

Jönsson, Kristian

Working Paper

Finite-Sample Stability of the KPSS Test

Working Paper, No. 2006:23

Provided in Cooperation with:

Department of Economics, School of Economics and Management, Lund University

Suggested Citation: Jönsson, Kristian (2006) : Finite-Sample Stability of the KPSS Test, Working Paper, No. 2006:23, Lund University, School of Economics and Management, Department of Economics, Lund

This Version is available at:

<https://hdl.handle.net/10419/259936>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Finite-Sample Stability of the KPSS Test

Kristian Jönsson^{*†}

December 14, 2006

Abstract

In the current paper, the finite-sample stability of various implementations of the KPSS test is studied. The implementations considered differ in how the so-called long-run variance is estimated under the null hypothesis. More specifically, the effects that the choice of kernel, the value of the bandwidth parameter and the application of a prewhitening filter have on the KPSS test are investigated. It is found that the finite-sample distribution KPSS test statistic can be very unstable when the Quadratic Spectral kernel is used and/or a prewhitening filter is applied. The instability manifests itself through making the small-sample distribution of the test statistic sensitive to the specific process that generates the data under the null hypothesis. This in turn implies that the size of the test can be hard to control. For the cases investigated in the current paper, it turns out that using the Bartlett kernel in the long-run variance estimation renders the most stable test. By supplying an empirical application, we illustrate the adverse effects that can occur when care is not taken in choosing what test implementation to employ when testing for stationarity in small-sample situations.

JEL Classification: C12; C13; C14; C15; C22; E21.

Keywords: Stationarity; Unit root; KPSS test; Size distortion; Long-run variance; Monte Carlo simulation; Private consumption; Permanent Income Hypothesis.

^{*}Department of Economics, Lund University, P.O. Box 7082, SE-22007 Lund, Sweden.
Email: kristian.jonsson@nek.lu.se Tel: +46 (0)46 2224970, fax: +46 (0)46 2224118.

[†]The author would like to thank David Edgerton, Thomas Elger, Joakim Westerlund for discussions on the topics covered in, and relating to, this paper. Financial support from *The Crafoord Foundation* and *The Jan Wallander and Tom Hedelius Foundation*, research grant number P2005-0117:1, is gratefully acknowledged.

1 Introduction

Following the contribution of Kwiatkowski et al. (1992), testing for stationarity has become an integral part of time-series econometrics. The test, often denoted the KPSS test, that was suggested by these authors tests the null hypothesis that a time series is level or trend stationary, i.e. integrated of order zero, $I(0)$, against the alternative hypothesis that the series is integrated of order one, i.e. is $I(1)$. The formulations of these hypotheses stands in contrast to the unit root test of Dickey and Fuller (1979), where the null and alternative hypothesis are reversed so that the series considered is $I(1)$ under the null, while it under the alternative is $I(0)$. Kwiatkowski et al. (1992) argue that it can be of interest to test both types of hypotheses when investigating the dynamic properties of a time series. Asymptotic distributions for the test statistics that take level and trend stationarity as the null hypothesis are derived by Kwiatkowski et al. (1992) and critical values are supplied in the article. If the covariance stationary stochastic component in the considered time series is allowed to follow a linear process, such as in Phillips and Solo (1992), an estimator for the long-run variance generally has to be employed in order to eliminate the nuisance parameters that serial dependence otherwise would bring into the asymptotic distribution of the test statistic. Most commonly, a class of non-parametric or semi-parametric long-run variance estimators is considered for this purpose. The use of an estimator from this class entails a choice of kernel, a selection of a bandwidth parameter and a decision regarding whether or not to employ a prewhitening filter in the variance estimation process. Provided that the bandwidth parameter is chosen appropriately with respect to the kernel and the sample size, the limiting distribution that applies for the KPSS test statistic under the null hypothesis is invariant to choices regarding the kernel, the bandwidth parameter and prewhitening. However, such invariance cannot be established analytically in small-sample situations.

Considering that many time series, especially annual macroeconomic time series, often are very limited when it comes to sample size, it is important to obtain knowledge about the small-sample properties of various estimators and tests. This applies also for the KPSS test. Since the finite-sample distribution of the latter test can be sensitive to the choice of kernel, the choice of bandwidth parameter and whether or not prewhitening is used, knowledge regarding the test's behavior along these three dimensions is important to obtain. If the test is shown to be highly dependent upon the choices made, care must be taken when performing the KPSS test in small samples, while finite-sample stability would ensure that test results are not as contingent upon the choices made when implementing the test. However, while several studies have documented the behavior of the KPSS test in fairly large samples (see e.g. Caner and Kilian, 2001; Hobijn et al., 2004; Sul et al., 2005; Carrion-i-Silvestre and Sansó, 2006), little evidence exists about

the small-sample stability of the KPSS test.

The current paper investigates the effects that the choice of kernel, the choice of bandwidth parameter and the choice of whether or not to apply a prewhitening filter have on the KPSS test in small samples. Two commonly used kernels are considered, the Quadratic Spectral (QS) kernel and the Bartlett kernel. In addition to studying the choice of kernel, it is investigated how the bandwidth parameter affects the small-sample distribution of the test statistic. As a final step, the paper investigates if prewhitening has any small-sample effects on the test distribution when the various kernels are used.

The main results indicate that using the Bartlett kernel, rather than the QS kernel, in the KPSS test can render a more stable test statistic in small-sample situations. More precisely, the distribution of the KPSS test can be less sensitive to the exact data generating process (DGP) under the null hypothesis when the Bartlett kernel is used, which makes it easier to control the size of the test. It is also found that the bandwidth parameter affects the small-sample distribution of the test regardless of which kernel that is used. However, the influence of the bandwidth parameter is smaller when the Bartlett kernel is considered than when the QS kernel is employed. Finally, prewhitening is found to be detrimental to the small-sample stability of the KPSS test, regardless of the choice of kernel. More specifically, prewhitening increases the test's sensitivity to various DGPs under the null hypothesis. The results indicate that it can be good practice to use the Bartlett kernel, and utilize the finite-sample critical values supplied by Hornok and Larsson (2000) and Jönsson (2006), when performing the KPSS test for level or trend stationarity in situations where samples are moderately sized.

To illustrate the main points of the paper, the KPSS test is employed in order to test the stationarity of private consumption for 22 OECD countries over the period 1960-2004. The main results from investigating these series are that the outcome of the test is highly dependent on which implementation of the Kwiatkowski et al. (1992) test that is used and that the finite-sample stability of the test is important to consider before performing the test in small-sample situations.

The remainder of this paper is organized as follows. In Section 2, the econometric model is introduced together with the Kwiatkowski et al. (1992) test for stationarity. Furthermore, some central issues in the test implementation are discussed in this section. Section 3, presents the main results concerning the finite-sample stability of the stationarity test, while Section 4 presents an empirical illustration. Finally, Section 5 offers some concluding remarks.

2 Testing for stationarity

2.1 The time-series model

The test for level or trend stationarity that was suggested by Kwiatkowski et al. (1992) takes the following unobserved components time-series model as point of departure:

$$y_t = \alpha_0 + \alpha_1 t + \xi_t + \varepsilon_t \quad (1)$$

$$\xi_t = \xi_{t-1} + \eta_t \quad (2)$$

The time-series y_t in (1) has three different components. First, α_0 and $\alpha_1 t$ are deterministic components in the form of a constant term and a linear time trend. Secondly, ξ_t is either a random walk component or a constant term, depending on whether the variance of η_t , denoted σ_η^2 , takes on a non-zero value or not.¹ Finally, ε_t is a disturbance term that in the current context is assumed to fulfill $\varepsilon_t = \Theta(L)\nu_t$, where $\Theta(L) = \sum_{i=0}^{\infty} \theta_i L^i$ such that $0 < \Theta(1) < \infty$ and $\sum_{i=1}^{\infty} i|\theta_i| < \infty$, while ν_t is independently and identically distributed with zero mean and $0 < \sigma_\nu^2 < \infty$, where $\sigma_\nu^2 = E(\nu_1^2)$. The assumptions on the error process amounts to assuming that the disturbance follow a linear process, as in e.g. Phillips and Solo (1992), with summable coefficients and homoscedastic disturbances in the infinite moving average representation. Under these assumptions, the test of Kwiatkowski et al. (1992) can be used to test whether or not y_t is I(0). The test is presented in the next subsection.

2.2 The KPSS test

In order to test the null hypothesis that y_t is covariance stationary, or I(0), Kwiatkowski et al. (1992) suggest that it should be tested whether or not σ_η^2 is equal to zero by employing the test statistic in (3).

$$LM = \frac{T^{-2} \sum_{t=1}^T S_t^2}{\hat{\sigma}^2} \quad (3)$$

In (3), S_t is the partial sum process $S_t = \sum_{i=1}^t e_i$, where e_t is the least squares residual obtained after detrending the time series y_t with either an intercept or an intercept and a time trend. Furthermore, $\hat{\sigma}^2$ in the denominator of (3) is a variance estimator that is intended to relieve the asymptotic distribution of the LM statistic from nuisance parameters under the null hypothesis. In order to see why such an estimator is necessary, it can be noted that the numerator of (3) weakly converges, as $T \rightarrow \infty$, either to

¹In the case where the variance equals zero, the random walk term will take on the value ξ_0 for all t , which will alter the constant term of y_t to $\alpha_0 + \xi_0$.

a scaled standard Brownian bridge, $\sigma^2 \int_0^1 V(r)^2 dr$, or a scaled second-level Brownian bridge, $\sigma^2 \int_0^1 V_2(r)^2 dr$, with the limiting behavior depending on whether and intercept, or an intercept and a time trend, are used to detrend y_t . The scaling parameter is given by $\sigma^2 = \sum_{i=-\infty}^{\infty} E(\varepsilon_t \varepsilon_{t-i})$, from which it is clear that σ^2 is affected by the serial dependence in ε_t . In order to relieve the asymptotic distribution from this nuisance parameter, often called the long-run variance of ε_t , a long-run variance estimator has to be employed. This issue is discussed next.

