Testing Stationarity in Small and Medium-Sized Samples when Disturbances are Serially Correlated

Jönsson, Kristian

2006

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Testing Stationarity in Small and Medium-Sized Samples when Disturbances are Serially Correlated

Kristian Jönsson* †

November 11, 2009

Abstract

In this paper, we study the size distortions of the KPSS test for stationarity when serial correlation is present and samples are small and medium-sized. It is argued that two distinct sources of the size distortions can be identified. The first source is the finite-sample distribution of the long-run variance estimator used in the KPSS test, while the second source of the size distortions is the serial correlation not captured by the long-run variance estimator due to a too narrow choice of truncation lag parameter. When the relative importance of the two sources is studied, it is found that the size of the KPSS test can be reasonably well controlled if the finite-sample distribution of the KPSS test statistic, conditional on the time-series dimension and the truncation lag parameter, is used. Hence, finite-sample critical values, that can be applied in order to reduce the size distortions of the KPSS test, are supplied. When the power of the test is studied, it is found that the price paid for the increased size control is a lower raw power against a non-stationary alternative hypothesis.

Keywords: Stationarity testing; Unit root; Finite-sample inference; Long-run variance; Monte Carlo simulation; Permanent Income Hypothesis; Private Consumption.
JEL classification: C12; C14; C15; C22; E21.

*Ministry of Finance, SE-103 33 Stockholm, Sweden. Email: kristian.jonsson@nance.ministry.se

†The author would like to thank Anindya Banerjee, three anonymous referees, Tommy Andersson, David Edgerton, Thomas Elger, Klas Fregert, Rolf Larsson, Johan Lyhagen, Joakim Westerlund and seminar participants at the Department of Economics, Lund University, for useful suggestions and discussions on the topics covered in, and relating to, this paper. The simulations in the current paper were carried out in Gauss on the LUNARC cluster. Financial support, from The Crafoord Foundation, The Royal Swedish Academy of Sciences and The Jan Wallander and Tom Hedelius Foundation, research grant number P2005-0117:1, is gratefully acknowledged. The views expressed in the current paper are solely the responsibility of the author.

1 Introduction

During the last three decades, one of the most investigated branches of econometrics in general, and time-series econometrics in particular, is that of unit root, stationarity and cointegration testing. The seminal contributions of Dickey and Fuller (1979), Engle and Granger (1987) and Kwiatkowski et al. (1992) have all made their mark on applied economics and are used as central tools for investigating various economic questions. When using the test of Kwiatkowski et al. (1992), the so-called KPSS test, to test if a series is stationary, I(0), against the alternative that the series contains a unit root, is I(1), several implementations of the test, all with attractive asymptotic behavior, are available.[1] The implementations differ (most commonly) in their estimation of the so-called long-run variance. The alternatives available involve the use of various kernels to estimate the long-run variance under the null hypothesis (see e.g. Hobijn et al., 2004), the use of automatic procedures for the selection of the truncation lag or bandwidth parameter (see e.g. Hobijn et al., 2004; Carrion-i-Silvestre and Sansó, 2006) and the application of a prewhitening filter in the long-run variance estimation (see e.g. Sul et al., 2005; Carrion-i-Silvestre and Sansó, 2006). Given that the choices made fulfill certain regularity conditions, the asymptotic distribution of the KPSS test statistic is the same regardless of what choices that are made. The appropriateness of the various implementations depends on how well the asymptotic approximation works for the specific sample size at hand.

In empirical applications, where it can be of interest to employ the KPSS test, sample sizes are always limited and in addition often small. This applies especially when post-war macroeconomic time series are investigated. Hence, when applied to investigate economic questions, the performance of the KPSS test relies to a large extent on how well the finite-sample distribution of the test statistic corresponds to the asymptotic distribution. Unfortunately, when serial dependence is present under the null hypothesis of stationarity, the asymptotic approximation can be poor, which causes problems relating to the size and power of the KPSS test (see e.g. Lee, 1996; Caner and Kilian, 2001; Hobijn et al., 2004; Müller, 2005). Methods to mitigate the size distortions within the framework of Kwiatkowski et al. (1992) have been suggested by Hobijn et al. (2004), Sul et al. (2005) and Carrion-i-Silvestre and Sansó (2006). However, a common feature among these suggestions is that their performance is investigated for rather large samples sizes, and that the suggested remedies may be inappropriate in small-sample situations (see e.g. Jönsson, 2006). Hence, the performance of the KPSS test in small samples, when serial

[1]By having the hypothesis that a time series is I(0) as the null hypothesis, the KPSS test differs from e.g. the Dickey and Fuller (1979) test which has as null hypothesis that the series under consideration is I(1). Hence, the KPSS test and the Dickey and Fuller (1979) tests can be considered as being complements to each other.

dependency is allowed for, is still largely unknown and possibly open for improvement.

In this paper, it is shown that the KPSS test for stationarity can be grossly oversized in small samples when serial correlation is allowed for. When the sources of these size distortions are studied, it is found that the long-run variance estimator is the main reason for the size distortion, while the actual serial correlation of the data series exerts little influence on the test once the truncation lag has been accounted for. A natural suggestion is then to proceed along the lines of Cheung et al. (1995) and Hornok and Larsson (2000) and supply finite-sample critical values for the KPSS test.

Within the unit root testing framework of Dickey and Fuller (1979) and Said and Dickey (1984), the use of small-sample critical values is well-established. In order to control the size of the so-called augmented Dickey-Fuller (ADF) test, Cheung and Lai (1995) suggest that critical values should be obtained from simulating time series that are generated as pure random walks, i.e. unit root processes with serially independent disturbances, and applying the ADF test to these time series. The authors argue that obtaining critical values for the ADF test under the assumption that errors are serially independent introduces nuisance parameters into the distribution of the test statistic when this assumption is violated and disturbances display serial dependence. However, it is also argued that the size distortions arising from these additional nuisance parameters are relatively small, and hence that the performance of the test can be improved by obtaining critical values under the simplifying assumption of serial independence. A similar line of reasoning is adopted in the current paper. The critical values supplied here are obtained by conditioning on the truncation lag and the sample size. Critical values are obtained for a wide range of truncation lags and sample sizes and are easily calculated from a supplied set of response surface parameters. When it is studied whether the size properties of the KPSS test are improved by using the supplied critical values, it is found that large size improvements become feasible by conditioning the critical values on the sample size and the truncation lag. However, the price that is paid for better size properties is a lower power against a non-stationary alternative hypothesis. In the smallest samples considered here, this loss of power can cause the test to be biased, i.e. make the test have a power that falls below the size. The bias indicates that caution should be taken whenever drawing conclusions based on the KPSS test applied to small samples. We also supply an empirical illustration that emphasizes the need to account for both the size of the sample and the truncation lag when performing the KPSS test for stationarity.

The rest of this paper is organized as follows. In Section 2, the stationarity test of Kwiatkowski et al. (1992) is presented. The properties of this test are, in Section 3, investigated when samples are small and when serial correlation is allowed for. In Section 4, it is studied what role the truncation lag, in the long-run variance estimator, plays for the finite-sample distribution of the KPSS test. Section 5 offers a suggestion of how to improve the finite-sample properties of the KPSS test, while the performance of this suggestion is studied in Section 6. The empirical illustration is supplied in Section 7. Finally, Section 8 presents some concluding remarks.

2 The stationarity test

The current paper is concerned with the stationarity properties of a time series y_t. The series y_t is assumed to consist of deterministic components, a potentially non-stationary component and a stationary disturbance term as in (1).

$$y_t = \delta_0 + \delta_1 t + \xi_t + \varepsilon_t \tag{1}$$

In (1), δ_0 and $\delta_1 t$ are the deterministic components taking the form of an intercept and a time trend. The non-stationary component is described by $\xi_t = \xi_{t-1} + \nu_t$, where ν_t is i.i.d. with zero mean and variance σ_ν^2, while $\xi_0 = 0$.[2] Finally, ε_t is the stationary disturbance term that fulfills e.g. the linear process assumptions of Phillips and Solo (1992), i.e. $\varepsilon_t = \Psi(L)\eta_t = \sum_{i=0}^{\infty} \psi_i \eta_{t-i}$, where η_t is i.i.d. with zero mean and variance σ_η^2, while it is assumed that $\sum_{i=0}^{\infty} i|\psi_i| < \infty$.[3]

To test for stationarity of the series y_t, Kwiatkowski et al. (1992) suggest that one should test the null hypothesis that H_0: $\sigma_\nu^2 = 0$ against the alternative hypothesis H_1: $\sigma_\nu^2 > 0$. The test statistic suggested by the authors is given in (2).

$$LM = \frac{T^{-2} \sum_{t=1}^{T} S_t^2}{s^2(l)} \tag{2}$$

In (2), S_t is the partial sum process, i.e. $S_t = \sum_{i=1}^{t} e_i$, where e_i is the least squares residual obtained after detrending y_t, while $s^2(l)$ is a variance estimator. The variance occurs in the denominator of (2) in order to relieve the asymptotic distribution of the test statistic from nuisance parameters.

To illustrate the role of the variance estimator, first consider the case where ε_t is distributed i.i.d. with mean and variance equal to $(0, \sigma_\varepsilon^2)$. For ease of exposition, assume that there are no deterministic terms, i.e. $e_t = \varepsilon_t$. Under this assumption, the numerator of the test statistic in (2) converges in distribution to $\sigma_\varepsilon^2 \int_0^1 W(r)^2 dr$ as $T \to \infty$.

[2] As long as we allow for an intercept in the series y_t, the assumption $\xi_0 = 0$ can be made without loss of generality.

[3] It can be noted that the asymptotics of the KPSS test also can be obtained under the strong-mixing conditions of Phillips and Perron (1988). However, since the current paper is concentrated on the case where disturbances may display serial correlation, we assume that the linear process assumption holds.

Hence, if $s^2(l)$ is equal to $T^{-1}\sum_{t=1}^{T} e_t^2$, an asymptotically pivotal test statistic is obtained from (2) since $T^{-1}\sum_{t=1}^{T} e_t^2 \xrightarrow{p} \sigma_\varepsilon^2$ as $T \to \infty$. However, the same variance estimator is not suitable when the i.i.d. assumption for ε_t is abandoned. To see this, let $\varepsilon_t = \sum_{i=0}^{\infty} \psi_i \eta_{t-i}$, with η_t being i.i.d. with mean and variance $(0, \sigma_\eta^2)$ and $\sum_{i=0}^{\infty} i|\psi_i| < \infty$. Under these assumptions, the numerator of (2) converges in distribution to $\sigma^2 \int_0^1 W(r)^2 dr$ as $T \to \infty$, where $\sigma^2 = (\Psi(1)\sigma_\eta)^2$. But $s^2(l) = T^{-1}\sum_{t=1}^{T} e_t^2$ now converges in probability to $\sigma_\varepsilon^2 = \sum_{i=0}^{\infty} \psi_i^2 \sigma_\eta^2$ as $T \to \infty$, and it becomes evident that this naive form of the variance estimator now fails to capture the correlation present between e_t and e_{t-i} for $i \neq 0$. The consequence will be that the LM statistic in (2) converges in distribution to $(\sum_{i=0}^{\infty} \psi_i^2 \sigma_\eta^2)^{-1}(\Psi(1)\sigma_\eta)^2 \int_0^1 W(r)^2 dr$, which clearly depends on nuisance parameters, related to the intertemporal dependence present in the disturbance process, even as $T \to \infty$. The solution that Kwiatkowski et al. (1992) suggested, to remedy this problem, is to employ the variance estimator of Newey and West (1987). This estimator consistently estimates the long-run variance $lim_{T\to\infty} T^{-1}E(\sum_{t=1}^{T} S_t)^2 = (\Psi(1)\sigma_\eta)^2$ as $T \to \infty$.[4] The variance estimator suggested by Newey and West (1987) is given in (3)-(4).

$$s^2(l) = T^{-1}\sum_{t=1}^{T} e_t^2 + 2T^{-1}\sum_{s=1}^{l} w(s,l) \sum_{t=s+1}^{T} e_t e_{t-s} \tag{3}$$

$$w(s,l) = 1 - \frac{s}{l+1} \tag{4}$$

By the virtue of the second term on the righthand side of (3), this variance estimator captures the non-contemporaneous correlation in e_t, i.e. the correlation between e_t and e_{t-i} for $i \neq 0$. However, in order to take the non-contemporaneous components into account, a truncation lag, l, has to be chosen in such a way that $l \to \infty$ as $T \to \infty$, while $l = o(T^{0.5})$. Kwiatkowski et al. (1992) suggest, following Schwert (1989), that l could be chosen according to $l = [k(T/100)^{0.25}]$, where $k \in \{4, 12\}$ is a truncation lag scaling parameter.

