1\leq t\leq n-k}(x_t + \varepsilon t^{\alpha}) > 0 | B_{1,k-1}, B_{2,n-k}))$$

$$\leq \sup_{k\geq 1} P(\min_{1\leq t\leq k}(x_t^* - \varepsilon t^{\alpha}) < 0 | B_{1k}) + \sup_{k\geq 1} P(\inf_{1\leq t\leq k}(x_t + \varepsilon t^{\alpha}) > 0 | B_{2k}) \equiv f_2(\varepsilon),$$

where $f_2(\varepsilon) \to 0$ as $\varepsilon \downarrow 0$ by Theorem 2 of Ritter (1981). $\square$

Next, I show that R_n is asymptotically determined by a finite but small number of observations close to τ_n.

Lemma 9. *Assume that Assumption 1 holds. Then for all $M < L$ and $L \geq 1$*

$$0 \leq \limsup_{n\to\infty} E \sum_{k=L+1}^{n-L} I(B_{1,k-1}, B_{2,n-k}, B_{3kA}, B_{4k\varepsilon}) \exp(-r \sum_{t=0}^{M} \exp(-x_t^* + c_{-t})) \exp(-r \sum_{t=1}^{M} \exp(x_t + c_t))$$

$$-E \sum_{k=L+1}^{n-L} I(B_{1,k-1}, B_{2,n-k}, B_{3kA}, B_{4k\varepsilon}) \exp(-r \sum_{t=0}^{k} \exp(-x_t^* + c_{-t})) \exp(-r \sum_{t=1}^{n-k} \exp(x_t + c_t))$$

$$\leq f_3(M, \varepsilon, A),$$

and for all $\varepsilon, A > 0$, $\lim_{M\to\infty} f_3(M, \varepsilon, A) = 0$.

Proof. Note that all possible values of k in the summation will exceed $M + 1$ and be less than $n - M$. Since for any sequences $a_t, b_t \geq 0$ and for all $k > M + 1$ and $k < n - M$,

$$0 \leq \exp(-\sum_{t=0}^{M} a_t - \sum_{t=1}^{M} b_t) - \exp(-\sum_{t=0}^{k} a_t - \sum_{t=1}^{n-k} b_t) \leq 1 - \exp(-\sum_{t=M+1}^{\infty} a_t - \sum_{t=M+1}^{\infty} b_t),$$

we can bound the expression from Lemma 9 by

$$\limsup_{n\to\infty} E \sum_{k=L+1}^{n-L} I(B_{1,k-1}, B_{2,n-k}, B_{3kA}, B_{4k\varepsilon})(1-\exp(-r \sum_{t=M+1}^{\infty} \exp(-x_t^* + c_{-t}) - r \sum_{t=M+1}^{\infty} \exp(x_t + c_t)))$$

$$\leq 1 - \exp(-2r \sum_{t=M+1}^{\infty} \exp(A(t+1)^{\phi} - \varepsilon t^{\alpha})) = f_3(M, \varepsilon, A),$$

and since we chose $\phi < \alpha < 1/2$, for all $\varepsilon, A > 0$, $\lim_{M\to\infty} f_3(M, \varepsilon, A) = 0$. $\square$

In order to limit dependence of c_t and be in a position to apply Lemma 4, we need the following result. Define $c_t^K = \sum_{j=-K}^{K} \alpha_j \varepsilon_{t+j}$, $c_{t+}^K = \sum_{j=-K}^{K} \alpha_j \varepsilon_{t+j} I(t + j \geq 1)$, and $c_{t-}^K = \sum_{-K}^{K} \alpha_j \varepsilon_{t+j} I(t + j \leq 0)$. Obviously, $c_{t+}^K + c_{t-}^K = c_t^K$. Also, in order to avoid unboundedness issues in the proof, define $e^C(x) = \exp(x)I(x \leq C) + \exp(C)I(x > C)$.

Lemma 10. *Assume that Assumption 1 holds. Let* $C_{A\varepsilon} = \sup_{t\geq 0}(A(t+1)^{\phi} - \varepsilon t^{\alpha})$. *Then for all* $L \geq 1$

$$\limsup_{n\to\infty} |E \sum_{k=L+1}^{n-L} I(B_{1,k-1}, B_{2,n-k}, B_{3kA}, B_{4k\varepsilon}) \exp(-r \sum_{t=0}^{M} e^{C_{A\varepsilon}}(-x_t^* + c_{-t}^K)) \exp(-r \sum_{t=1}^{M} e^{C_{A\varepsilon}}(x_t + c_t^K))$$

$$-E \sum_{k=L+1}^{n-L} I(B_{1,k-1}, B_{2,n-k}, B_{3kA}, B_{4k\varepsilon}) \exp(-r \sum_{t=0}^{M} \exp(-x_t^* + c_{-t})) \exp(-r \sum_{t=1}^{M} \exp(x_t + c_t))|$$

$$\leq f_4(\varepsilon, A, K, M),$$

and for all $M \geq 1$ *and* $\varepsilon, A > 0$, $\lim_{K\to\infty} f_4(\varepsilon, A, K, M) = 0$.