2.3 Long-run variance estimation

When calculating the test statistic presented in the previous subsection, an estimator for the long-run variance, $\sigma^2 = \sum_{i=-\infty}^{\infty} E(\varepsilon_t \varepsilon_{t-i})$, has to be chosen. Choosing an estimator within the non-parametric framework that was originally suggested by Kwiatkowski et al. (1992) entails a choice of a kernel, $w(i, l)$, and a choice of a bandwidth parameter, l .² Together, $w(i, l)$ and l give a scheme that describes how the i^{th} sample autocovariance, $\hat{\gamma}_i$, should be weighted in the estimation of the long-run variance. The general form for a long-run variance estimator from the non-parametric framework is given in (4) below.

$$\hat{\sigma}^2 = \hat{\gamma}_0 + \sum_{i=1}^{T-1} w(i, l) \hat{\gamma}_i \quad (4)$$

In their original article, Kwiatkowski et al. (1992) suggest that the long-run variance estimator of Newey and West (1987) should be used, which implies that the Bartlett kernel should be employed in the KPSS test. However, in practice, two kernels have become frequently used when performing the KPSS test. Besides the Bartlett kernel, the Quadratic Spectral kernel is often considered (see e.g. Hobijn et al., 2004; Sul et al., 2005). The two kernels are presented in (5) and (6) below.

$$\text{Bartlett:} \quad w(i, l) = \begin{cases} 1 - \frac{i}{l+1} & \text{if } i \leq l \\ 0 & \text{if } i > l \end{cases} \quad (5)$$

$$\text{Quadratic Spectral:} \quad w(i, l) = \frac{25}{12\pi^2(i/l)^2} \left[\frac{\sin(6\pi(i/l)/5)}{6\pi(i/l)/5} - \cos(6\pi(i/l)/5) \right] \quad (6)$$

While both render the same asymptotic distribution of the KPSS test, there are two main differences between the kernels presented above. The first difference is the number of sample autocovariances that are used in the estimation of the long-run variance. The

²It can be noted that, in addition to the non-parametric approach taken by Kwiatkowski et al. (1992), Leybourne and McCabe (1994, 1999), Saikkonen and Luukkonen (1993) and Lanne and Saikkonen (2003) have considered parametric versions of tests that takes covariance stationarity as the null hypothesis.

Quadratic Spectral kernel assigns a non-zero weight to every autocovariance that can be calculated from the sample. The Bartlett kernel, on the other hand, uses only some of the estimated autocovariances to estimate the long-run variance, more specifically only the first m autocovariances are used, where m is the largest integer such that $m \leq l$. To the other autocovariances that can be calculated from the sample at hand, i.e. to $\hat{\gamma}_i$ when $m < i < T$, the Bartlett kernel assigns a zero weight and hence neglects them in the long-run variance estimation.

The second difference between the kernels refers to how they are affected by a change in the bandwidth parameter. When l is changed for the Quadratic spectral kernel, the weighting scheme is altered for all the autocovariances. This stands in contrast to the Bartlett kernel, for which a change in the bandwidth parameter alters the weighting for a limited number of autocovariances and changes the number of autocovariances considered in the estimation only when the integer part of the bandwidth parameter changes. For the latter reason, the Bartlett kernel is often used in conjunction with integer-valued bandwidth parameters (see Kwiatkowski et al., 1992, p. 169), while the same does not apply for the QS kernel (see e.g. Hobijn et al., 2004), where instead any positive real value can be used.

The fact that the bandwidth parameter can take on any real positive value for the QS kernel, while it most commonly only takes on positive integer values for the Bartlett kernel, may be looked upon as a minor difference since both kernels can be used to obtain an asymptotically pivotal test statistic for the KPSS test. However, because the latter kernel most often only takes on integer values, it is possible to easily approximate the small-sample distribution of the LM statistic that applies this kernel by conditioning the distribution of the test statistic on the bandwidth parameter and the sample size, as done by e.g. Hornok and Larsson (2000) and Jönsson (2006). These authors argue that the KPSS test that employs the Bartlett kernel can be more sensitive to the choice of bandwidth than to the actual serial correlation in the disturbances, given that a sufficient bandwidth is chosen. Hence, the small-sample properties of the KPSS test can be enhanced in the same way as the properties of the Augmented Dickey-Fuller (ADF) test are enhanced by taking the number of augmentation terms into account when performing the test in small samples (see Cheung and Lai, 1995, for a treatment of the ADF test). The feasibility of such a small-sample improvement to the performance of the KPSS test have implications for the choice of kernel. If the KPSS test in general is sensitive to the choice of bandwidth parameter when samples are small and the QS kernel is used, the Bartlett kernel would be preferable to use since conditional distributions can be obtained more easily and since small-sample critical values are readily available for the implementation that considers this choice of kernel. However, if the small-sample distribution of the KPSS test statistic is insensitive to the bandwidth parameter when

the QS kernel is used, this kernel would be preferable since no small-sample critical values would be needed. Hence, to give guidance regarding the choice of kernel in small-sample situations, the finite-sample stability of the KPSS test will be investigated in the next section.

Besides using purely non-parametric estimators of the long-run variance, a semi-parametric estimator, that applies prewhitening, can be used. Employing a prewhitening filter in the estimation of the long-run variance can improve the properties of the estimator when there is serial dependence in ε_t . The prewhitening procedure is performed by fitting an ARMA(p,q) to the detrended series, applying the non-parametric long-run variance estimator to the ARMA-filtered residuals and then calculating the long-run variance based on the non-parametric variance estimate and the estimated ARMA coefficients. For convenient calculations, an AR(p), or even an AR(1), model is often used for prewhitening. The reason for using the non-parametric long-run variance estimator in the prewhitening procedure is that prewhitening is often performed, not to actually obtain white noise disturbances after having filtered the detrended series, but to obtain test statistics with enhanced properties in the presence of serial dependence in ε_t (see e.g. den Haan and Levin, 1997). Furthermore, even if serially uncorrelated disturbances are achieved through prewhitening the detrended series, the use of a long-run variance estimator can still be required in order to capture heteroscedasticity in the disturbance variance (see e.g. Carrion-i-Silvestre and Sansó, 2006). To avoid that the stationarity test becomes inconsistent and suffer from the power loss documented by Lee (1996), a boundary rule for the estimated ARMA coefficients has been suggested by Sul et al. (2005).

Since the prewhitening procedure does not affect the asymptotic distribution of the test under the null hypothesis, the critical values supplied by Kwiatkowski et al. (1992) can be used when samples are large. However, very little knowledge exists regarding the effects of prewhitening in small samples. Hence, it is important to consider the finite-sample stability of the KPSS test also when prewhitening is used.

In order to obtain more knowledge about the finite-sample robustness of the various implementations of the KPSS test, we set up and run Monte Carlo simulations. More specifically, we will study how the different test implementations perform under data-generating processes where various degrees of serial dependence are considered.

3 Finite-sample robustness

In this section, the properties of the KPSS test are investigated in a setting where samples are small or medium-sized, i.e. in a setting where it is not evident that the asymptotic

distributions can be used to describe the behavior of the KPSS test statistics. First, it is investigated how the choice of kernel and choice of bandwidth parameter affect the distribution of the KPSS test when serial correlation is absent. After that, the effects of prewhitening are studied. The robustness of the KPSS test under various degrees of serial correlation are considered next. Finally, the finite-sample size properties of the different implementations of the KPSS test are investigated under varying degrees of serial dependence. Since the finite-sample properties of the test are hard, if at all possible, to derive analytically even under stringent assumptions regarding the specific distribution of stochastic components, we resort to Monte Carlo simulations to investigate the behavior of the test.

3.1 Bandwidth dependence in small samples

As mentioned in Section 2.3, it can be of importance to consider the value of the bandwidth parameter when performing the KPSS test in finite-sample situations. This is especially important in situations where the small-sample distribution of the KPSS test is sensitive to the choice of bandwidth parameter. Under such circumstances it is also desirable to use a kernel for which the bandwidth parameter takes on only a finite number of values, since if this is the case, it is possible to easily condition the critical values on the choice of bandwidth parameter as well as on the sample size. For the situation where the distribution of the KPSS test exhibits no bandwidth dependence in small samples, it is of less importance to obtain finite-sample critical values, since the asymptotic critical values can be successfully used for inference about the stationarity hypothesis. Hence, as a first step, it is important to study to what extent the small-sample distribution of the KPSS test statistic depends on the bandwidth parameter for the various kernels that can be used in the test.

To illustrate why knowledge about bandwidth-dependence of the KPSS test statistic is vital, suppose that data is generated without any serial correlation in the disturbance term ε_t .³ If the small-sample distribution of the KPSS test statistic is altered by the bandwidth choice, then one set of critical values, such as the asymptotic critical values supplied by Kwiatkowski et al. (1992), would not suffice for stable inference since the test outcome under such circumstances very well could depend on the bandwidth parameter chosen when estimating the long-run variance. Instead, a minimum requirement for a robust test would be that critical values are available for each possible choice of

³Obviously, this situation makes it unnecessary to use a long-run variance estimator, the estimator in $\hat{\sigma}_\varepsilon^2 = \frac{\sum_{t=1}^T e_t^2}{T}$ would suffice. However, since the purpose here is to illustrate the possible bandwidth dependence of the test, long-run variance estimators will still be applied.

bandwidth parameter. However, the availability of such critical values would not be a sufficient condition for a stable stationarity test. Instead, an additional assumption that would ensure stability of the KPSS test across test implementations and across various DGPs under the null hypothesis would be that the distribution of the test statistic is relatively unaffected by various degrees of serial dependence under the null hypothesis once the time-series dimension and the bandwidth parameter have been taken into account. Hence, it is of interest to study two separate questions, the first being the issue regarding bandwidth dependence and the second being the robustness across DGPs given the the bandwidth parameter and the sample size. The former issue is probed next, while we later in the current paper will address the latter issue.

Figure 1 and Figure 2 show how the 10% finite-sample critical values of the KPSS test are affected by various bandwidth choices and time-series dimensions. The critical values in these figures are obtained by generating 25,000 different time series, y_t , with deterministic components, α_0 and $\alpha_1 t$, and stochastic disturbances, ϵ_t . More specifically, data is generated according to $y_t = \alpha_0 + \alpha_1 t + \epsilon_t$, where α_0 is distributed uniformly over the interval $[0,10]$, denoted $\alpha_0 \in U[0, 10]$, while $\alpha_1 = 0$ if only an intercept is considered and $\alpha_1 \in U[0, 2]$ if both an intercept and a linear time trend are present. The stochastic disturbance term is generated according to $\epsilon_t \sim i.i.d. N(0, 1)$. Time series are generated for the sample sizes $T \in \{20, 30, 40, 50\}$, while the bandwidth parameter is set as in (7) and (8), which is tantamount to the choice in e.g. Hobijn et al. (2004), with the bandwidth scaling parameter $k \in \{4, 8, 12\}$.^{4,5}

$$\text{Bartlett:} \quad l = \left[k \left(\frac{T-1}{100} \right)^{1/4} \right] \quad (7)$$

$$\text{Quadratic Spectral:} \quad l = \left[\frac{2k}{3} \left(\frac{T-1}{100} \right)^{2/9} \right] \quad (8)$$

When looking at Figure 1 and Figure 2, it is evident that the critical values for the KPSS test depend on both the sample size, T , and the bandwidth scaling parameter, k , regardless of which kernel that is used in the test and regardless of which deterministic components that are present. However, some distinct differences exist when comparing the behavior of the test statistic under various configurations of kernels and trend components.