In the next section, we go on by investigating the properties of the KPSS test when samples are small. More specifically, we will look into the size and power properties of the test when serially correlated disturbances are allowed for.

3 Properties of the KPSS test

Suppose that a researcher is interested in investigating the mean-reverting properties of annual post-war private consumption. The outcome of a stationarity test applied to such

[4] Müller (2005), footnote 1, p. 198, discusses the notation and description of the long-run variance. In the current paper, we stick to the notation of Kwiatkowski et al. (1992), while recognizing the point made by the former author.

a series is crucial for the conclusion drawn regarding the way private consumption is determined. If the consumption series follows a random walk, possibly with a drift, the permanent income hypothesis can be considered as a reasonable model for consumption determination (see e.g. Hall, 1978). However, this is not the case if stationarity applies and consumption is found to be mean-reverting. Under such circumstances, consumption could instead be determined as stated by the Keynesian consumption function with income being trend stationary. In order to gain knowledge about the empirical relevance of the permanent income hypothesis, the researcher collects somewhat over 50 time-series observations for private consumption, reflecting the amount of data commonly found in empirical applications, and applies the KPSS test for stationarity. To account for the deterministic components, the series is detrended, either by using an intercept only or by using both an intercept and a time trend.[5] Furthermore, in order to allow for serial correlation in the detrended series, the variance estimator in (3)-(4) is used in the KPSS test statistic in (2). Given the test result that is obtained, conclusions regarding private consumption behavior can be drawn, but how certain can we be that a good conclusion has been reached? In general, what is the probability that the KPSS test will reject the null hypothesis of stationarity given that the true data generating process (DGP) is mean-reverting? Furthermore, what is the probability that the KPSS test will fail to reject the null of stationarity given that the true DGP is non-stationary? Stated in another way, what are the size and the power properties of the KPSS test?

The size and power properties of the KPSS test obviously depend an several factors in finite samples. First and foremost, given that suitable deterministic components have been specified, it can be suspected that the size of the stationarity test is affected by some sort of distance between the DGP and the non-stationarity case, i.e. by the persistence of the series under the null hypothesis. Second, the choice of truncation lag can very well affect the test statistic under the null hypothesis, at least when samples are small. Finally, the rejection frequency under the alternative hypothesis is, in addition to just-mentioned factors, likely to depend on the relative variance of the permanent and transitory components η_t and ε_t.

Let us first consider the size of the KPSS test. In order to distinguish between the size effects caused by the DGP and the effects caused by altering the bandwidth parameter, Monte Carlo simulations based on various parameter setups can be used. To this end, data is generated according to $y_t = \alpha_0 + \alpha_1 t + u_t$. When only an intercept is considered, α_0 is set such that $\alpha_0 \in U[0, 10]$, while in addition $\alpha_1 \in U[0, 2]$ when both

[5] In the specific case of consumption, an intercept together with a time trend should be used since consumption is a trending variable.

an intercept and a time trend are considered.[6] Furthermore, u_t is generated by either an AR(1) process or a MA(1) process with the AR and MA parameter space given by $\{0.0, 0.2, \ldots, 0.8\}$. In both cases, the disturbance driving the AR or MA process is set to be normally, independently and identically distributed with zero mean and unit variance, i.i.d. $N(0, 1)$. Given this DGP, it can be investigated how many times the null hypothesis is (incorrectly) rejected on a given significance level, i.e. it is possible to investigate the size of the test. Furthermore, by considering the case where either the MA and AR parameter is equal to 0.0, it is possible to trace central determinants affecting the size of the test. The sample sizes considered are $T \in \{20, 30, 40, 50, 75, 100\}$, where in all cases T+100 time-series observations are generated for the AR and MA processes, while the first 100 of these are discarded in order to reduce the effects of the initial condition $u_1 = 0$. The size is calculated based on 5,000 replications for each parameter setup.

In Table 1, the size of the KPSS test, on the 10% significance level, is presented for the case where only an intercept is present.[7] Besides the different choices of AR and MA parameters, the size is investigated for various choices of truncation lag, more specifically truncation lags calculated using $k = \{4, 8, 12\}$.

The first thing that can be noticed in the table is that the test performs almost perfect when there is no serial correlation in the disturbance u_t. The exception is a slight upward size distortion when k=12 and T=20. As one moves down in Table 1, the performance of the test deteriorates. This applies especially for the case with AR(1) disturbances. However, it can be noticed that the larger the truncation lag, the more robust is the test to the increased serial correlation. The intuition behind this is that non-negligible higher-order autocorrelations are better taken into account by the long-run variance estimator in (3) and (4) when a larger truncation lag is employed. From Table 1 it can also be seen that, regardless of the choice of k, the test performs better as T increases. This is what we expect to see since the long-run variance is consistently estimated regardless of the choice of k. The overall conclusion, from having studied Table 1, is that the small-sample size distortions to the KPSS test, when allowing for an intercept only, are relatively small and mainly caused by failure to account for the serial correlation in the disturbances.

However, when attention is turned to the case where both an intercept and a time trend are allowed for, the picture becomes very different. In Table 2, the size for this case is presented. The first thing that stands out in comparison to the results in Table 1 is the poor performance that occurs when disturbances are set to be i.i.d.. More specifically, consider the case where there is no serial correlation in u_t, i.e. the cases when $\rho = 0$ and $\theta = 0$. Contrary to the results obtained from Table 1, the size of the KPSS test is heavily

[6]Here and in the rest of this paper, $U[a, b]$ denotes the uniform distribution on the interval $[a, b]$.

[7]The size is calculated using the asymptotic critical value supplied by Kwiatkowski et al. (1992).

affected by the choice of k even when no serial correlation is present. Stated differently, the choice of k will affect the distribution of the KPSS statistic, not only because of the capability of the long-run variance estimator to capture the serial correlation in u_t, but also because the small-sample distribution of the test depends to a large extent on k even in absence of serial correlation. If the true DGP is characterized by an intercept, a time trend and white noise disturbances, the KPSS test will over-reject the null hypothesis if k is large relative to T. If we consider the case where u_t is i.i.d., while k=12 and T=50, it can be seen that the test rejects the null hypothesis 25% of the time even though the null hypothesis is correct. The size distortion becomes even worse when $T < 50$ and the size is 100% on the 10% significance level when T=20 and k=12.

One obvious solution to the size problems encountered in the upper-most panels of Table 2 is to choose k in such a way that the truncation lag remains small. However, as the results in the four lower panels of Table 2 indicate, the solution to stick with a low value of k is not a viable way to proceed since this can cause large size distortions due to failure to account sufficiently well for the serial correlation when it is present. That is, choosing a low truncation lag makes the estimator in (3) and (4) incapable of capturing significant higher-order serial correlation and hence introduces a size distortion due to neglected autocorrelation. A better way to proceed would be to find a method that allows for a large value of k, i.e. $k \in \{8, 12\}$, without introducing size distortions when disturbances are serially independent. This solution will be explored later on in the current paper.

Besides the fact that k influences the distribution of the test statistic, the overall pattern, evident in Table 1, emerges also in Table 2. That is, more pronounced serial correlation increases the size distortions of the test for fixed T and k, while a wider truncation lag and a larger time-series dimension work in the opposite direction and reduces the size distortions for a given autoregressive parameter, ρ, or moving average parameter, θ.

The results presented in Table 1 and Table 2 extend beyond the choices of k, and the implied choices of l, considered here. If a researcher was to choose a truncation lag based on any other criterion, such as a data-dependent criterion, the pattern emerging from the tables is still likely to be present. A wider bandwidth will distort the size of the KPSS test regardless of whether or not serial correlation is present in the DGP. This applies in particular when a linear trend is allowed for.

Turning back to the researcher that wishes to investigate the empirical relevance of the permanent income hypothesis, we see that he or she should not be overly optimistic about properties of the KPSS test, and hence not put too much faith into the test results. Considering the possible linear trend in the natural logarithm of private consumption, there is a rather large probability that the null hypothesis is rejected even though private consumption can be described as a mean-reverting time-series process.

To be able to come to terms with the size distortions that arise in the KPSS test, it is important to find the source of the distortions. Since the results presented in this section indicate that the performance of the test depends on the truncation lag parameter, regardless of the degree of serial correlation that is present, it seems important to study the behavior of the long-run variance estimator, and the role that it plays for the behavior of the KPSS test statistic. This is done in the next subsection.

However, before possible solutions to the problem regarding the size distortions are investigated, the power of the KPSS test should be studied. In order to do this, data is generated as above, with the only difference that the series y_t now contains a random walk component, $\zeta_t = \zeta_{t-1} + \eta_t$, where η_t is i.i.d. $N(0, 1)$. Under this data generating process, the null hypothesis is clearly false. In Table 3 and Table 4, the power of the KPSS test is presented when the test is performed on the 10% significance level.[8]

As seen in Table 3, the power of the KPSS test is satisfactory when only an intercept is present. For the smallest sample size considered, T=20, the power never falls under 30% regardless of the choice of truncation lag scaling parameter and serial correlation. Hence, the power of the test is well above the size of the test, which was presented in Table 1. This is of course a nice property of the test since it implies that it is more likely that the test rejects the null hypothesis when it is false than when it is true. From Table 3, it can also be seen that the power of the test increases as T increases, while the power decreases with increasing truncation lag.

When the power of the KPSS test is investigated for the case where a linear trend is present, radically different conclusions are reached compared to when only an intercept is accounted for. The power of the test is reduced when both an intercept and a linear trend are present. When comparing the power in Table 4 to the size on Table 2, it can be seen that there are cases where the power of the test actually falls below the size of the test. For example, when T=20 and k=8, the size is 69.7% in absence of serial correlation. The corresponding power is 65.6%. When there is no serial correlation and (T,k)=(30,12), size and power are 71.8% and 61.8%, respectively. Similar results hold when disturbances are serially correlated. Hence, it can be seen that the KPSS test is biased when samples are small and serial correlation is allowed for. Furthermore, the bias arises without introducing any prewhitening as considered by e.g. Lee (1996).

Considering the small-sample size and power properties of the KPSS test, caution should be taken when drawing inference regarding the stationarity hypothesis, especially when a linear trend and serially correlated errors are allowed for. However, the results of this section indicate that there is a possible route for handling the size distortions that

[8] THe power presented is raw power, i.e. no size-adjustment is made.

arise. In order to explore this route, the current paper continues by investigating the sources of the distortions and elaborating on possible remedies.

