A robust version of the KPSS test, based on indicators

Robert M. de Jong   Christine Amsler   Peter Schmidt

Department of Economics
Michigan State University

August 2002

Abstract

This paper proposes a test of the null hypothesis of stationarity that is robust to the presence of fat-tailed errors. The test statistic is a modified version of the so-called KPSS statistic. The modified statistic uses the "sign" of the data minus the sample median, whereas KPSS used deviations from means. This "indicator" KPSS statistic has the same limit distribution as the standard KPSS statistic under the null, without relying on assumptions about moments, but a different limit distribution under unit root alternatives. The indicator test has lower power than standard KPSS when tails are thin, but higher power when tails are fat.

1 Introduction

In this paper we wish to test the null hypothesis that an observed series $\{x_t\}$ is stationary. We allow for non-zero level (mean or median) for x_t, but not for deterministic trend. This is often called the hypothesis of "level stationarity," and a standard test for this hypothesis is the $\hat{\eta}_\mu$ test of Kwiatkowski et al. (1992), which we will simply call the KPSS test. For a sample $x_1, \ldots, x_T$, define $\bar{x}_T$ as the sample mean, and the demeaned data $e_t = x_t - \bar{x}_T$ $(t = 1, \ldots, T)$. Then the KPSS statistic is a function of the e_t; the numerator is the sum of squares of the cumulations of the e_t, while the denominator is an estimate of their long run variance. The asymptotic distribution of the statistic is a functional of a Brownian bridge. This result depends on the series satisfying a short-memory condition and having finite variance.

In this paper we seek to relax the finite variance assumption. The motivation is an earlier paper, Amsler and Schmidt (2000), which considered the robustness (or lack of robustness) of the KPSS test to fat-tailed errors[1]. If the series follows a symmetric stable distribution with infinite variance, they showed that the KPSS statistic follows a different limit theory, based on the (demeaned) Lévy process. They also considered *local* departures from finite variance, in which case the series is assumed to be represented as follows: $x_t = x_{1t} + (c/T^{1/\alpha - 1/2})x_{2t}$, where x_{1t} has finite variance and x_{2t} has a symmetric stable distribution with parameter $\alpha < 2$. In this case the limit theory involves a mixture of the Wiener and Lévy processes. The asymptotic distribution depends on the "parameter" c which controls the weight given to the Lévy process. Thus, in terms of asymptotics, the KPSS test is not robust to even local departures from finite variance. Their simulations correspondingly show size distortions in finite samples, not only when the data have infinite variance, but also when they have finite variance but fat tails (e.g., student's t with three degrees of freedom).

In some sense it is obvious that to make the variance finite, or more generally to remove the effects of fat tails, we should trim the data. Strictly speaking, any given (not data dependent) trimming rule, like censoring the data at ± 100, should work, but intuitively it seems that a sensible trimming rule should depend on the location and scale of the data. Use of a data-dependent trimming rule will raise non-trivial questions about the asymptotic theory for the statistic. In this paper we choose a rather simple trimming rule: we replace the data by an indicator that equals plus or minus one, depending on whether the observation is above or below the sample median. Using these indicators, we have handled location sensibly but have sidestepped the perhaps more difficult question of scale. We then proceed to construct the KPSS statistic in the usual fashion, but from the transformed data. We will call this test the *indicator KPSS* test.

We show that under the null, whether or not the variance is finite, so long as a short memory condition holds, we have the same asymptotic distribution as the original KPSS test has under finite variance. Thus the test is robust to infinite variance (fat tails). Under the alternative, the asymptotic distribution of the indicator KPSS test is different from that of the KPSS test, even in the finite variance case, and for both tests it is different in the finite

[1] That paper also considers the modified rescaled range test of Lo(1991). We will not consider Lo's test in this paper, other than to note that it can be easily modified to be robust to fat tails in exactly the same way as we will do for KPSS.

variance case than in the infinite variance case. Thus we should expect power differences between the two tests depending on tail thickness. Our simulations show that this occurs. The indicator test is less powerful than the usual KPSS test if there are not fat tails, but it is more powerful (in terms of both power and size-adjusted power) if the tails are fat enough.

The plan of the paper is as follows. In Section 2, we provide the asymptotic theory for the indicator KPSS test. In Section 3, we report our simulations. Finally, Section 4 gives some concluding remarks and directions for future research. Proofs are given in an Appendix.

2 Asymptotic Theory

The data are $x_1, \ldots, x_T$. Let $m_T = \text{med}(x_1, x_2, \ldots, x_T)$. We now transform the data into the "indicator data" $\text{sgn}(x_t - m_T)$, where $\text{sgn}(x) = 1$ if $x > 0$, $\text{sgn}(x) = -1$ if $x < 0$, and $\text{sgn}(x) = 0$ if $x = 0$. Then the indicator KPSS statistic simply uses the indicator data $\text{sgn}(x_t - m_T)$ in the same way that the KPSS $\hat{\eta}_\mu$ statistic uses the demeaned data $x_t - \bar{x}_T$.

It may be worthwhile to motivate the use of the median in constructing the indicator data. The use of the median ensures that the indicator data sum to zero over the sample, just as the demeaned data do. This is why we can obtain the ususal KPSS asymptotics (based on a Brownian bridge), but without the need for assumptions about the moments of the data.

To be more explicit, let $\tilde{\mu}$ be the *population* median of the x_t, which will be assumed to be unique, and define

$$\sigma^2 = \lim_{T\to\infty} E(T^{-1/2} \sum_{t=1}^{T} \text{sgn}(x_t - \tilde{\mu}))^2. \tag{1}$$

Define the cumulations of the indicator data:

$$S_{Tt} = \sum_{j=1}^{t} \text{sgn}(x_j - m_T), \tag{2}$$

and let the HAC estimator $\hat{\sigma}^2$ be given by

$$\hat{\sigma}^2 = T^{-1} \sum_{i=1}^{T} \sum_{j=1}^{T} k((i-j)/\gamma_T)\text{sgn}(x_i - m_T)\text{sgn}(x_j - m_T). \tag{3}$$

The indicator KPSS statistic that is proposed in this paper is then

$$\hat{\sigma}^{-2} T^{-2} \sum_{t=1}^{T} S_{Tt}^2. \tag{4}$$

Note that the "standard" KPSS statistic $\hat{\eta}_\mu$ uses the same formula as above, but for the standard KPSS statistic, S_{Tt} is replaced by $\sum_{j=1}^{t}(x_j - \bar{x})$, and the HAC estimator $\hat{\sigma}^2$ would need to be modified accordingly. Also, define

$$W_T(\alpha) = (1/\sigma) T^{-1/2} \sum_{j=1}^{[\alpha T]} \text{sgn}(x_j - \tilde{\mu}). \tag{5}$$

In what follows, $F(.)$ denotes the distribution function of $x_j - \tilde{\mu}$. Note that by Theorem 3.1 of De Jong and Davidson (2000), under the conditions that will be imposed in Assumption 1, $W_T(\alpha) \Rightarrow W(\alpha)$ on [0,1], where W is Brownian motion and $X_T(\alpha) \Rightarrow X(\alpha)$ denotes weak convergence in $D[A]$, where A denotes the appropriate subspace of $\mathbb{R}^q$. The weak dependence assumptions on the x_j that we need are listed in Assumption 1 below.