Proof. Note that for all $t \geq 0$, $\exp(x_t + c_t)I(B_{3kA}, B_{4k\varepsilon}) = e^{C_{A\varepsilon}}(x_t + c_t)I(B_{3kA}, B_{4k\varepsilon})$ a.s. and $\exp(-x_t^* + c_{-t})I(B_{3kA}, B_{4k\varepsilon}) = e^{C_{A\varepsilon}}(-x_t^* + c_{-t})I(B_{3kA}, B_{4k\varepsilon})$ a.s.. Also, since $|\exp(-|x|) - \exp(-|y|)| \leq ||x| - |y||$ and because for all $x, y < 0$, $|e^{C_{A\varepsilon}}(x) - e^{C_{A\varepsilon}}(y)| \leq \exp(C_{A\varepsilon})|x - y|$,

$$|E \sum_{k=L+1}^{n-L} I(B_{1,k-1}, B_{2,n-k}, B_{3kA}, B_{4k\varepsilon}) \exp(-r \sum_{t=0}^{M} \exp(-x_t^* + c_{-t}^K)) \exp(-r \sum_{t=1}^{M} \exp(x_t + c_t^K))$$

$$-E \sum_{k=L+1}^{n-L} I(B_{1,k-1}, B_{2,n-k}, B_{3kA}, B_{4k\varepsilon}) \exp(-r \sum_{t=0}^{M} \exp(-x_t^* + c_{-t})) \exp(-r \sum_{t=1}^{M} \exp(x_t + c_t))|$$

$$\leq rE \sum_{k=L+1}^{n-L} I(B_{1,k-1}, B_{2,n-k}, B_{3kA}, B_{4k\varepsilon})| \sum_{t=0}^{M} e^{C_{A\varepsilon}}(-x_t^* + c_{-t}) + e^{C_{A\varepsilon}}(x_t + c_t)$$

$$-e^{C_{A\varepsilon}}(x_t^* + c_{-t}^K) - e^{C_{A\varepsilon}}(x_t + c_t^K)|$$

$$\leq 2\exp(C_{A\varepsilon})rE \sum_{k=L+1}^{n-L} I(B_{1,k-1}, B_{2,n-k}) \sum_{t=-M}^{M} |c_t^K - c_t|.$$

Now by the definition of conditional expectation given an event and Lemma 2, the last expression can be bounded by

$$2\exp(C_{A\varepsilon})r \sum_{k=L+1}^{n-L} P(B_{1,k-1}, B_{2,n-k}) \sum_{t=-M}^{M} E(|c_{t-}^K - c_{t-} + c_{t+}^K - c_{t+}||B_{1,k-1}, B_{2,n-k})$$

$$\leq 2\exp(C_{A\varepsilon})r \sum_{k=L+1}^{n-L} P(B_{1,k-1}, B_{2,n-k}) \sum_{t=-M}^{M} (E(|c_{t+}^K - c_{t+}||B_{2,n-k}) + E(|c_{t-}^K - c_{t-}||B_{1k}))$$

$$\leq 2\exp(C_{A\varepsilon})r\sum_{k=L+1}^{n-L}P(B_{1,k-1},B_{2,n-k})\sum_{j=K+1}^{\infty}|\alpha_j|\sum_{t=-M}^{M}(E(|\varepsilon_{t+j}||B_{2,n-k})+E(|\varepsilon_{t+j}||B_{1k}))$$

$$\leq 2\exp(C_{A\varepsilon})r\sum_{k=L+1}^{n-L}P(B_{1,k-1},B_{2,n-k})\sum_{j=K+1}^{\infty}(|\alpha_j|+|\alpha_{-j}|)(2M+1)2(c^1E|\varepsilon_0|+c^2E|\varepsilon_0|^2)$$

$$\leq 4\exp(C_{A\varepsilon})r\sum_{j=K+1}^{\infty}(|\alpha_j|+|\alpha-j|)(2M+1)(c^1E|\varepsilon_0|+c^2E|\varepsilon_0|^2)=f_4(\varepsilon,A,K,M),$$

and obviously for all $M\geq 1$ and $\varepsilon, A>0$, $\lim_{K\to\infty}f_4(\varepsilon,A,K,M)=0$. □

I now apply Lemma 4 and remove the need for working with conditional expectations:

Lemma 11. *Assume that Assumption 1 holds. Then for all* $L\geq 1$

$$\limsup_{n\to\infty}|E\sum_{k=L+1}^{n-L}I(B_{1,k-1},B_{2,n-k},B_{3kA},B_{4k\varepsilon})\exp(-r\sum_{t=0}^{M}e^{C_{A\varepsilon}}(-x_t^*+c_{-t}^K))\exp(-r\sum_{t=1}^{M}e^{C_{A\varepsilon}}(x_t+c_t^K))$$

$$-E\exp(-r\sum_{t=0}^{M}e^{C_{A\varepsilon}}(-\tilde{x}_t^*+\tilde{c}_{-t}^K))\exp(-r\sum_{t=1}^{M}e^{C_{A\varepsilon}}(\tilde{x}_t+\tilde{c}_t^K))|\leq f_5(K,M,L)$$

and for all $K,M\geq 1$, $\lim_{L\to\infty}f_5(K,M,L)=0$.

Proof. Define

$$\lim_{k\to\infty}\lim_{m\to\infty}E(\exp(-r\sum_{t=0}^{M}e^{C_{A\varepsilon}}(-x_t^*+c_{-t}^K))\exp(-r\sum_{t=1}^{M}e^{C_{A\varepsilon}}(x_t+c_t^K))|B_{1k},B_{2,m})$$

$$=E\exp(-r\sum_{t=0}^{M}e^{C_{A\varepsilon}}(-\tilde{x}_t^*+\tilde{c}_{-t}^K))\exp(-r\sum_{t=1}^{M}e^{C_{A\varepsilon}}(\tilde{x}_t+\tilde{c}_t^K))$$

$$=\psi_{MKC_{A\varepsilon}}(r)$$

and note that this limit exists by Lemma 4. Now by the definition of expectations conditioned on an event, the identity $\sum_{k=1}^{n}P(B_{1,k-1},B_{2,n-k})=1$ and Lemma 4,

$$|E\sum_{k=L+1}^{n-L}I(B_{1,k-1},B_{2,n-k})[\exp(-r\sum_{t=0}^{M}\exp(-x_t^*+c_{-t}^K))\exp(-r\sum_{t=1}^{M}\exp(x_t+c_t^K))-\psi_{LKC_{A\varepsilon}}(r)]|$$

$$\leq \sum_{k=L+1}^{n-L} P(B_{1,k-1}, B_{2,n-k}) |E(\exp(-r \sum_{t=0}^{M} \exp(-x_t^* + c_{-t}^K)) \exp(-r \sum_{t=1}^{M} \exp(x_t + c_t^K)) | B_{1,k-1}, B_{2,n-k})$$
$$- \psi_{MKC_{A\varepsilon}}(r)|$$
$$\leq \sup_{k \geq L} \sup_{m \geq L} |E(\exp(-r \sum_{t=0}^{M} \exp(-x_t^* + c_{-t}^K)) \exp(-r \sum_{t=1}^{M} \exp(x_t + c_t^K)) | B_{1k}, B_{2m}) - \psi_{LKC_{A\varepsilon}}(r)|$$
$$= f_5(K, M, L),$$

and $f_5(K, M, L) \to 0$ as $L \to \infty$ by Lemma 4. Also,

$$\psi_{LKC_{A\varepsilon}}(r) \sum_{k=L+1}^{n-L} P(B_{1,k-1}, B_{2,n-k}) \to \psi_{LKC_{A\varepsilon}}(r)$$

as $n \to \infty$. □

To complete the proof, now note that for $L \geq 1$

$$|E \exp(-r \sum_{t=1}^{n} \exp(y_t - m_n)) - E \exp(-r \sum_{t=0}^{\infty} \exp(-\tilde{x}_t^* + \tilde{c}_{-t})) \exp(-r \sum_{t=1}^{\infty} \exp(\tilde{x}_t + \tilde{c}_t))|$$

$$\leq |E \exp(-r \sum_{t=1}^{n} \exp(y_t - m_n)) - E \sum_{k=L+1}^{n-L} I(\tau_n = k) \exp(-r \sum_{t=1}^{n} \exp(y_t - x_k))|$$

$$+ |E \sum_{k=L+1}^{n-L} I(\tau_n = k) \exp(-r \sum_{t=1}^{n} \exp(y_t - x_k))$$
$$- E \sum_{k=L+1}^{n-L} I(B_{1,k-1}, B_{2,n-k}) \exp(-r \sum_{t=0}^{k-1} \exp(-x_t^* + c_{-t})) \exp(-r \sum_{t=1}^{n-k} \exp(x_t + c_t))|$$
$$+ |E \sum_{k=L+1}^{n-L} I(B_{1,k-1}, B_{2,n-k}) \exp(-r \sum_{t=0}^{k-1} \exp(-x_t^* + c_{-t})) \exp(-r \sum_{t=1}^{n-k} \exp(x_t + c_t))$$
$$- E \sum_{k=L+1}^{n-L} I(B_{1,k-1}, B_{2,n-k}, B_{3knA}) \exp(-r \sum_{t=0}^{k-1} \exp(-x_t^* + c_{-t})) \exp(-r \sum_{t=1}^{n-k} \exp(x_t + c_t))|$$