4 Effects of truncation lag and serial correlation

As seen in the previous section, the size of the KPSS test is affected by two main factors when we allow disturbances to be serially correlated. First and foremost, the truncation lag will affect how appropriate the asymptotic critical values will be in finite samples. Second, given the choice of truncation lag, the autocorrelation of the disturbances will affect the performance of the test. The later dependence is less severe for larger truncation lag, as evident from Tables 1 and Table 2. The fact that nuisance parameters, related to the truncation lag and serial correlation properties of the disturbances, enter the finite-sample distribution of the KPSS test can make inference unreliable. However, one key difference between the two nuisance parameter dimensions offers a potential remedy for the poor finite-sample size that was documented in the previous section.

In order to be able to improve on the size properties of the KPSS test, it would be ideal if the finite-sample distribution of the test depends more on the choice of truncation lag and less on the serial correlation of the disturbances. If such a situation was to be the case, it would be possible to, for each given sample size and truncation lag, obtain an approximate finite-sample distribution of the KPSS test with the prospect of working well even when serial correlation is present. The approximate distribution would be of great help since it would eliminate one of the two effects that work to distort the size of the stationarity test in small and medium-sized samples. By taking the truncation lag into account, the only effect unaccounted for would be that of uncaptured serial correlation. But given that the ideal situation entails that uncaptured serial correlation is a problem that can be regarded as minor once the truncation lag is taken into account, the properties of the test would still have been improved. In order to see if this is a feasible solution to reducing the size distortions of the KPSS test, the small-sample distribution of the test needs to be scrutinized. This can be done in two steps by first considering the behavior of the long-run variance estimator by itself and then considering the role that it plays for the KPSS test statistic.

Consider first the distribution of the long-run variance estimator. To be able to account for the effects that this estimator has on the KPSS test, the effects that arise as a consequence of serial correlation must be sorted out. Hence, the variance estimator is applied to an i.i.d. sequence of random variables. To this end, T i.i.d. $N(0,1)$ variables are generated and detrended with either an intercept or an intercept and a time trend. $\hat{s}(l)$ is then estimated for various choices of l. As above, we consider the case where $l =$

$[k(T/100)^{0.25}]$ with $k \in \{4, 8, 12\}$. Based on 5,000 replications, the empirical cumulative distribution function for the long-run variance estimator is obtained.

The empirical cumulative distribution functions (CDFs), for the cases where only an intercept is present and the sample sizes are T=20 and T=100, are given in Figure 1. In the upper panel of Figure 1, i.e. for the case where T=20, it can be seen that the long-run variance estimator is considerably downward biased. For the case where $k = 4$, about 70% of the probability mass falls below the true value of the long-run variance, which is equal to 1. Considering that the support of the CDF is approximately symmetric around 1, this implies a substantial downward bias in the long-run variance estimator. For $k = 8$ and $k = 12$, the corresponding figures for the probability mass are 75% and 80% respectively. However, the degree of bias displayed when $k = 4$ and $k = 8$ does not affect the distribution of the KPSS test to any larger extent. This was seen when the size was investigated in the upper panel of Table 1. For $k = 4$ and $k = 8$, the test only had small size distortions. However, for the case where $k = 12$, it was seen that the downward bias in the variance estimator induced an upward size distortion of about 4.5 percentage points. For the case where T=100, the bias for the case where $k = 12$ is smaller than the bias for the case where $k = 4$ and T=20. Since no noticeable size distortion was seen in the latter case, we would not expect to find any size distortion in the former either. When consulting Table 1, it can be seen that this assertion is indeed true.

Considering the fact that the downward bias of the distribution of the long-run variance estimator seems to affect the size of the KPSS test, it can be expected that the size distortions found in Table 2 are explained by a more severe bias in the variance estimator when both an intercept and a time trend are accounted for.

When we look at the empirical cumulative distribution functions in Figure 2, we see that the distribution of the long-run variance is indeed more biased when both an intercept and a time trend are present. In the case where $T = 20$ and $k = 4$, it can be seen that about 80% of the probability mass of the long-run variance density falls below the true value. This is approximately the same figure that was found in Figure 1 for $T = 20$ and $k = 12$. Hence, given that the shape of the empirical CDFs are similar in both cases, we would expect to find size distortions that are close in magnitude if the long-run variance estimator plays a central role in the finite-sample distribution of the KPSS test. If Table 1 and Table 2 are once again consulted, and the size when $T = 20$, $k = 12$ and an intercept is present is compared to the size when $T = 20$, $k = 4$ and both an intercept and a time trend are present, it can be seen that the distortions are of similar magnitude. In the former case the size is 14.5%, while in the latter case the size is 13.6%. As before, it can also be seen that the adverse effects of the bias in the long-run variance estimator increase with k and vanishes as T increases.

We are now in a position to say that one should take the bandwidth parameter into account when using the KPSS test to test for stationarity in finite-sample situations. More specifically, the finite-sample distribution of the KPSS statistic, conditional upon T and l, should be used when testing for stationarity. However, in order for this to be of any practical use, the distribution of the KPSS test statistic, conditional on the sample size and truncation lag, should depend relatively little on the remaining nuisance parameters, i.e. serial correlation of the disturbances. This will be investigated next.

In order to illustrate how the distribution of the KPSS test statistic depends on the nuisance parameters relating to the serial correlation and the truncation lag, data is generated as described in the previous section. But instead of calculating the size, we plot the kernel density estimates of the KPSS test. To separate out the influence of the truncation lag and the serial correlation, two sets of kernel densities are obtained. First, the serial correlation is set to zero and the density for the case where $k = 4$ is compared to the density when $k = 12$. This scenario is presented in Figure 3. Then, the bandwidth is kept fixed, equal to $k = 12$, and the kernel density estimates for the cases $\rho = 0.2$, $\rho = 0.4$ and $\rho = 0.6$ are compared. This comparison can be seen in Figure 4.[9]

As seen from Figure 3, the distribution of the KPSS test statistic depends heavily on the truncation lag, l. Of course this is expected given the bias found in the long-run variance estimator.[10] When we turn to Figure 4, on the other hand, it can be seen that once we have accounted for the truncation lag, the test statistic is not very dependent on the degree of serial correlation. This indicates that size distortions indeed can be reduced by taking into account the truncation lag when performing the stationarity test of Kwiatkowski et al. (1992).

From the findings of this section, the applied researcher could be better off by taking into account the bandwidth parameter when testing the consumption series for stationarity. Critical values for a limited choice of time-series dimensions and truncation lags are available in Hornok and Larsson (2000), but no extensive tabulation of critical values is currently available. In the next section, we provide approximate small-sample critical values for the KPSS test. These critical values will be obtained by conditioning the distribution of the test statistic on the sample size and the bandwidth parameter.

---

[9] To obtain the kernel density estimates we use the so-called triweight kernel with a bandwidth of 0.08.

[10] From the upper panel of Figure 3, it can be seen that the significance level chosen when investigating the size of the KPSS test affects how large the size distortions will be for various choices of k. Considering that the asymptotic critical value of the KPSS test is 0.347 on the 10% significance level, it is seen that this choice of significance level is likely to make the figures presented in Table 1 look favorable for the KPSS test. Hence, choosing another significance level should have made the main arguments of the current paper stronger from the outset.

5 Accounting for the finite sample

Based on the evidence found in the previous section, it would be desirable to have critical values available for each possible combination of sample size and truncation lag. Such critical values can be approximated numerically, as suggested by Hornok and Larsson (2000), and tabulated. Tabulations of such critical values would require a considerable amount of space. However, by fitting response surface regressions to simulated critical values, obtained for a wide range of T and l, critical values can be easily calculated for any given sample size and truncation lag.

As seen above, the truncation lag seems to affect the small-sample distribution of the stationarity test more than the serial correlation in the disturbances. Furthermore, the former nuisance parameter is integer-valued and effectively bounded from above by the sample size and from below by zero. The serial correlation, on the other hand, cannot easily be characterized by only a few nuisance parameters, since e.g. the entire class of stationary and invertible ARMA processes are candidates when modelling the disturbance term. The magnitude of this parameter space makes the truncation lag the primary candidate to condition the small-sample distribution on.

In order to obtain small-sample critical values for the KPSS test, data is generated according to $y_t = \alpha_0 + u_t$ or $y_t = \alpha_0 + \alpha_1 t + u_t$, with u_t being i.i.d. N(0,1). The series, y_t, is detrended using appropriate deterministic components and the KPSS test is applied to the detrended series. We obtain 100 critical values, on the 1%, 2.5%, 5% and the 10% significance levels, for each choice of T and l, based on 10,000 replications for each set of critical values. The time-series dimensions and truncation lags considered are $T \in \{20, 21, 22, \ldots, 100\}$ and $l \in \{1, 2, 3, \ldots, T-3\}$, which in total renders critical values for 4617 different combinations of T and l.[11] In order to reduce the influence of simulation variability, the mean over the 100 critical values is calculated for each parameter combination and significance level when running the response surface regressions. After some experimentation with the specification, the response surface regressions in (5) turned out to fit the critical values well.

$$CV_{T,l}^{\alpha} = \gamma_0 + \sum_{i=1}^{4} \gamma_{1,i} T^{-i/2} + \sum_{i=1}^{4} \gamma_{2,i} l^{-i/2} + \sum_{i=1}^{4} \gamma_{3,i} \left(\frac{T}{l}\right)^{-i/2} + \epsilon_{T,l} \tag{5}$$

In (5), $CV_{T,l}^{\alpha}$ is the mean over the 100 critical values, at the α% significance level, for a specific choice of T and l. When the regressions in (5) are estimated, we obtain the

[11] Although critical values can be obtained for truncation lags up to T-3, we recommend that the largest truncation lag used is T-8 in order to obtain a desirable finite-sample distribution of the KPSS test statistic. This will be of importance e.g. when employing the automatic truncation lag selection procedure investigated by Hobijn et al. (2004).

parameters in Table 5.[12] In Table 5, the R^2 measure of regression fit is also presented in order to assess the in-sample fit of the response surface regression.[13]

As seen in Table 5, the fit of the response surface regressions is excellent. The response surface regressions can hence be used to calculate critical values for a specific choice of time-series dimension and truncation lag.[14]

6 Finite-Sample performance

As we saw in the previous section, the response surface regressions have a good in-sample fit. The next step is to study how well the stationarity test performs when employing the finite-sample critical values and compare the performance to the case when asymptotic critical values are used for inference. From the results in Section 5, it can be expected that the use of finite-sample critical values reduces the size distortions that arise when disturbances are serially correlated under the null hypothesis. This assertion is studied both in the context of deterministic and data-dependent truncation lag determination. Moreover, the power of the finite-sample test is studied.

6.1 Size of the stationarity test

6.1.1 Deterministic choice of truncation lag

When a deterministic bandwidth selection procedure is employed, l is usually set as a function of T. In the previous literature, the selection rule $l = [k(T/100)^{0.25}]$, where $k \in \{4, 8, 12\}$, is frequently used, see e.g. Schwert (1989). As seen in Section 3, the KPSS test worked poorly when this rule was used and asymptotic critical values were employed. In order to see if the response surface regressions produce critical values that render better size properties, we begin by rerunning the size simulation of Section 3, with the modification that the finite-sample critical values, obtained from the parameters of Table 5, are employed instead of the asymptotic critical values supplied by Kwiatkowski

[12]The response surface regression in (5) is estimated using the OLS estimator. This estimator turned out to be more stable than the GMM estimator of Cragg (1983) as considered by e.g. MacKinnon (1994).

[13]Contrary to e.g. MacKinnon (1994), the sole purpose of the response surface parameters in Table 5 is to provide a good in-sample fit. Hence, the asymptotic critical values of the test will not, as an example, be represented by the intercepts.

[14]It should be noted that the critical values are to be calculated from the response surface regressions only when $l > 0$. When $l = 0$, no truncation lag is used and the critical values of Sephton (1995) can be applied.

et al. (1992). In Table 6 and Table 7, the resulting finite-sample size of the KPSS test is presented.