Assumption 1

1. *The x_j are stationary random variables, and $\tilde{\mu}$ is the unique population median of x_j.*
2. *x_j is strong (α-) mixing, and for some finite $r > 2$ and $C > 0$, and for some $\eta > 0$, $\alpha(m) \leq Cm^{-r/(r-2)-\eta}$.*
3. *$x_j - \tilde{\mu}$ has a continuous density $f(x)$ in a neighborhood $[-\eta, \eta]$ of 0 for some $\eta > 0$, and $\inf_{x \in [-\eta,\eta]} f(x) > 0$.*
4. *$\sigma^2 \in (0, \infty)$.*

For the kernel function $k(.)$ function in the HAC estimator, we need the following conditions:

Assumption 2

1. *$k(.)$ satisfies $\int_{-\infty}^{\infty} |\psi(\xi)| d\xi < \infty$, where*
$$\psi(\xi) = (2\pi)^{-1} \int_{-\infty}^{\infty} k(x) \exp(-i\xi x) dx. \tag{6}$$
2. *$k(.)$ is continuous at all but a finite number of points, $k(x) = k(-x)$, $|k(x)| \leq l(x)$ where $l(x)$ is nonincreasing and $\int_0^{\infty} |l(x)| dx < \infty$, and $k(0) = 1$.*
3. *$\gamma_T / T \to 0$, and $\gamma_T \to \infty$ as $T \to \infty$.*

Assumption 2 rules out the use of the uniform kernel function, but allows choices such as the Bartlett, Quadratic Spectral, and Parzen kernel. Under the above assumptions we can prove the following theorem.

Theorem 1 *Under Assumptions 1 and 2,*
$$T^{-2} \sum_{t=1}^{T} S_{Tt}^2 \xrightarrow{d} \sigma^2 \int_0^1 (W(\alpha) - \alpha W(1))^2 d\alpha, \tag{7}$$
and
$$\hat{\sigma}^2 \xrightarrow{p} \sigma^2. \tag{8}$$

From the above it follows that, analogously to standard KPSS,

$$\hat{\sigma}^{-2}T^{-2}\sum_{t=1}^{T}S_{Tt}^{2} \xrightarrow{d} \int_{0}^{1}(W(\alpha)-\alpha W(1))^{2}d\alpha, \tag{9}$$

KPSS (1992, p. 166) give critical values; see also MacNeill (1978, Table 2, p. 431).

Under the alternative that x_j is an I(1) process, we have the following result.

Theorem 2 *Suppose that x_j satisfies $T^{-1/2}x_{[\xi T]} \Rightarrow \lambda W(\xi)$ for some $\lambda \in (0,\infty)$. Then*

$$T^{-3}\sum_{t=1}^{T}S_{Tt}^{2} \xrightarrow{d} \lambda^{2}\int_{0}^{1}(\int_{0}^{\alpha}\operatorname{sgn}(W(\xi)-M))d\xi)^{2}d\alpha, \tag{10}$$

where M is such that $T^{-1/2}m_T \xrightarrow{d} M$, and

$$\gamma_T^{-1}\hat{\sigma}^{2} \xrightarrow{d} 2\lambda^{2}\int_{0}^{\infty}k(\alpha)d\alpha. \tag{11}$$

From Theorem 2, it follows that analogously to the "standard" KPSS statistic,

$$(\gamma_T T^{-1})\hat{\sigma}^{-2}T^{-2}\sum_{t=1}^{T}S_{Tt}^{2} \xrightarrow{d} D \tag{12}$$

for some random variable D that is distribution-free except for the multiplication factor $\int_0^\infty k(\alpha)d\alpha$. However, for our indicator-KPSS statistic the distribution under the alternative, as characterized by D, will be different from the distribution of the "standard" KPSS statistic under the alternative.

There is an important difference between the asymptotic distribution theory under the null of stationarity (Theorem 1) and under the unit root alternative (Theorem 2). The distribution theory under the null does not depend on any assumptions about the moments of the data x_t. However, the unit root alternative considered in Theorem 2 is that $x_{[\xi T]}$ follows an invariance principle for convergence to a Wiener process. This requires that x_t have zero mean and finite variance. The zero mean requirement would be easily relaxed, so long as the mean exists, but the finite variance condition is substantive. If x_t had infinite variance (e.g. if it followed a symmetric stable distribution with index $\alpha < 2$), we would expect a different limit theory to apply. Furthermore, this would also be different than the limit theory for the KPSS test in the infinite variance case.

The implications of these observations for the power of the test are as follows. In the finite variance case, the fact that the indicator KPSS statistic has a different limiting distribution than the usual KPSS statistic under the alternative implies that we should find differences in power even in large samples. Intuitively, transforming to indicators will cause a loss in power when tails are not fat, and this power loss will persist even asymptotically. However, the fact that the asymptotic distribution of the two tests under the alternative will also differ in the infinite variance case suggests that there may be the potential for gains in power from the use of indicators (or trimming more generally) in the case of fat tails. That is, the advantage of indicators may be added power, not just minimized size distortions, when tails are fat.

3 Simulations

In this section we report the results of simulations that investigate the size and power of the KPSS and indicator-KPSS tests. We are fundamentally interested in three questions. First, is the indicator-KPSS test robust to fat tails in finite samples? Second, how large is the loss in power in using indicators when the data do not have fat tails? Third, is there a gain in power, and of what size, from using indicators when the data do have fat tails?

Our simulations are performed on PCs in FORTRAN using the Lahey compiler. Normal random deviates are drawn using the random number generator GASDEV/RAN3 of Press *et al.* (1986). We vary tail thickness by considering students t distributions with differing degrees of freedom. Specifically, we consider normal (t_∞), t_5, t_3, t_2, and Cauchy (t_1). Note that t_2 has finite mean but infinite variance, while for Cauchy neither the mean nor the variance exist. These distributions are generated by the appropriate transformation of the standard normal deviates from GASDEV/RAN3. We also consider the local to finite variance (and also local to finite mean) alternative in which $x_t = x_{1t} + (c/T^{1/2})x_{2t}$, where x_1 is normal and x_2 is Cauchy. We choose $c = 1$ and call this the "local ($c = 1$)" case. We consider sample sizes from $T = 50$ to $T = 5000$. Our results are based on 20,000 replications.

In order to focus on the simplest and most understandable cases, most of our experiments are for the case that the data (under the null) or the innovations (under the alternative) are i.i.d.. We consider a few cases in which they are AR(1) with parameter equal to 0.5. Similarly, in most of our experiments we consider the case of no lags in the long run variance estimate, that is, the kernel function $k(.)$ in Equation (3) equals one for $i = j$ and zero otherwise. We will denote this as the case of $\gamma_T = \gamma_{T,0} = 0$. We also consider the case that the number of lags is $\gamma_T = \text{integer}[4(T/100)^{1/4}]$, which we denote as $\gamma_{T,1/4}$. Clearly our choices reflect a focus on the effects of tail thickness as opposed to short run dynamics.

