CAE Working Paper #05-08

A New Asymptotic Theory for Heteroskedasticity-Autocorrelation Robust Tests

by

Nicholas M. Kiefer
and
Timothy J. Vogelsang

January 2005

.

A New Asymptotic Theory for Heteroskedasticity-Autocorrelation Robust Tests

Nicholas M. Kiefer and Timothy J. Vogelsang*
Departments of Economics and Statistical Science, Cornell University

April 2002, Updated January, 2005

Abstract

A new first order asymptotic theory for heteroskedasticity-autocorrelation (HAC) robust tests based on nonparametric covariance matrix estimators is developed. The bandwidth of the covariance matrix estimator is modeled as a fixed proportion of the sample size. This leads to a distribution theory for HAC robust tests that explicitly captures the choice of bandwidth and kernel. This contrasts with the traditional asymptotics (where the bandwidth increases slower than the sample size) where the asymptotic distributions of HAC robust tests do not depend on the bandwidth or kernel. Finite sample simulations show that the new approach is more accurate than the traditional asymptotics. The impact of bandwidth and kernel choice on size and power of t-tests is analyzed. Smaller bandwidths lead to tests with higher power but greater size distortions and large bandwidths lead to tests with lower power but less size distortions. Size distortions across bandwidths increase as the serial correlation in the data becomes stronger. A new data dependent bandwidth is proposed in light of these results. Within a group of popular kernels, it shown that the Bartlett kernel has approximately the highest power and the quadratic spectral (QS) kernel has the lowest power regardless of the bandwidth. However, the Bartlett kernel gives the most size distorted tests whereas the QS kernels give the least size distorted tests. Overall, the results clearly indicate that for bandwidth and kernel choice there is a trade-off between size distortions and power.

Keywords: Covariance matrix estimator, inference, autocorrelation, truncation lag, prewhitening, generalized method of moments, functional central limit theorem.

*We thank Cliff Hurvich, Jeff Simonoff and seminar participants at NYU (Statistics) for helpful comments. We gratefully acknowledge financial support from the National Science Foundation through grant SES-0095211. We thank the Center for Analytic Economics at Cornell University. Corresponding Author: Tim Vogelsang, Department of Economics, Uris Hall, Cornell University, Ithaca, NY 14853-7601, Phone: 607-255-5108, Fax: 607-255-2818, email: tjv2@cornell.edu

1 Introduction

We provide a new and improved approach to the asymptotics of hypothesis testing in time series models with "arbitrary," i.e. unspecified, serial correlation and heteroskedasticity. Our results are general enough to apply to stationary generalized of method of moments (GMM) models. Heteroskedasticity and autocorrelation consistent (HAC) estimation and testing in these models involves calculating an estimate of the spectral density at zero frequency of the estimating equations or moment conditions defining the estimator. Important contributions to the development of these techniques include White (1984), Newey and West (1987), Gallant (1987), Gallant and White (1988), Andrews (1991), Andrews and Monahan (1992), Hansen (1992), Robinson (1998) and de Jong and Davidson (2000). We stress at the outset that we are not proposing new estimators or test statistics; rather we focus on improving the asymptotic distribution theory for existing techniques. Our results, however, do provide some guidance on the choice of HAC estimator.

Conventional asymptotic theory for HAC estimators is well established and has proved useful in providing practical formulas for estimating asymptotic variances. The ingenious "trick" is the assumption that the variance estimator depends on a fraction of sample autocovariances, with the number of sample autocovariances going to infinity, but the fraction going to zero as the sample size grows. Under this condition it has been shown that well-known HAC estimators of the asymptotic variance are consistent. Then, the asymptotic distribution of estimated coefficients can essentially be derived assuming the variance is known. That is, sampling variance of the variance estimator does not appear in the first order asymptotic distribution theory of test statistics regarding parameters of interest. While this is an extremely productive simplifying assumption that leads to standard asymptotic distribution theory for tests, the accuracy of the resulting asymptotic theory if often less than satisfactory. In particular there is a tendency for HAC robust tests to over reject (sometimes substantially) under the null hypothesis in finite samples; see Andrews (1991), Andrews and Monahan (1992), and the July 1996 special issue of *Journal of Business and Economic Statistics* for evidence.

There are two main sources of finite sample distortions. The first source is inaccuracy via the central limit theorem approximation to the sampling distribution of parameters of interest. This becomes a serious problem for data that has strong or persistent serial correlation. The second source is the sampling variability of the HAC estimate of the asymptotic variance and is the focus of this paper. This sampling variability can be substantial and depends on the choice of certain tuning parameters (kernel and bandwidth). Appealing to a consistency result for the asymptotic variance estimator does not capture this important source of sampling variability.

The assumption that the fraction of the sample autocovariances used in calculating the asymptotic variance goes to zero as the sample size goes to infinity is a clever technical assumption that substantially simplifies asymptotic calculations. However, in practice there is a given sample size and some fraction of sample autocovariances is used to estimate the asymptotic variance. Even if a practitioner chooses the fraction based on a rule such that the fraction goes to zero as the sample size grows, it does not change the fact that a positive fraction is being used for a particular data set. The implications of this simple observation have been eloquently summarized by Neave (1970, p.70) in the context of spectral density estimation:

> "When proving results on the asymptotic behavior of estimates of the spectrum of a stationary time series, it is invariably assumed that as the sample size T tends to infinity, so does the truncation point M, but at a slower rate, so that M/T tends to zero. This is a convenient assumption mathematically in that, in particular, it ensures consistency of the estimates, but it is unrealistic when such results are used as approximations to the finite case where the value of M/T cannot be zero".

Based on this observation, Neave (1970) derived an asymptotic approximation for the sampling variance of spectral density estimates under the assumption that M/T is a constant and showed that his approximation was more accurate than the standard approximation.

In this paper, we effectively generalize the approach of Neave (1970) for zero frequency nonparametric spectral density estimators (HAC estimators). We derive the entire asymptotic distribution (rather than just the variance) of these estimators under the assumption that $M = bT$ where $b \in (0, 1]$ is a constant. We show that under this assumption, asymptotic variance estimators converge to a limiting random matrix that is proportional to the unknown asymptotic variance and has a limiting distribution that depends on the kernel (through the second derivative of the kernel) and b. Under this alterative asymptotics, HAC robust test statistics computed in the usual way are shown to have limiting distributions that are pivotal but depend on the kernel and b.

While the assumption that the proportion of sample autocovariances remains fixed as the sample size grows is a better reflection of practice in reality, that alone does not justify the new asymptotic theory. In fact, our asymptotic theory leads to two important innovations for HAC robust testing. First, our asymptotic theory provides a better approximation of the sampling distribution of HAC robust tests by providing a more accurate approximation of the sampling distribution of the asymptotic variance estimator. Whereas the standard approach approximates the sampling variability of the asymptotic variance using the true variance by appealing to a consistency result, our approach approximates this sampling variability by a random variable that depends on the kernel and bandwidth. Second, our asymptotic theory permits a local asymptotic power approximation for HAC robust tests that depends on the kernel and bandwidth. We can theoretically analyze how the choices of kernel and bandwidth affect the power of HAC robust tests. As far as we know, we are the first authors to examine the theoretical relationship between test power and kernel/bandwidth choice. Such an analysis is not possible under the standard first order asymptotics because local asymptotic power does not depend on the choice of kernel or bandwidth. Because of this fact, the existing HAC robust testing literature has focused instead on minimizing the asymptotic truncated MSE of the asymptotic variance estimators when choosing the kernel and bandwidth. For the analysis of HAC robust tests, this is not a completely satisfying situation as noted by Andrews (1991, p.828):

> "If one wants to use [a covariance matrix variance estimator] in forming a test statistic involving [an estimator of the parameter of interest], however, the suitability of the truncated MSE criterion is less clear. A weak argument in its favor is that the asymptotics typically used with such test statistics treat the estimated covariance matrix as though it equals its probability limit. In consequence, in many cases the closer is the covariance matrix estimator to its probability limit, as measured, for example, by truncated MSE, the better is the asymptotic approximation. ... On the other hand, there are cases where the deviation of another part of a test statistic from its limiting behavior is offset by the deviation of another part of the statistic from its limiting behavior. In such cases, the argument above breaks down."

Additional discussion of this point is given by Cushing and McGarvey (1999, p. 80) and Simonoff (1993) provides an illustrative example.

Because we are approaching the distribution theory of HAC robust testing using a perspective that differs from the conventional approach, we think it is important to stress here that an important purpose of asymptotic theory is to provide approximations to sampling distributions. While sampling distributions can be obtained exactly under precise distributional assumptions by a change of variables, this approach often requires difficult if not impossible calculations that may be required on a case by case basis. A unifying approach, giving results applicable in a wide variety of settings, must involve approximations. The usual approach is to consider asymptotic approximations. Whether these are useful in any particular setting is a matter of how well the approximation mimics the exact sampling distribution. Although exact comparisons can be made is simple cases, typically, this approximation is most efficiently assessed by finite sample Monte Carlo experiments, and that is our approach here. This point is emphasized by Barndorff-Nielsen and Cox (1989, p. ix)

> "The approximate arguments are developed by supposing that some defining quantity, often a sample size but more generally an amount of information, becomes large: it must be stressed that this is a technical device for generating approximations whose adequacy always needs assessing, rather than a 'physical' limiting notion."

The main contributions of this paper can be outlined as follows.

1. The new distribution theory provides a better approximation to sampling distributions compared to the conventional asymptotics. Therefore, the new asymptotic approximations should be used for any choice of kernel and bandwidth when calculating HAC robust tests.

2. Finite sample simulations using the new asymptotic critical values show that size distortions are reduced when large bandwidths are used. This is especially true when the data has strong positive serial correlation. Based on this result we propose a new and sensible data dependent bandwidth. Among the Bartlett, Parzen and quadratic spectral (QS) kernels, the QS kernel leads to tests with the least size distortions and the Bartlett kernel leads to tests with the most size distortions.

3. A local asymptotic power analysis shows that power of the tests is higher for small bandwidths than for large bandwidths. Among a group of popular kernels, the Bartlett kernel is approximately the most powerful whereas the QS and Daniell kernels are the least powerful.

4. There is a clear trade-off in bandwidth and kernel choice between size and power.

The remainder of the paper is organized as follows. Section 2 lays out the GMM framework and reviews standard results. Section 3 introduces the new asymptotic theory. Section 4 analyzes the performance of the new asymptotic theory in terms of size distortions and local asymptotic power. The impact of the choice of bandwidth and kernel is analyzed. Section 5 gives concluding comments. Proofs and some formulas are given in two appendices.

The following notation is used throughout the paper. The symbol $\Rightarrow$ denotes weak convergence, $B_j(r)$ denotes a j vector of standard Brownian motions (Wiener processes) defined on $r \in [0,1]$, $\widetilde{B}_j(r) = B_j(r) - rB_j(1)$ denotes a j vector of standard Brownian bridges, and $[rT]$ denotes the integer part of rT for $r \in [0,1]$.

2 Inference in GMM Models: The Standard Approach

We present our results in the GMM framework noting that this covers maximum likelihood estimation (preferred when feasible) and estimating equation methods (Heyde, 1997). Since the influential work of Hansen (1982), GMM is widely used in virtually every field of economics. Heteroskedasticity or autocorrelation of unknown form is often an important specification issue especially in macroeconomics and financial applications. Typically the form of the correlation structure is not of direct interest (if it is, it should be modeled directly). What is desired is an inference procedure that is robust to the form of the heteroskedasticity and serial correlation. HAC covariance matrix estimators were developed for exactly this setting.

Suppose we are interesting in estimating the $p \times 1$ vector of parameters, $\theta \in \Theta \subset R^p$. Let θ_0 denote the true value of θ, and assume θ_0 is an interior point Θ. Let v_t denote a vector of observed data and assume that q moment conditions hold that can be written as

$$E[f(v_t, \theta_0)] = \mathbf{0}, \qquad t = 1, 2, ..., T, \tag{1}$$

where $f(\cdot)$ is a $q \times 1$ vector of functions with $q \geq p$. The moment conditions given by (1) are often derived from economic models and for fixed data, v_t, $f(\cdot)$ can be regarded as an injection $f : \Theta \to R^q$. The expectation is taken over the endogenous variables in v_t, and may be conditional on exogenous elements of v_t. There is no need in what follows to make this conditioning explicit in the notation. The idea behind GMM is to find a value of θ that satisfies as closely as possible the empirical analog of (1). Define

$$g_t(\theta) = T^{-1} \sum_{j=1}^{t} f(v_j, \theta),$$

where $g_T(\theta) = T^{-1} \sum_{t=1}^{T} f(v_t, \theta)$ is the sample analog to (1). When $q > p$ there is usually no solution, $\widehat{\theta}_T$, to the equation $g_T(\widehat{\theta}_T) = \mathbf{0}$, so instead we minimize the weighted sum of squares (and cross products) of the moments to define

$$\widehat{\theta}_T = \arg\min_{\theta \in \Theta} g_T(\theta)' W_T g_T(\theta) \tag{2}$$

where W_T is a $q \times q$ positive definite weighting matrix. Alternatively, $\widehat{\theta}_T$ can also be defined as an estimating equations estimator, the solution to the p first order conditions associated with (2)

$$G_T(\widehat{\theta}_T)' W_T g_T(\widehat{\theta}_T) = \mathbf{0}, \tag{3}$$

where $G_t(\theta) = T^{-1} \sum_{j=1}^{t} \partial f(v_j, \theta)/\partial \theta'$. Of course, when the model is exactly identified and $q = p$, an exact solution to $g_T(\widehat{\theta}_T) = \mathbf{0}$ is attainable and the weighting matrix W_T is irrelevant. Application of the mean value theorem implies that

$$g_t(\widehat{\theta}_T) = g_t(\theta_0) + G_t(\widehat{\theta}_T, \theta_0, \lambda_T)(\widehat{\theta}_T - \theta_0) \tag{4}$$

where $G_t(\widehat{\theta}_T, \theta_0, \lambda_T)$ denotes the $q \times p$ matrix whose i^{th} row is the corresponding row of $G_t(\overline{\theta}_T^{(i)})$ where $\overline{\theta}_T^{(i)} = \lambda_{i,T}\theta_0 + (1 - \lambda_{i,T})\widehat{\theta}_T$ for some $0 \leq \lambda_{i,T} \leq 1$ and λ_T is the $q \times 1$ vector with i^{th} element $\lambda_{i,T}$.