⁴In (7) and (8), $[a]$ denotes the integer part of a .

⁵The expressions in (7) and (8) use $T - 1$ in the numerator since, later on in this section, one time series observation is lost when prewhitening the detrended series. Hence, in order to assure comparability across the various implementations, the bandwidth parameter is calculated based on $T - 1$ observations even when prewhitening is not used.

The first thing that can be seen when comparing Figure 1 and Figure 2 is that the critical values are more dependent on T and k when both an intercept and a time trend are accounted for than when only an intercept is present. The same result was obtained by Jönsson (2006) for the KPSS test that employs the Bartlett kernel. The current result indicates that this feature also generalizes also to the case where the QS kernel is employed. Since both kernels exhibit sensitivity to the sample size and the bandwidth scaling parameter, there is no reason to prefer one over the other when it comes to the small-sample dependence of the test statistic with respect to these factors.

The second distinct feature that can be noticed when comparing the upper and lower panels of Figure 1 and Figure 2, respectively, is that the use of the Quadratic Spectral kernel, rather than the Bartlett kernel, renders a test statistic that is more sensitive to the sample size and bandwidth scaling parameter. This applies regardless of whether a constant or a constant and a time trend are considered in the test. From this perspective, failure to account for the finite-sample distribution of the KPSS test statistic is more detrimental to test performance when the QS kernel is employed than when the Bartlett kernel is used. Moreover, considering the fact that the QS kernel is a continuous function of a real variable that is not necessarily integer-valued, the use of the Bartlett kernel in small and medium-sized samples seems preferable, at least when the use of small-sample critical values are considered.

Having established that there is a finite-sample bandwidth dependence in the distribution of the KPSS test statistic regardless of whether the Bartlett kernel or the QS kernel is used in the test, the robustness of the implementations of the test, with respect to varying degrees of serial dependence, becomes relevant. This issue will be addressed below, but first, in the next subsection, the issue regarding the effects of applying a prewhitening filter will be investigated.

3.2 Effects of prewhitening

When estimating the long-run variance, σ^2 , it is possible to first filter the detrended series, e_t , through e.g. an ARMA(p,q) model in order to get better estimates of the long-run variance. However, since AR(p) models, such as the one in (9), are very simple to estimate they are often preferred over ARMA models. If \hat{u}_t denotes the AR-filtered residual, as in (10), the long-run variance estimator that relies on prewhitening can be calculated as in (9)-(11), where $\hat{\sigma}_u^2$ is the long-run variance estimate for the filtered series \hat{u}_t .

$$e_t - \phi_1 e_{t-1} - \dots - \phi_p e_{t-p} = \Phi_p(L)e_t = u_t \quad (9)$$

$$\hat{\Phi}(L)e_t = \hat{u}_t \quad (10)$$

$$\hat{\sigma}^2 = \frac{\hat{\sigma}_u^2}{\left(\hat{\Phi}_p(1)\right)^2} \quad (11)$$

In order to assure that the test implementation that utilizes prewhitening can be used also under the alternative hypothesis, where a unit root is present in y_t , boundary conditions are often employed for the lag polynomial. For example, Carrion-i-Silvestre and Sansó (2006) suggest, following Sul et al. (2005), that the boundary condition should be set in such a way that $\hat{\Phi}_p(1)$ does not exceed $1 - 1/\sqrt{T}$. This boundary rule is also used in the following simulations.

If the process of prewhitening renders a long-run variance estimator that makes the KPSS test more stable than the results of the previous subsections indicated, the use of small-sample critical values would become less urgent. To investigate if this is the case, data is once again generated as described above and the KPSS test is once again applied to the generated data but now with a prewhitening filter being employed in the long-run variance estimation. The 10% critical values are presented in Figure 3 and Figure 4.

As seen from Figure 3 and Figure 4, the use of a prewhitening filter in the long-run variance estimation renders a test statistic that is more sensitive to the sample size and the bandwidth parameter compared to the case where no prewhitening is applied. The same result is obtained regardless of whether the Bartlett kernel or the QS kernel is used. Moreover, it can also be seen from the figures that the use of the QS kernel makes the critical values more dependent on T and k than what is the case when the Bartlett kernel is used. Hence, applying prewhitening does not affect the choice of kernel when sample-size stability and bandwidth stability are considered to be desirable properties of the test. However, since it is easy to condition the small-sample distribution of the KPSS test on the sample size and the bandwidth when the Bartlett kernel is used, the prewhitening procedure can still be of interest if the KPSS test distribution is stable with regard to various degrees of serial dependence in data generating process under the null hypothesis of stationarity. The issue regarding stability of the different implementations of the KPSS test across various degrees of serial correlation in the DGP is investigated in the next subsection.

3.3 Robustness in the presence of serial dependence

One of the main advantages of conditioning the critical values used in the KPSS test on the sample size and bandwidth parameter is that when these two factors are controlled

for, the size properties of the test can be improved considerably across various data generating processes under the null hypothesis (see e.g. Hornok and Larsson, 2000; Jönsson, 2006). If a similar stability is present when prewhitening is used in the long-run variance estimation, critical values could in addition to being conditioned on T and k , also be conditioned on the use of a prewhitening filter. That is, if using the Bartlett kernel in conjunction with a prewhitening filter does not affect the distribution of the test statistic under varying serial correlation patterns, prewhitening could perhaps improve the size properties of the stationarity test. But if this implementation of the test becomes highly dependent on the specific form of serial dependence under the null hypothesis, it would be intractable to use prewhitening in small-sample situations. For this reason, we will study how the distribution of the KPSS test statistic is affected by various degrees of serial correlation both when prewhitening is used and when it is not.

To see whether the LM statistic is sensitive to the actual data generating process under the null hypothesis, various degrees of autocorrelation are now allowed for in the stationary disturbance term. To make the situation as favorable as possible for the prewhitening filter, an AR(1) process is assumed, while an AR(1) prewhitening filter is employed. Data is hence generated as described in Section 3.1, but now with the modification that $\epsilon_t = \rho\epsilon_{t-1} + \vartheta_t$, with ϑ_t being *i.i.d.* $N(0, 1)$ and $\rho \in \{0.0, 0.2, 0.4, 0.6\}$. By altering ρ it is possible to see how the distribution of the KPSS test statistic is affected by various DGPs under the null hypothesis.

In Figure 5 and Figure 6, the kernels density estimates for the distribution of the KPSS test statistic are presented for various choices of ρ and for different combinations of kernels and deterministic components when no prewhitening is used. These graphs are constructed for two reasons. First, the effects that serial dependence has for the distribution of the test statistic that does not apply prewhitening can be analyzed from them. Secondly, they serve as a background that facilitate an analysis of the effects that prewhitening has on test stability. The density estimates are based on 25,000 test statistics generated as described above and estimated using the Triweight kernel with the smoothing parameter to 0.08.

The distribution of the test statistic when only an intercept is present is relatively stable across the various values of ρ , as seen in Figure 5. This applies regardless of whether the Bartlett kernel or the Quadratic kernel is considered. Hence, it seems like small-sample critical values could be obtained for either one of the kernels and testing could, by conditioning on the bandwidth parameter, be successfully performed using these critical values. However, when both an intercept and a time trend are present, the stability of the KPSS statistic breaks down when the QS kernel is used. Instead of retaining stability, the test statistic now becomes sensitive to the actual value of ρ . This, of course, makes the Quadratic Spectral kernel undesirable in small-sample situations

since the size performance of the test will vary with the (unobserved) DGP under the null hypothesis. However, the test based on the Bartlett kernel retains its stability, which is a feature that can make it an attractive candidate in finite-sample situations.

The next step is to investigate what effects prewhitening has on the distribution of the test statistics. To this end, the same DGPs are considered and the KPSS tests performed using a long-run variance estimator that employs an AR(1) prewhitening filter. The kernel density estimates for this simulation setup are presented in in Figure 7 and Figure 8.

As seen from Figure 7 and Figure 8, the distribution of the test statistic becomes very sensitive to the autoregressive parameter, ρ , when prewhitening is used. This applies both for the Bartlett kernel and the Quadratic Spectral kernel and regardless of whether an intercept or an intercept and a time trend are included in the model. Hence, it is not desirable to try to control the size of the KPSS test by using finite-sample critical values obtained for the case where $\rho = 0.0$, i.e. for the case where no serial correlation is present.

The results presented in this section indicate two things. First, if critical values are to be obtained by conditioning on the bandwidth parameter, the Bartlett kernel should be used in the test. The main reason for this is that the distribution of the KPSS test statistic is more robust to changes in the DGP under the null hypothesis when this kernel is used, at least when considering the DGPs investigated in the current paper. Moreover, the Bartlett kernel takes integer-valued arguments, which makes it easier to condition the distribution of the test statistic on the bandwidth. The second main result is that prewhitening can have detrimental effects in small samples. Hence, one way to maintain control of the size of the KPSS test in small samples could be to use finite-sample critical values obtained by applying the Bartlett kernel without prewhitening.

As a final note, it must be recognized that the computational suggestion of Carrion-i-Silvestre and Sansó (2006) still applies when samples become large. This is illustrated by kernel density estimates of the KPSS statistic in Figure 9 and Figure 10. The distributions in these figures are obtained for $T=100$, i.e. the smallest sample size considered by Carrion-i-Silvestre and Sansó (2006), and for the case where the Quadratic Spectral kernel is used. From the figures it is evident that the dependence on the DGP becomes less pronounced as the sample size increases. This applies for the case when prewhitening is used as well as for the case where no prewhitening is applied. The results presented in the current paper should hence be seen as a remainder that care should be taken when testing for stationarity in small-sample situations.