We first consider the size of the tests. Our results are given in Table 1. KPSS and IKPSS refer to the usual and the indicator versions of KPSS, respectively. Nominal size is 5%. The upper panel is for the case that the data are i.i.d. and we choose $\gamma_{T,0}$, so that there are no short run dynamics and there is no allowance made for short run dynamics either. The indicator KPSS test has size very close to 0.05 for all cases, and would appear to be robust to fat tails even in samples as small as $T = 50$. The usual KPSS test is not robust to the Cauchy or local to finite variance distributions; it rejects too seldom. It is reasonably robust to the other fat tailed distributions, such as t_3 or even t_2, though there is a (slight) tendency to reject too seldom in these cases when T is small. Its robustness to t_2 is perhaps surprising because this is an infinite-variance distribution.

The lower panel of Table 1 is for the case of $\gamma_{T,1/4}$ lags in the long run variance calculation. Here we report results only for the normal, t_3 and Cauchy cases, but we also have some cases in which the data are AR(1) with parameter 0.5. In the normal and t_3 cases with iid data, there is a slight problem of underrejection for the smaller sample sizes, with no particular difference between the KPSS and indicator KPSS tests. In the Cauchy case the two tests are quite different from each other, as they were in the case of $\gamma_{T,0}$, with the indicator KPSS test being far more robust. When the data are AR(1) with parameter 0.5, we have a problem of overrejection. Again there is not much difference between the two tests in the finite variance cases. In the Cauchy case, the usual KPSS test has smaller size distortions because its tendency to underreject in the presence of fat tails counteracts its tendency to overreject in the presence of positive autocorrelation.

The results in Table 1 are easy to summarize. Both the usual KPSS test and the indicator KPSS test are quite robust to finite variance fat tailed data, such as t_5 or t_3. However, the indicator test is also robust to distributions without finite mean and variance.

In Tables 2 and 3 we report the power and size-adjusted power of the tests, for the case of $\gamma_{T,0}$ (no lags in the long run variance estimate). Power and size-adjusted power will differ non-trivially only in the cases in which we found size distortions in Table 1, namely the Cauchy and local to finite variance cases. We will discuss only size-adjusted power (Table 3). Comparing power instead of size-adjusted power would make the comparisons more favorable to the indicator KPSS test for the Cauchy and local to finite variance cases.

We parameterize the unit root alternative in the same way as in KPSS (1992). That is, we have the components representation $x_t = r_t + \varepsilon_t$, where $r_0 = 0$ and $r_t = r_{t-1} + u_t$, $t = 1, \ldots, T$. The innovations ε_t and u_t are iid and independent of each other. Thus x_t is the sum of a random walk and a white noise process, and $\lambda = \sigma_u^2/\sigma_\varepsilon^2$ measures the relative importance of the random walk component. In all case, ε_t and u_t both have the same distribution (normal, t_5, or whatever).

Consider first the normal case (the first two columns of results in Table 3). The KPSS test is clearly more powerful than the indicator KPSS test. The loss in power from using indicators instead of the untransformed data is not huge, but it is also non-trivial. For example, when the power of KPSS is approximately 0.30, the power of indicator KPSS is about 0.23; when the power of KPSS is approximately 0.6, the power of indicator KPSS is about 0.5; and so forth.

With fat tails, this situation reverses. For the t_5 case, the difference in power is very close to zero, and in fact the comparisons are mixed. For the other cases with even fatter tails, the indicator KPSS test is clearly more powerful. For the t_3 case, the extent to which indicator KPSS is more powerful than KPSS is roughly comparable to the extent to which KPSS was more powerful than indicator KPSS for the normal case. For the t_2, Cauchy and local to finite variance cases, the superiority of the indicator test is larger; for the Cauchy case it is much larger.

Table 4 gives the results for some cases with $\gamma_{T,1/4}$, and it includes cases with iid innovations and also cases when ε_t (but not u_t) is AR(1). The main result is the same as in the previous tables: the usual KPSS test is more powerful than the indicator KPSS test under normality, but the indicator KPSS test is more powerful than the KPSS test when the tails are fat (t_3 and Cauchy).

4 Concluding Remarks

In this paper we have constructed an "indicator KPSS" test that has the usual KPSS asymptotic distribution under the null, without requiring finite mean or variance for the data. It is based on the transformation of the data into an indicator of whether the observation is above or below the median. The resulting test is quite robust to fat tailed errors. It is less powerful than the usual KPSS test under normality or near normality, but more powerful if the errors have sufficiently fat tails.

Our test, like the KPSS $\hat{\eta}_\mu$ test, is a test of "level stationarity" as opposed to "trend stationarity." We could also seek to generalize the KPSS $\hat{\eta}_\tau$ test, which allows for linear deterministic trend. We conjecture that this could be done by defining an indicator of whether one is above or below the LAD regression of x_t on t. However, this leads to more challenging asymptotic theory than we have needed in this paper.

A general research agenda into which this paper fits is to try to find appropriate mechanisms for trimming the data, so that one ends up with robustness to fat tails, minimal power loss under normality, and good power properties when the tails are fat. Our results so far are (we hope) a successful demonstration project indicating that this can be done. However, transformation of the data into indicators is a rather extreme form of trimming, and was chosen largely so that the asymptotic theory was tractable. We might expect that less severe trimming rules might be preferable, at least in minimizing the power loss under normality. These trimming rules would have to involve measures of scale as well as location, and are the subject of ongoing research.

References

Amsler, C. and P. Schmidt (2000), Tests of short memory with thick tailed errors, unpublished manuscript, Michigan State University.

Davidson, J. (1994), *Stochastic limit theory.* Oxford: Oxford University Press.

De Jong, R.M. and J. Davidson (1998), The functional central limit theorem and weak convergence to stochastic integrals I: weakly dependent and fractionally integrated processes, Econometric Theory 16, 621-642.

De Jong, R.M. and J. Davidson (2000), Consistency of kernel estimators of heteroscedastic and autocorrelated covariance matrices, Econometrica 68, 407-423.

Kim, J. and D. Pollard (1990), Cube root asymptotics, Annals of Statistics 18, 191-219.

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

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

Pollard, D. (1990), Empirical processes, theory and applications. CBMS regional conference series in probability and statistics, Volume 2, Institute of Mathematical Statistics, Hayward, CA.

Press, W.H., B.P. Flannery, S.A. Teukolsky and W.T. Vetterling (1986), *Numerical recipes: the art of scientific computing.* Cambridge University Press, Cambridge.

Appendix: Proofs

For the proof of Theorem 1, we will need the following lemmas.

Lemma 1 *For strong (α-) mixing random variables $y_{Tt} \in \mathbb{R}$ for which the α-mixing coefficients satisfy $\alpha(m) \le Cm^{-r/(r-2)-\eta}$ for some $\eta > 0$,*

$$E \max_{1 \le i \le T} (\sum_{t=1}^{i} (y_{Tt} - Ey_{Tt}))^2 \le C' \sum_{t=1}^{T} \parallel y_{Tt} \parallel_r^2$$

for constants C,$C' > 0$.