In order to focus on the new asymptotic theory for tests, we avoid listing primitive assumptions and make rather high-level assumptions on the GMM estimator $\widehat{\theta}_T$. Lists of sufficient conditions for these to hold can be found in Hansen (1982) and Newey and McFadden (1994). Our assumptions are:

Assumption 1 $p\lim \widehat{\theta}_T = \theta_0$.

Assumption 2 *$T^{-1/2}\sum_{t=1}^{[rT]} f(v_t,\theta_0) = T^{1/2}g_{[rT]}(\theta_0) \Rightarrow \Lambda B_q(r)$ where $\Lambda\Lambda' = \Omega = \sum_{j=-\infty}^{\infty}\Gamma_j$, $\Gamma_j = E[f(v_t,\theta_0), f(v_{t-j},\theta_0)']$ and Λ can be any matrix square root of Ω.*

Assumption 3 *$p\lim G_{[rT]}(\widehat{\theta}_T) = rG_0$ and $p\lim G_{[rT]}(\widehat{\theta}_T,\theta_0,\lambda_T) = rG_0$ uniformly in $r \in [0,1]$ where $G_0 = E[\partial f(v_t,\theta_0)/\partial\theta']$.*

Assumption 4 *W_T is positive semi-definite and $p\lim W_T = W_\infty$ where W_∞ is a matrix of constants.*

While we will not claim these assumptions are weak, they hold in wide generality for the models seen in economics, and with the exception of Assumption 2 they are fairly standard. Assumption 2 requires that a functional central limit theorem hold for $T^{1/2}g_t(\theta_0)$. This is stronger than the central limit theorem for $T^{1/2}g_T(\theta_0)$ that is required for asymptotic normality of $\widehat{\theta}_T$. However, consistent estimation of the asymptotic variance of $\widehat{\theta}_T$ requires an estimate of Ω. Conditions for consistent estimation of Ω are typically stronger than Assumption 2 and often imply Assumption 2. For example, Andrews (1991) requires that $f(v_t,\theta_0)$ is a mean zero fourth order stationary process that is $\alpha - mixing$. Phillips and Durlauf (1986) show that Assumption 2 holds under the weaker assumption that $f(v_t,\theta_0)$ is a mean zero, $2+\delta$ order stationary process (for some $\delta > 0$) that is $\alpha - mixing$. Thus our assumptions are slightly weaker than those usually given for asymptotic testing in HAC-estimated GMM models.

Under our assumptions $\widehat{\theta}_T$ is asymptotically normally distributed, as recorded in the following lemma which is proved in the appendix.

Lemma 1 *Under Assumptions 1 - 4, as $T \to \infty$,*

$$T^{1/2}(\widehat{\theta}_T - \theta_0) \Rightarrow -(G_0'W_\infty G_0)^{-1}\Lambda^* B_p(1) \sim N(0,V),$$

where $\Lambda^*\Lambda^{*\prime} = G_0'W_\infty\Lambda\Lambda'W_\infty G_0$ and $V = (G_0'W_\infty G_0)^{-1}\Lambda^*\Lambda^{*\prime}(G_0'W_\infty G_0)^{-1}$.

In practice the hard part involved in using this result for inference is the need for a consistent estimator of $\Lambda\Lambda' = \Omega$. The other pieces of V can be easily estimated (i.e. $(G_0'W_\infty G_0)^{-1}$ can be estimated by $[G_T(\widehat{\theta}_T)'W_T G_T(\widehat{\theta}_T)]^{-1}$). If a consistent estimator, $\widehat{\Omega}$, of Ω can be found, then V can be consistently estimated by

$$\widehat{V} = [G_T(\widehat{\theta}_T)'W_T G_T(\widehat{\theta}_T)]^{-1} G_T(\widehat{\theta}_T)'W_T\widehat{\Omega}W_T G_T(\widehat{\theta}_T)[G_T(\widehat{\theta}_T)'W_T G_T(\widehat{\theta}_T)]^{-1}. \quad (5)$$

The HAC literature builds on the time series literature of consistent estimation of spectral densities to suggest consistent estimators of Ω, and therefore of V. The widely used class of nonparametric estimators of Ω take the form

$$\widehat{\Omega} = \sum_{j=-(T-1)}^{T-1} k(j/M)\widehat{\Gamma}_j \quad (6)$$

with

$$\widehat{\Gamma}_j = T^{-1} \sum_{t=j+1}^{T} f(v_t, \widehat{\theta}_T) f(v_{t-j}, \widehat{\theta}_T)' \; for \; j \geq 0,$$

$$\widehat{\Gamma}_j = T^{-1} \sum_{t=-j+1}^{T} f(v_{t+j}, \widehat{\theta}_T) f(v_t, \widehat{\theta}_T)' \; for \; j < 0,$$

where $k(x)$ is a kernel function $k : R \to R$ satisfying $k(x) = k(-x)$, $k(0) = 1$, $|k(x)| \leq 1$, $k(x)$ continuous at $x = 0$ and $\int_{-\infty}^{\infty} k^2(x)dx < \infty$. Often $k(x) = 0$ for $|x| > 1$ so M "trims" the sample autocovariances and acts as a truncation lag. Some popular kernel functions do not truncate, and M is often called a bandwidth parameter in those cases. For kernels that truncate, the cutoff at $|x| = 1$ is arbitrary and is essentially a normalization. For kernels that do not truncate, an implicit normalization must be made since the weights generated by the kernel $k(x)$ and bandwidth, M are the same as those generated by kernel $k(ax)$ with bandwidth aM. Therefore, there is an interaction between bandwidth and kernel choice. We focus on kernels that yield positive definite $\widehat{\Omega}$ for the obvious practical reasons.

Standard asymptotic analysis can proceed under the additional assumption that $M \to \infty$ and $M/T \to 0$ as $T \to \infty$. This assumption on the rate at which M grows has little to do with econometric practice; rather it is an ingenious technical assumption allowing an estimable asymptotic approximation to the asymptotic distribution of $\widehat{\theta}_T$ to be calculated. Under this assumption, $\widehat{\Omega}$ has been shown to be a consistent estimator which in turn delivers a consistent estimator of V through (5). The difficulty in practice for this approach is that any choice of M for a given sample size, T, can be made consistent with the above rate requirement. Although the rate requirement can been refined if one is interested in minimizing the MSE of $\widehat{\Omega}$ (e.g. M must increase at rate $T^{1/3}$ for the Bartlett kernel), these refinements do not deliver specific choices for M. This fact has long been recognized in the spectral density and HAC literatures and data dependent methods for choosing M have been proposed. See Andrews (1991) and Newey and West (1994). The basic idea in those papers is to choose M to minimize the truncated MSE of $\widehat{\Omega}$. Because the MSE of $\widehat{\Omega}$ depends on the serial correlation structure of $f(v_t, \theta_0)$, the practitioner must estimate the serial correlation structure of $f(v_t, \theta_0)$ either nonparametrically or with an "approximate" parametric model. While data dependent methods are a significant improvement over the basic case for empirical implementation, the practitioner is still faced with either a choice of approximating parametric model or the choice of bandwidth in a preliminary nonparametric estimation problem. See den Haan and Levin (1997) for details and additional practical challenges.

An alternative to the nonparametric approach has been advocated by den Haan and Levin (1997,1998). They propose estimating Ω by fitting a VAR model to $f(v_t, \widehat{\theta}_T)$. If the VAR, specifically the lag length, is correctly specified, the resulting estimator of Ω is consistent. A "more robust" approach, achieving essentially the same generality as the nonparametric approach, can be obtained by constructing an asymptotic theory that has the lag length increasing with the sample size at a suitable rate. Of course, consistency in this case requires assumptions on the growth in lag length with respect to sample size; a result as theoretically delicate as the nonparametric results based on the growth in M as T grows.

We conclude this section by pointing out a simple fact about the standard approach to HAC robust testing that helps motivate the remainder of the paper. Because the focus in the standard theory is on obtaining consistent estimates of Ω, the asymptotic approximation used for tests of θ_0 is the same regardless of what specific kernel or bandwidth is used in practice. It is well known that the choice of kernel and, especially, bandwidth can greatly affect the sampling behavior of $\widehat{\Omega}$. Hence, these choices can greatly affect the sampling distribution of tests of θ_0. Therefore, it is desirable to have an asymptotic approximation that explicitly reflects the choice of kernel and bandwidth as such an asymptotics is likely to be more accurate. In the next section we show that such an asymptotic theory can be developed.

3 A New Asymptotic Theory

3.1 Distribution of $\widehat{\Omega}$ when $M = bT$

Rather than focus on the rate at which M must increase for $\widehat{\Omega}$ to be consistent or the rate that M must increase to minimize $MSE(\widehat{\Omega})$, we instead take M and T as given and show that there exists a useful asymptotic approximation for the sampling distribution of $\widehat{\Omega}$ that explicitly reflects M given T. Suppose the asymptotic behavior of $\widehat{\Omega}$ is derived under the assumption that $b = M/T$ is held constant as T goes to ∞. In other words, suppose the bandwidth is modeled as $M = bT$ where $b \in (0, 1]$ is fixed. To avoid any confusion in later developments, we denote by $\widehat{\Omega}_{M=bT}$ estimators of Ω given by (6) with $M = bT$. The corresponding estimator of the asymptotic variance of $\widehat{\theta}_T$ follows from (5) and is denoted by $\widehat{V}_{M=bT}$. The limiting distribution of $\widehat{\Omega}_{M=bT}$ can be written in terms of $Q_i(b)$, an $i \times i$ random matrix that takes on one of three forms depending on the second derivative of the kernel. The following definition gives the forms of $Q_i(b)$.

Definition 1 *Let the $i \times i$ random matrix, $Q_i(b)$ be defined as follows. If $k(x)$ is twice continuously differentiable everywhere,*

$$Q_i(b) = -\int_0^1 \int_0^1 \frac{1}{b^2} k''(\frac{r-s}{b}) \widetilde{B}_i(r) \widetilde{B}_i(s)' dr ds.$$

If $k(x)$ is continuous, $k(x) = 0$ for $|x| \geq 1$, and $k(x)$ is twice continuously differentiable everywhere *except for* $|x| = 1$,

$$Q_i(b) = -\iint_{|r-s|<b} \frac{1}{b^2} k''(\frac{r-s}{b}) \widetilde{B}_i(r) \widetilde{B}_i(s)' dr ds + \frac{k'_-(1)}{b} \int_0^{1-b} \left(\widetilde{B}_i(r+b) \widetilde{B}_i(r)' + \widetilde{B}_i(r) \widetilde{B}_i(r+b)' \right) dr.$$

where $k'_-(1) = \lim_{h \longrightarrow 0} \left[(k(1) - k(1-h))/h \right]$, *i.e.* $k'_-(1)$ *is the derivative of* $k(x)$ *from the left at* $x = 1$. *If* $k(x)$ *is the Bartlett kernel (see the formula appendix)*

$$Q_i(b) = \frac{2}{b} \int_0^1 \widetilde{B}_i(r) \widetilde{B}_i(r)' dr - \frac{1}{b} \int_0^{1-b} \left(\widetilde{B}_i(r+b) \widetilde{B}_i(r)' + \widetilde{B}_i(r) \widetilde{B}_i(r+b)' \right) dr.$$

For clarity, we first consider the asymptotic distribution of $\widehat{\Omega}_{M=bT}$ for the case of exactly identified models.

Theorem 1 *(Exactly Identified Models) Suppose that* $q = p$. *Let* $b \in (0, 1]$ *be a constant. Let* $Q_p(b)$ *be given by Definition 1 for* $i = p$. *Then, under Assumptions 1-4, as* $T \to \infty$,

$$\widehat{\Omega}_{M=bT} \Rightarrow \Lambda Q_p(b) \Lambda'.$$

Several useful observations can be made regarding this theorem. As expected, $\widehat{\Omega}_{M=bT}$ is not a consistent estimator. However, $\widehat{\Omega}_{M=bT}$ converges to a matrix of random variables (rather than constants) that is proportional to Ω through Λ and Λ'. These limiting random variables depend on the kernel through $k''(x)$ and $k'_-(1)$ and on the bandwidth through b but are otherwise nuisance parameter free. Because of the asymptotic proportionality to Ω, nuisance parameter free asymptotic distributions can be obtained for tests of θ_0 when using $\widehat{\Omega}_{M=bT}$ (details are given below). Therefore, it is possible to obtain a first order asymptotic distribution theory that explicitly captures the choice of kernel and bandwidth. Note that Theorem 1 generalizes results obtained by Kiefer and Vogelsang (2002)b and Kiefer and Vogelsang (2002)a where the focus was $b = 1$.

When $q > p$ and the model is overidentified, the limiting expressions for $\widehat{\Omega}_{M=bT}$ are more complicated and asymptotic proportionality to Ω no longer holds. This was established for the special case of $b = 1$ by Vogelsang (2003). This does not mean, however, that valid testing is not possible when using $\widehat{\Omega}_{M=bT}$ in overidentified models because the required asymptotic proportionality does holds for $G_T(\widehat{\theta}_T)' W_T \widehat{\Omega}_{M=bT} W_T G_T(\widehat{\theta}_T)$, the middle term in $\widehat{V}_{M=bT}$. The following theorem provides the relevant result.

Theorem 2 *(Over-identified Models) Suppose that* $q > p$. *Let* $b \in (0, 1]$ *be a constant. Let* $Q_p(b)$ *be given by Definition 1 for* $i = p$. *Let* $\Lambda^* = G_0' W_\infty \Lambda$. *Under Assumptions 1-4, as* $T \to \infty$,

$$G_T(\widehat{\theta}_T)' W_T \widehat{\Omega}_{M=bT} W_T G_T(\widehat{\theta}_T) \Rightarrow \Lambda^* Q_p(b) \Lambda^{*\prime}.$$

This theorem shows that $G_T(\widehat{\theta}_T)' W_T \widehat{\Omega}_{M=bT} W_T G_T(\widehat{\theta}_T)$ is asymptotically proportional to $\Lambda^*\Lambda^{*\prime}$ and otherwise only depends on the random matrix $Q_p(b)$. It directly follows that $\widehat{V}$ is asymptotically proportional to V, and asymptotically pivotal tests can be obtained.