3.4 Effects on size

Having established that the finite-sample distribution of the KPSS test can be considerably more stable when the Bartlett kernel is used without prewhitening than when the same kernel is used in conjunction with prewhitening or when the QS kernel is used with or without prewhitening, it is interesting to see how the size of the KPSS test is affected by the instability. To do this, we consider the same DGP as in the proceeding subsections and study the size of the KPSS test for the cases where $T \in \{20, 30, 40, 50, 100, 250, 500\}$, while $k = 12$. As above, we consider the cases where $\rho \in \{0.0, 0.2, 0.4, 0.6\}$. In order to get a picture of the stability of the KPSS test, and show that it is important to ensure that the test implementation applied is stable, the finite-sample critical values for the situation where $\rho = 0.0$ are used to calculate the size of the test for all the implementations that are considered and for all of the various DGPs that are studied. The size properties, on the 10% significance level, for the various implementations of the test are presented in Table 1.

Before studying the size properties in Table 1 in greater detail, it can be interesting to turn back to the results from the previous subsections. The kernel density plots that were analyzed previously give clear indications of the qualitative properties that we can expect to find when the size is studied. First and foremost, for all cases where prewhitening is used, the size of the stationarity test is expected to be downward distorted for the case where serial correlation is present. Figure 7 and Figure 8 show that the distribution of the KPSS test statistic shifts leftward as the serial correlation increases when prewhitening is applied, regardless of which kernel that is used. Furthermore, looking at Figure 6, it can be seen that the same effect should be expected when a linear trend is allowed for even when no prewhitening is used. From this figure, it is also evident that the size distortions should be more severe for the test that applies the QS kernel than for the test that applies the Bartlett kernel. When only an intercept is accounted for in the series under consideration, Figure 5 indicates that only small size distortions should occur as the serial correlation increases. These figures indicate the direction of the size distortions that can be expected, but they say nothing, however, about the magnitude of the size distortions that will occur. To get a picture of the magnitude of the size distortions, Table 1 has to be consulted.

From the upper panel of Table 1, where the size is considered for a level stationary process, it can be seen that there are only small size distortions to the stationarity test that applies the Bartlett kernel without prewhitening. For the lower values of the autoregressive coefficient, i.e. for $\rho = 0.2$ and $\rho = 0.4$, the size distortions is about 2-3 percentage points. For the other test, i.e. for the tests that apply the QS kernel and/or prewhitening, the corresponding size distortions are 3-5 percentage points. Hence, there

are only small differences between the size performance of the various test implementations. When $\rho = 0.6$, the size distortion is about 5 percentage points regardless of what test implementation that is considered.

When attention is turned to the case where both an intercept and a linear trend is allowed for in the time series, the picture becomes somewhat different. The size distortions in the KPSS test that applies the Bartlett kernel without prewhitening remain at the same level as in the case where only an intercept was allowed for. However, when the other implementations of the test are considered, the size distortions become much more severe. When prewhitening is used, the size on the 10% significance level can drop as much as 9 percentage points. The direction of this size distortion was predicted from the figures above, but it now turns out that the magnitude is rather large. Also when no prewhitening is applied and the QS kernel is used, there are large size distortions to the test. These results support the conjecture that, from a stability point of view, the KPSS tests that applies the Bartlett kernel without prewhitening can be preferable to the implementation that rely on the QS kernel and/or prewhitening in the long-run variance estimation.

Before we continue with an empirical illustration that shows the effects that can occur when care is not taken when selecting an implementation of the KPSS test, a few remarks on the above results are in place. First, it can be noticed that the size results presented in Table 1 give an indication of the stability of the KPSS test. The results says nothing about the general performance of the various test implementations. One can easily imagine a situation where an upward size distortion eliminates the downward size distortions observed in Table 1, and as a consequence alters the relative attractiveness of the various test implementations. For example, if asymptotic critical values were to be used in the investigation of the size, it could be imagined that the size for the case $\rho = 0.0$ would be upward distorted. When the cases where $\rho > 0$ are considered, the downward size distortions of the tests that use the QS kernel and/or prewhitening could eliminate the upward size distortions for some specific ρ , and hence render a situation where these test implementations get good size properties. However, since the applied research never knows the DGP that has generated the data at hand, it would be highly unlikely to encounter a situation like this when applying the KPSS test, and it would be better to rely on stability across DGPs rather than to rely on pure luck.

Second, it should be emphasized that the relative performance of the tests can change if a bandwidth that does not capture the serial correlation good enough is used. If the bandwidth parameter is given a value that is too small, the Bartlett kernel that does not apply prewhitening could suffer from this choice since too many significant higher-order autocorrelations are neglected in the variance estimation. However, it is not necessarily the case that this deficiency changes the relative performance of this test implementation

compared to the implementations that utilize prewhitening. The reason for this is that if a misspecified prewhitening filter is applied in the test implementation that utilizes prewhitening, the test would also suffer from this.⁶ Moreover, the tests that apply prewhitening would be negatively affected through two channels. First, choosing a too narrow bandwidth leaves serial correlation in disturbances that would not be picked up by the long-run variance estimator. But a second channel, introduced by biased estimators for the parameters of the AR model as a consequence of autocorrelated disturbances, could also worsen the size properties. The overall effects could hence be ambiguous, depending on the specific case considered.

Finally, it can be noted that it might be that case that increases in test stability come at the cost of a lower power against an $I(1)$ alternative hypothesis. Stated differently, it is possible that a higher power can be bought at the price of less stable finite-sample distribution of the test under the null hypothesis. This, of course, highlights the ever-present tradeoff between size and power properties.

In order to further point out the relevance of the findings presented above, the next section supplies an empirical illustration.

4 Empirical illustration

As seen from the previous sections, the KPSS test can be very sensitive to the choice of kernel, the choice of bandwidth parameter and the choice of whether or not to use prewhitening in the long-run variance estimation process. To further stress the relevance of these findings, we supply an empirical illustration in this section.

As shown by Hall (1978), private consumption can follow a random walk if the permanent income hypothesis is true. If, on the other hand, private consumption is covariance stationary, then the permanent income model with a quadratic instantaneous utility function is not supported by the data. In order to test whether this is the case, the KPSS test can be employed. In this paper, we use the per capita private consumption data set of Jönsson (2006) to show how the lack of finite-sample stability of behalf of the KPSS test can cause problems when using the test in empirical applications.⁷

To test for stationarity in the private consumption series, the KPSS test is applied under various choices regarding the kernel used and the choice of whether or not to

⁶Remember that the prewhitening filter used in the current paper is correctly specified with respect to the DGP. Hence, the test implementations that apply prewhitening are given a slight advantage since the effects of model misspecification are eliminated from the simulations.

⁷The per capita private consumption series are gathered from the OECD Economic Outlook Database, Vol 2005 release 02, and contains data for 22 OECD countries.

prewhiten the detrended series previous to variance estimation. The bandwidth scaling parameter is set as $k = 12$, while an intercept and a time trend are used as deterministic components. The bandwidth is then calculated for the Bartlett kernel and the QS kernel as in (7) and (8) above. Furthermore, the KPSS test is calculated both for the case where prewhitening has been used and for the case where it has not been used. The test results are presented in Table 2.

From Table 2 it can be seen that the conclusion drawn, based on the various implementations of the KPSS test, varies to a large extent with the choices made when calculating the test statistic. However, some general patterns can be easily seen from the table.

Looking at the test results in Table 2, it can be noticed that the KPSS test that employs the Bartlett kernel without prewhitening rejects the null hypothesis of stationarity for 13 out of the 22 countries.⁸ The test that employs the QS kernel, but still does not utilize prewhitening in the variance estimation process, rejects the null hypothesis for 11 of the 22 countries. It can be seen that the the latter test fails to reject the null hypothesis for Austria and Japan, countries for which the KPSS test utilizing the Bartlett kernel reject the null. Finally, the tests that rely on prewhitening do not reject the null hypothesis of stationarity for any of the countries in the sample.

The test results in Table 2 can be interpreted in at least two ways. The first possible interpretation is that the tests that do not apply prewhitening has an upward size distortion that is larger than what is the case for the tests that apply prewhitening. The consequence, if this were to be the case, would be that the former test implementations reject the null hypothesis more often than what is warranted by the significance level. Hence, the test results would be an indication of the fact that applying a prewhitening filter in the variance estimation process enhances the performance of the KPSS test. However, the simulation evidence presented in the previous sections offers another explanation for the results.

The results from the simulations that are presented in the previous sections render two important conclusions. The first conclusion is that using the Bartlett kernel in the KPSS test can give rise to a test that is more stable than what is the case when the QS kernel is used. More specifically, using the QS kernel when a positive first-order serial correlation is present could alter the distribution of the test statistic in such a way that the test become downward size-distorted (see Figure 6 and Table 1). The second conclusion is that using prewhitening also can introduce a downward size distortion, but

⁸It should be noted that we have used asymptotic critical values for the stationarity test, rather than the finite-sample critical values of Hornok and Larsson (2000) or Jönsson (2006), in order to assure comparability across test implementations.

now regardless of what kernel that is used (see Figure 7, Figure 8 and Table 1). These results offer an reasonable explanation of the results in Table 2.

Turing back to the test results of Table 2, it can be seen that the test outcomes very well can be explained by the finding that the test implementations that relies on the QS kernel and/or prewhitening can incur considerable downward size distortions when samples are small. Moreover, based on the results above, it was seen that such size distortions are smallest when Bartlett kernel is used without prewhitening, largest when prewhitening is used, with the performance of the test implementation using the QS kernel without prewhitening falling in between the former two. These results are perfectly in line with the varying test results presented in Table 2. The test implementation with the smallest downward size distortions, $KPSS_{\text{Bartlett}}^{\text{No PW}}$, rejects the null more often than the test with the second best size properties (and the second largest downward size distortions), i.e. the $KPSS_{\text{QS}}^{\text{No PW}}$ tests. Finally, the test implementations with the worst size properties (and the largest downward size distortions), the $KPSS_{\text{Bartlett}}^{\text{PW}}$ and the $KPSS_{\text{QS}}^{\text{PW}}$ test, fail to reject the null most often of all test implementations.

Which of the two explanations above that is most likely to explain the results of Table 2 is of course hard to tell. Nevertheless, the empirical illustration presented in this section emphasizes one very urgent aspect of stationarity testing, namely the importance of considering test stability when choosing an implementation of the KPSS test to test for stationarity of a time series.

5 Concluding remarks

In the current paper, the finite-sample stability of the KPSS test is investigated. More specifically, the stability of the test is studied when various kernels are used in the long-run variance estimation, when various choices of bandwidth parameters are considered and when prewhitening is used in the variance estimation process.

The main results are that the distribution of the KPSS test can depend critically on the choice of kernel when samples are small or medium-sized. Furthermore, applying a prewhitening filter in the estimation of the long-run variance need not to improve the performance of the test, rather it can make the test more sensitive to the specific DGP under the null hypothesis. The main conclusion from these results is that it can be recommendable to use the Bartlett kernel without prewhitening, together with the currently available small-sample critical values, when performing the KPSS test in small samples, at least under circumstances that resembles those investigated in the current paper.