Proof of Lemma 1:

First note that by Theorem 17.5 of Davidson (1994), y_{Tt} is a mixingale of size -1/2 under Assumption 1 with mixingale magnitude indices $\parallel y_{Tt} \parallel_r$. Therefore by 16.10 from Davidson (1994),

$$E \max_{1 \le i \le T} (\sum_{t=1}^{i} (y_{Tt} - Ey_{Tt}))^2 \le C \sum_{t=1}^{T} \parallel y_{Tt} \parallel_r^2$$

for some constant $C > 0$, as asserted. $\square$

Lemma 2 *Let*

$$y_j(\phi) = \mathrm{sgn}(x_j - \tilde{\mu} - \phi T^{-1/2}) - \mathrm{sgn}(x_j - \tilde{\mu}).$$

Then under Assumption 1, for all $K, \varepsilon > 0$,

$$\lim_{\delta \to 0} \limsup_{T \to \infty} P(\sup_{\phi, \phi' \in [-K,K] : |\phi - \phi'| < \delta} T^{-1/2} \sum_{j=1}^{T} |y_j(\phi) - y_j(\phi') - Ey_j(\phi) + Ey_j(\phi')| > \varepsilon) = 0. \quad (13)$$

Proof of Lemma 2:

For T large enough such that $KT^{-1/2} \le \eta$,

$$\sup_{\phi, \phi' : |\phi - \phi'| < \delta} T^{-1/2} \sum_{j=1}^{T} |Ey_j(\phi) - Ey_j(\phi')|$$

$$= 2 \sup_{\phi, \phi' : |\phi - \phi'| < \delta} T^{-1/2} \sum_{j=1}^{T} |F(\phi T^{-1/2}) - F(\phi' T^{-1/2})|$$

$$\leq 2 \sup_{\phi,\phi':|\phi-\phi'|<\delta} T^{-1/2} \sum_{j=1}^{T} \sup_{x\in[-\eta,\eta]} f(x) T^{-1/2}|\phi-\phi'| \leq 2\delta \sup_{x\in[-\eta,\eta]} f(x).$$

This establishes equicontinuity of $T^{-1/2}\sum_{j=1}^{T}|Ey_j(\phi)-Ey_j(\phi')|$. Next, note that $y_j(\phi)$ is nondecreasing in ϕ, and therefore

$$\sup_{\phi,\phi'\in[-K,K]:|\phi-\phi'|<\delta} T^{-1/2}\sum_{j=1}^{T}|y_j(\phi)-y_j(\phi')|$$

$$= \sup_{-[K/\delta]-1\leq i\leq[K/\delta]} \sup_{\phi,\phi'\in[i\delta,(i+2)\delta]\cap[-K,K]} T^{-1/2}\sum_{j=1}^{T}|y_j(\phi)-y_j(\phi')|$$

$$\leq \sup_{-[K/\delta]-1\leq i\leq[K/\delta]} T^{-1/2}\sum_{j=1}^{T}|y_j(i\delta)-y_j((i+2)\delta)|$$

$$\xrightarrow{p} \sup_{-[K/\delta]-1\leq i\leq[K/\delta]} T^{-1/2}\sum_{j=1}^{T}E|y_j(i\delta)-y_j((i+2)\delta)|$$

$$\leq \sup_{\phi,\phi'\in[-K,K]:|\phi-\phi'|\leq 2\delta} T^{-1/2}\sum_{j=1}^{T}|Ey_j(\phi)-Ey_j(\phi')|, \tag{14}$$

and the last term was earlier shown to be equicontinuous. Therefore, by the triangle inequality, we have now demonstrated the result of Equation (13); the convergence in probability result leading up to Equation (14) holds because for every value of i and for T large enough such that $KT^{-1/2}\leq\eta$,

$$E(T^{-1/2}\sum_{j=1}^{T}(E|y_j(i\delta)-y_j((i+2)\delta)|-E|y_j(i\delta)-y_j((i+2)\delta)|))^2$$

$$\leq CT^{-1}\sum_{j=1}^{T}\sup_{\phi\in[-K,K]}\parallel y_j(\phi)\parallel_r^2$$

$$= CT^{-1}\sum_{j=1}^{T}\sup_{\phi\in[-K,K]}\parallel \operatorname{sgn}(x_j-\tilde{\mu}-\phi T^{-1/2})-\operatorname{sgn}(x_j-\tilde{\mu})\parallel_r^2$$

$$\leq C'|F(KT^{-1/2})-F(-KT^{-1/2})|^{2/r}\leq C'(2\sup_{x\in[-\eta,\eta]}f(x)KT^{-1/2})^{2/r}\longrightarrow 0$$

as $T\to\infty$ for constants C,$C'>0$, where the first inequality follows from Lemma 1. □

The key lemma of this section, which will allow a Taylor-type expansion of the objective function and is also used to determine the limit behavior of the HAC estimator $\hat{\sigma}^2$, is the following.

Lemma 3 *Let $y_j(\phi)$ be as before, and let*

$$G_T(\alpha,\phi) = T^{-1/2}\sum_{j=1}^{[\alpha T]} y_j(\phi).$$

Then under Assumption 1, for any $K > 0$,

$$\sup_{\alpha\in[0,1]}\ \sup_{-K\le\phi\le K} |G_T(\alpha,\phi) - EG_T(\alpha,\phi)| \xrightarrow{p} 0.$$

Proof of Lemma 3:

We can conclude, e.g. from Pollard (1990), that

$$\sup_{\phi\in[-K,K]}\ \sup_{\alpha\in[0,1]} |G_T(\alpha,\phi)| \xrightarrow{p} 0$$

because of the compactness of the index set for ϕ, finite-dimensional convergence for each $\phi \in [-K, K]$, and stochastic equicontinuity of $\sup_{\alpha\in[0,1]} |G_T(\alpha,\phi)|$. Finite-dimensional convergence follows because for every $\phi \in [-K, K]$, by Lemma 1, for T large enough such that $KT^{-1/2} \le \eta$,

$$E\sup_{\alpha\in[0,1]} |G_T(\alpha,\phi) - EG_T(\alpha,\phi)|^2$$

$$\le CT^{-1}\sum_{j=1}^{T} \parallel \mathrm{sgn}(x_j - \tilde{\mu} - \phi T^{-1/2}) - \mathrm{sgn}(x_j - \tilde{\mu}) \parallel_r^2$$

$$\le C'|F(KT^{-1/2}) - F(-KT^{-1/2})|^{2/r} \le C''(\sup_{x\in[-\eta,\eta]} f(x)2KT^{-1/2})^{2/r} \to 0$$

by Lemma 1, and constants $C, C', C'' > 0$. Stochastic equicontinuity follows because

$$|\sup_{\alpha\in[0,1]} |G_T(\alpha,\phi) - EG_T(\alpha,\phi)| - \sup_{\alpha\in[0,1]} |G_T(\alpha,\phi') - EG_T(\alpha,\phi')||$$

$$\le \sup_{\alpha\in[0,1]} |G_T(\alpha,\phi) - EG_T(\alpha,\phi) - G_T(\alpha,\phi') + EG_T(\alpha,\phi')|$$

$$\le T^{-1/2}\sum_{j=1}^{T} |y_j(\phi) - y_j(\phi') - Ey_j(\phi) + Ey_j(\phi')|,$$

and stochastic equicontinuity follows from Lemma 2 now. □

Next, we derive the behavior of the median m_T.