The first thing that stands out when the results of Table 6 and Table 7 are investigated is that the size is close to the 10% level when no serial correlation is present, i.e. when $\rho = 0$ and $\theta = 0$. This is of course an expected property since the critical values where obtained especially for these circumstances. Nevertheless, the size distortions introduced by the mere application of a long-run variance estimator are eliminated, and hence there is room for possible improvement of the performance of the stationarity test even when serial correlation is present, i.e. when $\rho \neq 0$ or $\theta \neq 0$. As we move down in Table 6 and Table 7, it can be seen that as the degree of serial dependence, as described by the AR and MA parameters, increases, so do the size distortions of the test. However, the most important feature that can be observed when looking at the simulation results in the tables, and comparing them to the results of Table 1 and Table 2, is that the test that employs the finite-sample critical values outperforms the test that employs the asymptotic critical values of Kwiatkowski et al. (1992).

To further illustrate the effects of accounting for the finite sample distribution of the KPSS test, data is also generated under the null hypothesis while letting the stationary disturbance term follow an AR(1) process. The AR parameter takes the values $\rho \in \{0.0, 0.1, \ldots, 0.9\}$, while, as before, the innovation driving the AR process is distributed i.i.d. $N(0, 1)$. Data is generated for $T \in \{20, 30, \ldots, 100\}$ and 50,000 data sets are generated for each parameter setup. The KPSS test is then applied to the simulated data sets, while letting $k = 12$, and the 10% size is calculated using both finite-sample and asymptotic critical values.[15]

In Figure 5, a size comparison is plotted for the case where only an intercept is present. From Figure 5, it can be seen that the size of the KPSS test, when based on finite-sample critical values, is better than the size of the test based on asymptotic critical values when T=20. For larger values of T, the two tests perform equally well with the differences in performance being attributable to simulation error. These results are in line with what we would expect from the results in Table 1, where it was seen that size distortions were present only for T=20. If we look at Table 2, on the other hand, it is expected that there are more to gain in terms of reduced size distortions when both an intercept and a time trend is allowed for.

In Figure 6, the size comparison is presented for the case where both an intercept and a time trend is present. As expected from Table 2, the KPSS test employing asymptotic critical values is heavily size distorted. However, when using finite-sample critical values

[15]The main conclusions regarding the size are unaltered when the truncation lag scaling parameter is set to $k = 4$ or $k = 8$.

it can be seen that the size distortion, to a large extent, can be reduced. Contrary to the case where only an intercept is present, the size performance of the KPSS test can be enhanced for sample sizes up to T=90 when both an intercept and a trend are present. From studying Figure 6 it is seen that the finite-sample critical values are preferable over the asymptotic critical values for the situations where $T \leq 90$ and for all ρ considered. For the case where $T = 100$ the tests perform equally well across the various values of the autoregressive parameter.

The overall conclusion, when considering deterministic bandwidth selection, is that the size distortions, that arise when performing the KPSS test with asymptotic critical values, can be reduced by employing the finite-sample critical values. The reduced size distortions are achieved by accounting for the detrimental effects of the long-run variance estimator through the use of finite-sample critical values that depend, not only on the sample size, but also on the truncation lag.

The results presented in this subsection rely on a bandwidth choice that is a deterministic function of T. In practice this can induce some problems since we do not know a priori which choice of k that is appropriate for a specific empirical situation. To come to terms with this problem, the researcher can, following Newey and West (1994), let the truncation lag selection depend on the characteristics of the data material at hand.

6.1.2 Data-dependent choice of truncation lag

In the previous subsection, it was seen that applying finite-sample critical values improved the performance of the KPSS test compared to the case where asymptotic critical values were used. The use of deterministic truncation lag selection only requires tabulations of critical values for a restricted set of truncation lag parameters since it is deterministically related to the sample size. However, the response surface parameters in Table 5 can be employed to obtain critical values, not only for the values of l implied by $k \in \{4, 8, 12\}$, but for every $T \in \{20, 21, \ldots, 100\}$ and $l \in \{1, 2, \ldots, T-8\}$. Hence, the data-dependent selection procedure of Newey and West (1994) can be implemented without restricting the truncation lag parameter space and, perhaps, without disturbing the size of the KPSS test. The latter issue will be investigated next.

In order to study the size of the KPSS test when using a data-dependent method for truncation lag selection, data is generated as in the previous subsection. The KPSS test is then performed while setting l according to (6)-(10) below (see e.g. Newey and West, 1994; Hobijn et al., 2004).

$$l = min\left([\hat{\gamma}T^{1/3}], \left[12\left(\frac{T}{100}\right)^{0.25}\right]\right) \tag{6}$$

$$\hat{\gamma} = 1.1447\left(\left(\frac{\hat{s}^{(1)}}{\hat{s}^{(0)}}\right)^2\right)^{1/3} \tag{7}$$

$$\hat{s}^{(0)} = T^{-1}\sum_{t=1}^{T} e_t^2 + 2T^{-1}\sum_{s=1}^{n_T}\sum_{t=s+1}^{T} e_t e_{t-s} \tag{8}$$

$$\hat{s}^{(1)} = 2T^{-1}\sum_{s=1}^{n_T} s \sum_{t=s+1}^{T} e_t e_{t-s} \tag{9}$$

$$n_T = \left[12\left(\frac{T}{100}\right)^{2/9}\right] \tag{10}$$

As noted from (6) above, we do not allow the truncation lag to take values above $[12(T/100)^{0.25}]$. This restriction ensures that the stationarity test is consistent, i.e. has an asymptotic power that is greater than the size of the test (see e.g. Carrion-i-Silvestre and Sansó, 2006).[16]

In Figure 7 and Figure 8 we depict the size of the KPSS test when the data-dependent truncation lag selection procedure is applied. As above, the size is given for different sample sizes and autoregressive parameters.

If we start looking at the size of the test including an intercept only, we see from Figure 7 that the size is upward biased regardless of whether finite-sample or asymptotic critical values are used. The figure also indicates that there are relatively small gains from using finite-sample critical values when an intercept is the only deterministic component that occurs in the data-generating process and the test. This result is not expected to apply when the time series contains both an intercept and a trend.

As expected from the experience of the previous sections, we see in Figure 8 that the principal pattern from Figure 7 is reinforced even more when both an intercept and a time trend are allowed for. The size of the KPSS test turns out to have a severe upward bias when asymptotic critical values are used, while the test based on finite-sample critical values performs better.

From this section we have seen that the KPSS test has an upward size distortion in small samples when serial correlation is allowed for even though no serial correlation is present. We have also seen that the size performance of the KPSS test can be enhanced by taking this fact into account and use critical values that are conditioned on, not only the sample size, but also the truncation lag used in the test.

[16]The author is grateful to an anonymous referee for pointing this out.

In the following subsection, we will investigate how the power of the KPSS test is affected when critical values are conditioned on the time-series dimension and the truncation lag.

6.2 Power of the stationarity test

As seen from Section 3, the KPSS test has an upward size distortion in small samples when asymptotic critical values are used and serially correlated disturbances are allowed for. It was also seen that the size distortions could be so large that the power fell below the size and the KPSS test became biased. In this subsection, we will investigate the power properties of the KPSS test when small-sample critical values are used.

To investigate the power of the KPSS test, we generate data under the alternative hypothesis (as described in Section 3) and apply the KPSS test using the finite sample critical values of Section 5. The power of the tests on the 10% significance level is presented in Table 8 and Table 9.[17]

As seen from the Table 8, the power of the KPSS test is well above the size for all situations considered when an intercept is the only deterministic component included. However, as was the case with the power of the KPSS test utilizing asymptotic critical values, the power of the test falls below the size when T is small, k is large and an intercept and a time trend are included in the econometric model. As seen from Table 9, the power falls below the size when $T = 20$ with $k \in \{8, 12\}$ and $T = 30$ with $k = 12$. However, for all the other parameter combinations considered, the KPSS test, utilizing finite-sample critical values calculated from the response surface regression parameters in Table 5, is unbiased.

Having studied the behavior of the KPSS test with respect to how well the asymptotic approximation works in small samples when disturbances are allowed to be serially correlated under the null hypothesis, it is interesting to see whether or not accounting for the finite-sample has any implications for empirical application of the test. This issue is investigated in the next section.

7 Empirical illustration

In order to illustrate how inappropriate use of asymptotic critical values affect inference about the stationarity hypothesis, we set out to investigate the permanent income

[17]The finite-sample critical values are used when investigating the power. Hence, the power is not size-adjusted for the cases where serial correlation is present.

hypothesis. More specifically, assume that we are to investigate whether or not the permanent income hypothesis has any empirical support. To do this, it is first assumed that a representative agent maximizes the expected lifetime utility, $\mathcal{U}_t$, given in (11).

$$\mathcal{U}_t = \sum_{i=0}^{\infty} \beta^i E_t u(c_{t+i}) \tag{11}$$

In (11), $\beta = 1/(1+\delta)$ is a subjective discount factor, with δ being the subjective discount rate. Furthermore, $u(\cdot)$ is the instantaneous utility function, which is assumed to be quadratic in private consumption, c_t, i.e. $u(c_t) = -0.5(\bar{c} - c_t)^2$.[18] If it is assumed that the permanent income hypothesis holds and assumed that the subjective discount rate is equal to the interest rate, private consumption can, following Hall (1978), be shown to follow a random walk.

To test the whether or not the permanent income hypothesis has any empirical support, we can use the KPSS test for stationarity as put forward in the previous sections. Under the null hypothesis, the stationarity test stipulates consumption is stationary and that the permanent income implications discussed above has no empirical support. If the null hypothesis is rejected on the other hand, the stationarity test indicates that the predictions of the permanent income hypothesis are not violated.

In Table 10, the stationarity test results for real per capita consumption in 22 OECD countries are presented.[19] The stationarity tests are performed using a linear trend, while the truncation lag is determined either by the data-dependent rule in (6)-(7) with $l = min\left([\hat{\gamma}T^{1/3}], \left[k\,(T/100)^{0.25}\right]\right)$ and $n_T = \left[k\,(T/100)^{2/9}\right]$ or by the deterministic rule according to $l = [k(T/100)^{0.25}]$, where in all cases $k = 18$.[20]

As seen in Table 10, the test results are mixed for the various countries. When using the deterministic bandwidth selection procedure asymptotic critical values, the null hypothesis of stationarity is rejected for all countries except Australia. Hence, the results indicate that, for almost all countries, the permanent income hypothesis seems to be plausible. However, when using the finite-sample critical values, the conclusion regarding the permanent income hypothesis is reversed for nine of the countries. The results for almost half of the countries are hence dependent of the fact that the finite-sample behavior of the KPSS test can differ from the asymptotic behavior.

This result is even more accentuated when the deterministic bandwidth selection procedure is used. Under these circumstances, the conclusion is reversed for all but one of the countries in the sample.

[18] $\bar{c}$ is used to denote the bliss level of consumption.

[19] The data is obtained from OECD Economic Outlook Database, No 78, Vol 2005, release 02.

[20] The value for k is chosen for illustrative purposes only.

The results from the empirical illustration show a considerable heterogeneity between the countries investigated with respect to conclusions regarding private consumption determination. More importantly, however, the results indicate that the failure to take the finite-sample distribution of the KPSS test into account can have large effects on conclusions drawn about economic hypotheses.