$$
\begin{aligned}
&+|E\sum_{k=L+1}^{n-L} I(B_{1,k-1},B_{2,n-k},B_{3knA})\exp(-r\sum_{t=0}^{k-1}\exp(-x_t^*+c_{-t}))\exp(-r\sum_{t=1}^{n-k}\exp(x_t+c_t))\\
&\quad -E\sum_{k=L+1}^{n-L} I(B_{1,k-1},B_{2,n-k},B_{3knA},B_{4kn\varepsilon})\exp(-r\sum_{t=0}^{k-1}\exp(-x_t^*+c_{-t}))\exp(-r\sum_{t=1}^{n-k}\exp(x_t+c_t))|\\
&+|E\sum_{k=L+1}^{L} I(B_{1,k-1},B_{2,n-k},B_{3knA},B_{4kn\varepsilon})\exp(-r\sum_{t=0}^{k-1}\exp(-x_t^*+c_{-t}))\exp(-r\sum_{t=1}^{n-k}\exp(x_t+c_t))\\
&\quad -E\sum_{k=L+1}^{n-L} I(B_{1,k-1},B_{2,n-k},B_{3knA},B_{4kn\varepsilon})\exp(-r\sum_{t=0}^{M}\exp(-x_t^*+c_{-t}))\exp(-r\sum_{t=1}^{M}\exp(x_t+c_t))|\\
&+|E\sum_{k=L+1}^{n-L} I(B_{1,k-1},B_{2,n-k},B_{3knA},B_{4kn\varepsilon})\exp(-r\sum_{t=0}^{M}\exp(-x_t^*+c_{-t}))\exp(-r\sum_{t=1}^{M}\exp(x_t+c_t))\\
&\quad -E\sum_{k=L+1}^{n-L} I(B_{1,k-1},B_{2,n-k},B_{3knA},B_{4kn\varepsilon})\exp(-r\sum_{t=0}^{M}\exp(-x_t^*+c_{-t}^K))\exp(-r\sum_{t=1}^{M}\exp(x_t+c_t^K))|\\
&+|E\sum_{k=L+1}^{n-L} I(B_{1,k-1},B_{2,n-k},B_{3knA},B_{4kn\varepsilon})\exp(-r\sum_{t=0}^{M}\exp(-x_t^*+c_{-t}^K))\exp(-r\sum_{t=1}^{M}\exp(x_t+c_t^K))\\
&\quad -E\exp(-r\sum_{t=0}^{M}e^{C_{A\varepsilon}}(-\tilde{x}_t^*+\tilde{c}_{-t}^K))\exp(-r\sum_{t=1}^{M}e^{C_{A\varepsilon}}(\tilde{x}_t+\tilde{c}_t^K))|\\
&+|E\exp(-r\sum_{t=0}^{M}e^{C_{A\varepsilon}}(-\tilde{x}_t^*+\tilde{c}_{-t}^K))\exp(-r\sum_{t=1}^{M}e^{C_{A\varepsilon}}(\tilde{x}_t+\tilde{c}_t^K))\\
&\quad -E\exp(-r\sum_{t=0}^{\infty}\exp(-\tilde{x}_t^*+\tilde{c}_{-t}))\exp(-r\sum_{t=1}^{\infty}\exp(\tilde{x}_t+\tilde{c}_t))|\\
&\le o(1)+f_1(A)+f_2(\varepsilon)+f_3(M,\varepsilon,A)+f_4(\varepsilon,A,K,M)+f_5(K,M,L)\\
&+|E\exp(-r\sum_{t=0}^{M}e^{C_{A\varepsilon}}(-\tilde{x}_t^*+\tilde{c}_{-t}^K))\exp(-r\sum_{t=1}^{M}e^{C_{A\varepsilon}}(\tilde{x}_t+\tilde{c}_t^K))\\
&\quad -E\exp(-r\sum_{t=0}^{\infty}\exp(-\tilde{x}_t^*+\tilde{c}_{-t}))\exp(-r\sum_{t=1}^{\infty}\exp(\tilde{x}_t+\tilde{c}_t))|.
\end{aligned}
$$