3.2 Inference using $\widehat{\Omega}_{M=bT}$

We now examine the limiting null distributions of tests regarding θ_0 when the bandwidth is modeled as $M = bT$. Consider the hypotheses

$$H_0 : r(\theta_0) = 0$$
$$H_1 : r(\theta_0) \neq 0$$

where $r(\theta)$ is an $m \times 1$ vector ($m \leq p$) of continuously differentiable functions with first derivative matrix, $R(\theta) = \partial r(\theta)/\partial\theta'$. Applying the delta method to Lemma 1 we obtain

$$T^{1/2} r(\widehat{\theta}_T) \Rightarrow -R(\theta_0) V^{1/2} B_p(1) \equiv N(\mathbf{0}, V_R), \tag{7}$$

where $V_R = R(\theta_0) V R(\theta_0)'$. Using (7) one can construct a Wald-type test of the null hypothesis or a t-test in the case of $m = 1$. We consider the Wald-type test

$$F_b^* = T r(\widehat{\theta}_T)' \left(R(\widehat{\theta}_T) \widehat{V}_{M=bT} R(\widehat{\theta}_T)' \right)^{-1} r(\widehat{\theta}_T)/m,$$

which, except for the m in the denominator, is the Wald-type statistic that would usually be used in practice. When $m = 1$ a t-statistic can be computed as

$$t_b^* = \frac{T^{1/2} r(\widehat{\theta}_T)}{\sqrt{R(\widehat{\theta}_T) \widehat{V}_{M=bT} R(\widehat{\theta}_T)'}}.$$

Often, the significance of individual statistics are of interest which leads to t-statistics of the form

$$t_b^* = \frac{\widehat{\theta}_{iT}}{se(\widehat{\theta}_{iT})},$$

where $se(\widehat{\theta}_i) = \sqrt{T^{-1} \widehat{V}_{M=bT}^{ii}}$ and $\widehat{V}_{M=bT}^{ii}$ is the i^{th} diagonal element of the $\widehat{V}_{M=bT}$ matrix. The formulas for t_b^* are the same as in the standard approach.

Note that some kernels, including the Tukey-Hanning, allow negative variance estimates. In this case some convention must be adopted in calculating the denominator of the test statistics. Equally arbitrary conventions include reflection of negative values through the origin or setting negatives to a small positive value. We see no merit in using a kernel allowing negative estimated variances absent a compelling argument in a specific case. Nevertheless, we have experimented with the Tukey-Hanning and trapezoid kernels and results not reported here do not support their consideration over a kernel guaranteeing positive variance estimates.

The following theorem provides the asymptotic null distributions of F_b^* and t_b^*.

Theorem 3 *Let $b \in (0, 1]$ be a constant and suppose $M = bT$. Let $Q_i(b)$ be given by Definition 1 for $i = m$. Then, under Assumptions 1-4 and H_0, as $T \to \infty$,*

$$F_b^* \Rightarrow B_m(1)' Q_m(b)^{-1} B_m(1)/m,$$

if $m = 1$,

$$t_b^* \Rightarrow \frac{B_1(1)}{\sqrt{Q_1(b)}}.$$

Theorem 3 shows that when the bandwidth is modeled as a constant proportion of the sample size, asymptotically pivotal tests are obtained. And, the asymptotic distributions reflect the choices of kernel and bandwidth. This contrasts asymptotic results under the standard approach where F_b^* would have a limiting χ_m^2/m distribution and t_b^* would have a $N(0, 1)$ limiting distribution regardless of the choice of M and $k(x)$. It is natural to expect that the new asymptotics given by Theorem 3 provide a more accurate approximation in finite samples than the traditional asymptotics. Heuristically, the traditional asymptotics approximates the random variable, $\widehat{\Omega}$, by a constant whereas our asymptotics approximates it with the random variable $Q_m(b)$. While this new approximation is not guaranteed to be better, it usually is better as we show in Section 4.

3.3 Asymptotic Critical Values

The limiting distributions given by Theorem 3 are nonstandard. Analytical forms of the densities are not available with the exception of t_b^* for the case of the Bartlett kernel with $b = 1$ (see Abadir and Paruolo, 2002 and Kiefer and Vogelsang, 2002b). However, because the limiting distributions are simple functions of standard Brownian motions, critical values are easily obtained using simulations. We provide critical values for the t_b^* statistic for a selection of popular kernels (see the formula appendix for formulas for the kernels). To save space, additional critical values for the F_b^* test will be made available in a follow-up paper.

For each kernel we give right tail critical values for t_b^* (left tail critical values follow by symmetry around zero) for $b = 0.02, 0.04, ...0.98, 1.0$. The critical values are tabulated in Tables I through V. The critical values were calculated via simulation methods using 50,000 replications. Normalized partial sums of 1,000 *i.i.d.* $N(0, 1)$ random deviates were used to approximate the standard Brownian motions in the respective distributions given by Theorem 3. In practice, given the kernel and bandwidth, M, we recommend that the critical value corresponding to $b = M/T$ be used. Critical values can be interpolated for values of b that fall between the values on the grid.

Two patterns in the critical values are worth noting. First, for small values of b, the critical values are close to the critical values of a standard normal for each of the kernels. This is not surprising given that the standard asymptotics has b going to zero. Second, as b increases, the critical values increase in magnitude suggesting that sampling variation in t_b^* increases as the bandwidth grows. This suggests that power may be lower with larger bandwidths. This intuition is confirmed in Section 4.

4 Choice of Kernel and Bandwidth and Performance

In this section we analyze the choice of kernel and bandwidth on the performance of HAC robust tests. We focus on accuracy of the asymptotic approximation under the null and on local asymptotic power. We focus on simple models for clarity. As far as we know, our analysis is the first to theoretically explore the effects of kernel and bandwidth choice on power of HAC robust tests.

4.1 Accuracy of the Asymptotic Approximation under the Null

The way to evaluate the accuracy of an asymptotic approximation to a null distribution, or indeed any approximation, is to compare the approximate distribution to the exact distribution. Sometimes this can be done analytically; more commonly the comparison can be made by simulation. We argued above that our approximation to the distribution of HAC robust tests was likely to be better than the usual approximation, since ours takes into account the randomness in the estimated variance. However, as noted, that argument is unconvincing in the absence of evidence on the approximation's performance. We provide results for three popular positive definite kernels: Bartlett, Parzen and QS. Results for the Bohman and Daniell kernel are similar and are not reported here.

The simulations were based on the following two simple regression models

$$y_t = \theta_1 + u_t, \tag{8}$$

$$y_t = \theta_1 + \theta_2 x_t + u_t, \tag{9}$$

where $u_t = \rho u_{t-1} + \xi_t$, $\xi_t \sim i.i.d.\ N(0,1)$, $u_0 = 0$, x_t is a scalar $AR(1)$ process given by $x_t = 0.5x_{t-1} + \varepsilon_t$, $\varepsilon_t \sim i.i.d.\ N(0,1)$, $x_0 = 0$. ξ_t and ε_t are assumed to be uncorrelated with each other. Model (8) is sometimes called a simple location model. This model provides the simplest environment in which to do the analysis. We generated data according to models (8) and (9) with $\theta_1 = 0$ and $\theta_2 = 0$, and we analyzed t-statistics for testing $H_0 : \theta_1 \leq 0$ in model (8) and testing $H_0 : \theta_2 \leq 0$ in model (9). We report results for sample sizes $T = 50, 100, 200$ and $\rho = -0.8, -0.5, -0.3, 0.0, 0.3, 0.5, 0.7, 0.9, 0.95$. We report results without prewhitening and with $AR(1)$ prewhitening. In all cases 10,000 replications were used.

For each kernel, several different bandwidth choices were used. The first set of bandwidths are deterministic and are given by $M = bT$ for $b = 0.1, 0.25, 0.35, 0.5, 0.65, 0.75, 0.9, 1.0$. The second bandwidth is the data dependent bandwidth proposed by Andrews (1991) based on the VAR(1) plug-in formula. For a given kernel, we denote this bandwidth by $\widehat{M}$ and the corresponding value for b is denoted by $\widehat{b} = \widehat{M}/T$. The third bandwidth is a new data dependent bandwidth defined as

$$\widetilde{M} = \min(|\widehat{\rho}|, 1)T,$$

where $\widehat{\rho}$ is obtained from the regression of $\widehat{u}_t$ on $\widehat{u}_{t-1}$ and $\widehat{u}_t$ are the least squares residuals. This bandwidth leads to a value of b denoted by $\widetilde{b} = \widetilde{M}/T = \min(|\widehat{\rho}|, 1)$. The motivation for this bandwidth is given below.

A t-statistic computed using $\widehat{M}$ with rejection probabilities based on $N(0,1)$ critical values is denoted by $t_{\widehat{b}}$. This statistic is the benchmark for the standard approach. We can also use the new asymptotic critical values for this t-statistic and we denote that test by $t^*_{\widehat{b}}$. For each replication, the critical value used for $t^*_{\widehat{b}}$ corresponds to the realized value of $\widehat{b}$. A t-statistic computed using $\widetilde{M}$ with rejection probabilities based on the new asymptotic critical values is denoted by $t^*_{\widetilde{b}}$. For each replication, the critical value used is for $t^*_{\widetilde{b}}$ corresponds to the realized value of $\widetilde{b}$. The t-statistics computed using fixed values of b are denoted by t^*_b with rejection probabilities computed using the new asymptotic critical values.

The empirical null rejection probabilities are given in Tables VI through VIII for the simple location model (8) and in Tables IX through XI for the two variable regression model (9). In all cases, the nominal level is 0.05. The third and fifth columns of each table report the average, across the 10,000 replications, of the data dependent bandwidths measured as proportions of the sample sizes, i.e. averages of $\widehat{b}$ and $\widetilde{b}$.

Several interesting patterns can be seen in the tables. First, when the errors have positive serial correlation ($\rho > 0$), rejection probabilities are closer to 0.05 for $t^*_{\widehat{b}}$ than $t_{\widehat{b}}$. The differences can be quite large as seen for the Bartlett kernel in the simple location model (Table VIII). When the serial correlation is negative ($\rho < 0$), $t^*_{\widehat{b}}$ tends to under-reject relative to $t_{\widehat{b}}$ although this is no longer true when prewhitening is used. These patterns strongly suggest that even if the bandwidth is chosen using traditional methods, the new asymptotic approximation should be used as it is more accurate for the economically relevant case of positive serial correlation.

Second, use of the new data dependent bandwidth, $\widetilde{M}$ further improves the accuracy of the approximation as seen by the results for $t^*_{\widetilde{b}}$. The reason (and motivation) for why $\widetilde{b}$ delivers a more accurate test can be seen in the patterns of the tests with fixed b. Notice that, regardless of kernel, empirical rejection probabilities tend to get closer to 0.05 as b increases. This is especially true for strong positive serial correlation. The new data dependent bandwidth is designed to exploit this pattern. When serial correlation is weak ($\widehat{\rho}$ close to zero), $\widetilde{b}$ does not have to be large to give an accurate test. But, when serial correlation is strong ($\widehat{\rho}$ close to one), $\widetilde{b}$ is much bigger, thus reducing the tendency to over-reject. Because it is usually not known in practice just how strong the serial correlation is, and because it is not obvious how one would generalize the formula for $\widetilde{M}$ for unknown error structure, a conservative approach with respect to minimizing the tendency to over-reject is to use $b = 1$.

Third, comparison across kernels indicates that the QS kernel delivers tests with the most accurate asymptotic approximation. This is generally true regardless of model, sample size or bandwidth choice. This may be related to the fact that the QS kernel minimizes the MSE of $\widehat{\Omega}$. However, the MSE optimality of the QS kernel only holds when used with the appropriate bandwidth rule. Here, the QS kernel performs well regardless of the bandwidth choice.

Fourth, whereas prewhitening tends to reduce the over-rejection problem of $t_{\widehat{b}}$, it does not always improve matters for $t^*_{\widehat{b}}$ and $t^*_{\widetilde{b}}$. See, for example, the Bartlett and QS kernels in the simple location model (Tables VIII and IX). This seemingly paradoxical result is easy to explain. If we focus on the fixed b statistics, we see that prewhitening reduces the tendency to over-reject across the board (both models, all kernels and all T). However, for small values of b, the tendency to over-reject stays relatively large with prewhitening. Because prewhitening removes most of serial correlation, $\widehat{b}$ and $\widetilde{b}$ tend to be close to zero and the tests over-reject. What is needed is a larger bandwidth when using prewhitening.

Fifth, all of the asymptotic approximations tend to become less accurate as ρ approaches one. This is expected because the accuracy of the central (and functional central) limit theorem approximation deteriorates as the model approaches a nonstationary border. But, this source of distortion cannot explain the large differences in rejection probabilities between $t_{\widehat{b}}$ and $t^*_{\widehat{b}}$ when $\rho \geq 0.7$. The reason is that there is an additional source of distortion for $t_{\widehat{b}}$ that is not present for $t^*_{\widehat{b}}$. Recall that the asymptotic critical values become larger for t^*_b as b approaches one. This suggests that sampling variability in HAC robust t-tests increases as the bandwidth increases. Notice from the tables that for a given sample size, the Andrews (1991) data dependent bandwidth increases as ρ approaches one. For the Parzen kernel, it can become quite large. Because under the traditional approach, the critical values are the same regardless of bandwidth used in practice, rejection probabilities will increase as the bandwidth increases. Therefore, part of the over-rejection problem of the traditional approach is caused by the invariance of the traditional first order asymptotics with respect to bandwidth choice.

4.2 Local Asymptotic Power

For the purposes of keeping size distortions small, the preceding subsection suggested that *a)* large bandwidths should be used and *b)* the QS kernel dominates the Bartlett and Parzen kernel. The other suitable metric by which to judge the choice of kernel and bandwidth is power. In this section we compare power of HAC robust t-tests using a local asymptotic power analysis. Our analysis permits comparison of power across bandwidths and across kernels. Such a comparison is not possible using the traditional first order asymptotics because local asymptotic power is the same for all bandwidths and kernels.

For clarity, we restrict attention to linear regression models. Given the results in Theorem 1, the derivations in this section are very simple extensions of results given by Kiefer and Vogelsang (2002)b. Therefore, details are kept to a minimum. Consider the regression model

$$y_t = x_t'\theta_0 + u_t \tag{10}$$

with θ_0 and x_t $p \times 1$ vectors. In terms of the general model we have $f(v_t, \theta_0) = x_t\,(y_t - x_t'\theta_0)$. Without loss of generality, we focus on θ_{i0}, one element of θ, and consider null and alternative hypotheses

$$H_0 : \theta_{i0} \leq 0$$
$$H_1 : \theta_{i0} = cT^{-1/2}$$

where $c > 0$ is a constant. If the regression model satisfies Assumptions 1 through 4, then we can use the results of Theorem 1 and results from Kiefer and Vogelsang (2002)b to easily establish that under the local alternative, H_1, as $T \to \infty$,

$$t_b^* \Rightarrow \frac{\delta + B_1(1)}{\sqrt{Q_1(b)}}, \tag{11}$$

where $\delta = c/\sqrt{V^{ii}}$, V^{ii} is the i^{th} diagonal element of V, and $Q_1(b)$ is given by Definition 1 for $i = 1$.