Besides pointing out important factors that must be taken into account when per-

forming the KPSS test in finite-sample situations, the results also point out dimensions along which future research can extend. For example, the size results reported in the current paper, apply for the case where disturbances are serially correlated. However, it can very well be the case that the conclusions drawn apply equally well when various forms of heteroscedasticity are considered. Investigation of this issue is left for future research.

Finally, it can be noted that the results in this paper also have implications for the use of stationarity tests in panel data. Since so-called mean group stationarity tests employ a standardized average over a set of time-series stationarity statistics, the failure to account for the finite-sample distribution of the KPSS test induce an error also in the panel data framework of e.g. Hadri (2000). Since the error will accumulate over cross-sections, panel stationarity tests can be expected to have even worse small-sample properties, more specifically small-T properties, than the time-series tests studied in this paper. As in the case with the univariate stationarity test, considerable improvements of the small-sample size properties of the panel stationarity tests should be feasible by taking into account the small-sample distribution of the KPSS test.

References

- Caner, M. and Kilian, L. (2001). Size Distortions of Tests of the Null Hypothesis of Stationarity: Evidence and Implications for the PPP debate. *Journal of International Money and Finance*, 20:639–657.
- Carrion-i-Silvestre, J. L. and Sansó, A. (2006). A Guide to the Computation of Stationarity Tests. *Empirical Economics*, 31:433–448.
- Cheung, Y.-W. and Lai, K. S. (1995). Lag Order and Critical Values of the Augmented Dickey-Fuller Test. *Journal of Business and Economic Statistics*, 13(3):277–280.
- den Haan, W. J. and Levin, A. (1997). A Practitioner’s Guide to Robust Covariance Matrix Estimation. In Maddala, G. S. and Rao, C. R., editors, *Handbook of Statistics*, volume 15. North-Holland.
- Dickey, D. A. and Fuller, W. A. (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association*, 74:427–431.
- Hadri, K. (2000). Testing for Stationarity in Heterogeneous Panel Data. *Econometrics Journal*, 3:148–161.

- Hall, R. E. (1978). Stochastic Implications of the Life Cycle Permanent Income Hypothesis: Theory and Evidence. *Journal of Political Economy*, 86:971–987.
- Hobijn, B., Franses, P. H., and Ooms, M. (2004). Generalizations of the KPSS-test for Stationarity. *Statistica Neerlandica*, 58(4):483–502.
- Hornok, A. and Larsson, R. (2000). The Finite Sample Distribution of the KPSS test. *Econometrics Journal*, 3:108–121.
- Jönsson, K. (2006). Testing Stationarity in Small and Medium-Sized Samples when Disturbances are Serially Correlated. Working Paper No. 2006:20, Lund University.
- Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., and Shin, Y. (1992). Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root. *Journal of Econometrics*, 54:159–178.
- Lanne, M. and Saikkonen, P. (2003). Reducing Size Distortions of Parametric Stationarity Tests. *Journal of Time Series Analysis*, 24(4):423–439.
- Lee, J. (1996). On the Power of Stationarity Tests Using Optimal Bandwidth Estimates. *Economics Letters*, 51:131–137.
- Leybourne, S. J. and McCabe, B. P. M. (1994). A Consistent Test for a Unit Root. *Journal of Business & Economic Statistics*, 12(2):157–166.
- Leybourne, S. J. and McCabe, B. P. M. (1999). Modified Stationarity Tests with Data-Dependent Model-Selection Rules. *Journal of Business & Economic Statistics*, 17(2):264–270.
- Newey, W. K. and West, K. D. (1987). A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, 55(3):703–708.
- Phillips, P. C. B. and Solo, V. (1992). Asymptotics for Linear Processes. *The Annals of Statistics*, 20(2):971–1001.
- Saikkonen, P. and Luukkonen, R. (1993). Testing for a Moving Average Unit Root in Autoregressive Integrated Moving Average Models. *Journal of the American Statistical Association*, 88(422):596–601.
- Sul, D., Phillips, P. C. B., and Choi, C.-Y. (2005). Prewhitening Bias in HAC Estimation. *Oxford Bulletin of Economics and Statistics*, 67(4):517–546.

Table 1: Size effects of various choices in the KPSS test.

Intercept								
T	$\rho = 0.0$				$\rho = 0.2$			
	Bartlett		QS		Bartlett		QS	
	No PW	PW	No PW	PW	No PW	PW	No PW	PW
20	0.100	0.100	0.100	0.100	0.113	0.077	0.074	0.075
30	0.100	0.100	0.100	0.100	0.115	0.084	0.092	0.081
40	0.100	0.100	0.100	0.100	0.112	0.087	0.100	0.090
50	0.100	0.100	0.100	0.100	0.110	0.091	0.099	0.088
100	0.100	0.100	0.100	0.100	0.107	0.095	0.102	0.096
250	0.100	0.100	0.100	0.100	0.110	0.102	0.107	0.104
500	0.100	0.100	0.100	0.100	0.105	0.100	0.103	0.101

T	$\rho = 0.4$				$\rho = 0.6$			
	Bartlett		QS		Bartlett		QS	
	No PW	PW	No PW	PW	No PW	PW	No PW	PW
20	0.121	0.058	0.056	0.056	0.146	0.050	0.051	0.046
30	0.135	0.074	0.089	0.069	0.156	0.061	0.090	0.053
40	0.128	0.074	0.101	0.074	0.153	0.063	0.122	0.064
50	0.127	0.084	0.109	0.084	0.147	0.069	0.130	0.069
100	0.120	0.090	0.108	0.090	0.143	0.087	0.131	0.088
250	0.115	0.097	0.108	0.100	0.132	0.097	0.126	0.098
500	0.115	0.101	0.108	0.103	0.126	0.096	0.117	0.098

Intercept and trend								
T	$\rho = 0.0$				$\rho = 0.2$			
	Bartlett		QS		Bartlett		QS	
	No PW	PW	No PW	PW	No PW	PW	No PW	PW
20	0.100	0.100	0.100	0.100	0.095	0.042	0.061	0.053
30	0.100	0.100	0.100	0.100	0.093	0.041	0.058	0.054
40	0.100	0.100	0.100	0.100	0.092	0.049	0.061	0.060
50	0.100	0.100	0.100	0.100	0.101	0.062	0.079	0.068
100	0.100	0.100	0.100	0.100	0.108	0.086	0.095	0.086
250	0.100	0.100	0.100	0.100	0.107	0.095	0.101	0.095
500	0.100	0.100	0.100	0.100	0.110	0.101	0.105	0.103

T	$\rho = 0.4$				$\rho = 0.6$			
	Bartlett		QS		Bartlett		QS	
	No PW	PW	No PW	PW	No PW	PW	No PW	PW
20	0.083	0.015	0.031	0.024	0.072	0.006	0.016	0.011
30	0.081	0.014	0.029	0.022	0.056	0.004	0.013	0.008
40	0.095	0.026	0.037	0.034	0.086	0.014	0.028	0.018
50	0.108	0.039	0.070	0.048	0.108	0.021	0.074	0.032
100	0.121	0.075	0.100	0.078	0.142	0.063	0.117	0.065
250	0.119	0.091	0.106	0.092	0.136	0.086	0.124	0.086
500	0.121	0.100	0.108	0.100	0.139	0.098	0.123	0.096

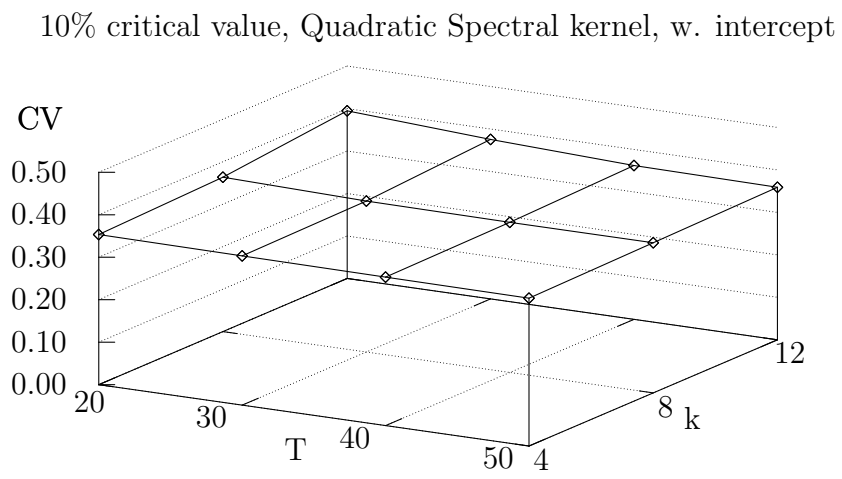
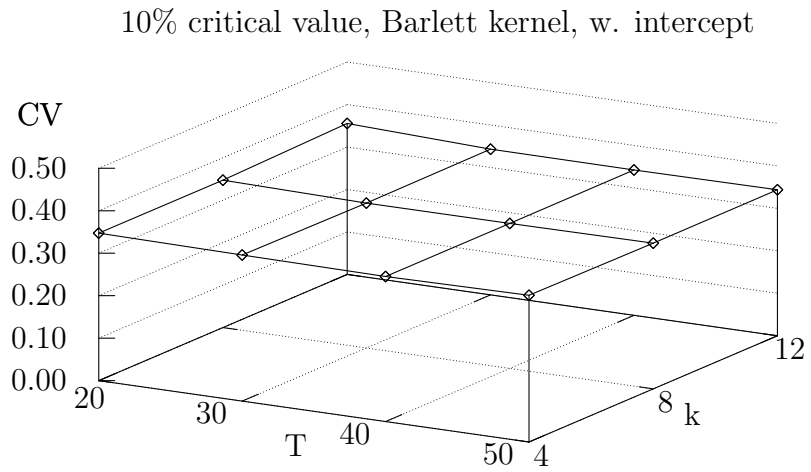
Table 2: Empirical application of the KPSS test.^{a,b}

	Sample period	Asy. CV	Test statistics			
			$KPSS_{Bartlett}^{No\ PW}$	$KPSS_{Bartlett}^{PW}$	$KPSS_{QS}^{No\ PW}$	$KPSS_{QS}^{PW}$
Australia	1960-2004	0.119	0.122	0.034	0.120	0.039
Austria	1960-2004	0.119	0.126	0.034	0.108	0.030
Belgium	1960-2004	0.119	0.128	0.040	0.124	0.036
Canada	1961-2004	0.119	0.099	0.026	0.080	0.017
Denmark	1966-2004	0.119	0.118	0.074	0.107	0.074
Finland	1960-2004	0.119	0.108	0.027	0.083	0.017
France	1963-2004	0.119	0.132	0.078	0.132	0.066
Greece	1960-2004	0.119	0.125	0.060	0.129	0.052
Iceland	1960-2004	0.119	0.117	0.064	0.113	0.061
Ireland	1960-2004	0.119	0.147	0.079	0.152	0.084
Italy	1960-2004	0.119	0.136	0.046	0.124	0.041
Japan	1960-2004	0.119	0.126	0.034	0.108	0.032
Luxembourg	1960-2004	0.119	0.153	0.062	0.126	0.041
Netherlands	1960-2003	0.119	0.103	0.053	0.103	0.049
New Zealand	1962-2004	0.119	0.129	0.029	0.121	0.028
Norway	1960-2004	0.119	0.106	0.021	0.089	0.020
Portugal	1960-2004	0.119	0.084	0.028	0.075	0.023
Spain	1960-2004	0.119	0.091	0.032	0.082	0.024
Sweden	1960-2004	0.119	0.112	0.029	0.083	0.016
Switzerland	1965-2003	0.119	0.149	0.064	0.143	0.064
UK	1960-2004	0.119	0.166	0.090	0.174	0.087
USA	1960-2004	0.119	0.153	0.051	0.143	0.043

Notes: ^a The tests are performed on the 10% significance level.