Next, we let $L \to \infty$ first (which makes f_5 vanish), then $K \to \infty$ (which makes f_4 vanish), then $M \to \infty$ (which makes f_3 vanish), and finally $A \to \infty$ and $\varepsilon \downarrow 0$ (which makes f_1 and f_2 vanish). Also, by the dominated convergence theorem,

$$\lim_{A\to\infty}\lim_{M\to\infty}\lim_{K\to\infty}|E\exp(-r\sum_{t=0}^{M}e^{C_{A\varepsilon}}(-\tilde{x}_t^*+\tilde{c}_{-t}^K))\exp(-r\sum_{t=1}^{M}e^{C_{A\varepsilon}}(\tilde{x}_t+\tilde{c}_t^K))$$
$$-E\exp(-r\sum_{t=0}^{\infty}\exp(-\tilde{x}_t^*+\tilde{c}_{-t}))\exp(-r\sum_{t=1}^{\infty}\exp(\tilde{x}_t+\tilde{c}_t))|=0.$$

This completes the proof of the convergence of the Laplace transform of R_n. By Feller's (1968, page 408) Theorem 2, $\lim_{n\to\infty} E\exp(-rR_n)$ is the transform of a possibly defective distribution $F(\cdot)$, and the limit $F(\cdot)$ is not defective if $\lim_{n\to\infty} E\exp(-rR_n) \to 1$ as $r \downarrow 0$. To show the limit is not defective, first note that by the earlier argument

$$\liminf_{n\to\infty} E\exp(-rR_n)$$
$$\geq \liminf_{n\to\infty} E\sum_{k=L+1}^{n-L} I(B_{1,k-1},B_{2,n-k},B_{3knA},B_{4kn\varepsilon})\exp(-r\sum_{t=0}^{k-1}\exp(-x_t^*+c_{-t}))\exp(-r\sum_{t=1}^{n-k}\exp(x_t+c_t))$$
$$-f_1(A)-f_2(\varepsilon)$$

where f_1 and f_2 do not depend on r. Also,

$$E\sum_{k=L+1}^{n-L} I(B_{1,k-1},B_{2,n-k},B_{3knA},B_{4kn\varepsilon})\exp(-r\sum_{t=0}^{k-1}\exp(-x_t^*+c_{-t}))\exp(-r\sum_{t=1}^{n-k}\exp(x_t+c_t))$$
$$\geq E\sum_{k=L+1}^{n-L} I(B_{1,k-1},B_{2,n-k},B_{3knA},B_{4kn\varepsilon})\exp(-2r\sum_{t=0}^{\infty}\exp(A(t+1)^{\phi}-\varepsilon t^{\alpha}))$$
$$\geq (1-o(1))\exp(-2r\sum_{t=0}^{\infty}\exp(A(t+1)^{\phi}-\varepsilon t^{\alpha})).$$

Therefore, for all $\varepsilon > 0$ and $A > 0$,

$$\liminf_{n\to\infty} E\exp(-rR_n) \geq -f_1(A)-f_2(\varepsilon)+\exp(-2r\sum_{t=0}^{\infty}\exp(A(t+1)^{\phi}-\varepsilon t^{\alpha})),$$

implying that

$$\liminf_{r\downarrow 0}\liminf_{n\to\infty} E\exp(-rR_n) \geq 1-f_1(\varepsilon)-f_2(A).$$

Since ε and A were arbitrary and because the Laplace transform is bounded from above by 1, the result now follows. □

Proof of Theorem 2

Since $z_{nt} \geq 0$ and because $|k(\cdot)| \leq k^*$, by the result of Equation (5),

$$\sum_{t=1}^{n}\sum_{s=1}^{n} z_{nt}z_{ns}k((t-s)/\gamma_n)I(|t-\tau_n| > M)$$

$$\leq v_{n1M} \equiv k^*(\sum_{s=1}^{n} z_{ns})(\sum_{t=1}^{n} z_{nt}I(|t-\tau_n| > M)) \xrightarrow{d} v_{1M}.$$

Obviously as $M \to \infty$, the last term vanishes. Also,

$$\sum_{t=1}^{n}\sum_{s=1}^{n} z_{nt}z_{ns}k((t-s)/\gamma_n)I(|t-\tau_n| \leq M)I(|s-\tau_n| > M) \leq v_{n1M}.$$

Defining

$$v_{n2M} = |\sum_{t=1}^{n}\sum_{s=1}^{n} z_{nt}z_{ns}k((t-s)/\gamma_n) - (\sum_{t=1}^{n} z_{nt}I(|t-\tau_n| \leq M))^2|,$$

it now follows, since $z_t \geq 0$ and $\exp(z_{nt}) = \exp(x_t + c_t - m_n) \leq \exp(c_t)$, that

$$v_{n2M} \leq 2v_{n1M} + \sum_{t=1}^{n}\sum_{s=1}^{n} z_{nt}z_{ns}|k((t-s)/\gamma_n) - 1|I(|t-\tau_n| \leq M)I(|s-\tau_n| \leq M)$$

$$\leq 2v_{n1M} + M^2 \sup_{|z|\leq 2M} |k(z/\gamma_n) - 1| \exp(2 \sup_{|t-\tau_n|\leq M} |c_t|) \xrightarrow{d} 2v_{1M}.$$