Lemma 4 *Under Assumption 1,*

$$T^{1/2}(m_T - \tilde{\mu}) = 2^{-1} f(0)^{-1} \sigma W_T(1) + o_P(1).$$

This lemma can be proven by showing that $T^{1/2}(m_T - \tilde{\mu}) = O_P(1)$, and then using the result of Lemma 3. $m_T - \tilde{\mu} = O_P(T^{-1/2})$ because for $T \geq K^2\eta^{-2}$,

$$\inf_{\phi > K} T^{-1/2} \sum_{j=1}^{T} \text{sgn}(x_j - \tilde{\mu} - \phi T^{-1/2})$$

$$\geq T^{-1/2} \sum_{j=1}^{T} \text{sgn}(x_j - \tilde{\mu} - KT^{-1/2}))$$

$$\xrightarrow{p} T^{-1/2} \sum_{j=1}^{T} (1 - 2F(KT^{-1/2}))$$

$$\geq 2K \inf_{-\eta \leq x \leq \eta} f(x),$$

implying that $\limsup_{T\to\infty} P(T^{1/2}(m_T - \tilde{\mu}) > K)$ can be made arbitrary small by choosing K large enough. For $P(T^{1/2}(m_T - \tilde{\mu}) < -K)$ a similar result can be derived, which proves that $m_T - \tilde{\mu} = O_P(T^{-1/2})$.
Next, note that since $T^{1/2}(m_T - \tilde{\mu}) = O_P(1)$, with arbitrary large probability we can assume that $T^{1/2}|m_T - \tilde{\mu}| \leq K$, and if that happens,

$$O(T^{-1/2}) = T^{-1/2} \sum_{j=1}^{T} \text{sgn}(x_j - m_T)$$

$$= (G_T(1, T^{1/2}(m_T - \tilde{\mu})) - EG_T(1, T^{1/2}(m_T - \tilde{\mu})))$$

$$+ T^{-1/2} \sum_{j=1}^{T} (\text{sgn}(x_j - \tilde{\mu}) - E\text{sgn}(x_j - \tilde{\mu})) + EG_T(1, T^{1/2}(m_T - \tilde{\mu}))$$

$$= o_P(1) + T^{-1/2} \sum_{j=1}^{T} \text{sgn}(x_j - \tilde{\mu}) + T^{1/2}(1 - 2F(m_T - \tilde{\mu}))$$

and therefore, under the maintained assumptions, by Taylor's theorem and consistency of m_T,

$$T^{1/2}(m_T - \tilde{\mu}) = o_P(1) + 2^{-1} f(0)^{-1} T^{-1/2} \sum_{j=1}^{T} \text{sgn}(x_j - \tilde{\mu})$$

$$= o_P(1) + \sigma 2^{-1} f(0)^{-1} W_T(1),$$

which is the desired result. □

Proof of Theorem 1:

Assume that $-K \le T^{1/2}(m_T - \tilde{\mu}) \le K$, which for large K will happen with arbitrary large probability by Lemma 4. Then

$$T^{-1/2} S_{T,[\alpha T]} = T^{-1/2} \sum_{j=1}^{[\alpha T]} \operatorname{sgn}(x_j \le m_T)$$

$$= (G_T(\alpha, T^{1/2}(m_T - \tilde{\mu})) - EG_T(\alpha, T^{1/2}(m_T - \tilde{\mu})))$$

$$+ T^{-1/2} \sum_{j=1}^{[\alpha T]} \operatorname{sgn}(x_j - \tilde{\mu}) - 2T^{-1/2}[\alpha T](m_T - \tilde{\mu}) f(\tilde{m}_T - \tilde{\mu})$$

$$= o_P(1) + \sigma W_T(\alpha) - \alpha \sigma W_T(1)$$

where $\tilde{m}_T$ is on the line between $\tilde{\mu}$ and m_T, and the $o_P(1)$ term is uniformly in α. Next, note that the last expression converges weakly in α to $\sigma W(\alpha) - \sigma \alpha W(1)$ under Assumption 1, and therefore, by the continuous mapping theorem,

$$T^{-2} \sum_{t=1}^{T} S_{Tt}^2 = \int_0^1 (T^{-1} S_{T,[\alpha T]+1}^2) d\alpha \xrightarrow{d} \sigma^2 \int_0^1 (W(\alpha) - \alpha W(1))^2 d\alpha.$$

To prove the consistency of $\hat{\sigma}^2$, note that

$$\operatorname{sgn}(x_j - m_T) = y_j(T^{1/2}(m_T - \tilde{\mu})) - Ey_j(T^{1/2}(m_T - \tilde{\mu}))$$

$$+ Ey_j(T^{1/2}(m_T - \tilde{\mu})) + \operatorname{sgn}(x_j - \tilde{\mu})$$

$$= (y_j(T^{1/2}(m_T - \tilde{\mu})) - Ey_j(T^{1/2}(m_T - \tilde{\mu}))) + (1 - 2F(m_T - \tilde{\mu})) + \operatorname{sgn}(x_j - \tilde{\mu})$$

$$= a_{Tj} + b_T + c_j,$$

say. Also note that $\hat{\sigma}^2$ now equals

$$\hat{\sigma}^2 = T^{-1} \sum_{t=1}^{T} \sum_{s=1}^{T} k((t-s)/\gamma_T)(a_{Tt} + b_T + c_t)(a_{Ts} + b_T + c_s).$$

We will show that $\hat{\sigma}^2$ is asymptotically equivalent to

$$T^{-1} \sum_{t=1}^{T} \sum_{s=1}^{T} k((t-s)/\gamma_T) c_t c_s.$$

To show this, note that

$$T^{-1} \sum_{t=1}^{T} \sum_{s=1}^{T} k((t-s)/\gamma_T) b_T a_{Tt}$$

$$\leq T^{-3/2}\sum_{t=1}^{T}|a_{Tt}|\sum_{j=-T}^{T}k(j/\gamma_T)\times O_P(1)=O_P(\gamma_T/T),$$

and that

$$T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T}k((t-s)/\gamma_T)b_Tc_{Tt}$$

$$\leq T^{-3/2}\sum_{t=1}^{T}c_{Tt}\sum_{s=1}^{T}k((t-s)/\gamma_T)\times O_P(1),$$

and

$$E(T^{-3/2}\sum_{t=1}^{T}c_{Tt}\sum_{s=1}^{T}k((t-s)/\gamma_T))^2$$

$$\leq CT^{-3}\sum_{t=1}^{T}\parallel c_{Tt}\sum_{s=1}^{T}k((t-s)/\gamma_T)\parallel_r^2$$

$$\leq C'T^{-3}\sum_{t=1}^{T}(\sum_{s=1}^{T}k((t-s)/\gamma_T))^2$$

$$\leq C''T^{-3}\sum_{t=1}^{T}(\sum_{j=-T}^{T}k(j/\gamma_T))^2$$

$$\leq C''T^{-2}\gamma_T^2.$$

Also,

$$b_T^2T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T}k((t-s)/\gamma_T)=O_P(T^{-1}\gamma_T).$$