Asymptotic power curves can be computed for given bandwidths and kernels by simulating the asymptotic distribution of t_b^* based on (11) for a range of values for δ and computing rejection probabilities with respect to the relevant null critical value. Using the same simulation methods as for the asymptotic critical values, local asymptotic power was computed for $\delta = 0, 0.2, 0.4, ..., 4.8, 5.0$ using 5% asymptotic null critical values.

The power results are reported in two ways. Figures 1-8 plot power across the kernels for a given value of b. Figures 9-13 plot power across values of b for a given kernel. Figures 1-8 show that for small bandwidths, power is essentially the same across kernels. As b increases, it becomes clear that the Bartlett kernel has the highest power while the QS and Daniell kernels have the lowest power. If power is the criterion used to choose a test, then the Bartlett kernel is the best choice within this set of five kernels. If we compare the Bartlett, Parzen and QS kernels, we see that the power ranking of these kernels is the reverse of their ranking based on accuracy of the asymptotic approximation under the null.

Figures 9-13 show how the choice of bandwidth affects power. Regardless of the kernel, power is highest for small bandwidths and lowest for large bandwidths and power is decreasing in b. These figures also show that power of the Bartlett kernel is least sensitive to b whereas power of the QS and Daniell kernels is the most sensitive to b. Again, power rankings of b are the opposite of rankings of b based on accuracy of the asymptotic approximation under the null.

5 Conclusions

We have provided a new approach to the asymptotic theory of HAC robust testing. Our results are general enough to apply to stationary models estimated by GMM. In our approach, the ratio of bandwidth to sample size is held constant when deriving the asymptotic behavior of the relevant covariance matrix estimator (i.e. zero frequency spectral density estimator) . In standard asymptotics, this ratio is sent to zero. Our approach improves upon two well known problems with the standard approach. First, as has been well documented in the literature, the standard asymptotic approximation of the sampling behavior of tests is often poor. Second, the kernel and bandwidth choice do not appear in the approximate distribution, leaving the standard theory silent on the choice of kernel and bandwidth with respect to properties of the tests. Our theory leads to approximate distributions that explicitly depend on the kernel and bandwidth. The new approximation performs much better and gives insight into the choice of kernel and bandwidth with respect to test behavior.

The new approximations should be used for HAC robust test statistics for any choice of kernel and bandwidth. Our approximation is an unambiguous improvement over the standard approximation in most cases considered. We show that size distortions are reduced when large bandwidths are used, but so is asymptotic power. Generally there is a trade-off in bandwidth and kernel choice between size (the accuracy of the approximation) and power. Among a group of popular kernels, the QS kernel leads to the least size distortion, while the Bartlett kernel leads to tests with highest power (and generally acceptable size distortion when large bandwidths are used). We also give an alternative simple rule for data dependent bandwidth selection. While this new bandwidth rule is as arbitrary as any other, it has intuitive appeal and appears to perform well.

6 Appendix: Proofs

We first define some relevant functions and derive preliminary results before proving the lemma and theorems. Define the functions

$$k^*(x) = k\left(\frac{x}{b}\right),$$

$$K_{ij} = k\left(\frac{i-j}{bT}\right) = k^*\left(\frac{i-j}{T}\right),$$

$$\Delta^2 K_{ij} = (K_{ij} - K_{i,j+1}) - (K_{i+1,j} - K_{i+1,j+1}),$$

$$D_T^*(r) = T^2\left[\left(k^*\left(\frac{[rT]+1}{T}\right) - k^*\left(\frac{[rT]}{T}\right)\right) - \left(k^*\left(\frac{[rT]}{T}\right) - k^*\left(\frac{[rT]-1}{T}\right)\right)\right].$$

Notice that

$$T^2\Delta^2 K_{ij} = -D_T^*\left(\frac{i-j}{T}\right).$$

Because $k(r)$ is an even function around $r = 0$, $D_T^*(-r) = D_T^*(r)$. If $k^{*\prime\prime}(r)$ exists then $\lim_{T\to\infty} D_T^*(r) = k^{*\prime\prime}(r)$ by the definition of the second derivative. If $k^{*\prime\prime}(r)$ is continuous, then $D_T^*(r)$ converges to $k^{*\prime\prime}(r)$ uniformly in r. Define the stochastic process

$$X_T(r) = G_T(\widehat{\theta}_T)' W_T T^{1/2} g_{[rT]}(\theta_0).$$

It directly follows from Assumptions 2, 3 and 4 that

$$X_T(r) \Rightarrow G_0' W_\infty \Lambda B_q(r) \equiv \Lambda^* B_p(r). \tag{12}$$

Proof of Lemma 1: Setting $t = T$, multiplying both sides of (4) by $G_T(\widehat{\theta}_T)' W_T$, and using the first order condition $G_T(\widehat{\theta}_T)' W_T g_T(\widehat{\theta}_T) = 0$ gives

$$0 = G_T(\widehat{\theta}_T)' W_T g_T(\theta_0) + G_T(\widehat{\theta}_T)' W_T G_T(\widehat{\theta}_T, \theta_0, \lambda_T)(\widehat{\theta}_T - \theta_0). \tag{13}$$

Solving (13) for $(\widehat{\theta}_T - \theta_0)$ and scaling by $T^{1/2}$ gives

$$\begin{aligned} T^{1/2}(\widehat{\theta}_T - \theta_0) &= -[G_T'(\widehat{\theta}_T) W_T G_T(\widehat{\theta}_T, \theta_0, \lambda_T)]^{-1} G_T'(\widehat{\theta}_T) W_T T^{1/2} g_T(\theta_0) \\ &= -[G_T(\widehat{\theta}_T)' W_T G_T(\widehat{\theta}_T, \theta_0, \lambda_T)]^{-1} X_T(1). \end{aligned} \tag{14}$$

Because $p\lim G_T(\widehat{\theta}_T)' W_T G_T(\widehat{\theta}_T, \theta_0, \lambda_T) = G_0' W_\infty G_0$ by Assumptions 3 and 4, it follows from (12) that

$$T^{1/2}(\widehat{\theta}_T - \theta_0) \Rightarrow -\left(G_0' W_\infty G_0\right)^{-1} \Lambda^* B_p(r).$$

The proof of Theorem 1 follows the same arguments as the proof for Theorem 2 and is omitted.

Proof of Theorem 2: Define the random process

$$\widetilde{X}_T(r) = G_T(\widehat{\theta}_T)' W_T T^{1/2} g_{[rT]}(\widehat{\theta}_T).$$

Plugging in for $g_{[rT]}(\widehat{\theta}_T)$ using (4) gives

$$\begin{aligned} \widetilde{X}_T(r) &= X_T(r) + G_T(\widehat{\theta}_T)' W_T G_{[rT]}(\widehat{\theta}_T, \theta_0, \lambda_T) T^{1/2}(\widehat{\theta}_T - \theta_0) \\ &= X_T(r) - G_T(\widehat{\theta}_T)' W_T G_{[rT]}(\widehat{\theta}_T, \theta_0, \lambda_T) \left[G_T(\widehat{\theta}_T)' W_T G_T(\widehat{\theta}_T, \theta_0, \lambda_T)\right]^{-1} X_T(1), \end{aligned}$$

using (14). It directly follows from Assumptions 3 and 4 and (12) that

$$\begin{aligned} \widetilde{X}_T(r) &\Rightarrow \Lambda^* B_p(r) - r G_0' W_\infty G_0 \left(G_0' W_\infty G_0\right)^{-1} \Lambda^* B_p(1) \\ &= \Lambda^* \left(B_p(r) - r B_p(1)\right) \equiv \Lambda^* \widetilde{B}_p(r). \end{aligned} \tag{15}$$

Straightforward algebra gives

$$\widehat{\Omega}_{M=bT} = \sum_{j=-(T-1)}^{T-1} k\left(\frac{j}{bT}\right)\widehat{\Gamma}_j = T^{-1}\sum_{i=1}^{T}\sum_{j=1}^{T} f(v_j,\widehat{\theta}_T)K_{ij}f(v_j,\widehat{\theta}_T)'.$$

Using algebraic arguments similar to those used by Kiefer and Vogelsang (2002)b, it is straightforward to show that

$$\widehat{\Omega}_{M=bT} = T^{-1}\sum_{i=1}^{T-1}\sum_{j=1}^{T-1}\Delta^2 K_{ij}Tg_i(\widehat{\theta}_T)Tg_j(\widehat{\theta}_T)' + g_T(\widehat{\theta}_T)\sum_{i=1}^{T-1}(K_{Ti}-K_{T,i+1})Tg_i(\widehat{\theta}_T)' + \left(\sum_{j=1}^{T} f(v_j,\widehat{\theta}_T)K_{jT}\right)g_T(\widehat{\theta}_T)'. \tag{16}$$

Using (16) it directly follows that

$$\begin{aligned}
&G_T(\widehat{\theta}_T)'W_T\widehat{\Omega}_{M=bT}W_TG_T(\widehat{\theta}_T)\\
&= T^{-1}\sum_{i=1}^{T-1}\sum_{j=1}^{T-1}\Delta^2 K_{ij}G_T(\widehat{\theta}_T)'W_TTg_i(\widehat{\theta}_T)Tg_j(\widehat{\theta}_T)'W_TG_T(\widehat{\theta}_T)\\
&= T^{-2}\sum_{i=1}^{T-1}\sum_{j=1}^{T-1}T^2\Delta^2 K_{ij}G_T(\widehat{\theta}_T)'W_TT^{1/2}g_i(\widehat{\theta}_T)T^{1/2}g_j(\widehat{\theta}_T)'W_TG_T(\widehat{\theta}_T)\\
&= T^{-2}\sum_{i=1}^{T-1}\sum_{j=1}^{T-1}-D_T^*\left(\frac{i-j}{T}\right)G_T(\widehat{\theta}_T)'W_TT^{1/2}g_i(\widehat{\theta}_T)T^{1/2}g_j(\widehat{\theta}_T)'W_TG_T(\widehat{\theta}_T)
\end{aligned} \tag{17}$$

where the second and third terms of (16) vanish because from (3) we have

$$\begin{aligned}
G_T(\widehat{\theta}_T)'W_TTg_T(\widehat{\theta}_T) &= 0,\\
Tg_T(\widehat{\theta}_T)'W_TG_T(\widehat{\theta}_T) &= 0.
\end{aligned}$$

The rest of proof is divided into three cases.

Case 1: $k(x)$ is twice continuously differentiable. Using (17) it follows that

$$\begin{aligned}
&G_T(\widehat{\theta}_T)'W_T\widehat{\Omega}_{M=bT}W_TG_T(\widehat{\theta}_T)\\
&= -\int_0^1\int_0^1 D_T^*(r-s)G_T(\widehat{\theta}_T)'W_TT^{1/2}g_{[rT]}(\widehat{\theta}_T)T^{1/2}g_{[sT]}(\widehat{\theta}_T)'W_TG_T(\widehat{\theta}_T)drds\\
&= -\int_0^1\int_0^1 D_T^*(r-s)\widetilde{X}_T(r)\widetilde{X}_T(s)'drds\\
&\Rightarrow -\Lambda^*\int_0^1\int_0^1 k^{*\prime\prime}(r-s)\widetilde{B}_p(r)\widetilde{B}_p(s)'drds\Lambda^{*\prime},
\end{aligned}$$

using the continuous mapping theorem. The final expression is obtained using $k^{*\prime\prime}(x) = \frac{1}{b^2}k''(\frac{x}{b})$.

Case 2: $k(x)$ is continuous, $k(x) = 0$ for $|x| \geq 1$, and $k(x)$ is twice continuously differentiable everywhere except for $|x| = 1$. Let $1(\bullet)$ denote the indicator function. Noting that $\Delta^2 K_{ij} = 0$ for $|i-j| > [bT]$ and $\Delta^2 K_{ij} = -k^*\left(b - \frac{1}{T}\right)$ for $|i-j| = [bT]$, break up the double sum in the second line of the expression for $G_T(\widehat{\theta}_T)'W_T\widehat{\Omega}_{M=bT}W_TG_T(\widehat{\theta}_T)$ into three pieces corresponding to $|i-j| < [bT]$, $|i-j| = [bT]$, and $|i-j| > [bT]$ to obtain

$$G_T(\widehat{\theta}_T)'W_T\widehat{\Omega}_{M=bT}W_TG_T(\widehat{\theta}_T) =$$

$$T^{-2}\sum_{i=1}^{T-1}\sum_{j=1}^{T-1} 1\left(|i-j| < [bT]\right) T^2\Delta^2 K_{ij} G_T(\widehat{\theta}_T)'W_T T^{1/2} g_i(\widehat{\theta}_T) T^{1/2} g_j(\widehat{\theta}_T)'W_T G_T(\widehat{\theta}_T)$$
$$- Tk^*\left(b - \frac{1}{T}\right) T^{-1} \sum_{i=1}^{T-[bT]-1} G_T(\widehat{\theta}_T)'W_T T^{1/2} g_{i+[bT]}(\widehat{\theta}_T) T^{1/2} g_i(\widehat{\theta}_T)'W_T G_T(\widehat{\theta}_T)$$
$$- Tk^*\left(b - \frac{1}{T}\right) T^{-1} \sum_{j=1}^{T-[bT]-1} G_T(\widehat{\theta}_T)'W_T T^{1/2} g_j(\widehat{\theta}_T) T^{1/2} g_{j+[bT]}(\widehat{\theta}_T)'W_T G_T(\widehat{\theta}_T).$$

It directly follows that

$$G_T(\widehat{\theta}_T)'W_T\widehat{\Omega}_{M=bT}W_TG_T(\widehat{\theta}_T) =$$
$$-\iint_{|r-s|<b} D_T^*(r-s)\widetilde{X}_T(r)\widetilde{X}_T(s)'drds - Tk^*\left(b - \frac{1}{T}\right)\int_0^{1-b}\left(\widetilde{X}_T(r+b)\widetilde{X}_T(r)' + \widetilde{X}_T(r)\widetilde{X}_T(r+b)'\right)dr.$$