It was used here that $\sup_{|t-\tau_n|\leq M} |c_t| = O_p(1)$ because

$$E \sup_{|t-\tau_n|\leq M} |c_t| \leq \sum_{k=1}^{n} P(\tau_n = k) \sum_{|t-k|\leq M} E(|c_t||B_{1,k-1}, B_{2,n-k}) = O(1)$$

by Lemma 3. In addition,

$$0 \leq v_{n3M} = n^{-1}\sum_{t=1}^{n}(\sum_{j=1}^{t} z_{nj})^2 - n^{-1}\sum_{t=1}^{n}(\sum_{j=1}^{t} z_{nj}I(|j-\tau_n| \leq M))^2$$

$$\leq 2n^{-1}\sum_{t=1}^{n}(\sum_{j=1}^{t} z_{nj})(\sum_{j=1}^{n} z_{nj}I(|j-\tau_n| > M) = 2v_{n1M}/k^*.$$

Also, observe that

$$n^{-1}\sum_{t=\tau_n-M}^{\tau_n+M-1}(\sum_{j=1}^{t} z_{nj}I(|j-\tau_n|\leq M))^2 \leq n^{-1}2M(\sum_{j=1}^{n} z_{nj})^2 = O_p(n^{-1}),$$

$$n^{-1}\sum_{t=1}^{\tau_n-M-1}(\sum_{j=1}^{t} z_{nj}I(|j-\tau_n|\leq M))^2 = 0,$$

and

$$n^{-1}\sum_{t=\tau_n+M}^{n}(\sum_{j=1}^{t} z_{nj}I(|j-\tau_n|\leq M))^2 = n^{-1}(n-\tau_n-M)(\sum_{j=1}^{n} z_{nj}I(|\tau_n-j|\leq M))^2,$$

implying that

$$\frac{KPSS_1}{n^{-1}(n-\tau_n)} = \frac{n^{-1}\sum_{t=1}^{n}(\sum_{j=1}^{t} z_{nj})^2}{n^{-1}(n-\tau_n)\sum_{t=1}^{n}\sum_{s=1}^{n} z_{nt}z_{ns}k((t-s)/\gamma_n)}$$

$$\leq \frac{n^{-1}(n-\tau_n-M)((\sum_{t=1}^{n} z_{nt}I(|t-\tau_n|\leq M))^2 + v_{n3M} + 2v_{n4M} + O_p(n^{-1}))}{n^{-1}(n-\tau_n)((\sum_{t=1}^{n} z_{nt}I(|t-\tau_n|\leq M))^2 - v_{n2M})}$$

$$\xrightarrow{d} v_{4M}$$

for some limit v_{4M} such that $v_{4M} \xrightarrow{d} 1$ as $M \to \infty$. A similar lower bound can be obtained as well. This implies that

$$\frac{KPSS_1}{n^{-1}(n-\tau_n)} \xrightarrow{p} 1$$

and therefore that $KPSS_1 \xrightarrow{d} 1-\tau \overset{d}{=} \tau$. For the numerator of $KPSS_2$, a similar reasoning gives

$$n^{-1}\sum_{t=1}^{n}(\sum_{j=1}^{t} z_{nj} - t\bar{z})^2 = n^{-1}\sum_{t=1}^{n}[(\sum_{j=1}^{t} z_{nj})^2 - 2t\bar{z}\sum_{j=1}^{t} z_{nj} + t^2\bar{z}^2]$$

$$\leq o_p(1) + v_{n5M} + (\sum_{j=1}^{n} z_{nj}I(|\tau_n-j|\leq M))^2((1-\tau_n/n) - 2\int_{\tau_n/n}^{1} ydy + (1/3))$$

$$= o_p(1) + v_{n5M} + (\sum_{j=1}^{n} z_{nj} I(|\tau_n - j| \leq M))^2((\tau_n/n)^2 - (\tau_n/n) + (1/3))$$

for some random sequence v_{n5M} such that $v_{n5M} \xrightarrow{d} v_{5M}$ where $v_{5M} \xrightarrow{d} 0$ as $M \to \infty$. A similar lower bound can also be obtained. For the denominator of $KPSS_2$, note that

$$|\sum_{t=1}^{n}\sum_{s=1}^{n}(z_{nt} - \bar{z})(z_{ns} - \bar{z})k((t-s)/\gamma_n) - \sum_{t=1}^{n}\sum_{s=1}^{n} z_{nt}z_{ns}k((t-s)/\gamma_n)| = o_p(1)$$

because

$$\sum_{t=1}^{n}\sum_{s=1}^{n} k((t-s)/\gamma_n)z_{nt}\bar{z} \leq 2n^{-1}\sum_{j=0}^{n} k(j/\gamma_n)(\sum_{t=1}^{n} z_{nt})^2 = O_p(\gamma_n/n) = o_p(1)$$

and

$$\sum_{t=1}^{n}\sum_{s=1}^{n} k((t-s)/\gamma_n)(\bar{z})^2 \leq n^{-2}\sum_{t=1}^{n}\sum_{s=1}^{n} k((t-s)/\gamma_n)(\sum_{t=1}^{n} z_{nt})^2 = O_p(\gamma_n/n) = o_p(1).$$