Therefore under Assumption 1, $\hat{\sigma}^2$ is asymptotically equivalent to

$$T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T}k((t-s)/\gamma_T)(a_{Tt}+c_{Tt})(a_{Ts}+c_{Ts}).$$

Furthermore,

$$T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T}k((t-s)/\gamma_T)a_{Tt}c_s$$

$$= \int_{-\infty}^{\infty} (T^{-1} \sum_{t=1}^{T} \sum_{s=1}^{T} a_{Tt} c_s \exp(i\xi(t-s)/\gamma_T)\psi(\xi) d\xi$$

$$\leq T^{-1/2} \sum_{t=1}^{T} |a_{Tt}| \int_{-\infty}^{\infty} (T^{-1/2} \sum_{s=1}^{T} c_s \exp(-is\xi/\gamma_T))\psi(\xi) d\xi,$$

and $T^{-1/2} \sum_{t=1}^{T} |a_{Tt}|$ is $o_P(1)$ by Lemma 3 under Assumption 1, and the second term is $O_P(1)$ because

$$E| \int_{-\infty}^{\infty} (T^{-1/2} \sum_{s=1}^{T} c_s \exp(-is\xi/\gamma_T))\psi(\xi) d\xi|$$

$$\leq \int_{-\infty}^{\infty} |\psi(\xi)| d\xi \sup_{\xi \in \mathbb{R}} \parallel T^{-1/2} \sum_{s=1}^{T} c_s \exp(-is\xi/\gamma_T)) \parallel_2 < \infty.$$

Finally,

$$T^{-1} \sum_{t=1}^{T} \sum_{s=1}^{T} k((t-s)/\gamma_T) a_{Tt} a_{Ts} \leq \int_{-\infty}^{\infty} |\psi(\xi)| d\xi (T^{-1/2} \sum_{t=1}^{T} |a_{Tt}|)^2,$$

and the last term is $o_P(1)$ by Lemma 3 under Assumptions 1 and 2. Therefore,

$$\hat{\sigma}^2 - T^{-1} \sum_{t=1}^{T} \sum_{s=1}^{T} k((t-s)/\gamma_T) c_t c_s \xrightarrow{p} 0,$$

and by Theorem 2.1 of de Jong and Davidson (2000), under Assumptions 1 and 2,

$$T^{-1} \sum_{t=1}^{T} \sum_{s=1}^{T} k((t-s)/\gamma_T) c_t c_s \xrightarrow{p} \sigma^2,$$

which completes the argument. □

Proof of Theorem 2:

First, we prove that $T^{-1/2} m_T \xrightarrow{d} M$ for some random variable M. This is done in two steps; first, we show that $T^{-1/2} m_T$ is $O_P(1)$; second, we derive the limit distribution of $T^{-1/2} m_T$. To show that $T^{-1/2} m_T$ is $O_P(1)$, note that

$$\sup_{m > K} T^{-1} \sum_{t=1}^{T} \text{sgn}(T^{-1/2} x_t - m)$$

$$= T^{-1}\sum_{t=1}^{T}\text{sgn}(T^{-1/2}x_t - K)$$

$$\xrightarrow{d} \int_0^1 \text{sgn}(W(\xi) - K)d\xi = T(K),$$

and therefore

$$P(T^{-1/2}m_T > K) = P(\sup_{m>K} T^{-1}\sum_{t=1}^{T}\text{sgn}(T^{-1/2}x_t - m) \geq 0)$$

$$\rightarrow P(T(K) \geq 0),$$

and because $T(K) \xrightarrow{p} -1$ as $K \rightarrow \infty$, it now follows that

$$\limsup_{K\rightarrow\infty}\limsup_{T\rightarrow\infty} P(T^{-1/2}m_T > K) = 0.$$

A similar argument shows that

$$\limsup_{K\rightarrow\infty}\limsup_{T\rightarrow\infty} P(T^{-1/2}m_T < -K) = 0,$$

thereby establishing that $T^{-1/2}m_T$ is $O_P(1)$. By Theorem 2.7 of Kim and Pollard (1990), it now follows that $T^{-1/2}m_T$ converges to $\text{argmin}_m Q(m)$, where $Q(m)$ is such that

$$Q_T(m) = |T^{-1}\sum_{t=1}^{T}\text{sgn}(T^{-1/2}x_t - m)| \Rightarrow Q(m).$$

Note that formally we cannot use Kim and Pollard's Theorem 2.7 directly because it requires that $|Q(\phi)| \rightarrow \infty$ as $|\phi| \rightarrow \infty$; but this can be fixed easily by considering e.g. $\Phi^{-1}(Q(.))$.) Next, note that for any $m \in \mathbb{R}$

$$Q_T(m) \xrightarrow{d} \int_0^1 \text{sgn}(W(\xi) - m)d\xi = Q(m). \tag{15}$$

Note that this cannot be directly derived from the continuous mapping theorem because of the discontinuity of the "sgn" function; however, the above result can be easily shown to hold following the arguments in Park and Phillips (1999). Also, $Q_T(m)$ is shown below to be stochastically equicontinuous on $[-K, K]$ for any finite K, thereby establishing that $Q_T(m) \Rightarrow Q(m)$ on $[-K, K]$. Stochastic equicontinuity is $Q_T(m)$ holds because for $m < m'$,

$$\sup_{m\in[-K,K]}\sup_{m':|m-m'|<\delta}|Q_T(m) - Q_T(m')|$$

$$= \sup_{-[K/\delta]-1\leq j\leq[K/\delta]}\sup_{m\in[j\delta,(j+1)\delta]}\sup_{m':|m-m'|<\delta}|Q_T(m) - Q_T(m')|$$

$$\leq \sup_{-[K/\delta]-1\leq j\leq [K/\delta]} T^{-1}\sum_{t=1}^{T} I(j\delta \leq T^{-1/2}x_t \leq (j+1)\delta)$$

$$\xrightarrow{d} \sup_{-[K/\delta]-1\leq j\leq [K/\delta]} \int_0^1 I(j\delta \leq W(\xi) \leq (j+1)\delta)d\xi,$$

and by the occupation time formula (see Park and Phillips (1999, p. 271)), the last expression equals

$$\sup_{-[K/\delta]-1\leq j\leq [K/\delta]} \int_{j\delta}^{(j+1)\delta} L(1,s)ds,$$

where $L(t,s)$ is a random process called "local time" that is jointly continuous; see Park and Phillips (1999) for more details and references regarding the local time process $L(.,.)$. Therefore, $\sup_{s\in[-K,K]}|L(1,s)|$ is a well-defined random variable, and the last expression therefore can be bounded by

$$\sup_{s\in[-K,K]} |L(1,s)|\delta,$$

which completes the proof of stochastic equicontinuity of $Q_T(.)$ on $[-K,K]$. Next, note that

$$T^{-3}\sum_{t=1}^{T}(\sum_{j=1}^{t}\text{sgn}(x_j-m_T))^2$$

$$= T^{-1}\sum_{t=1}^{T}(T^{-1}\sum_{j=1}^{t}\text{sgn}(T^{-1/2}x_j - T^{-1/2}m_T))^2,$$

and it can also be shown that

$$T^{-1}\sum_{j=1}^{[\alpha T]}\text{sgn}(x_j-m_T) \Rightarrow \int_0^{\alpha}\text{sgn}(W(\xi)-M)d\xi.$$