8 Concluding remarks

In this paper we have studied the small-sample behavior of the KPSS test for stationarity. We have investigated the case where samples are small or medium-sized and serial correlation is allowed for. It was found that the finite-sample size distortions that arise in the stationarity test are by and large a consequence of the poor properties of long-run variance estimator when applied to small samples. This applies specifically when a linear trend is present. We find that the size distortions can be controlled in small and medium-sized samples by conditioning the distribution of the KPSS test on the sample size and the choice of truncation lag. However, a considerable loss of power against a non-stationary alternative follows in the path of the controlled size. Sometimes the loss of power can be so severe that the rejection frequencies under the alternative are below the rejection frequencies under the null, i.e. the test becomes biased.[21] In order to restore the power of the KPSS test, a viable way for future research could perhaps be to find a long-run variance estimator that is in some sense optimal specifically for the KPSS test, more precise in the estimation of the long-run variance and more insensitive to the choice of truncation lag. Finally, by applying the main findings of this paper in an empirical testing problem, the adverse effects of failing to control for a limited sample size, while allowing for serially correlated errors, are illustrated.

As a final remark, it should be noted that the many of the results of the current paper are obtained by varying k when setting l according to $l = [k(T/100)^a]$. Exactly the same results apply when selecting l by varying a for a fixed k or when simultaneously varying k and a. The main factor driving the results is the value of l.[22]

[21] All of the conclusions presented here are obtained with the minimum time-series dimensions being T=20. Hence, caution should be taken when considering even shorter time series.

[22] The author thanks an anonymous referee for accentuating this aspect of the lag selection procedure.

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., Chinn, M. D., and Tran, T. (1995). How Sensitive are Trends to Data Definitions? Results for East Asian and G5 Countries. *Applied Economics Letters*, 2:1–6.

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.

Cragg, J. G. (1983). More Efficient Estimation in the Presence of Heteroscedasticity of Unknown Form. *Econometrica*, 51(3):751–764.

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(366):437–431.

Engle, R. F. and Granger, C. W. J. (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica*, 55(2):251–276.

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). Finite-Sample Stability of the KPSS Test. Working Paper 2006:23, Department of Economics, Lund University, Sweden.

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.

Lee, J. (1996). On the Power of Stationarity Tests Using Optimal Bandwith Estimates. *Economics Letters*, 51:131–137.

MacKinnon, J. G. (1994). Approximate Asymptotic Distribution Functions for Unit-Root and Cointegration Tests. *Journal of Business and Economic Statistics*, 12(2):167–176.

Müller, U. K. (2005). Size and Power of Tests of Stationarity in Highly Autocorrelated Time Series. *Journal of Econometrics*, 128:195–213.

Newey, W. K. and West, K. D. (1987). A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, 55(3):703–708.

Newey, W. K. and West, K. D. (1994). Automatic Lag Selection in Covariance Matrix Estimation. *Review of Economic Studies*, 61:631–653.

Phillips, P. C. B. and Perron, P. (1988). Testing for a Unit Root in Time Series Regression. *Biometrica*, 75:335–346.

Phillips, P. C. B. and Solo, V. (1992). Asymptotics for Linear Processes. *The Annals of Statistics*, 20(2):971–1001.

Said, S. E. and Dickey, D. A. (1984). Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order. *Biometrika*, 71(3):599–607.

Schwert, W. G. (1989). Tests for Unit Roots: A Monte Carlo Investigation. *Journal of Business and Economic Statistics*, 7(2):147–159.

Sephton, P. S. (1995). Response Surface Estimates of the KPSS Stationarity Test. *Economics Letters*, 47:255–261.

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 of KPSS test, asymptotic CV, intercept only.[a]

	$\rho = 0$				$\theta = 0$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.095	0.095	0.145	20	0.095	0.095	0.145
30	0.102	0.104	0.100	30	0.102	0.104	0.100
40	0.096	0.091	0.095	40	0.096	0.091	0.095
50	0.097	0.100	0.099	50	0.097	0.100	0.099
75	0.103	0.099	0.102	75	0.103	0.099	0.102
100	0.098	0.097	0.099	100	0.098	0.097	0.099
	$\rho = 0.2$				$\theta = 0.2$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.129	0.113	0.160	20	0.117	0.109	0.157
30	0.134	0.117	0.112	30	0.123	0.114	0.109
40	0.113	0.106	0.103	40	0.106	0.103	0.099
50	0.125	0.112	0.112	50	0.116	0.108	0.108
75	0.116	0.102	0.100	75	0.106	0.099	0.097
100	0.120	0.109	0.105	100	0.113	0.105	0.102
	$\rho = 0.4$				$\theta = 0.4$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.173	0.129	0.172	20	0.130	0.113	0.162
30	0.186	0.139	0.124	30	0.137	0.120	0.113
40	0.155	0.128	0.115	40	0.116	0.108	0.105
50	0.166	0.135	0.125	50	0.127	0.112	0.111
75	0.156	0.120	0.110	75	0.116	0.102	0.099
100	0.152	0.123	0.115	100	0.120	0.108	0.105
	$\rho = 0.6$				$\theta = 0.6$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.254	0.168	0.184	20	0.135	0.117	0.165
30	0.284	0.185	0.150	30	0.144	0.123	0.116
40	0.233	0.166	0.145	40	0.119	0.111	0.107
50	0.245	0.174	0.147	50	0.133	0.116	0.114
75	0.229	0.153	0.129	75	0.119	0.105	0.099
100	0.213	0.156	0.135	100	0.122	0.110	0.105
	$\rho = 0.8$				$\theta = 0.8$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.411	0.252	0.219	20	0.138	0.119	0.168
30	0.466	0.294	0.211	30	0.148	0.123	0.116
40	0.401	0.277	0.212	40	0.121	0.112	0.105
50	0.414	0.287	0.212	50	0.135	0.118	0.116
75	0.432	0.260	0.196	75	0.122	0.106	0.099
100	0.387	0.251	0.198	100	0.123	0.111	0.106

Notes: [a] ρ and θ are the AR and MA parameters, respectively.

Table 2: Size of KPSS test, asymptotic CV, intercept and trend.[a]

	$\rho = 0$				$\theta = 0$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.136	0.697	1.000	20	0.136	0.697	1.000
30	0.109	0.198	0.718	30	0.109	0.198	0.718
40	0.115	0.164	0.400	40	0.115	0.164	0.400
50	0.108	0.126	0.262	50	0.108	0.126	0.262
75	0.107	0.109	0.139	75	0.107	0.109	0.139
100	0.097	0.099	0.110	100	0.097	0.099	0.110
	$\rho = 0.2$				$\theta = 0.2$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.163	0.665	1.000	20	0.150	0.672	1.000
30	0.140	0.208	0.672	30	0.126	0.206	0.680
40	0.134	0.160	0.385	40	0.125	0.156	0.388
50	0.138	0.140	0.269	50	0.129	0.135	0.270
75	0.134	0.128	0.150	75	0.124	0.121	0.148
100	0.130	0.119	0.124	100	0.123	0.114	0.123
	$\rho = 0.4$				$\theta = 0.4$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.214	0.609	1.000	20	0.164	0.649	1.000
30	0.205	0.222	0.627	30	0.140	0.209	0.664
40	0.178	0.179	0.374	40	0.137	0.164	0.384
50	0.190	0.163	0.269	50	0.140	0.143	0.270
75	0.191	0.149	0.160	75	0.133	0.127	0.151
100	0.170	0.137	0.136	100	0.130	0.119	0.124
	$\rho = 0.6$				$\theta = 0.6$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.306	0.579	1.000	20	0.171	0.637	1.000
30	0.324	0.252	0.589	30	0.146	0.212	0.655
40	0.272	0.218	0.353	40	0.141	0.165	0.383
50	0.285	0.208	0.267	50	0.147	0.146	0.272
75	0.315	0.190	0.180	75	0.140	0.131	0.151
100	0.266	0.176	0.157	100	0.136	0.122	0.126
	$\rho = 0.8$				$\theta = 0.8$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.434	0.586	1.000	20	0.177	0.630	1.000
30	0.514	0.347	0.561	30	0.148	0.211	0.653
40	0.464	0.320	0.370	40	0.144	0.166	0.385
50	0.507	0.325	0.318	50	0.151	0.145	0.272
75	0.563	0.320	0.246	75	0.141	0.132	0.152
100	0.505	0.314	0.236	100	0.138	0.123	0.127

Notes: [a] ρ and θ are the AR and MA parameters, respectively.

Table 3: Power of KPSS test, asymptotic CV, intercept only.[a]

	$\rho = 0.0$				$\theta = 0.0$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.599	0.465	0.366	20	0.599	0.465	0.366
30	0.708	0.578	0.483	30	0.708	0.578	0.483
40	0.734	0.629	0.551	40	0.734	0.629	0.551
50	0.778	0.667	0.573	50	0.778	0.667	0.573
75	0.878	0.739	0.654	75	0.878	0.739	0.654
100	0.890	0.766	0.693	100	0.890	0.766	0.693
	$\rho = 0.2$				$\theta = 0.2$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.594	0.455	0.359	20	0.593	0.455	0.359
30	0.702	0.577	0.481	30	0.702	0.579	0.481
40	0.713	0.615	0.539	40	0.714	0.615	0.539
50	0.776	0.670	0.570	50	0.776	0.669	0.570
75	0.879	0.729	0.647	75	0.878	0.729	0.647
100	0.895	0.779	0.707	100	0.895	0.779	0.707
	$\rho = 0.4$				$\theta = 0.4$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.587	0.449	0.348	20	0.582	0.450	0.351
30	0.693	0.568	0.472	30	0.693	0.572	0.477
40	0.708	0.608	0.531	40	0.709	0.612	0.535
50	0.768	0.665	0.566	50	0.770	0.666	0.568
75	0.873	0.720	0.645	75	0.874	0.723	0.645
100	0.891	0.778	0.704	100	0.893	0.778	0.706
	$\rho = 0.6$				$\theta = 0.6$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.580	0.432	0.342	20	0.571	0.439	0.343
30	0.684	0.547	0.456	30	0.682	0.563	0.470
40	0.694	0.592	0.521	40	0.703	0.606	0.531
50	0.756	0.650	0.553	50	0.764	0.664	0.564
75	0.863	0.709	0.635	75	0.870	0.721	0.644
100	0.882	0.767	0.698	100	0.890	0.777	0.705
	$\rho = 0.8$				$\theta = 0.8$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.594	0.437	0.332	20	0.555	0.426	0.337
30	0.676	0.529	0.432	30	0.670	0.555	0.461
40	0.678	0.570	0.489	40	0.695	0.602	0.527
50	0.742	0.625	0.523	50	0.757	0.659	0.562
75	0.841	0.685	0.612	75	0.865	0.716	0.642
100	0.866	0.751	0.677	100	0.887	0.774	0.701

Notes: [a] ρ and θ are the AR and MA parameters, respectively.

Table 4: Power of KPSS test, asymptotic CV, intercept and trend.[a]

	$\rho = 0.0$				$\theta = 0.0$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.448	0.656	1.000	20	0.448	0.656	1.000
30	0.590	0.452	0.618	30	0.590	0.452	0.618
40	0.632	0.491	0.497	40	0.632	0.491	0.497
50	0.733	0.550	0.483	50	0.733	0.550	0.483
75	0.857	0.659	0.524	75	0.857	0.659	0.524
100	0.889	0.734	0.589	100	0.889	0.734	0.589
	$\rho = 0.2$				$\theta = 0.2$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.443	0.631	1.000	20	0.441	0.633	1.000
30	0.591	0.457	0.630	30	0.589	0.457	0.631
40	0.621	0.471	0.490	40	0.620	0.473	0.490
50	0.710	0.528	0.474	50	0.710	0.529	0.474
75	0.853	0.647	0.514	75	0.854	0.647	0.515
100	0.881	0.725	0.587	100	0.882	0.726	0.586
	$\rho = 0.4$				$\theta = 0.4$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.437	0.618	1.000	20	0.429	0.631	1.000
30	0.585	0.445	0.625	30	0.573	0.449	0.630
40	0.606	0.458	0.488	40	0.608	0.465	0.488
50	0.699	0.514	0.469	50	0.699	0.520	0.472
75	0.841	0.634	0.505	75	0.846	0.640	0.510
100	0.873	0.719	0.579	100	0.878	0.721	0.583
	$\rho = 0.6$				$\theta = 0.6$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.444	0.599	1.000	20	0.412	0.629	1.000
30	0.585	0.430	0.608	30	0.555	0.440	0.632
40	0.597	0.450	0.478	40	0.591	0.457	0.488
50	0.686	0.498	0.458	50	0.684	0.511	0.470
75	0.828	0.616	0.493	75	0.834	0.633	0.505
100	0.862	0.702	0.574	100	0.872	0.716	0.579
	$\rho = 0.8$				$\theta = 0.8$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.483	0.592	1.000	20	0.395	0.628	1.000
30	0.624	0.439	0.586	30	0.530	0.433	0.636
40	0.614	0.450	0.466	40	0.577	0.452	0.491
50	0.688	0.489	0.438	50	0.670	0.504	0.466
75	0.817	0.587	0.463	75	0.822	0.624	0.502
100	0.845	0.669	0.544	100	0.866	0.713	0.577

Notes: [a] ρ and θ are the AR and MA parameters, respectively.