The same reasoning as applied to $KPSS_1$ now gives

$$KPSS_2 \xrightarrow{d} \tau^2 - \tau + (1/3),$$

which completes the proof. □

Appendix 2: Simulation results

Table 2: Simulated quantiles and mean of the distribution of $Q_n(1)$; ε_t standardnormal

n	0.05	0.10	0.25	0.5	0.75	0.90	0.95	mean
50	1.74	2.23	3.17	4.47	6.13	7.93	9.12	4.82
100	1.99	2.50	3.50	4.96	6.88	9.03	10.5	5.43
500	2.40	2.93	4.03	5.71	7.98	10.6	12.5	6.34
1000	2.52	3.05	4.17	5.90	8.24	11.0	12.9	6.56
5000	2.71	3.24	4.37	6.15	8.60	11.5	13.5	6.87
10000	2.76	3.28	4.43	6.22	8.69	11.6	13.7	6.95
50000	2.82	3.34	4.50	6.30	8.80	11.7	13.8	7.04
100000	2.83	3.36	4.51	6.32	8.84	11.8	13.9	7.08

Table 3: Simulated quantiles and mean of the distribution of $Q_n(2)$; ε_t standardnormal

n	0.05	0.10	0.25	0.5	0.75	0.90	0.95	mean
50	0.98	1.20	1.60	2.24	3.10	4.08	4.76	2.45
100	1.09	1.28	1.69	2.37	3.28	4.35	5.09	2.62
500	1.22	1.40	1.84	2.55	3.54	4.70	5.51	2.84
1000	1.25	1.43	1.87	2.59	3.60	4.77	5.61	2.90
5000	1.30	1.48	1.92	2.65	3.68	4.88	5.74	2.97
10000	1.31	1.48	1.94	2.67	3.69	4.90	5.77	2.99

Table 4: Simulated quantiles and mean of the distribution of $Q_n(5)$; ε_t standardnormal

n	0.05	0.10	0.25	0.5	0.75	0.90	0.95	mean
50	0.18	1.00	1.04	1.23	1.68	2.18	2.54	1.38
100	0.82	1.00	1.05	1.27	1.74	2.24	2.62	1.44
500	1.00	1.01	1.07	1.32	1.80	2.33	2.71	1.51
1000	1.00	1.01	1.08	1.33	1.82	2.35	2.74	1.52
5000	1.00	1.02	1.09	1.35	1.84	2.37	2.77	1.55
10000	1.01	1.02	1.09	1.35	1.84	2.37	2.77	1.55

Table 5: Simulated quantiles and mean of the distribution of $Q_n(10)$; ε_t standardnormal

n	0.05	0.10	0.25	0.5	0.75	0.90	0.95	mean
50	0.02	1.00	1.00	1.03	1.24	1.67	1.92	1.12
100	0.37	1.00	1.00	1.04	1.27	1.71	1.95	1.16
500	1.00	1.00	1.00	1.05	1.31	1.75	1.99	1.20
1000	1.00	1.00	1.00	1.05	1.32	1.76	1.99	1.22
5000	1.00	1.00	1.00	1.06	1.33	1.78	2.01	1.23
10000	1.00	1.00	1.00	1.06	1.34	1.78	2.01	1.23

Table 6: Simulated quantiles and mean of the distribution of $Q_n(1)$; ε_t uniform

n	0.05	0.10	0.25	0.5	0.75	0.90	0.95	mean
50	1.74	2.19	3.03	4.23	5.76	7.42	8.53	4.56
100	1.97	2.44	3.35	4.69	6.48	8.47	9.82	5.13
500	2.38	2.85	3.85	5.40	7.51	9.97	11.7	6.00
1000	2.50	2.97	3.99	5.58	7.77	10.3	12.1	6.22
5000	2.68	3.14	4.18	5.83	8.11	10.8	12.7	6.51
10000	2.72	3.18	4.23	5.88	8.20	10.9	12.8	6.58