This last result is derived by noting that finite-dimensional convergence for each $\alpha \in [0,1]$ holds because of a similar argument as used in Equation (15), and stochastic equicontinuity holds trivially here. From this, it follows that

$$T^{-3}\sum_{t=1}^{T}(\sum_{j=1}^{t}\text{sgn}(x_j-m_T))^2 \xrightarrow{d} \int_0^1(\int_0^{\alpha}\text{sgn}(W(\xi)-M)d\xi)^2 d\alpha.$$

To establish the result for $\hat{\sigma}^2$, note that

$$\gamma_T^{-1}\hat{\sigma}^2 = \gamma_T^{-1}T^{-1}\sum_{t=1}^{T}\text{sgn}(x_t-m_T)^2$$

$$+2\gamma_T^{-1}T^{-1}\sum_{j=1}^{T} k(j/\gamma_T)\sum_{t=1}^{T-j}\text{sgn}(X_T((t+j)/T)-m_T)\text{sgn}(X_T(t/T)-m_T)$$

$$= o_P(1) + 2\int_0^{T/\gamma_T}\int_0^1 k([\alpha\gamma_T]/\gamma_T)\text{sgn}(X_T(\xi+\alpha\gamma_T/T)-m_T)\text{sgn}(X_T(\xi)-m_T)d\xi d\alpha$$

$$= o_P(1) + 2\int_0^{\infty} k(\alpha)d\alpha\int_0^1 \text{sgn}(X_T(\xi)-m_T)^2 d\xi$$

$$= o_P(1) + 2\int_0^{\infty} k(\alpha)d\alpha,$$

where we carried out the substitutions $t/T = \xi$ and $j/\gamma_T = \alpha$. □

TABLE 1

SIZE, KPSS AND INDICATOR KPSS: $\gamma_T = \gamma_{T,0} = 0$

	NORMAL		t_5		t_3		t_2		LOCAL ($c = 1$)		CAUCHY	
T	KPSS	IKPSS	KPSS	IKPSS	KPSS	IKPSS	KPSS	IKPSS	KPSS	IKPSS	KPSS	IKPSS
50	.044	.048	.049	.049	.046	.051	.043	.051	.036	.047	.025	.053
100	.049	.051	.050	.052	.048	.049	.043	.050	.038	.051	.029	.052
200	.049	.050	.051	.051	.048	.049	.043	.051	.035	.050	.028	.051
500	.049	.049	.049	.051	.048	.050	.048	.051	.038	.049	.026	.047
1000	.050	.051	.049	.051	.047	.051	.045	.048	.036	.051	.026	.051
2000	.048	.050	.051	.050	.049	.050	.050	.051	.035	.048	.028	.050
5000	.050	.049	.049	.051	.051	.052	.045	.051	.038	.050	.029	.050

SIZE, KPSS AND INDICATOR KPSS: $\gamma_T = \gamma_{T,1/4}$

	NORMAL		t_3		CAUCHY		NORMAL, $\rho = .5$		t_3, $\rho = .5$		CAUCHY, $\rho = .5$	
T	KPSS	IKPSS	KPSS	IKPSS	KPSS	IKPSS	KPSS	IKPSS	KPSS	IKPSS	KPSS	IKPSS
50	.037	.037	.039	.040	.020	.043	.093	.095	.097	.103	.074	.129
100	.044	.045	.042	.043	.024	.045	.091	.085	.092	.090	.071	.117
200	.047	.048	.044	.046	.026	.046	.097	.090	.096	.091	.075	.116
500	.048	.049	.046	.049	.026	.047	.088	.082	.088	.085	.065	.099
1000	.050	.050	.047	.051	.025	.050	.077	.072	.076	.075	.049	.087
2000	.047	.048	.049	.049	.027	.050	.071	.065	.072	.071	.050	.084
5000	.050	.050	.051	.052	.029	.050	.069	.063	.070	.070	.044	.075

TABLE 2

POWER, KPSS AND INDICATOR KPSS: $\gamma_T = \gamma_{T,0} = 0$

		NORMAL		t_5		t_3		t_2		LOCAL ($c = 1$)		CAUCHY	
λ	T	KPSS	IKPSS	KPSS	IKPSS	KPSS	IKPSS	KPSS	IKPSS	KPSS	IKPSS	KPSS	IKPSS
.0001	50	.053	.066	.052	.062	.049	.065	.051	.075	.075	.106	.087	.234
	100	.060	.064	.060	.068	.060	.074	.066	.104	.110	.140	.139	.479
	200	.098	.084	.099	.102	.102	.129	.120	.230	.187	.230	.231	.784
	500	.309	.233	.304	.301	.316	.393	.333	.590	.386	.489	.414	.982
	1000	.605	.503	.610	.596	.608	.692	.608	.859	.589	.735	.580	.999
	2000	.866	.793	.863	.859	.862	.913	.846	.976	.776	.914	.731	1.000
	5000	.991	.979	.992	.991	.990	.996	.981	1.000	.917	.995	.880	1.000
.001	50	.082	.083	.079	.089	.083	.105	.101	.158	.148	.186	.189	.458
	100	.170	.132	.166	.168	.181	.220	.203	.335	.261	.315	.306	.762
	200	.398	.313	.400	.385	.406	.479	.418	.638	.445	.551	.469	.949
	500	.786	.699	.787	.776	.787	.846	.767	.930	.713	.855	.678	.999
	1000	.954	.915	.955	.951	.950	.972	.933	.992	.861	.968	.812	1.000
	2000	.995	.990	.996	.996	.995	.998	.987	1.000	.934	.998	.902	1.000
	5000	1.000	1.000	1.000	1.000	1.000	1.000	.999	1.000	.975	1.000	.960	1.000
.01	50	.296	.229	.294	.282	.309	.342	.332	.446	.360	.417	.395	.711
	100	.593	.495	.585	.559	.586	.634	.587	.733	.564	.675	.559	.924
	200	.852	.776	.850	.830	.848	.875	.828	.928	.765	.886	.726	.993
	500	.989	.973	.988	.983	.989	.991	.974	.997	.908	.991	.871	1.000
	1000	1.000	.998	1.000	.999	.999	1.000	.995	1.000	.958	.999	.935	1.000
	2000	1.000	1.000	1.000	1.000	1.000	1.000	.999	1.000	.980	1.000	.968	1.000
.1	50	.731	.632	.726	.675	.722	.707	.709	.758	.667	.746	.644	.860
	100	.924	.869	.923	.893	.913	.908	.898	.932	.833	.921	.794	.974
	200	.992	.977	.990	.982	.989	.986	.979	.991	.929	.989	.894	.998
	500	1.000	1.000	1.000	.999	1.000	1.000	.998	1.000	.971	1.000	.955	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.987	1.000	.979	1.000
1	50	.925	.862	.925	.867	.923	.875	.910	.887	.863	.893	.828	.917
	100	.989	.968	.990	.969	.987	.971	.980	.976	.943	.921	.921	.987
	200	1.000	.997	.999	.996	.999	.997	.998	.998	.977	.998	.963	.999
	500	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.991	1.000	.985	1.000