Let $k_-^{*\prime}(b)$ denote the first derivative of $k^*(x)$ from the left at $x = b$. By definition

$$\frac{k_-'(1)}{b} = k_-^{*\prime}(b) = \lim_{T\to\infty}\left[-Tk^*\left(b - \frac{1}{T}\right)\right].$$

Therefore, by the continuous mapping theorem

$$G_T(\widehat{\theta}_T)'W_T\widehat{\Omega}_{M=bT}W_TG_T(\widehat{\theta}_T)dr \Rightarrow$$
$$\Lambda^*\left[-\iint_{|r-s|<b} k^{*\prime\prime}(r-s)\widetilde{B}_p(r)\widetilde{B}_p(s)'drds + k_-^{*\prime}(b)\int_0^{1-b}\left(\widetilde{B}_p(r+b)\widetilde{B}_p(r)' + \widetilde{B}_p(r)\widetilde{B}_p(r+b)'\right)dr\right]\Lambda^{*\prime}.$$

Case 3: $k(x)$ is the Bartlett kernel. It is easy to calculate that for the Bartlett kernel, $\Delta^2 K_{ij} = \frac{2}{bT}$ for $|i-j| = 0$, $\Delta^2 K_{ij} = -\frac{1}{bT}$ for $|i-j| = [bT]$ and $\Delta^2 K_{ij} = 0$ otherwise. Therefore we have

$$G_T(\widehat{\theta}_T)'W_T\widehat{\Omega}_{M=bT}W_TG_T(\widehat{\theta}_T) =$$

$$
\begin{aligned}
&\frac{2}{bT}\sum_{i=1}^{T-1} G_T(\widehat{\theta}_T)'W_T T^{1/2} g_i(\widehat{\theta}_T) T^{1/2} g_i(\widehat{\theta}_T)' W_T G_T(\widehat{\theta}_T) \\
&\quad - \frac{1}{bT}\sum_{i=1}^{T-[bT]-1} G_T(\widehat{\theta}_T)'W_T T^{1/2} g_{i+[bT]}(\widehat{\theta}_T) T^{1/2} g_i(\widehat{\theta}_T)' W_T G_T(\widehat{\theta}_T) \\
&\quad - \frac{1}{bT}\sum_{j=1}^{T-[bT]-1} G_T(\widehat{\theta}_T)'W_T T^{1/2} g_j(\widehat{\theta}_T) T^{1/2} g_{j+[bT]}(\widehat{\theta}_T)' W_T G_T(\widehat{\theta}_T)
\end{aligned}
$$

$$
= \frac{2}{b}\int_0^1 \widetilde{X}_T(r)\widetilde{X}_T(r)'dr - \frac{1}{b}\int_0^{1-b} \left(\widetilde{X}_T(r+b)\widetilde{X}_T(r)' + \widetilde{X}_T(r)\widetilde{X}_T(r+b)'\right) dr
$$

$$
\Rightarrow \Lambda^* \left[\frac{2}{b}\int_0^1 \widetilde{B}_p(r)\widetilde{B}_p(r)'dr - \frac{1}{b}\int_0^{1-b} \left(\widetilde{B}_p(r+b)\widetilde{B}_p(r)' + \widetilde{B}_p(r)\widetilde{B}_p(r+b)'\right) dr\right] \Lambda^{*\prime}.
$$

Proof of Theorem 3: We only give the proof for F^* as the proof for t^* follows using similar arguments. Applying the delta method to the result in Lemma 1 and using the fact that $B_q(1)$ is a vector of independent standard normal random variables gives

$$
\begin{aligned}
T^{1/2} r(\widehat{\theta}_T) &\Rightarrow -R(\theta_0)\left(G_0'W_\infty G_0\right)^{-1} G_0'W_\infty \Lambda B_q(1) \\
&\equiv -R(\theta_0)\left(G_0'W_\infty G_0\right)^{-1} \Lambda^* B_p(1) \\
&\equiv \Lambda^{**} B_m(1), \qquad (18)
\end{aligned}
$$

where Λ^{**} is the matrix square root of $R(\theta_0)\left(G_0'W_\infty G_0\right)^{-1}\Lambda^*\Lambda^{*\prime}\left(G_0'W_\infty G_0\right)^{-1} R(\theta_0)'$. Using the results in Theorem 2, it directly follows that

$$
\begin{aligned}
R(\widehat{\theta}_T)\widehat{V}_{M=bT}R(\widehat{\theta}_T)' &= R(\widehat{\theta}_T)\left[G_T(\widehat{\theta}_T)'W_T G_T(\widehat{\theta}_T, \theta_0, \lambda_T)\right]^{-1} G_T(\widehat{\theta}_T)'W_T \widehat{\Omega}_{M=T} W_T G_T(\widehat{\theta}_T) \\
&\qquad \times \left[G_T(\widehat{\theta}_T)'W_T G_T(\widehat{\theta}_T, \theta_0, \lambda_T)\right]^{-1} R(\widehat{\theta}_T)' \\
&\Rightarrow R(\theta_0)\left(G_0'W_\infty G_0\right)^{-1}\Lambda^* Q_p(b)\Lambda^{*\prime}\left(G_0'W_\infty G_0\right)^{-1} R(\theta_0)' \\
&\equiv \Lambda^{**} Q_m(b) \Lambda^{**\prime}, \qquad (19)
\end{aligned}
$$

where we use the fact that

$$
\begin{aligned}
R(\theta_0)\left(G_0'W_\infty G_0\right)^{-1}\Lambda^*\widetilde{B}_p(r) &= R(\theta_0)\left(G_0'W_\infty G_0\right)^{-1}\Lambda^*\left(B_p(r) - rB_p(1)\right) \\
&\equiv \Lambda^{**}\left(B_m(r) - rB_m(1)\right) \\
&= \Lambda^{**}\widetilde{B}_m(r).
\end{aligned}
$$

Using (18) and (19) it directly follows that

$$
\begin{aligned}
F^* &= Tr(\widehat{\theta}_T)' \left(R(\widehat{\theta}_T)\widehat{V}_{M=bT}R(\widehat{\theta}_T)' \right)^{-1} r(\widehat{\theta}_T)/m \\
&= T^{1/2}r(\widehat{\theta}_T)' \left(R(\widehat{\theta}_T)\widehat{V}_{M=T}R(\widehat{\theta}_T)' \right)^{-1} T^{1/2}r(\widehat{\theta}_T)/m \\
&\Rightarrow \left(\Lambda^{**}B_m(1)\right)' \left(\Lambda^{**}Q_m(b)\Lambda^{**\prime}\right)^{-1} \left(\Lambda^{**}B_m(1)\right)/m \\
&\equiv B_m(1)'Q_m(b)^{-1}B_m(1)/m,
\end{aligned}
$$

which completes the proof.

7 Appendix: Kernel Formulas

The formulas for the kernels analyzed in this paper are

$$
\begin{aligned}
&\text{Bartlett} \quad & k(x) &= \begin{cases} 1-|x| & \text{for } |x| \le 1, \\ 0 & \textit{otherwise}, \end{cases} \\
&\text{Parzen} \quad & k(x) &= \begin{cases} 1-6x^2+6|x|^3 & \text{for } |x| \le \frac{1}{2}, \\ 2(1-|x|)^3 & \text{for } \frac{1}{2} \le |x| \le 1 \\ 0 & \textit{otherwise}, \end{cases} \\
&\text{Bohman} \quad & k(x) &= \begin{cases} (1-|x|)\cos(\pi x) + \sin(\pi\,|x|)/\pi & \text{for } |x| \le 1, \\ 0 & \textit{otherwise}, \end{cases} \\
&\text{Quadratic Spectral (QS)} \quad & k(x) &= \frac{25}{12\pi^2x^2}\left(\frac{\sin(6\pi x/5)}{6\pi x/5} - \cos(6\pi x/5)\right), \\
&\text{Daniell} \quad & k(x) &= \frac{\sin(\pi x)}{\pi x}.
\end{aligned}
$$

References

Abadir, K. M. and Paruolo, P. (2002), Simple Robust Testing of Regression Hypotheses: a Comment, *Econometrica* **70**, 2097–2099.

Andrews, D. W. K. (1991), Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation, *Econometrica* **59**, 817–854.

Andrews, D. W. K. and Monahan, J. C. (1992), An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator, *Econometrica* **60**, 953–966.

Barndorff-Nielsen, O. and Cox, D. (1989), *Asymptotic Techniques for Use in Statistics*, Chapman and Hall, London.

Cushing, M. J. and McGarvey, M. G. (1999), Covariance Matrix Estimation, *in* L. Matyas (ed.), *Generalized Method of Moments Estimation*, Cambridge University Press, New York.

de Jong, R. M. and Davidson, J. (2000), Consistency of Kernel Estimators of Heteroskedastic and Autocorrelated Covariance Matrices, *Econometrica* **68**, 407–424.

den Haan, W. J. and Levin, A. (1997), A Practictioner's Guide to Robust Covariance Matrix Estimation, *in* G. Maddala and C. Rao (eds), *Handbook of Statistics: Robust Inference*, Volume 15, Elsevier, New York, pp. 291–341.

den Haan, W. J. and Levin, A. (1998), Vector Autoregressive Covariance Matrix Estimation, working paper, International Finance Division, FED Board of Governors.

Gallant, A. (1987), *Nonlinear Statistical Models*, Wiley, New York.

Gallant, A. and White, H. (1988), *A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models*, Basil Blackwell, New York.

Hansen, B. E. (1992), Consistent Covariance Matrix Estimation for Dependent Heterogenous Processes, *Econometrica* **60**, 967–972.

Hansen, L. P. (1982), Large Sample Properties of Generalized Method of Moments Estimators, *Econometrica* **50**, 1029–1054.

Heyde, C. (1997), *Quasi-Likelihood and its Application. A General Approach to Optimal Parameter Estimation*, Springer, New York.

Kiefer, N. M. and Vogelsang, T. J. (2002)a, Heteroskedasticity-Autocorrelation Robust Standard Errors Using the Bartlett Kernel Without Truncation, *Econometrica* **70**, 2093–2095.

Kiefer, N. M. and Vogelsang, T. J. (2002)b, Heteroskedasticity-Autocorrelation Robust Testing Using Bandwidth Equal to Sample Size, *Econometric Theory* **18**, 1350–1366.

Neave, H. R. (1970), An Improved Formula for the Asymptotic Variance of Spectrum Estimates, *Annals of Mathematical Statistics* **41**, 70–77.

Newey, W. K. and McFadden, D. L. (1994), Large Sample Estimation and Hypothesis Testing, *in* R. Engle and D. L. McFadden (eds), *Handbook of Econometrics*, Vol. 4, Elsevier Science Publishers, Amseterdam, The Netherlands, pp. 2113–2247.

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

Newey, W. K. and West, K. D. (1994), Automatic Lag Selection in Covariance Estimation, *Review of Economic Studies* **61**, 631–654.

Phillips, P. C. B. and Durlauf, S. N. (1986), Multiple Regression with Integrated Processes, *Review of Economic Studies* **53**, 473–496.

Robinson, P. (1998), Inference-Without Smoothing in the Presence of Nonparametric Autocorrelation, *Econometrica* **66**, 1163–1182.

Simonoff, J. (1993), The Relative Importance of Bias and Variability in the Estimation of the Variance of a Statistic, *The Statistician* **42**, 3–7.

Vogelsang, T. J. (2003), Testing in GMM Models Without Truncation, *Advances in Econometrics, Maximum Likelihood Estimation of Misspecified Models: Twenty Years Later*, Vol. 17, pp. 199–233.

White, H. (1984), *Asymptotic Theory for Econometricians*, Academic Press, New York.

Table I: Asymptotic Critical Values for t_b^* Using Bartlett Kernel

$b =$	0.02	0.04	0.06	0.08	0.1	0.12	0.14	0.16	0.18	0.2
90%	1.323	1.343	1.368	1.390	1.414	1.442	1.469	1.498	1.529	1.563
95%	1.690	1.731	1.772	1.813	1.861	1.902	1.944	1.988	2.030	2.081
97.5%	2.018	2.072	2.125	2.179	2.235	2.296	2.355	2.417	2.481	2.553
99%	2.377	2.459	2.537	2.627	2.709	2.792	2.882	2.961	3.051	3.140
$b =$	0.22	0.24	0.26	0.28	0.3	0.32	0.34	0.36	0.38	0.4
90%	1.587	1.615	1.644	1.674	1.709	1.737	1.767	1.796	1.830	1.865
95%	2.124	2.179	2.222	2.274	2.324	2.367	2.412	2.459	2.505	2.556
97.5%	2.617	2.678	2.736	2.805	2.878	2.930	2.999	3.067	3.129	3.192
99%	3.229	3.315	3.385	3.476	3.580	3.707	3.791	3.858	3.942	4.038
$b =$	0.42	0.44	0.46	0.48	0.50	0.52	0.54	0.56	0.58	0.6
90%	1.901	1.931	1.963	1.994	2.022	2.053	2.086	2.117	2.149	2.183
95%	2.601	2.651	2.696	2.739	2.781	2.828	2.872	2.913	2.956	3.007
97.5%	3.253	3.318	3.382	3.447	3.514	3.567	3.636	3.684	3.740	3.783
99%	4.111	4.212	4.306	4.399	4.480	4.567	4.645	4.711	4.762	4.831
$b =$	0.62	0.64	0.66	0.68	0.70	0.72	0.74	0.76	0.78	0.80
90%	2.211	2.242	2.272	2.297	2.325	2.359	2.389	2.418	2.450	2.477
95%	3.048	3.082	3.124	3.162	3.198	3.245	3.291	3.330	3.367	3.408
97.5%	3.834	3.880	3.934	3.995	4.054	4.101	4.153	4.208	4.260	4.312
99%	4.912	4.981	5.041	5.124	5.190	5.279	5.333	5.376	5.445	5.493
$b =$	0.82	0.84	0.86	0.88	0.90	0.92	0.94	0.96	0.98	1.0
90%	2.506	2.533	2.560	2.591	2.618	2.649	2.678	2.706	2.733	2.740
95%	3.444	3.494	3.537	3.579	3.616	3.654	3.692	3.727	3.764	3.764
97.5%	4.367	4.417	4.470	4.524	4.568	4.617	4.664	4.713	4.764	4.771
99%	5.554	5.609	5.672	5.724	5.782	5.868	5.933	5.998	6.058	6.090

Notes: The critical values for $b = 1$ are analytical (see Abadir and Paruolo, 2002 and Kiefer and Vogelsang, 2002b). The remaining critical values were calculated via simulation methods using 50,000 replications. Normalized partial sums of 1,000 *i.i.d.* $N(0,1)$ random deviates were used to approximate the standard Brownian motions in the respective distributions given by Theorem 3. The bandwidth is given by $M = bT$.