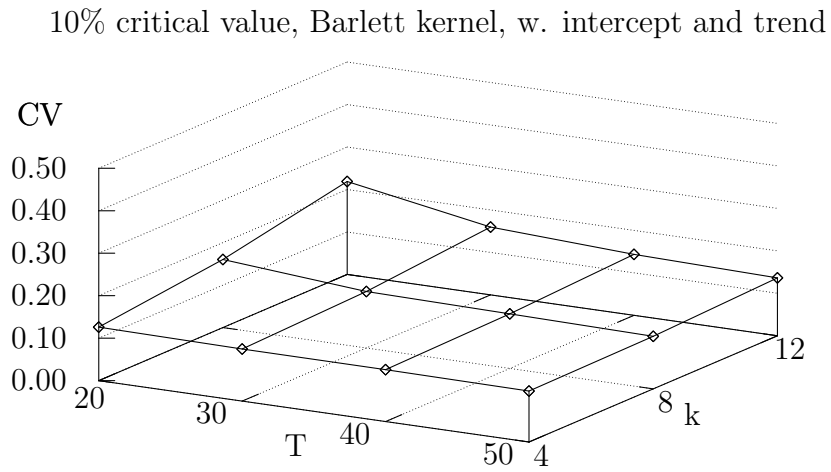
^b Boldface numbers indicate rejection of the null hypothesis of stationarity.

Figure 1: Critical values for the KPSS test, intercept only, no prewhitening.^{a,b}

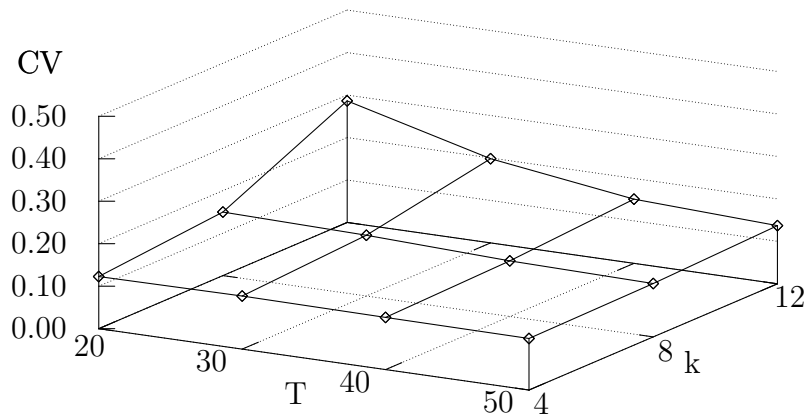


Notes: ^aCritical value as a function of the sample-size, T, and the bandwidth scaling, k.
^b Long-run variance estimated without prewhitening.

Figure 2: Critical values for the KPSS test, intercept and trend, no prewhitening.^{a,b}



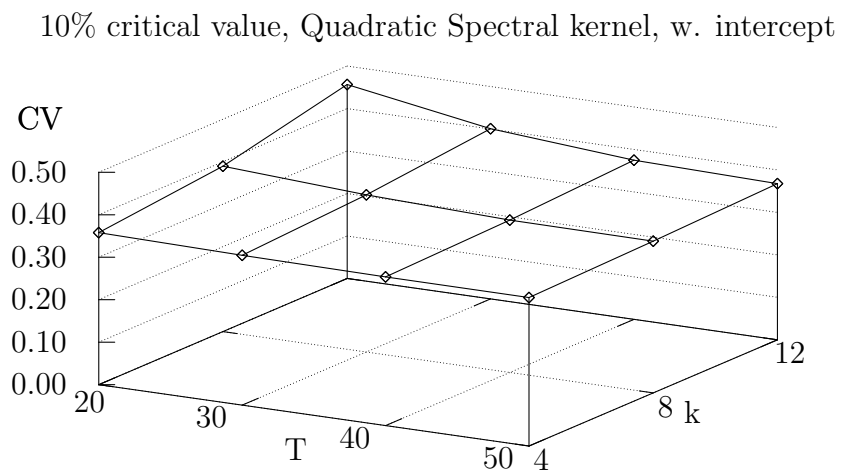
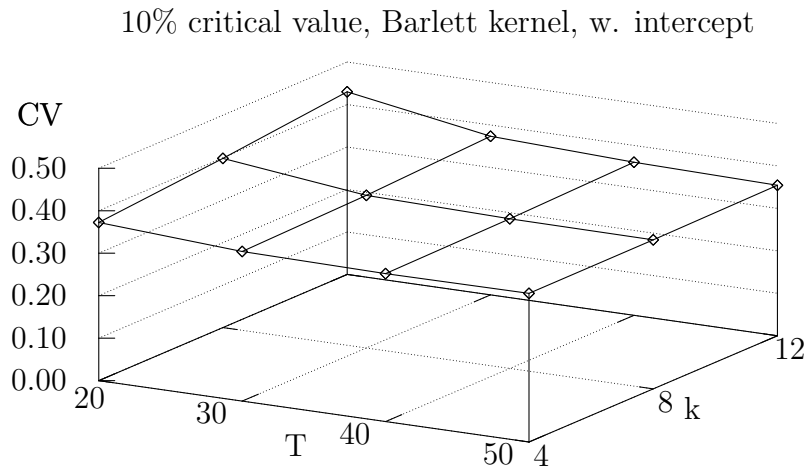
10% critical value, Quadratic Spectral kernel, w. intercept and trend



Notes: ^aCritical value as a function of the sample-size, T, and the bandwidth scaling, k.

^b Long-run variance estimated without prewhitening.

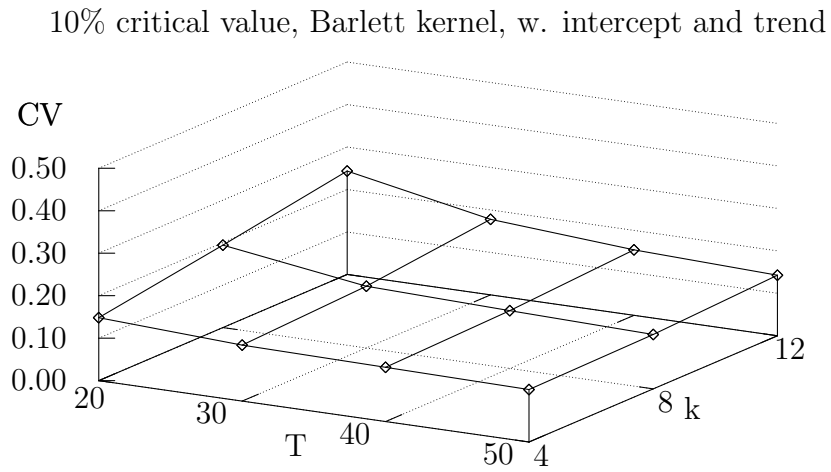
Figure 3: Critical values for the KPSS test, intercept only, prewhitening.^{a,b}



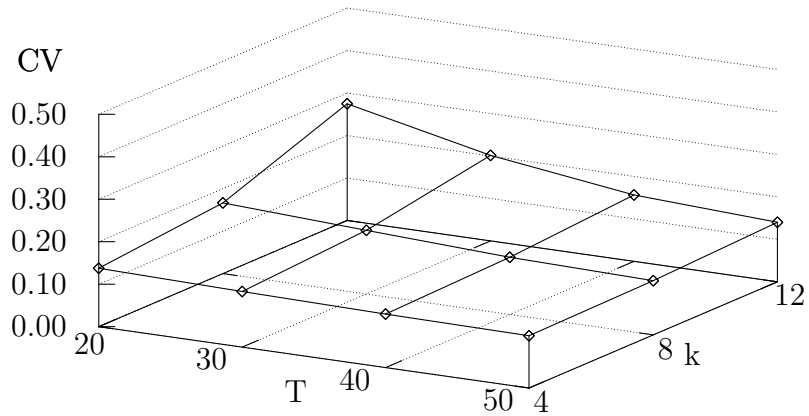
Notes: ^aCritical value as a function of the sample-size, T, and the bandwidth scaling, k.

^b Long-run variance estimated with an AR(1) prewhitening filter.