Table 5: Estimated response surface parameters.[a]

Intercept														
α	γ_0	$\gamma_{1,1}$	$\gamma_{1,2}$	$\gamma_{1,3}$	$\gamma_{1,4}$	$\gamma_{2,1}$	$\gamma_{2,2}$	$\gamma_{2,3}$	$\gamma_{2,4}$	$\gamma_{3,1}$	$\gamma_{3,2}$	$\gamma_{3,3}$	$\gamma_{3,4}$	R^2
1%	-0.1609	-6.6240	25.2718	-58.4332	62.0085	4.1571	-8.6749	9.0484	-3.5124	4.3474	-13.6537	18.0637	-8.0172	0.9875
2.5%	-0.0088	-4.0376	14.9380	-32.6497	32.6291	2.5905	-5.4183	5.6538	-2.1893	3.0214	-9.3154	12.0964	-5.2534	0.9822
5%	0.1274	-2.2115	8.2267	-17.5773	17.3330	1.4319	-3.0234	3.1663	-1.2251	1.8007	-5.4890	7.0521	-2.9745	0.9874
10%	0.2749	-0.4687	1.8420	-3.0531	2.3207	0.3199	-0.7192	0.7597	-0.2900	0.4540	-1.5017	2.0453	-0.7776	0.9978

Intercept and trend														
α	γ_0	$\gamma_{1,1}$	$\gamma_{1,2}$	$\gamma_{1,3}$	$\gamma_{1,4}$	$\gamma_{2,1}$	$\gamma_{2,2}$	$\gamma_{2,3}$	$\gamma_{2,4}$	$\gamma_{3,1}$	$\gamma_{3,2}$	$\gamma_{3,3}$	$\gamma_{3,4}$	R^2
1%	0.8504	4.0178	-15.5209	37.2153	-32.3373	-2.6506	6.1716	-7.3830	3.1784	-2.5998	2.2661	2.1137	-2.2215	0.9957
2.5%	0.4268	1.6583	-6.6597	17.4434	-14.4266	-1.1164	2.8416	-3.6649	1.6543	-0.7457	-1.2694	5.2507	-3.2130	0.9975
5%	0.1087	-0.1841	0.3283	1.7458	-0.0718	0.0807	0.2222	-0.7199	0.4398	0.5784	-3.6398	7.1982	-3.7626	0.9987
10%	-0.1655	-1.8063	6.5980	-12.6246	13.4028	1.1392	-2.1348	1.9671	-0.6799	1.6451	-5.3563	8.3673	-3.9757	0.9996

Notes: [a] In the table, α denotes the significance level.

Table 6: Size of KPSS test, finite-sample CV, intercept only.[a]

	$\rho = 0.0$				$\theta = 0.0$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.107	0.108	0.097	20	0.107	0.108	0.097
30	0.099	0.096	0.093	30	0.099	0.096	0.093
40	0.100	0.097	0.099	40	0.100	0.097	0.099
50	0.097	0.098	0.097	50	0.097	0.098	0.097
75	0.103	0.106	0.105	75	0.103	0.106	0.105
100	0.091	0.090	0.090	100	0.091	0.090	0.090
	$\rho = 0.2$				$\theta = 0.2$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.124	0.102	0.103	20	0.115	0.098	0.099
30	0.135	0.118	0.113	30	0.121	0.114	0.110
40	0.126	0.118	0.115	40	0.116	0.114	0.112
50	0.123	0.116	0.113	50	0.115	0.108	0.109
75	0.129	0.118	0.109	75	0.119	0.114	0.108
100	0.121	0.110	0.112	100	0.114	0.107	0.109
	$\rho = 0.4$				$\theta = 0.4$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.173	0.122	0.111	20	0.127	0.104	0.103
30	0.185	0.137	0.131	30	0.137	0.119	0.113
40	0.164	0.139	0.130	40	0.126	0.119	0.118
50	0.162	0.137	0.125	50	0.123	0.114	0.111
75	0.170	0.133	0.123	75	0.132	0.117	0.110
100	0.154	0.128	0.123	100	0.119	0.110	0.112
	$\rho = 0.6$				$\theta = 0.6$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.254	0.157	0.124	20	0.134	0.107	0.108
30	0.274	0.182	0.152	30	0.144	0.121	0.115
40	0.243	0.186	0.157	40	0.132	0.123	0.118
50	0.239	0.176	0.149	50	0.127	0.118	0.112
75	0.249	0.169	0.145	75	0.135	0.119	0.113
100	0.219	0.162	0.147	100	0.123	0.113	0.112
	$\rho = 0.8$				$\theta = 0.8$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.411	0.246	0.142	20	0.138	0.109	0.111
30	0.455	0.293	0.212	30	0.147	0.123	0.117
40	0.403	0.286	0.221	40	0.133	0.123	0.120
50	0.416	0.291	0.215	50	0.130	0.118	0.113
75	0.454	0.269	0.205	75	0.137	0.120	0.114
100	0.385	0.255	0.203	100	0.125	0.113	0.113

Notes: [a] ρ and θ are the AR and MA parameters, respectively.

Table 7: Size of KPSS test, finite-sample CV, intercept and trend.[a]

	$\rho = 0.0$				$\theta = 0.0$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.105	0.079	0.129	20	0.105	0.079	0.129
30	0.109	0.104	0.070	30	0.109	0.104	0.070
40	0.101	0.107	0.076	40	0.101	0.107	0.076
50	0.101	0.109	0.096	50	0.101	0.109	0.096
75	0.095	0.102	0.112	75	0.095	0.102	0.112
100	0.085	0.096	0.118	100	0.085	0.096	0.118
	$\rho = 0.2$				$\theta = 0.2$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.129	0.068	0.117	20	0.118	0.074	0.120
30	0.144	0.116	0.068	30	0.129	0.115	0.070
40	0.137	0.120	0.079	40	0.126	0.116	0.079
50	0.123	0.111	0.091	50	0.113	0.108	0.091
75	0.124	0.117	0.125	75	0.113	0.113	0.123
100	0.104	0.101	0.115	100	0.095	0.098	0.112
	$\rho = 0.4$				$\theta = 0.4$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.174	0.051	0.111	20	0.128	0.069	0.121
30	0.208	0.125	0.056	30	0.148	0.118	0.070
40	0.186	0.138	0.076	40	0.138	0.122	0.082
50	0.174	0.135	0.098	50	0.124	0.112	0.095
75	0.173	0.137	0.134	75	0.124	0.116	0.125
100	0.145	0.120	0.128	100	0.104	0.102	0.116
	$\rho = 0.6$				$\theta = 0.6$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.251	0.032	0.096	20	0.133	0.069	0.125
30	0.319	0.150	0.041	30	0.157	0.118	0.071
40	0.278	0.168	0.072	40	0.145	0.123	0.081
50	0.272	0.179	0.101	50	0.128	0.116	0.096
75	0.292	0.174	0.149	75	0.129	0.121	0.124
100	0.232	0.155	0.151	100	0.107	0.104	0.118
	$\rho = 0.8$				$\theta = 0.8$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.372	0.020	0.071	20	0.138	0.069	0.126
30	0.503	0.222	0.019	30	0.160	0.118	0.072
40	0.467	0.259	0.072	40	0.147	0.127	0.082
50	0.497	0.302	0.132	50	0.132	0.116	0.097
75	0.551	0.291	0.211	75	0.132	0.122	0.123
100	0.464	0.274	0.222	100	0.109	0.106	0.119

Notes: [a] ρ and θ are the AR and MA parameters, respectively.

Table 8: Power of KPSS test, finite-sample CV, intercept only.[a]

	$\rho = 0.0$				$\theta = 0.0$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.608	0.460	0.249	20	0.608	0.460	0.249
30	0.694	0.580	0.479	30	0.694	0.580	0.479
40	0.728	0.622	0.541	40	0.728	0.622	0.541
50	0.769	0.666	0.572	50	0.769	0.666	0.572
75	0.876	0.732	0.649	75	0.876	0.732	0.649
100	0.892	0.760	0.690	100	0.892	0.760	0.690
	$\rho = 0.2$				$\theta = 0.2$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.600	0.458	0.249	20	0.600	0.460	0.249
30	0.684	0.566	0.464	30	0.683	0.569	0.463
40	0.718	0.615	0.538	40	0.718	0.615	0.537
50	0.773	0.671	0.584	50	0.774	0.672	0.584
75	0.869	0.735	0.659	75	0.869	0.735	0.660
100	0.889	0.769	0.696	100	0.890	0.768	0.696
	$\rho = 0.4$				$\theta = 0.4$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.591	0.439	0.247	20	0.589	0.447	0.247
30	0.675	0.552	0.455	30	0.677	0.559	0.461
40	0.709	0.606	0.527	40	0.711	0.611	0.533
50	0.767	0.665	0.579	50	0.771	0.668	0.581
75	0.864	0.731	0.657	75	0.867	0.733	0.658
100	0.886	0.764	0.695	100	0.888	0.766	0.697
	$\rho = 0.6$				$\theta = 0.6$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.579	0.422	0.241	20	0.575	0.434	0.246
30	0.662	0.535	0.434	30	0.666	0.549	0.453
40	0.696	0.591	0.517	40	0.704	0.602	0.526
50	0.758	0.651	0.564	50	0.764	0.663	0.579
75	0.855	0.721	0.649	75	0.862	0.731	0.657
100	0.878	0.754	0.688	100	0.885	0.765	0.695
	$\rho = 0.8$				$\theta = 0.8$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.585	0.420	0.230	20	0.559	0.419	0.241
30	0.662	0.523	0.419	30	0.651	0.540	0.444
40	0.679	0.567	0.487	40	0.695	0.595	0.524
50	0.745	0.634	0.534	50	0.757	0.657	0.572
75	0.846	0.704	0.624	75	0.855	0.728	0.654
100	0.864	0.743	0.669	100	0.883	0.762	0.692

Notes: [a] ρ and θ are the AR and MA parameters, respectively.