Table II: Asymptotic Critical Values for t_b^* Using Parzen Kernel

$b=$	0.02	0.04	0.06	0.08	0.1	0.12	0.14	0.16	0.18	0.2
90%	1.312	1.334	1.353	1.373	1.393	1.410	1.433	1.451	1.475	1.498
95%	1.678	1.716	1.748	1.773	1.811	1.846	1.885	1.922	1.962	1.994
97.5%	2.000	2.046	2.089	2.135	2.180	2.226	2.274	2.323	2.374	2.433
99%	2.352	2.422	2.487	2.547	2.629	2.706	2.777	2.847	2.931	3.012
$b=$	0.22	0.24	0.26	0.28	0.3	0.32	0.34	0.36	0.38	0.4
90%	1.522	1.545	1.568	1.596	1.623	1.649	1.676	1.700	1.727	1.755
95%	2.035	2.071	2.112	2.151	2.193	2.231	2.277	2.322	2.371	2.425
97.5%	2.488	2.544	2.596	2.659	2.734	2.793	2.856	2.925	2.997	3.065
99%	3.114	3.190	3.272	3.349	3.432	3.512	3.619	3.709	3.814	3.916
$b=$	0.42	0.44	0.46	0.48	0.50	0.52	0.54	0.56	0.58	0.6
90%	1.783	1.812	1.844	1.875	1.905	1.939	1.968	2.000	2.030	2.064
95%	2.471	2.518	2.561	2.607	2.655	2.709	2.766	2.824	2.876	2.936
97.5%	3.130	3.195	3.262	3.331	3.401	3.476	3.559	3.638	3.723	3.805
99%	4.037	4.157	4.277	4.402	4.519	4.640	4.769	4.886	4.997	5.123
$b=$	0.62	0.64	0.66	0.68	0.70	0.72	0.74	0.76	0.78	0.80
90%	2.101	2.134	2.171	2.203	2.238	2.271	2.302	2.337	2.377	2.412
95%	2.997	3.049	3.112	3.166	3.220	3.276	3.332	3.394	3.454	3.515
97.5%	3.879	3.957	4.051	4.132	4.216	4.301	4.390	4.468	4.570	4.673
99%	5.259	5.407	5.587	5.734	5.867	5.981	6.138	6.300	6.431	6.556
$b=$	0.82	0.84	0.86	0.88	0.90	0.92	0.94	0.96	0.98	1.0
90%	2.450	2.490	2.525	2.561	2.605	2.642	2.681	2.723	2.764	2.807
95%	3.587	3.650	3.711	3.775	3.833	3.903	3.960	4.031	4.098	4.179
97.5%	4.764	4.846	4.932	5.019	5.130	5.233	5.324	5.422	5.543	5.649
99%	6.686	6.841	7.012	7.181	7.332	7.495	7.667	7.863	7.991	8.162

Notes: The critical values were calculated via simulation methods using 50,000 replications. Normalized partial sums of 1,000 *i.i.d.* $N(0,1)$ random deviates were used to approximate the standard Brownian motions in the respective distributions given by Theorem 3. The bandwidth is given by $M = bT$.

Table III: Asymptotic Critical Values for t_b^* Using Bohman Kernel

$b =$	0.02	0.04	0.06	0.08	0.1	0.12	0.14	0.16	0.18	0.2
90%	1.238	1.326	1.354	1.380	1.400	1.422	1.444	1.467	1.493	1.518
95%	1.589	1.708	1.748	1.787	1.827	1.866	1.906	1.948	1.986	2.027
97.5%	1.902	2.040	2.097	2.147	2.198	2.250	2.300	2.357	2.421	2.483
99%	2.232	2.413	2.498	2.575	2.664	2.740	2.815	2.904	2.992	3.101
$b =$	0.22	0.24	0.26	0.28	0.3	0.32	0.34	0.36	0.38	0.4
90%	1.542	1.569	1.599	1.628	1.655	1.684	1.712	1.742	1.772	1.805
95%	2.069	2.114	2.156	2.201	2.243	2.296	2.349	2.400	2.456	2.508
97.5%	2.544	2.603	2.671	2.747	2.812	2.886	2.958	3.038	3.108	3.179
99%	3.193	3.274	3.366	3.453	3.557	3.653	3.754	3.874	4.008	4.143
$b =$	0.42	0.44	0.46	0.48	0.50	0.52	0.54	0.56	0.58	0.6
90%	1.838	1.873	1.909	1.940	1.975	2.007	2.044	2.083	2.120	2.158
95%	2.557	2.604	2.656	2.713	2.774	2.839	2.904	2.966	3.023	3.087
97.5%	3.247	3.334	3.409	3.487	3.582	3.660	3.752	3.846	3.922	4.027
99%	4.261	4.388	4.528	4.651	4.778	4.905	5.046	5.184	5.371	5.541
$b =$	0.62	0.64	0.66	0.68	0.70	0.72	0.74	0.76	0.78	0.80
90%	2.195	2.230	2.268	2.305	2.347	2.387	2.427	2.469	2.504	2.544
95%	3.147	3.209	3.266	3.332	3.396	3.462	3.535	3.610	3.677	3.747
97.5%	4.122	4.209	4.305	4.390	4.495	4.599	4.694	4.790	4.881	4.979
99%	5.683	5.835	5.998	6.145	6.288	6.428	6.575	6.740	6.918	7.078
$b =$	0.82	0.84	0.86	0.88	0.90	0.92	0.94	0.96	0.98	1.0
90%	2.587	2.629	2.674	2.720	2.762	2.811	2.856	2.899	2.936	2.982
95%	3.811	3.878	3.951	4.023	4.103	4.175	4.246	4.314	4.389	4.460
97.5%	5.075	5.170	5.280	5.405	5.521	5.646	5.756	5.875	5.997	6.111
99%	7.264	7.433	7.608	7.796	7.956	8.118	8.285	8.471	8.637	8.778

Notes: The critical values were calculated via simulation methods using 50,000 replications. Normalized partial sums of 1,000 *i.i.d.* $N(0,1)$ random deviates were used to approximate the standard Brownian motions in the respective distributions given by Theorem 3. The bandwidth is given by $M = bT$.

Table IV: Asymptotic Critical Values for t_b^* Using Daniels Kernel

$b =$	0.02	0.04	0.06	0.08	0.1	0.12	0.14	0.16	0.18	0.2
90%	1.325	1.360	1.387	1.424	1.460	1.492	1.530	1.572	1.614	1.662
95%	1.700	1.754	1.807	1.869	1.930	1.982	2.054	2.117	2.191	2.281
97.5%	2.036	2.102	2.184	2.261	2.356	2.461	2.566	2.669	2.784	2.902
99%	2.408	2.529	2.638	2.775	2.921	3.069	3.265	3.408	3.561	3.799
$b =$	0.22	0.24	0.26	0.28	0.3	0.32	0.34	0.36	0.38	0.40
90%	1.708	1.753	1.802	1.859	1.920	1.974	2.032	2.092	2.157	2.212
95%	2.360	2.466	2.549	2.642	2.749	2.856	2.951	3.059	3.161	3.260
97.5%	3.039	3.185	3.328	3.487	3.614	3.770	3.944	4.107	4.267	4.403
99%	4.005	4.210	4.425	4.702	4.964	5.202	5.484	5.837	6.134	6.450
$b =$	0.42	0.44	0.46	0.48	0.50	0.52	0.54	0.56	0.58	0.6
90%	2.276	2.333	2.402	2.473	2.544	2.614	2.685	2.751	2.833	2.911
95%	3.366	3.480	3.594	3.720	3.844	3.975	4.105	4.239	4.369	4.505
97.5%	4.579	4.751	4.959	5.139	5.357	5.580	5.802	6.031	6.285	6.493
99%	6.802	7.117	7.430	7.769	8.136	8.459	8.805	9.083	9.401	9.826
$b =$	0.62	0.64	0.66	0.68	0.70	0.72	0.74	0.76	0.78	0.80
90%	2.988	3.059	3.138	3.215	3.302	3.399	3.495	3.576	3.672	3.769
95%	4.634	4.761	4.896	5.040	5.167	5.313	5.470	5.630	5.816	5.970
97.5%	6.723	6.974	7.220	7.451	7.663	7.873	8.106	8.356	8.610	8.837
99%	10.181	10.489	10.950	11.281	11.583	11.996	12.419	12.734	13.248	13.658
$b =$	0.82	0.84	0.86	0.88	0.90	0.92	0.94	0.96	0.98	1.0
90%	3.866	3.961	4.059	4.156	4.257	4.349	4.443	4.546	4.653	4.750
95%	6.138	6.281	6.447	6.625	6.779	6.935	7.129	7.327	7.508	7.680
97.5%	9.095	9.341	9.597	9.827	10.068	10.367	10.642	10.904	11.167	11.386
99%	14.015	14.487	14.875	15.400	15.820	16.256	16.805	17.363	17.862	18.225

Notes: The critical values were calculated via simulation methods using 50,000 replications. Normalized partial sums of 1,000 *i.i.d.* $N(0,1)$ random deviates were used to approximate the standard Brownian motions in the respective distributions given by Theorem 3. The bandwidth is given by $M = bT$.

Table V: Asymptotic Critical Values for t_b^* Using QS Kernel

$b =$	0.02	0.04	0.06	0.08	0.1	0.12	0.14	0.16	0.18	0.2
90%	1.329	1.363	1.401	1.438	1.473	1.516	1.562	1.607	1.655	1.706
95%	1.708	1.761	1.826	1.893	1.960	2.027	2.093	2.173	2.261	2.341
97.5%	2.041	2.115	2.197	2.285	2.388	2.485	2.598	2.714	2.844	2.989
99%	2.409	2.533	2.666	2.806	2.958	3.138	3.324	3.455	3.650	3.851
$b =$	0.22	0.24	0.26	0.28	0.3	0.32	0.34	0.36	0.38	0.4
90%	1.760	1.814	1.866	1.931	1.995	2.057	2.122	2.193	2.266	2.330
95%	2.442	2.539	2.627	2.732	2.856	2.966	3.074	3.189	3.297	3.417
97.5%	3.127	3.252	3.407	3.573	3.729	3.892	4.067	4.239	4.420	4.609
99%	4.075	4.335	4.576	4.797	5.092	5.403	5.690	6.010	6.351	6.656
$b =$	0.42	0.44	0.46	0.48	0.50	0.52	0.54	0.56	0.58	0.6
90%	2.396	2.471	2.547	2.624	2.707	2.782	2.863	2.945	3.033	3.118
95%	3.553	3.670	3.816	3.947	4.081	4.217	4.358	4.498	4.646	4.797
97.5%	4.788	4.979	5.178	5.404	5.657	5.875	6.133	6.365	6.653	6.882
99%	6.991	7.337	7.661	8.038	8.357	8.724	9.117	9.448	9.785	10.210
$b =$	0.62	0.64	0.66	0.68	0.70	0.72	0.74	0.76	0.78	0.80
90%	3.203	3.281	3.370	3.454	3.550	3.651	3.747	3.843	3.948	4.054
95%	4.935	5.073	5.210	5.369	5.531	5.685	5.864	6.039	6.219	6.398
97.5%	7.139	7.366	7.648	7.921	8.154	8.399	8.649	8.886	9.130	9.376
99%	10.590	10.998	11.373	11.781	12.178	12.679	13.110	13.514	13.962	14.447
$b =$	0.82	0.84	0.86	0.88	0.90	0.92	0.94	0.96	0.98	1.0
90%	4.158	4.260	4.364	4.474	4.575	4.682	4.781	4.892	5.000	5.112
95%	6.576	6.742	6.915	7.095	7.271	7.451	7.642	7.854	8.058	8.245
97.5%	9.660	9.932	10.211	10.481	10.760	11.044	11.359	11.632	11.928	12.195
99%	14.892	15.382	15.838	16.344	16.776	17.259	17.848	18.394	18.909	19.516

Notes: The critical values were calculated via simulation methods using 50,000 replications. Normalized partial sums of 1,000 *i.i.d.* $N(0,1)$ random deviates were used to approximate the standard Brownian motions in the respective distributions given by Theorem 3. The bandwidth is given by $M = bT$.