Figure 4: Critical values for the KPSS test, intercept and trend, prewhitening.^{a,b}



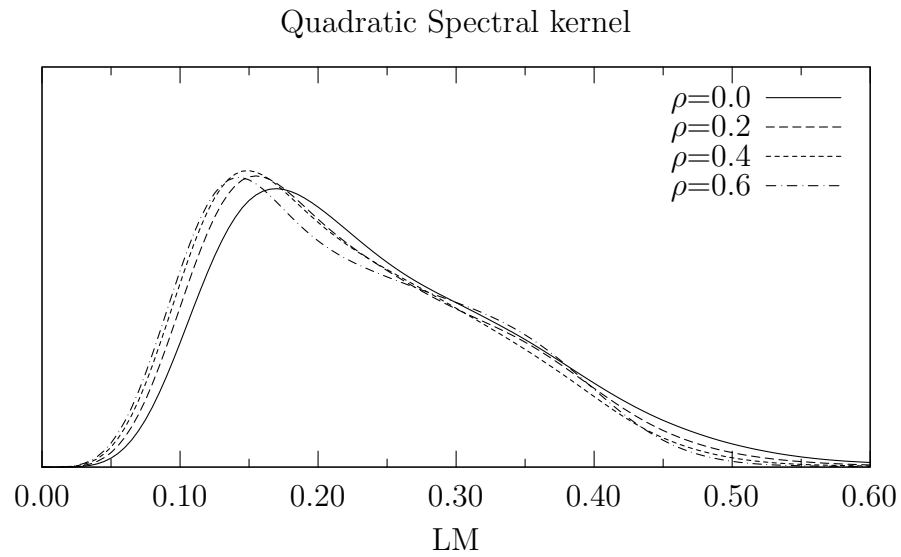
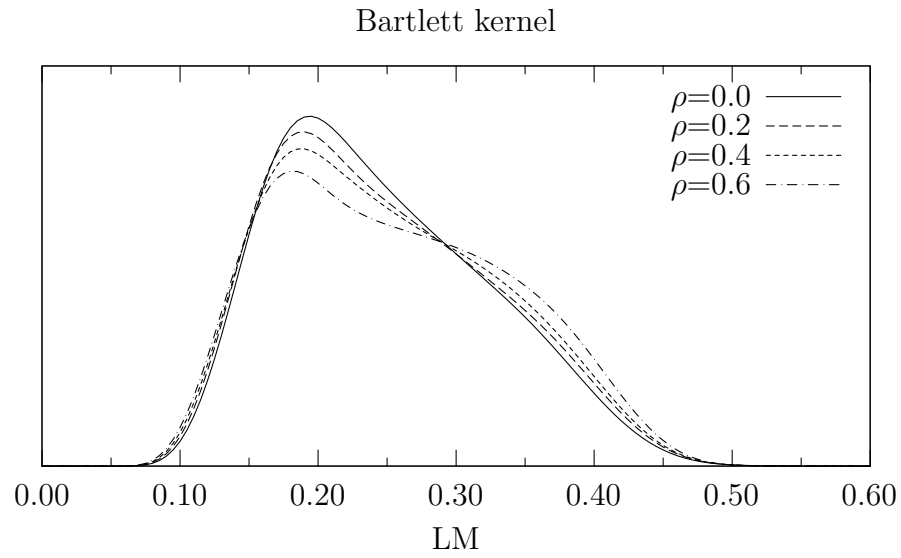
10% critical value, Quadratic Spectral kernel, w. intercept and trend



Notes: ^aCritical value as a function of the sample-size, T, and the bandwidth scaling, k.

^b Long-run variance estimated with an AR(1) prewhitening filter.

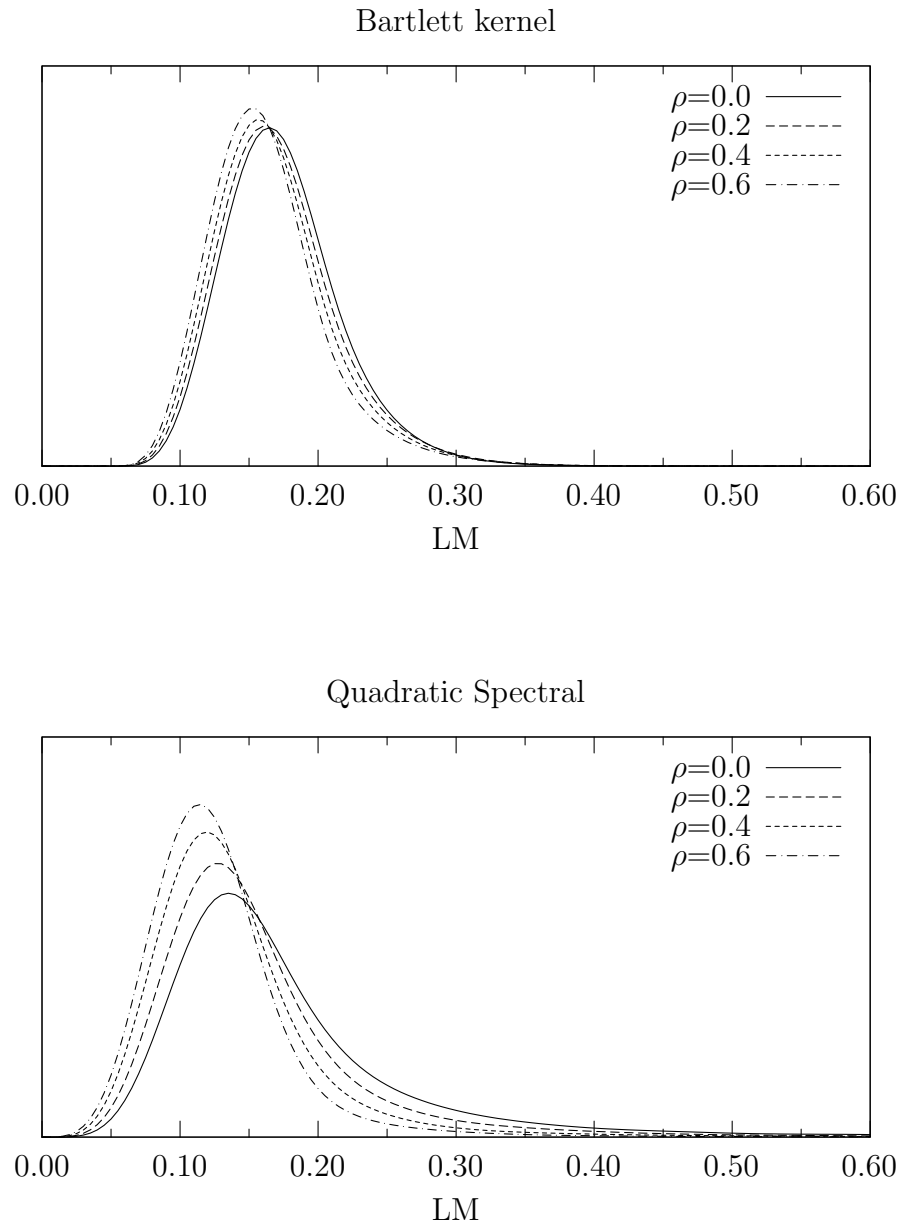
Figure 5: Kernel density estimates, w. intercept, no prewhitening.^{a,b}



Notes: ^aThe density estimates are obtained for $T=20$ and $k=12$.

^b Long-run variance estimated without prewhitening.

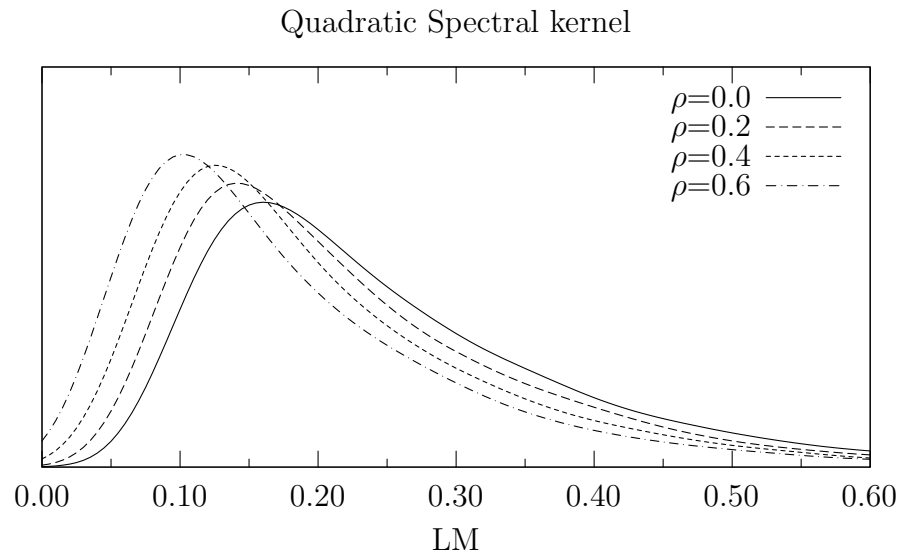
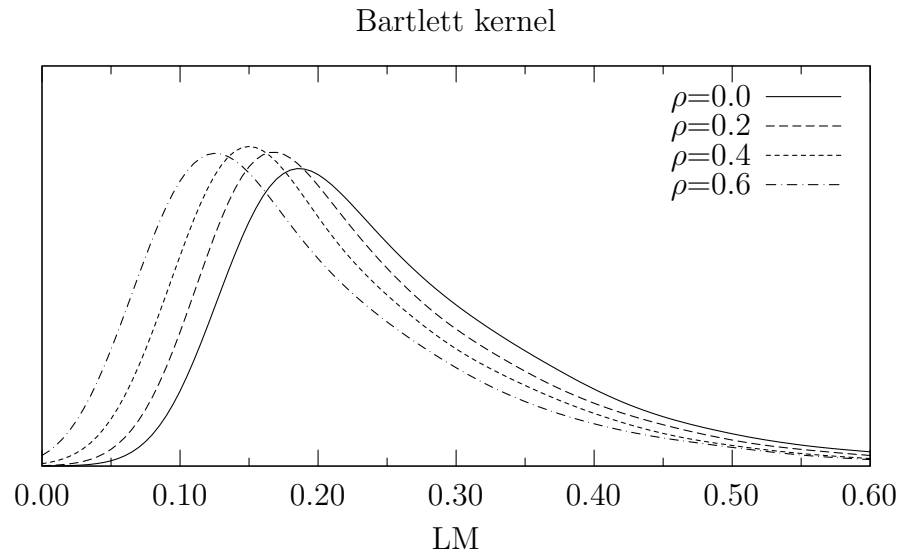
Figure 6: Kernel density estimates, w. intercept and trend, no prewhitening.^{a,b}



Notes: ^aThe density estimates are obtained for $T=20$ and $k=12$.

^b Long-run variance estimated without prewhitening.