Table 9: Power of KPSS test, finite-sample CV, intercept and trend.[a]

	$\rho = 0.0$				$\theta = 0.0$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.387	0.024	0.060	20	0.387	0.024	0.060
30	0.594	0.341	0.017	30	0.594	0.341	0.017
40	0.634	0.411	0.114	40	0.634	0.411	0.114
50	0.727	0.512	0.276	50	0.727	0.512	0.276
75	0.852	0.637	0.490	75	0.852	0.637	0.490
100	0.873	0.703	0.572	100	0.873	0.703	0.572
	$\rho = 0.2$				$\theta = 0.2$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.391	0.025	0.052	20	0.388	0.026	0.053
30	0.570	0.326	0.019	30	0.569	0.327	0.019
40	0.626	0.426	0.125	40	0.627	0.427	0.125
50	0.718	0.521	0.279	50	0.720	0.521	0.279
75	0.843	0.644	0.493	75	0.845	0.645	0.493
100	0.863	0.696	0.573	100	0.864	0.698	0.574
	$\rho = 0.4$				$\theta = 0.4$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.382	0.022	0.062	20	0.372	0.026	0.058
30	0.564	0.317	0.020	30	0.556	0.319	0.021
40	0.613	0.409	0.120	40	0.613	0.417	0.123
50	0.700	0.509	0.271	50	0.706	0.515	0.277
75	0.833	0.630	0.487	75	0.837	0.635	0.490
100	0.854	0.690	0.567	100	0.861	0.693	0.571
	$\rho = 0.6$				$\theta = 0.6$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.391	0.019	0.060	20	0.354	0.028	0.063
30	0.571	0.306	0.019	30	0.539	0.311	0.023
40	0.598	0.392	0.113	40	0.598	0.408	0.120
50	0.684	0.488	0.262	50	0.689	0.504	0.272
75	0.821	0.609	0.474	75	0.828	0.628	0.487
100	0.838	0.670	0.558	100	0.852	0.689	0.567
	$\rho = 0.8$				$\theta = 0.8$		
T	k=4	k=8	k=12	T	k=4	k=8	k=12
20	0.430	0.014	0.053	20	0.336	0.031	0.065
30	0.610	0.315	0.014	30	0.520	0.303	0.025
40	0.609	0.385	0.105	40	0.582	0.402	0.119
50	0.692	0.466	0.244	50	0.673	0.494	0.266
75	0.810	0.587	0.451	75	0.816	0.624	0.482
100	0.821	0.643	0.529	100	0.845	0.683	0.564

Notes: [a] ρ and θ are the AR and MA parameters, respectively.

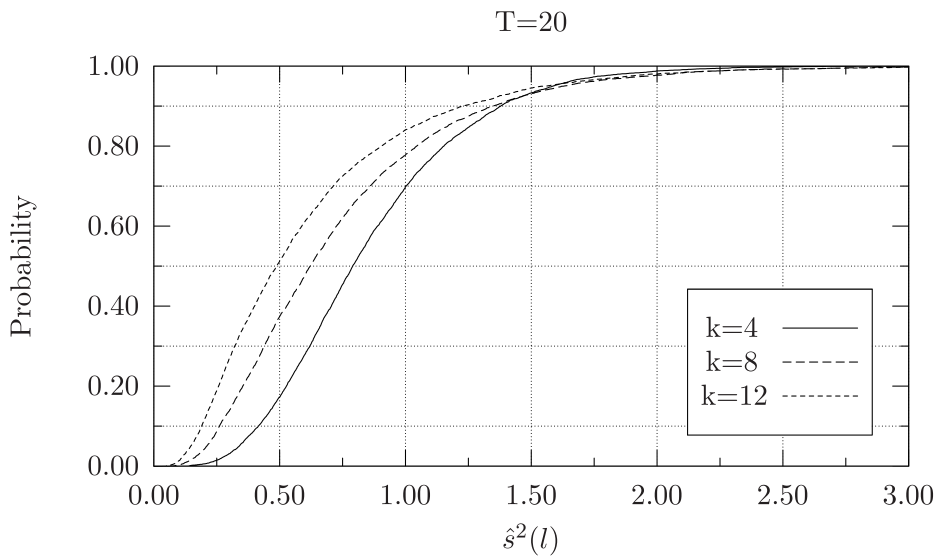
Table 10: Testing stationarity for private consumption.[a,b]

		Data-dependent bandwidth					Deterministic bandwidth				
Country	Sample	Statistic	l	Asy. CV	F-S CV	Concl. alt.?	Statistic	l	Asy. CV	F-S CV	Concl. alt.?
Australia	1960-2004	0.109	14	0.119	0.181	No	0.109	14	0.119	0.181	No
Austria	1960-2004	0.250	4	0.119	0.121	No	0.146	14	0.119	0.181	Yes
Belgium	1960-2004	0.189	6	0.119	0.124	No	0.141	14	0.119	0.181	Yes
Canada	1961-2004	0.157	6	0.119	0.125	No	0.132	14	0.119	0.185	Yes
Denmark	1966-2004	0.143	14	0.119	0.208	Yes	0.143	14	0.119	0.208	Yes
Finland	1960-2004	0.177	6	0.119	0.124	No	0.142	14	0.119	0.181	Yes
France	1963-2004	0.175	6	0.119	0.126	No	0.142	14	0.119	0.193	Yes
Greece	1960-2004	0.205	5	0.119	0.122	No	0.139	14	0.119	0.181	Yes
Iceland	1960-2004	0.205	5	0.119	0.122	No	0.136	14	0.119	0.181	Yes
Ireland	1960-2004	0.124	14	0.119	0.181	Yes	0.124	14	0.119	0.181	Yes
Italy	1960-2004	0.292	3	0.119	0.120	No	0.151	14	0.119	0.181	Yes
Japan	1960-2004	0.174	7	0.119	0.127	No	0.155	14	0.119	0.181	Yes
Luxembourg	1960-2004	0.150	9	0.119	0.137	No	0.142	14	0.119	0.181	Yes
Netherlands	1960-2003	0.128	10	0.119	0.146	Yes	0.129	14	0.119	0.185	Yes
New Zealand	1962-2004	0.129	14	0.119	0.189	Yes	0.129	14	0.119	0.189	Yes
Norway	1960-2004	0.123	14	0.119	0.181	Yes	0.123	14	0.119	0.181	Yes
Portugal	1960-2004	0.133	14	0.119	0.181	Yes	0.133	14	0.119	0.181	Yes
Spain	1960-2004	0.130	14	0.119	0.181	Yes	0.130	14	0.119	0.181	Yes
Sweden	1960-2004	0.135	10	0.119	0.144	Yes	0.137	14	0.119	0.181	Yes
Switzerland	1965-2003	0.152	9	0.119	0.148	No	0.152	14	0.119	0.208	Yes
UK	1960-2004	0.160	5	0.119	0.122	No	0.148	14	0.119	0.181	Yes
USA	1960-2004	0.143	14	0.119	0.181	Yes	0.143	14	0.119	0.181	Yes

Notes: [a]Statistic denotes the calculated value for the test statistic, l denotes the selected bandwidth, Asy. CV denotes asymptotic critical values, F-S CV denotes finite-sample critical values, while Concl. alt.? denotes whether the conclusion of the test is altered when using finite-sample critical values.

[b]The tests are performed on the 10% significance level.

Figure 1: Empirical distribution of long-run variance, intercept only.



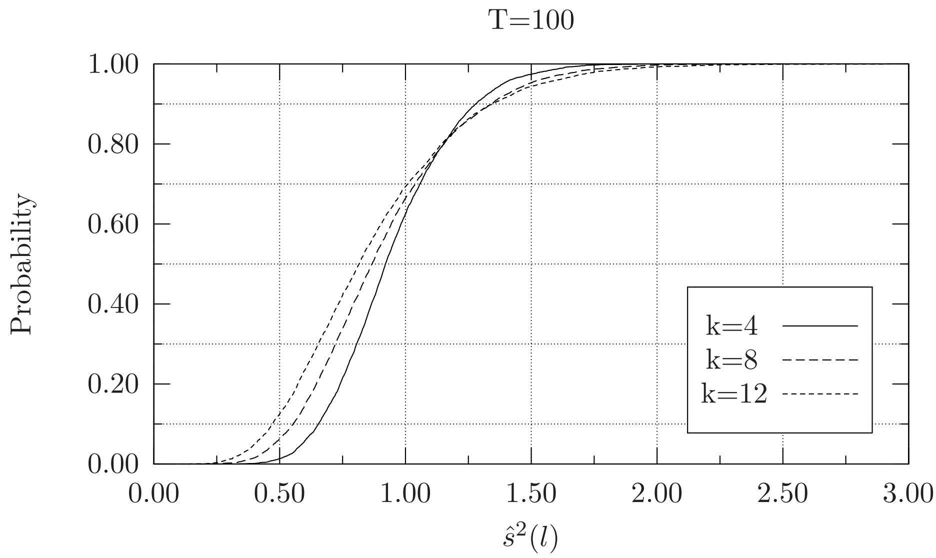
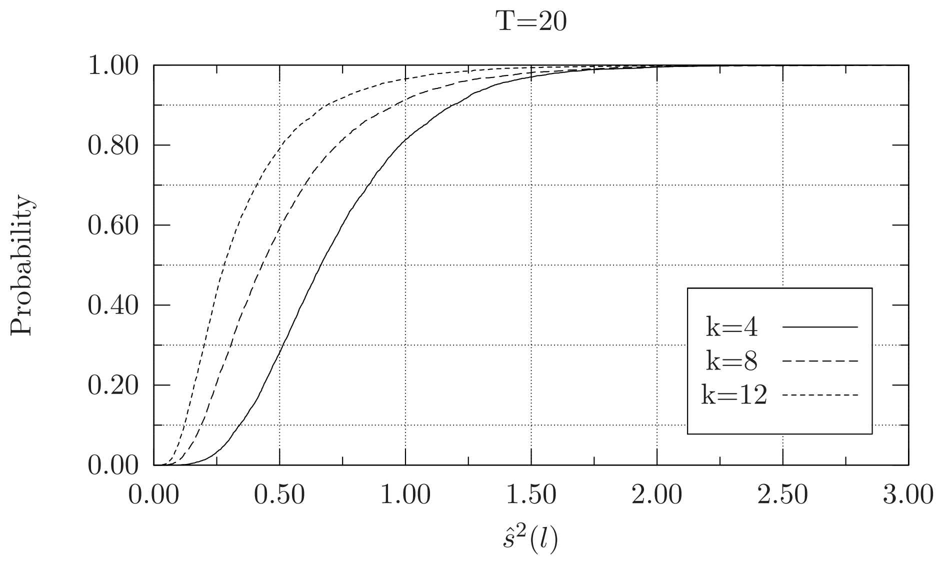


Figure 2: Empirical distribution of long-run variance, intercept and trend.