Table VI: Empirical Null Rejection Probabilities in Simple Location Model
5% Nominal Level, 10,000 Replications

$y_t = \theta_1 + u_t$, $u_t = \rho u_{t-1} + \xi_t$, $\xi_t \sim i.i.d.\ N(0,1)$, $u_0 = 0$

$H_0 : \theta_1 \leq 0$, $H_1 : \theta_1 > 0$

Panel A: Bartlett Kernel, No Prewhitening, $T = 50$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.051	.019	.23	.023	.77	.008	.018	.020	.022	.023	.022	.023	.023
-.5	.049	.031	.10	.037	.49	.029	.037	.037	.038	.038	.039	.038	.039
-.3	.050	.036	.07	.042	.30	.039	.042	.042	.043	.044	.043	.043	.043
.0	.060	.053	.03	.053	.12	.049	.048	.048	.048	.050	.049	.048	.049
.3	.085	.074	.06	.060	.27	.064	.056	.057	.056	.056	.055	.055	.055
.5	.104	.082	.10	.065	.45	.080	.066	.065	.064	.063	.062	.062	.062
.7	.135	.099	.15	.079	.64	.116	.084	.081	.081	.078	.078	.078	.078
.9	.220	.157	.29	.135	.81	.234	.159	.144	.135	.132	.134	.136	.136
.95	.288	.204	.36	.183	.85	.303	.224	.202	.187	.183	.182	.183	.183

Panel B: Parzen Kernel, No Prewhitening, $T = 50$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.049	.036	.09	.046	.77	.036	.046	.047	.045	.045	.046	.046	.045
-.5	.051	.038	.08	.047	.49	.041	.046	.049	.048	.048	.048	.047	.046
-.3	.052	.040	.07	.046	.30	.044	.047	.049	.048	.049	.049	.048	.047
.0	.061	.052	.06	.054	.12	.051	.049	.050	.049	.049	.049	.050	.048
.3	.081	.065	.12	.057	.27	.067	.054	.052	.050	.051	.051	.049	.049
.5	.097	.070	.19	.059	.45	.087	.062	.058	.055	.053	.052	.052	.052
.7	.129	.079	.32	.065	.64	.133	.082	.072	.065	.063	.060	.060	.059
.9	.225	.116	.62	.099	.81	.261	.171	.144	.120	.106	.099	.094	.091
.95	.291	.146	.73	.131	.85	.328	.243	.207	.172	.151	.140	.129	.123

Panel C: QS Kernel, No Prewhitening, $T = 50$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.049	.037	.04	.046	.77	.048	.045	.045	.046	.046	.047	.045	.045
-.5	.050	.038	.04	.047	.49	.049	.049	.048	.048	.047	.047	.047	.047
-.3	.050	.040	.04	.046	.30	.049	.048	.048	.048	.048	.048	.049	.049
.0	.060	.051	.03	.053	.12	.050	.050	.049	.048	.048	.048	.049	.049
.3	.079	.066	.06	.055	.27	.056	.051	.051	.049	.049	.051	.051	.050
.5	.094	.071	.09	.056	.45	.067	.054	.053	.051	.052	.053	.053	.053
.7	.123	.079	.16	.059	.64	.092	.064	.060	.058	.057	.054	.055	.056
.9	.221	.115	.34	.084	.81	.201	.118	.098	.087	.084	.082	.081	.079
.95	.294	.141	.43	.105	.85	.275	.173	.138	.116	.107	.105	.101	.099

Table VI: Continued

Panel D: Bartlett Kernel, $AR(1)$ Prewhitening, $T = 50$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.067	.061	.03	.060	.09	.056	.053	.051	.048	.046	.046	.043	.042
-.5	.063	.059	.02	.057	.06	.053	.052	.050	.050	.048	.048	.047	.047
-.3	.064	.060	.01	.058	.04	.053	.052	.051	.050	.049	.049	.048	.047
.0	.066	.064	.01	.062	.02	.053	.052	.051	.051	.050	.049	.048	.048
.3	.071	.068	.01	.065	.04	.055	.052	.052	.051	.051	.050	.049	.049
.5	.075	.071	.02	.067	.06	.059	.055	.054	.053	.052	.051	.050	.050
.7	.088	.083	.02	.077	.08	.071	.063	.060	.059	.058	.056	.055	.055
.9	.138	.132	.03	.121	.10	.115	.098	.092	.089	.086	.086	.085	.084
.95	.180	.174	.03	.158	.10	.154	.131	.123	.117	.112	.108	.105	.103

Panel E: Parzen Kernel, $AR(1)$ Prewhitening, $T = 50$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.068	.059	.05	.061	.09	.057	.055	.054	.050	.048	.047	.043	.042
-.5	.064	.057	.05	.059	.06	.054	.052	.053	.051	.050	.049	.047	.045
-.3	.065	.058	.04	.060	.04	.054	.052	.052	.050	.049	.050	.048	.045
.0	.066	.062	.03	.063	.02	.055	.051	.051	.049	.049	.050	.049	.046
.3	.070	.066	.04	.066	.04	.056	.052	.052	.049	.049	.049	.048	.045
.5	.075	.068	.05	.068	.06	.061	.053	.053	.050	.049	.049	.048	.047
.7	.087	.080	.05	.079	.08	.073	.063	.058	.054	.052	.051	.050	.049
.9	.137	.128	.06	.123	.10	.120	.100	.090	.081	.076	.074	.069	.067
.95	.178	.169	.06	.162	.10	.161	.135	.123	.110	.099	.093	.085	.081

Panel F: QS Kernel, $AR(1)$ Prewhitening, $T = 50$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.067	.059	.03	.059	.09	.055	.051	.048	.044	.037	.033	.028	.026
-.5	.063	.056	.02	.057	.06	.054	.050	.050	.046	.042	.039	.034	.031
-.3	.064	.058	.02	.059	.04	.053	.050	.050	.046	.043	.040	.036	.033
.0	.066	.062	.02	.061	.02	.052	.050	.050	.047	.044	.042	.038	.036
.3	.070	.066	.02	.064	.04	.053	.049	.050	.046	.045	.042	.039	.037
.5	.074	.069	.02	.066	.06	.055	.050	.050	.047	.046	.043	.039	.037
.7	.087	.081	.03	.075	.08	.065	.054	.052	.050	.048	.045	.041	.038
.9	.137	.128	.03	.114	.10	.108	.081	.073	.066	.060	.055	.049	.044
.95	.178	.170	.03	.150	.10	.145	.110	.096	.081	.069	.063	.055	.050

Table VII: Empirical Null Rejection Probabilities in Simple Location Model
5% Nominal Level, 10,000 Replications

$y_t = \theta_1 + u_t,\ u_t = \rho u_{t-1} + \xi_t,\ \xi_t \sim i.i.d.\ N(0,1),\ u_0 = 0$

$H_0 : \theta_1 \leq 0,\ H_1 : \theta_1 > 0$

Panel A: Bartlett Kernel, No Prewhitening, $T = 100$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.048	.025	.14	.034	.79	.024	.030	.032	.034	.034	.033	.034	.034
-.5	.046	.034	.07	.044	.50	.038	.041	.044	.046	.045	.044	.044	.044
-.3	.047	.038	.04	.046	.30	.044	.044	.047	.049	.048	.047	.047	.046
.0	.056	.053	.02	.051	.08	.049	.048	.050	.052	.051	.050	.050	.050
.3	.079	.070	.04	.055	.28	.058	.053	.054	.056	.053	.053	.053	.054
.5	.091	.075	.06	.059	.47	.067	.058	.058	.059	.057	.057	.055	.057
.7	.112	.085	.10	.065	.67	.084	.068	.066	.066	.065	.064	.064	.064
.9	.174	.123	.20	.099	.86	.172	.112	.105	.099	.097	.098	.100	.100
.95	.232	.163	.29	.137	.90	.247	.168	.151	.138	.136	.137	.136	.137

Panel B: Parzen Kernel, No Prewhitening, $T = 100$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.044	.036	.05	.051	.79	.046	.050	.047	.049	.050	.050	.050	.050
-.5	.047	.040	.05	.050	.50	.047	.050	.048	.050	.051	.051	.051	.050
-.3	.049	.042	.04	.049	.30	.048	.050	.049	.050	.051	.051	.051	.050
.0	.057	.051	.03	.051	.08	.052	.051	.050	.051	.051	.052	.051	.050
.3	.073	.061	.07	.054	.28	.058	.052	.051	.051	.051	.052	.052	.051
.5	.084	.063	.11	.054	.47	.066	.055	.052	.053	.053	.053	.052	.051
.7	.102	.071	.19	.058	.67	.092	.064	.059	.058	.057	.057	.055	.054
.9	.170	.094	.45	.073	.86	.195	.116	.097	.082	.077	.074	.072	.071
.95	.236	.120	.63	.097	.90	.273	.179	.152	.125	.111	.104	.096	.093

Panel C: QS Kernel, No Prewhitening, $T = 100$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.047	.040	.03	.048	.79	.050	.049	.051	.051	.049	.049	.049	.048
-.5	.046	.041	.02	.050	.50	.051	.050	.053	.051	.050	.049	.048	.049
-.3	.047	.042	.02	.050	.30	.051	.051	.052	.051	.050	.050	.050	.049
.0	.056	.052	.02	.051	.08	.051	.051	.052	.051	.050	.051	.051	.051
.3	.073	.063	.04	.053	.28	.053	.052	.053	.052	.051	.050	.051	.051
.5	.082	.064	.06	.054	.47	.056	.053	.054	.053	.051	.051	.052	.051
.7	.100	.071	.09	.054	.67	.069	.058	.058	.054	.053	.052	.052	.052
.9	.165	.092	.22	.066	.86	.140	.081	.073	.069	.065	.064	.064	.064
.95	.230	.119	.34	.082	.90	.212	.124	.102	.089	.085	.082	.082	.081

Table VII: Continued

Panel D: Bartlett Kernel, $AR(1)$ Prewhitening, $T = 100$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.059	.056	.01	.055	.06	.053	.051	.053	.052	.049	.048	.048	.047
-.5	.058	.056	.01	.054	.04	.051	.050	.053	.053	.050	.049	.049	.049
-.3	.057	.056	.01	.055	.03	.051	.051	.053	.053	.051	.051	.049	.049
.0	.058	.057	.01	.057	.01	.051	.050	.052	.052	.051	.050	.049	.049
.3	.060	.059	.01	.057	.03	.052	.050	.051	.053	.050	.049	.049	.049
.5	.064	.062	.01	.059	.04	.054	.051	.052	.052	.050	.050	.049	.050
.7	.071	.069	.01	.064	.06	.058	.053	.054	.053	.051	.051	.050	.050
.9	.104	.101	.01	.091	.07	.083	.072	.070	.068	.064	.064	.063	.062
.95	.141	.138	.01	.126	.07	.117	.100	.095	.089	.085	.085	.083	.083

Panel E: Parzen Kernel, $AR(1)$ Prewhitening, $T = 100$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.060	.055	.03	.054	.06	.054	.054	.051	.052	.050	.050	.048	.045
-.5	.059	.055	.02	.056	.04	.053	.053	.051	.053	.052	.051	.051	.049
-.3	.058	.055	.02	.056	.03	.053	.052	.052	.052	.051	.051	.050	.050
.0	.058	.056	.01	.057	.01	.053	.052	.052	.051	.051	.051	.050	.050
.3	.060	.057	.02	.058	.03	.053	.052	.051	.051	.050	.050	.050	.049
.5	.064	.060	.02	.059	.04	.053	.052	.051	.051	.050	.050	.050	.049
.7	.071	.067	.03	.065	.06	.058	.054	.052	.052	.050	.050	.049	.048
.9	.103	.099	.03	.093	.07	.086	.072	.067	.062	.059	.057	.052	.051
.95	.140	.136	.03	.130	.07	.122	.103	.093	.082	.076	.072	.068	.066

Panel F: QS Kernel, $AR(1)$ Prewhitening, $T = 100$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.059	.055	.01	.054	.06	.054	.052	.052	.047	.043	.039	.034	.030
-.5	.058	.055	.01	.055	.04	.054	.053	.052	.050	.048	.045	.039	.035
-.3	.057	.055	.01	.055	.03	.053	.053	.053	.051	.048	.046	.040	.037
.0	.058	.056	.01	.057	.01	.052	.052	.052	.050	.048	.046	.042	.039
.3	.060	.058	.01	.057	.03	.051	.051	.052	.050	.047	.045	.041	.039
.5	.064	.060	.01	.058	.04	.051	.051	.052	.050	.046	.044	.040	.038
.7	.070	.067	.01	.063	.06	.054	.050	.051	.048	.045	.043	.038	.036
.9	.104	.099	.01	.086	.07	.075	.062	.058	.051	.047	.046	.043	.040
.95	.140	.137	.02	.121	.07	.108	.083	.073	.065	.059	.054	.048	.044

Table VIII: Empirical Null Rejection Probabilities in Simple Location Model
5% Nominal Level, 10,000 Replications

$y_t = \theta_1 + u_t,\ u_t = \rho u_{t-1} + \xi_t,\ \xi_t \sim i.i.d.\ N(0,1),\ u_0 = 0$

$H_0 : \theta_1 \leq 0,\ H_1 : \theta_1 > 0$

Panel A: Bartlett Kernel, No Prewhitening, $T = 200$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.042	.029	.09	.038	.79	.029	.036	.038	.038	.038	.038	.038	.039
-.5	.042	.036	.04	.044	.50	.042	.044	.044	.044	.044	.045	.046	.046
-.3	.044	.039	.03	.045	.30	.045	.046	.045	.046	.046	.046	.047	.046
.0	.051	.049	.01	.049	.06	.048	.048	.047	.047	.047	.048	.049	.048
.3	.067	.061	.03	.050	.29	.053	.050	.049	.049	.049	.049	.049	.049
.5	.075	.066	.04	.052	.49	.055	.051	.051	.051	.052	.050	.051	.051
.7	.087	.071	.06	.054	.68	.064	.056	.056	.055	.055	.054	.055	.055
.9	.129	.096	.14	.073	.88	.109	.078	.076	.072	.073	.073	.073	.074
.95	.173	.124	.22	.097	.93	.169	.113	.105	.100	.099	.098	.097	.098

Panel B: Parzen Kernel, No Prewhitening, $T = 200$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.038	.035	.03	.045	.80	.046	.048	.047	.047	.045	.045	.046	.046
-.5	.045	.041	.03	.047	.50	.047	.048	.048	.048	.046	.045	.046	.046
-.3	.046	.043	.03	.048	.30	.047	.048	.048	.049	.046	.046	.047	.047
.0	.052	.050	.02	.050	.06	.048	.048	.048	.048	.046	.047	.047	.046
.3	.062	.055	.04	.049	.29	.050	.048	.049	.049	.047	.047	.047	.047
.5	.069	.057	.07	.050	.49	.054	.049	.049	.049	.047	.047	.048	.047
.7	.079	.062	.11	.049	.68	.063	.052	.051	.050	.048	.049	.049	.048
.9	.119	.076	.28	.057	.88	.123	.075	.066	.060	.057	.057	.057	.057
.95	.169	.093	.46	.070	.93	.190	.116	.097	.081	.074	.072	.070	.069

Panel C: QS Kernel, No Prewhitening, $T = 200$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.044	.041	.01	.046	.79	.048	.047	.045	.046	.046	.047	.046	.045
-.5	.045	.042	.01	.047	.50	.049	.047	.046	.047	.046	.046	.045	.046
-.3	.046	.042	.01	.046	.30	.049	.048	.047	.048	.048	.047	.046	.047
.0	.051	.050	.01	.048	.06	.049	.048	.047	.048	.047	.048	.047	.047
.3	.062	.056	.02	.051	.29	.050	.047	.047	.049	.049	.049	.048	.048
.5	.068	.058	.03	.049	.49	.050	.047	.048	.050	.049	.049	.049	.050
.7	.078	.062	.06	.051	.68	.053	.048	.049	.051	.049	.050	.051	.050
.9	.117	.077	.14	.055	.88	.087	.058	.056	.058	.056	.056	.055	.055
.95	.162	.092	.23	.064	.93	.136	.079	.070	.067	.068	.066	.064	.064