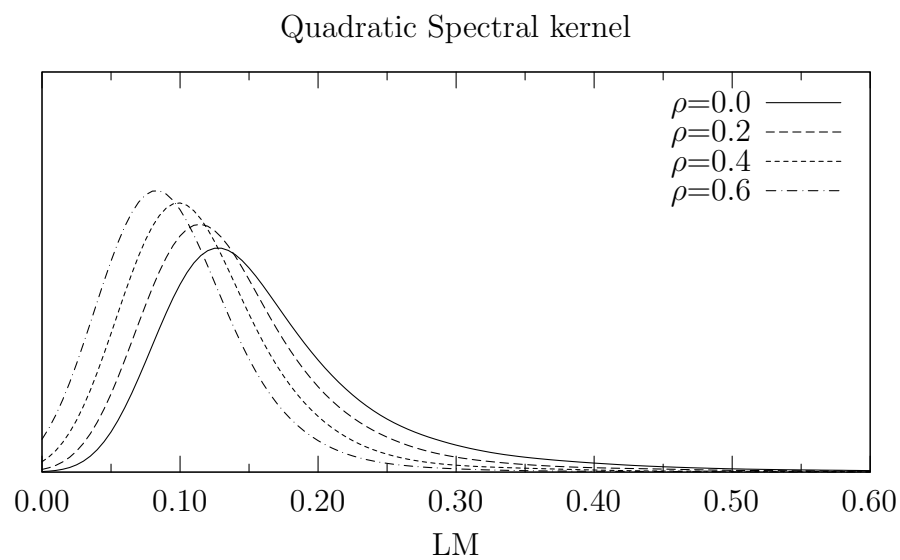
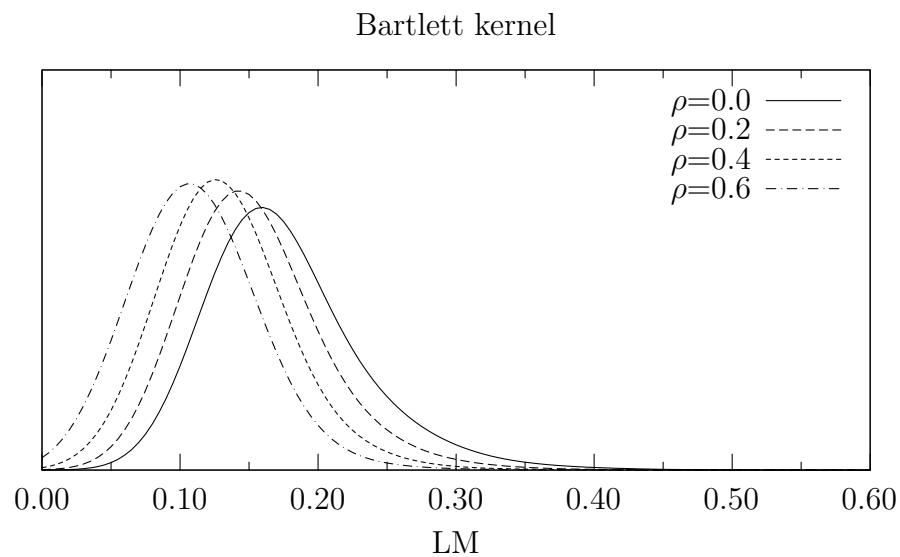
Figure 7: Kernel density estimates, w. intercept, prewhitening.^{a,b}



Notes: ^aThe density estimates are obtained for $T=20$ and $k=12$.

^b Long-run variance estimated with an AR(1) prewhitening filter.

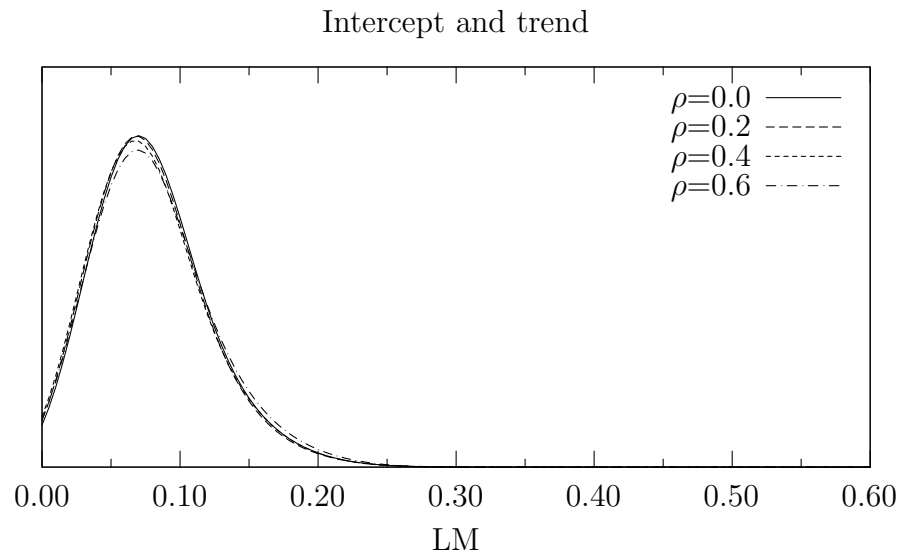
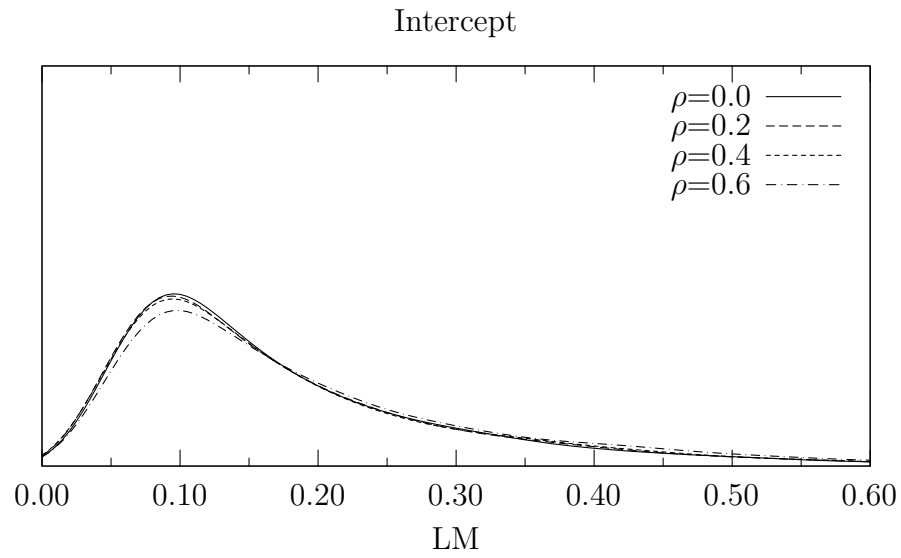
Figure 8: Kernel density estimates, w. intercept and trend, prewhitening.^{a,b}



Notes: ^aThe density estimates are obtained for $T=20$ and $k=12$.

^b Long-run variance estimated with an AR(1) prewhitening filter.

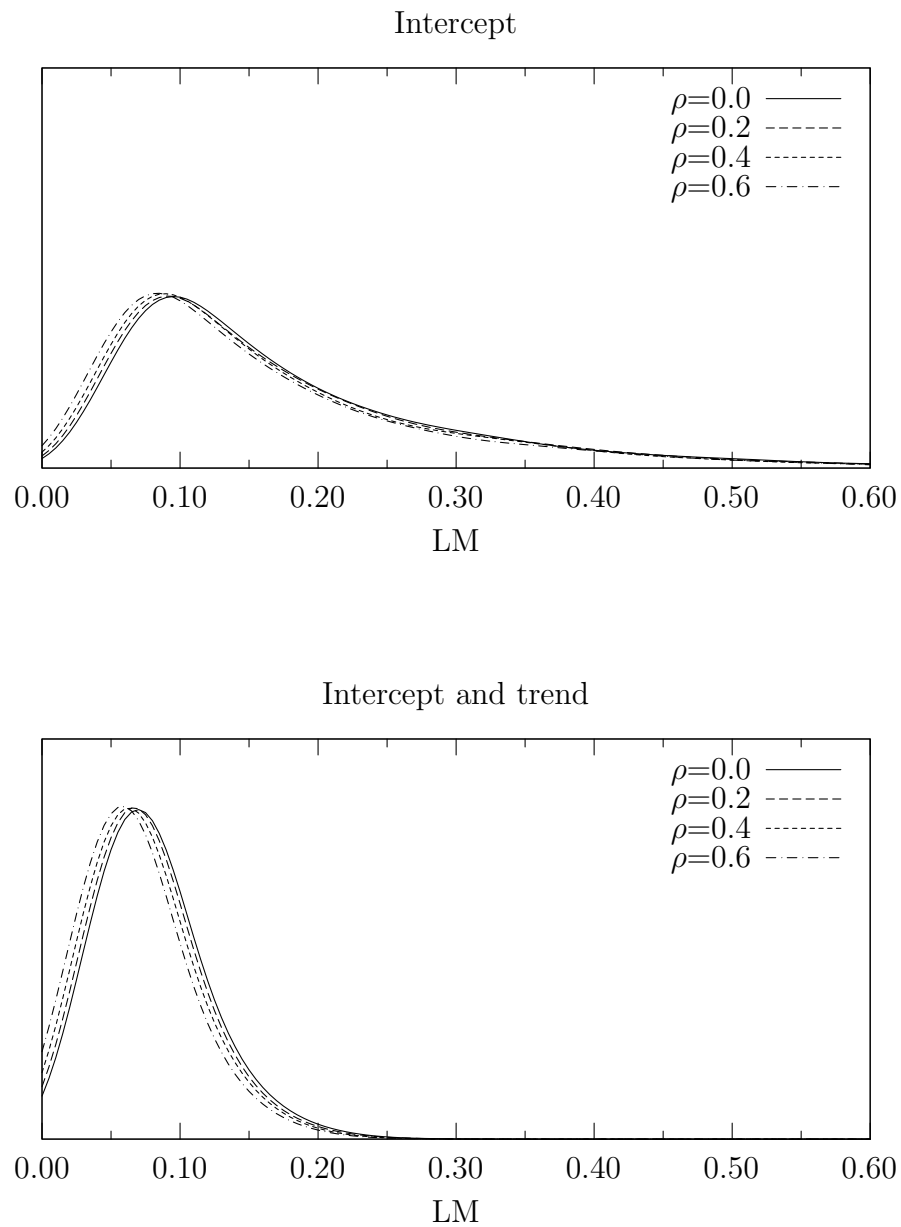
Figure 9: Kernel density estimates, Quadratic Spectral kernel, no prewhitening.^{a,b}



Notes: ^aThe density estimates are obtained for $T=100$ and $k=12$.

^b Long-run variance estimated without prewhitening.

Figure 10: Kernel density estimates, Quadratic Spectral kernel, prewhitening.^{a,b}



Notes: ^aThe density estimates are obtained for $T=100$ and $k=12$.

^b Long-run variance estimated with an AR(1) prewhitening filter.