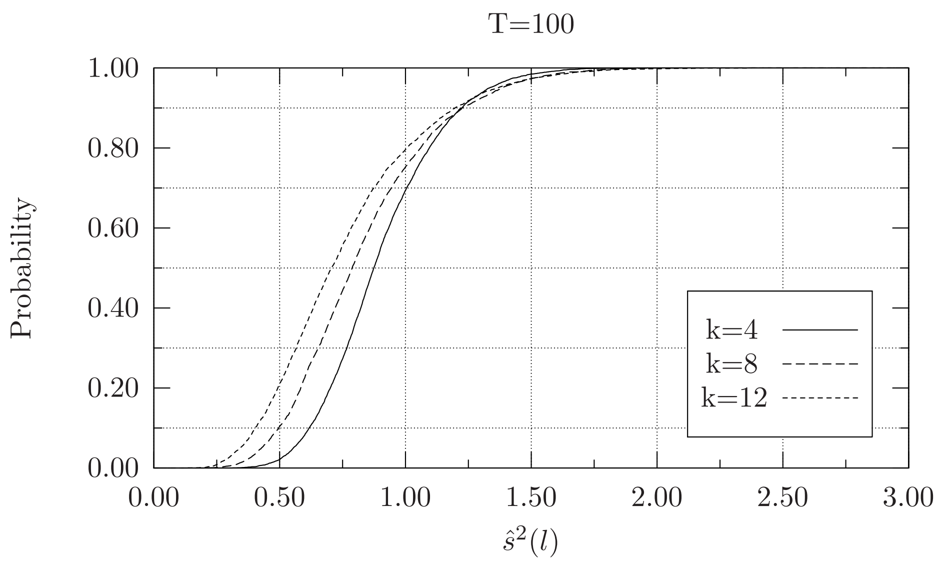
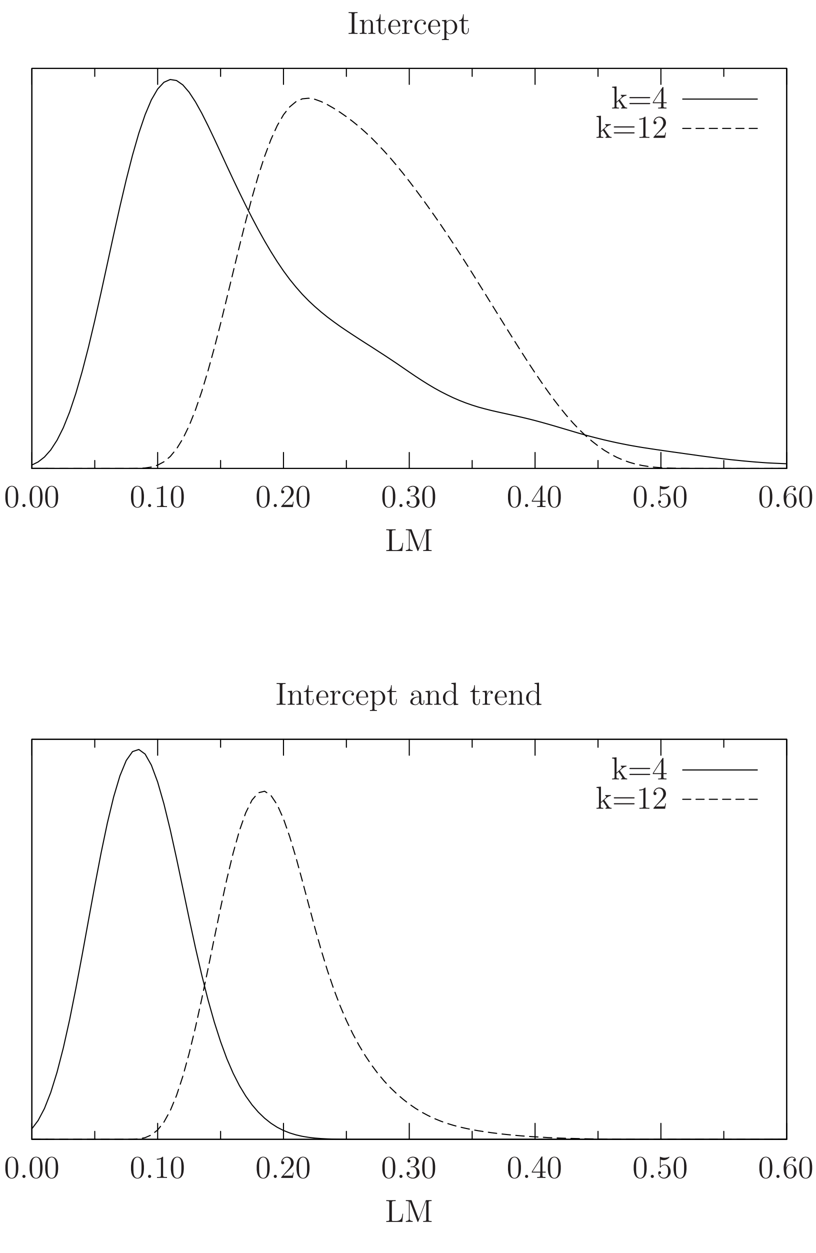
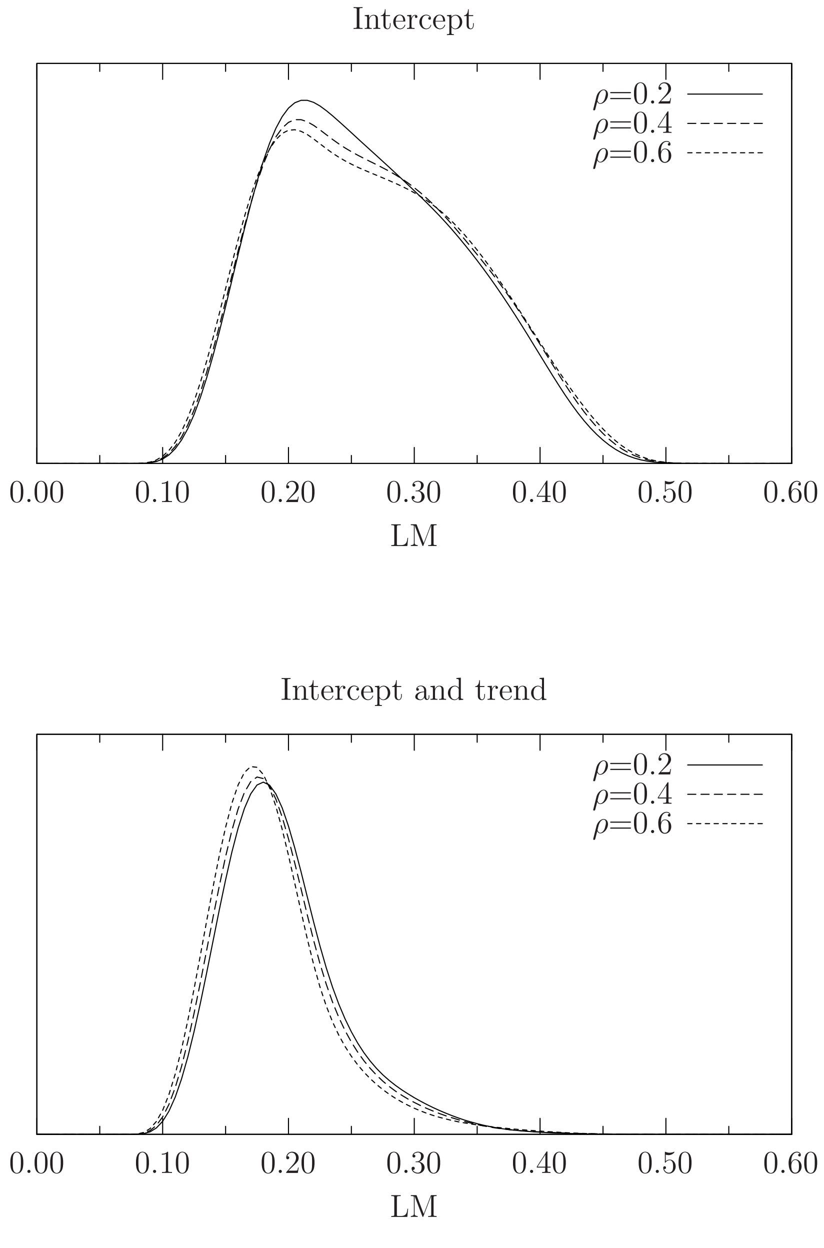


Figure 3: Kernel density estimates, fixed $\rho = 0.0$, various k.[a]



Notes: [a]The kernel densities are obtained for the case where T=20.

Figure 4: Kernel density estimates, fixed $k = 12$, various ρ.[a]



Notes: [a]The kernel densities are obtained for the case where T=20.

Figure 5: Size comparison, intercept.

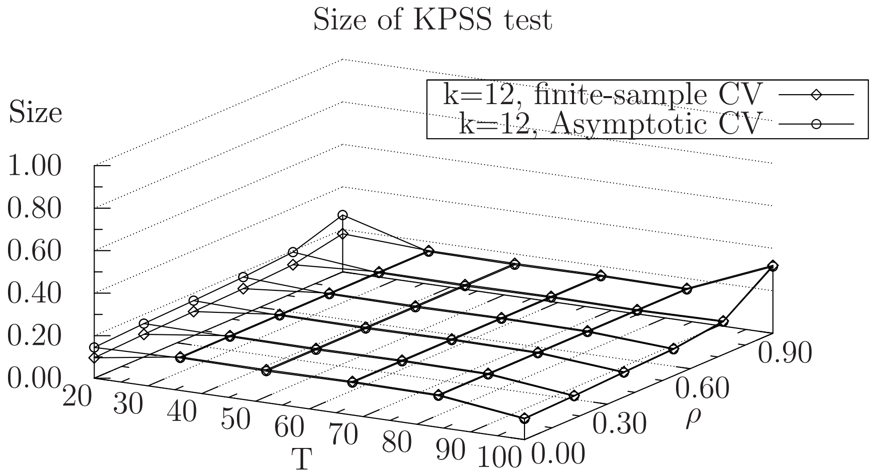


Figure 6: Size comparison, intercept and trend.

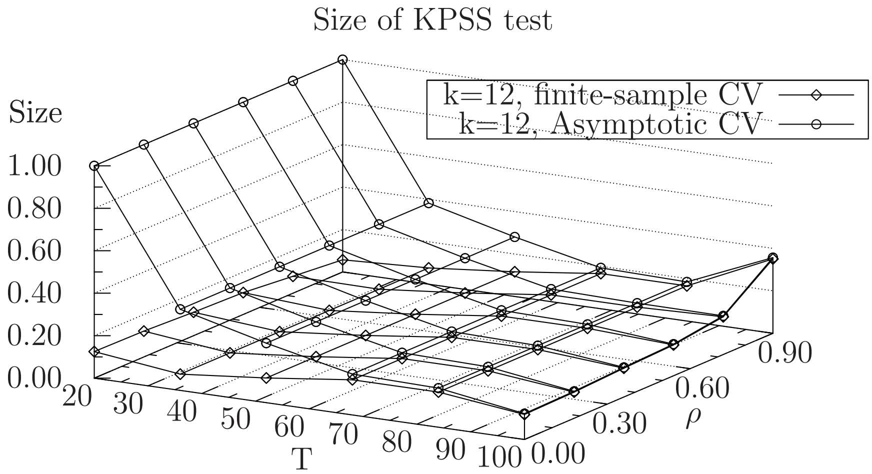


Figure 7: Size comparison, intercept.

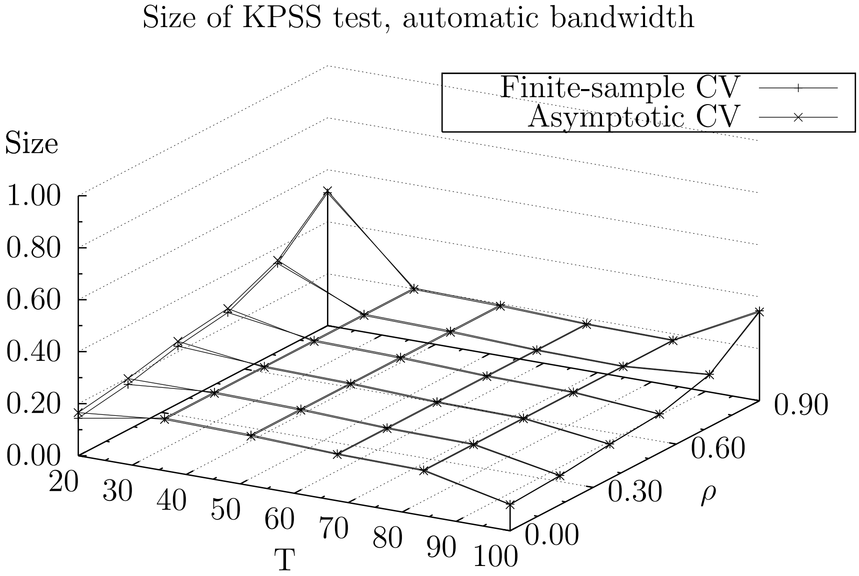


Figure 8: Size comparison, intercept and trend.