Table 7: Simulated quantiles and mean of the distribution of $Q_n(2)$; ε_t uniform

n	0.05	0.10	0.25	0.5	0.75	0.90	0.95	mean
50	0.86	1.18	1.51	2.08	2.85	3.74	4.36	2.27
100	1.11	1.25	1.60	2.20	3.03	3.99	4.67	2.43
500	1.22	1.36	1.73	2.37	3.26	4.32	5.07	2.64
1000	1.25	1.39	1.76	2.41	3.32	4.40	5.16	2.69
5000	1.29	1.43	1.80	2.46	3.39	4.49	5.27	2.76
10000	1.29	1.43	1.81	2.48	3.41	4.52	5.30	2.78

Table 8: Simulated quantiles and mean of the distribution of $Q_n(5)$; ε_t uniform

n	0.05	0.10	0.25	0.5	0.75	0.90	0.95	mean
50	0.14	1.00	1.02	1.17	1.58	2.04	2.38	1.32
100	0.69	1.00	1.03	1.20	1.63	2.10	2.45	1.36
500	1.00	1.01	1.05	1.25	1.69	2.18	2.54	1.43
1000	1.00	1.01	1.06	1.26	1.71	2.20	2.57	1.45
5000	1.00	1.01	1.06	1.27	1.73	2.22	2.59	1.47
10000	1.00	1.01	1.06	1.27	1.73	2.23	2.60	1.48

Table 9: Simulated quantiles and mean of the distribution of $Q_n(10)$; ε_t uniform

n	0.05	0.10	0.25	0.5	0.75	0.90	0.95	mean
50	0.01	1.00	1.00	1.02	1.18	1.60	1.86	1.10
100	1.00	1.00	1.00	1.02	1.21	1.64	1.89	1.13
500	1.00	1.00	1.00	1.03	1.25	1.68	1.95	1.17
1000	1.00	1.00	1.00	1.03	1.26	1.69	1.93	1.19
5000	1.00	1.00	1.00	1.04	1.27	1.71	1.95	1.20
10000	1.00	1.00	1.00	1.04	1.27	1.71	1.95	1.20

Table 10: Simulated quantiles and mean of the distribution of $Q_n(1)$; ε_t Rademacher

n	0.05	0.10	0.25	0.5	0.75	0.90	0.95	mean
50	1.72	2.19	2.89	4.06	5.72	7.69	9.09	4.56
100	2.00	2.40	3.19	4.53	6.46	8.83	10.5	5.16
500	2.39	2.77	3.69	5.24	7.56	10.5	12.6	6.06
1000	2.49	2.87	3.81	5.42	7.82	10.8	13.1	6.29
5000	2.64	3.02	4.00	5.68	8.19	11.4	13.7	6.60
10000	2.68	3.05	4.04	5.73	8.28	11.5	13.9	6.67

Table 11: Simulated quantiles and mean of the distribution of $Q_n(2)$; ε_t Rademacher

n	0.05	0.10	0.25	0.5	0.75	0.90	0.95	mean
50	0.63	1.29	1.41	1.82	2.94	4.27	5.31	2.38
100	1.16	1.32	1.49	1.98	3.14	4.71	5.77	2.57
500	1.32	1.35	1.55	2.30	3.55	5.19	6.41	2.82
1000	1.32	1.36	1.57	2.40	3.62	5.28	6.54	2.89
5000	1.34	1.38	1.61	2.47	3.75	5.42	6.71	2.97
10000	1.34	1.39	1.62	2.48	3.77	5.45	6.75	2.99

Table 12: Simulated quantiles and mean of the distribution of $Q_n(5)$; ε_t Rademacher

n	0.05	0.10	0.25	0.5	0.75	0.90	0.95	mean
50	0.02	1.01	1.01	1.03	2.04	3.05	4.05	1.70
100	1.01	1.01	1.02	1.04	2.05	3.07	4.07	1.80
500	1.01	1.01	1.02	1.05	2.06	4.05	5.05	1.92
1000	1.01	1.01	1.02	1.05	2.07	4.05	5.06	1.96
5000	1.01	1.01	1.02	1.07	2.08	4.05	5.07	2.01
10000	1.01	1.01	1.02	1.07	2.08	4.05	5.07	2.02

Table 13: Simulated quantiles and mean of the distribution of $Q_n(10)$; ε_t Rademacher

n	0.05	0.10	0.25	0.5	0.75	0.90	0.95	mean
50	0.00	1.00	1.00	1.00	2.00	3.00	4.00	1.67
100	1.00	1.00	1.00	1.00	2.00	3.00	4.00	1.76
500	1.00	1.00	1.00	1.00	2.00	4.00	5.00	1.89
1000	1.00	1.00	1.00	1.00	2.00	4.00	5.00	1.93
5000	1.00	1.00	1.00	1.00	2.00	4.00	5.00	1.96
10000	1.00	1.00	1.00	1.00	2.00	4.00	5.00	1.98