TABLE 3

SIZE-ADJUSTED POWER, KPSS AND INDICATOR KPSS: $\gamma_T = \gamma_{T,0} = 0$

		NORMAL		t_5		t_3		t_2		LOCAL ($c = 1$)		CAUCHY	
λ	T	KPSS	IKPSS	KPSS	IKPSS	KPSS	IKPSS	KPSS	IKPSS	KPSS	IKPSS	KPSS	IKPSS
.0001	50	.059	.070	.053	.063	.053	.064	.059	.074	.092	.111	.115	.229
	100	.062	.063	.060	.066	.064	.075	.075	.104	.128	.138	.168	.475
	200	.099	.084	.099	.101	.105	.132	.131	.226	.208	.230	.266	.782
	500	.310	.236	.306	.300	.321	.395	.337	.588	.410	.491	.450	.983
	1000	.605	.502	.612	.595	.615	.690	.618	.863	.612	.734	.617	.999
	2000	.868	.793	.863	.859	.864	.913	.846	.976	.792	.915	.752	1.000
	5000	.991	.979	.992	.991	.990	.996	.982	1.000	.922	.995	.891	1.000
.001	50	.091	.087	.080	.091	.089	.104	.110	.156	.169	.192	.223	.452
	100	.173	.131	.166	.165	.186	.221	.215	.336	.283	.313	.337	.760
	200	.400	.314	.400	.384	.413	.485	.433	.635	.470	.550	.503	.949
	500	.787	.701	.789	.775	.791	.847	.771	.930	.731	.856	.707	.999
	1000	.954	.915	.955	.950	.951	.971	.936	.993	.872	.969	.831	1.000
	2000	.996	.990	.996	.996	.995	.998	.987	1.000	.939	.997	.912	1.000
	5000	1.000	1.000	1.000	1.000	1.000	1.000	.999	1.000	.977	1.000	.964	1.000
.01	50	.312	.235	.297	.284	.318	.340	.345	.443	.388	.425	.433	.706
	100	.595	.494	.585	.556	.590	.635	.600	.734	.588	.674	.589	.923
	200	.852	.776	.849	.830	.851	.878	.837	.926	.781	.885	.749	.993
	500	.989	.974	.988	.983	.987	.991	.975	.997	.914	.991	.885	1.000
	1000	1.000	.998	1.000	.999	1.000	1.000	.996	1.000	.961	.999	.942	1.000
	2000	1.000	1.000	1.000	1.000	1.000	1.000	.999	1.000	.981	1.000	.972	1.000
.1	50	.742	.643	.728	.678	.729	.705	.720	.756	.691	.751	.677	.856
	100	.925	.869	.923	.890	.915	.908	.904	.932	.845	.920	.814	.974
	200	.992	.977	.990	.982	.989	.986	.981	.991	.933	.989	.906	.998
	500	1.000	1.000	1.000	.999	1.000	1.000	.998	1.000	.973	1.000	.960	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.988	1.000	.981	1.000
1	50	.931	.866	.927	.869	.926	.874	.916	.886	.876	.896	.849	.914
	100	.989	.967	.990	.968	.988	.971	.983	.976	.949	.978	.930	.987
	200	1.000	.997	.999	.996	.999	.997	.998	.998	.979	.998	.968	.999
	500	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.992	1.000	.987	1.000

TABLE 4

SIZE-ADJUSTED POWER, KPSS AND INDICATOR KPSS: $\gamma_T = \gamma_{T,1/4}$

		NORMAL		t_3		CAUCHY		NORMAL, $\rho = .5$		t_3, $\rho = .5$		CAUCHY, $\rho = .5$	
λ	T	KPSS	IKPSS	KPSS	IKPSS	KPSS	IKPSS	KPSS	IKPSS	KPSS	IKPSS	KPSS	IKPSS
.0001	50	.049	.064	.054	.070	.108	.207	.054	.055	.051	.050	.082	.089
	100	.055	.058	.062	.075	.158	.437	.053	.053	.052	.054	.107	.166
	200	.094	.080	.104	.128	.255	.741	.061	.061	.063	.064	.166	.372
	500	.300	.229	.312	.383	.444	.965	.124	.115	.133	.158	.301	.772
	1000	.589	.487	.598	.672	.605	.995	.301	.271	.308	.375	.450	.942
	2000	.857	.783	.853	.902	.749	1.000	.607	.569	.605	.679	.606	.995
	5000	.988	.974	.987	.994	.890	1.000	.907	.890	.906	.940	.796	1.000
.001	50	.075	.076	.083	.103	.205	.396	.060	.060	.060	.060	.135	.146
	100	.150	.116	.174	.207	.314	.669	.079	.071	.084	.089	.205	.324
	200	.376	.295	.395	.457	.482	.897	.182	.150	.176	.196	.331	.625
	500	.757	.674	.759	.813	.693	.990	.484	.448	.494	.550	.542	.923
	1000	.931	.891	.931	.951	.822	.995	.762	.731	.766	.816	.704	.986
	2000	.991	.983	.990	.995	.908	1.000	.941	.929	.939	.957	.826	.999
	5000	1.000	1.000	1.000	1.000	.963	1.000	.998	.997	.997	.999	.927	1.000
.01	50	.257	.193	.278	.303	.387	.582	.122	.110	.131	.128	.256	.262
	100	.514	.428	.520	.559	.537	.781	.282	.244	.288	.305	.388	.524
	200	.784	.712	.791	.809	.715	.939	.548	.511	.557	.581	.570	.790
	500	.964	.945	.963	.966	.873	.994	.869	.847	.870	.884	.772	.966
	1000	.994	.989	.994	.992	.936	.999	.971	.962	.970	.973	.878	.993
	2000	.999	.999	1.000	1.000	.970	1.000	.997	.995	.996	.997	.939	.999
.1	50	.594	.512	.601	.575	.593	.655	.399	.342	.397	.363	.447	.403
	100	.766	.715	.771	.755	.734	.811	.626	.580	.619	.592	.599	.633
	200	.929	.900	.930	.914	.873	.946	.840	.811	.842	.819	.772	.854
	500	.991	.984	.991	.985	.952	.993	.972	.964	.973	.963	.910	.977
	1000	.998	.996	.998	.996	.977	.998	.994	.991	.994	.991	.956	.995
1	50	.705	.631	.716	.648	.730	.668	.579	.526	.604	.517	.610	.486
	100	.824	.768	.832	.784	.838	.812	.748	.684	.742	.672	.743	.665
	200	.947	.917	.951	.921	.937	.944	.905	.870	.908	.864	.884	.870
	500	.993	.984	.993	.985	.981	.992	.983	.972	.984	.969	.962	.979