Table VIII: Continued

Panel D: Bartlett Kernel, $AR(1)$ Prewhitening, $T = 200$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.051	.050	.01	.049	.05	.048	.048	.048	.048	.047	.048	.047	.047
-.5	.052	.051	.01	.050	.03	.049	.048	.047	.048	.048	.048	.047	.047
-.3	.052	.051	.01	.049	.02	.049	.048	.048	.047	.047	.048	.048	.047
.0	.053	.053	.01	.052	.01	.049	.048	.048	.047	.048	.048	.048	.048
.3	.055	.054	.01	.052	.02	.048	.048	.048	.047	.047	.048	.048	.048
.5	.057	.055	.01	.051	.03	.048	.048	.047	.047	.047	.048	.047	.047
.7	.059	.057	.01	.054	.04	.049	.049	.048	.048	.046	.047	.047	.047
.9	.080	.079	.02	.070	.05	.061	.056	.054	.052	.051	.053	.051	.051
.95	.103	.101	.01	.092	.05	.081	.068	.066	.063	.063	.062	.061	.060

Panel E: Parzen Kernel, $AR(1)$ Prewhitening, $T = 200$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.052	.049	.01	.050	.04	.048	.050	.049	.049	.047	.045	.045	.044
-.5	.051	.050	.01	.050	.03	.048	.050	.050	.048	.047	.046	.046	.045
-.3	.052	.050	.01	.050	.02	.048	.048	.049	.049	.047	.047	.046	.045
.0	.053	.052	.01	.052	.01	.048	.049	.049	.049	.047	.046	.046	.046
.3	.055	.054	.01	.052	.02	.049	.049	.049	.048	.047	.046	.046	.045
.5	.057	.055	.01	.053	.03	.049	.048	.049	.048	.047	.046	.046	.046
.7	.058	.057	.01	.055	.04	.049	.048	.048	.047	.046	.046	.045	.045
.9	.079	.077	.01	.072	.05	.064	.055	.053	.049	.048	.047	.046	.044
.95	.102	.100	.01	.095	.05	.085	.069	.064	.058	.054	.053	.050	.049

Panel F: QS Kernel, $AR(1)$ Prewhitening, $T = 200$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.051	.050	.01	.049	.05	.051	.049	.046	.045	.043	.041	.037	.035
-.5	.052	.050	.01	.050	.03	.050	.049	.047	.046	.045	.044	.041	.040
-.3	.051	.050	.01	.050	.02	.050	.049	.047	.047	.045	.045	.042	.041
.0	.053	.052	.01	.052	.01	.051	.048	.048	.047	.046	.044	.042	.042
.3	.055	.054	.01	.051	.02	.050	.048	.048	.047	.045	.044	.042	.042
.5	.056	.055	.01	.051	.03	.050	.048	.048	.047	.045	.044	.041	.041
.7	.059	.057	.01	.053	.04	.049	.048	.048	.047	.044	.042	.041	.039
.9	.079	.077	.01	.069	.05	.057	.049	.048	.045	.044	.043	.041	.040
.95	.103	.101	.01	.089	.05	.075	.057	.053	.050	.046	.045	.041	.040

Table IX: Empirical Null Rejection Probabilities in Simple Regression Model
5% Nominal Level, 10,000 Replications

$y_t = \theta_1 + \theta_2 x_t + u_t$, $x_t = 0.5x_{t-1} + \varepsilon_t$, $\varepsilon_t \sim i.i.d.\ N(0,1)$, $x_0 = 0$, $u_t = \rho u_{t-1} + \xi_t$, $\xi_t \sim i.i.d.$ $N(0,1)$, $u_0 = 0$ $H_0 : \theta_1 \leq 0$, $H_1 : \theta_1 > 0$

Panel A: Bartlett Kernel, No Prewhitening, $T = 50$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.048	.035	.07	.040	.35	.036	.039	.040	.041	.041	.040	.040	.041
-.5	.054	.043	.05	.047	.24	.045	.049	.050	.051	.051	.051	.050	.051
-.3	.061	.051	.04	.055	.18	.053	.055	.056	.055	.055	.055	.055	.055
.0	.073	.065	.03	.065	.12	.062	.061	.064	.062	.062	.062	.062	.063
.3	.087	.081	.03	.073	.14	.073	.068	.068	.067	.068	.067	.068	.068
.5	.099	.090	.04	.077	.18	.080	.072	.072	.070	.071	.072	.073	.073
.7	.110	.096	.05	.079	.24	.085	.079	.077	.077	.077	.077	.077	.077
.9	.122	.104	.07	.087	.31	.099	.086	.085	.084	.082	.083	.082	.082
.95	.127	.109	.07	.087	.33	.100	.086	.084	.085	.083	.083	.083	.084

Panel B: Parzen Kernel, No Prewhitening, $T = 50$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.050	.041	.08	.046	.35	.042	.048	.048	.046	.048	.048	.049	.047
-.5	.056	.046	.07	.052	.24	.049	.054	.055	.054	.054	.054	.054	.054
-.3	.063	.053	.06	.057	.18	.055	.058	.058	.057	.055	.055	.055	.055
.0	.074	.065	.06	.066	.12	.063	.062	.062	.061	.060	.060	.060	.059
.3	.086	.075	.07	.073	.14	.073	.068	.067	.063	.064	.064	.063	.062
.5	.095	.083	.09	.077	.18	.082	.072	.069	.066	.064	.064	.063	.063
.7	.105	.085	.11	.078	.24	.088	.075	.074	.070	.069	.069	.068	.068
.9	.119	.092	.14	.082	.31	.103	.082	.078	.075	.074	.073	.073	.072
.95	.122	.094	.14	.081	.33	.106	.082	.077	.074	.073	.073	.074	.073

Panel C: QS Kernel, No Prewhitening, $T = 50$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.049	.041	.04	.047	.35	.048	.046	.048	.048	.048	.048	.049	.050
-.5	.053	.045	.03	.050	.24	.053	.052	.054	.055	.055	.055	.055	.056
-.3	.061	.052	.03	.056	.18	.058	.056	.054	.054	.056	.056	.056	.055
.0	.072	.064	.03	.063	.12	.063	.061	.060	.061	.061	.060	.062	.062
.3	.085	.075	.04	.070	.14	.070	.063	.063	.062	.064	.064	.065	.065
.5	.095	.083	.04	.071	.18	.074	.064	.064	.062	.063	.064	.064	.064
.7	.103	.086	.06	.073	.24	.076	.069	.068	.067	.067	.067	.067	.067
.9	.116	.092	.07	.077	.31	.086	.074	.074	.073	.072	.072	.072	.072
.95	.119	.095	.07	.078	.33	.087	.073	.074	.074	.071	.069	.068	.067

Table IX: Continued

Panel D: Bartlett Kernel, $AR(1)$ Prewhitening, $T = 50$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.060	.057	.02	.057	.05	.051	.049	.048	.049	.049	.049	.047	.048
-.5	.065	.062	.01	.062	.03	.057	.056	.056	.057	.057	.056	.055	.055
-.3	.072	.069	.01	.068	.03	.062	.060	.061	.059	.060	.060	.059	.058
.0	.078	.075	.01	.073	.02	.066	.064	.066	.063	.063	.062	.062	.062
.3	.083	.081	.01	.079	.02	.072	.067	.068	.066	.066	.066	.064	.064
.5	.089	.086	.01	.083	.03	.074	.071	.069	.068	.067	.067	.068	.068
.7	.090	.087	.01	.085	.03	.075	.072	.072	.071	.070	.069	.069	.068
.9	.097	.093	.01	.090	.04	.081	.077	.077	.075	.074	.075	.073	.074
.95	.102	.097	.02	.092	.05	.082	.078	.077	.077	.074	.074	.072	.072

Panel E: Parzen Kernel, $AR(1)$ Prewhitening, $T = 50$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.061	.055	.04	.058	.05	.052	.051	.050	.048	.049	.049	.049	.047
-.5	.065	.060	.03	.062	.03	.057	.057	.056	.054	.055	.055	.054	.052
-.3	.072	.066	.03	.069	.03	.062	.061	.060	.058	.056	.056	.055	.054
.0	.077	.073	.03	.074	.02	.067	.064	.063	.063	.061	.060	.058	.057
.3	.083	.078	.03	.080	.02	.071	.068	.066	.064	.063	.063	.061	.060
.5	.088	.083	.03	.084	.03	.075	.072	.069	.065	.063	.062	.060	.060
.7	.090	.084	.04	.085	.03	.076	.071	.070	.068	.068	.065	.064	.064
.9	.096	.090	.04	.090	.04	.082	.076	.074	.073	.070	.071	.069	.067
.95	.101	.093	.04	.093	.05	.084	.077	.074	.072	.070	.069	.069	.067

Panel F: QS Kernel, $AR(1)$ Prewhitening, $T = 50$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.060	.055	.02	.057	.05	.051	.048	.049	.049	.046	.046	.045	.042
-.5	.064	.060	.02	.062	.03	.056	.054	.055	.054	.053	.052	.048	.045
-.3	.072	.066	.02	.068	.03	.061	.057	.055	.054	.052	.050	.047	.045
.0	.077	.074	.01	.073	.02	.065	.062	.060	.059	.057	.056	.053	.051
.3	.083	.079	.01	.078	.02	.070	.064	.063	.060	.060	.059	.057	.054
.5	.088	.083	.02	.082	.03	.072	.064	.063	.060	.059	.059	.056	.054
.7	.089	.084	.02	.084	.03	.073	.067	.066	.065	.065	.061	.057	.055
.9	.095	.090	.02	.089	.04	.078	.071	.070	.069	.066	.064	.060	.058
.95	.100	.094	.02	.091	.05	.077	.071	.072	.069	.063	.060	.056	.054

Table X: Empirical Null Rejection Probabilities in Simple Regression Model
5% Nominal Level, 10,000 Replications
$y_t = \theta_1 + \theta_2 x_t + u_t$, $x_t = 0.5x_{t-1} + \varepsilon_t$, $\varepsilon_t \sim i.i.d.\ N(0,1)$, $x_0 = 0$, $u_t = \rho u_{t-1} + \xi_t$, $\xi_t \sim i.i.d.$ $N(0,1)$, $u_0 = 0$ $H_0 : \theta_1 \leq 0$, $H_1 : \theta_1 > 0$

Panel A: Bartlett Kernel, No Prewhitening, $T = 100$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.044	.034	.05	.045	.37	.038	.043	.043	.045	.045	.045	.045	.045
-.5	.052	.046	.03	.051	.24	.050	.050	.050	.050	.051	.051	.051	.052
-.3	.055	.050	.02	.053	.16	.054	.053	.052	.053	.051	.051	.051	.052
.0	.064	.060	.02	.058	.09	.057	.056	.056	.053	.054	.054	.054	.055
.3	.078	.072	.02	.064	.13	.063	.060	.060	.060	.059	.059	.059	.059
.5	.085	.078	.03	.067	.20	.068	.064	.065	.064	.064	.063	.063	.063
.7	.095	.084	.04	.070	.28	.071	.069	.068	.067	.067	.067	.069	.069
.9	.099	.088	.05	.070	.37	.076	.070	.070	.070	.070	.070	.069	.070
.950	.109	.092	.05	.073	.39	.082	.071	.071	.072	.073	.072	.072	.073

Panel B: Parzen Kernel, No Prewhitening, $T = 100$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.046	.039	.05	.049	.37	.045	.048	.049	.048	.047	.048	.048	.048
-.5	.055	.049	.04	.055	.24	.055	.055	.054	.052	.052	.053	.052	.052
-.3	.056	.051	.04	.056	.16	.057	.056	.054	.052	.051	.052	.052	.051
.0	.064	.059	.03	.059	.09	.059	.058	.056	.054	.053	.052	.053	.051
.3	.075	.067	.04	.065	.13	.063	.061	.057	.056	.056	.056	.056	.056
.5	.081	.072	.06	.067	.20	.068	.063	.063	.061	.061	.060	.060	.059
.7	.090	.074	.07	.067	.28	.070	.066	.065	.065	.063	.063	.061	.061
.9	.094	.079	.09	.067	.37	.076	.067	.065	.063	.064	.064	.064	.062
.95	.100	.082	.09	.068	.39	.082	.070	.066	.064	.065	.064	.064	.062

Panel C: QS Kernel, No Prewhitening, $T = 100$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.046	.039	.02	.049	.37	.048	.047	.048	.050	.050	.050	.050	.050
-.5	.054	.049	.02	.053	.24	.055	.052	.052	.052	.052	.052	.051	.051
-.3	.055	.050	.02	.054	.16	.057	.051	.052	.052	.052	.054	.055	.054
.0	.064	.058	.02	.058	.09	.057	.053	.052	.051	.050	.051	.051	.050
.3	.075	.068	.02	.062	.13	.061	.055	.057	.055	.055	.055	.055	.055
.5	.081	.073	.03	.066	.20	.064	.060	.061	.059	.058	.058	.057	.057
.7	.089	.075	.04	.066	.28	.066	.062	.062	.060	.059	.060	.059	.058
.9	.093	.079	.04	.064	.37	.068	.062	.063	.063	.062	.062	.061	.063
.95	.099	.082	.05	.065	.39	.070	.063	.064	.062	.062	.062	.062	.062

Table X: Continued

Panel D: Bartlett Kernel, $AR(1)$ Prewhitening, $T = 100$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.055	.053	.01	.052	.04	.049	.049	.050	.051	.049	.049	.049	.049
-.5	.060	.060	.01	.059	.02	.056	.054	.053	.053	.054	.053	.054	.054
-.3	.062	.061	.01	.061	.02	.057	.055	.055	.054	.053	.053	.053	.053
.0	.067	.066	.01	.065	.01	.059	.057	.056	.054	.055	.054	.053	.053
.3	.071	.069	.01	.067	.01	.062	.061	.060	.060	.058	.058	.058	.058
.5	.073	.072	.01	.070	.02	.063	.062	.063	.063	.061	.061	.061	.062
.7	.074	.072	.01	.070	.03	.063	.063	.065	.063	.063	.062	.063	.063
.9	.077	.075	.01	.074	.04	.066	.062	.062	.063	.065	.063	.063	.063
.95	.079	.077	.01	.074	.04	.068	.064	.064	.065	.066	.065	.064	.064

Panel E: Parzen Kernel, $AR(1)$ Prewhitening, $T = 100$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.056	.051	.02	.054	.04	.050	.050	.049	.049	.048	.048	.047	.047
-.5	.061	.058	.02	.059	.02	.058	.056	.055	.053	.053	.053	.053	.051
-.3	.063	.060	.01	.061	.02	.059	.057	.055	.053	.053	.052	.052	.051
.0	.067	.064	.01	.065	.01	.060	.058	.058	.055	.052	.051	.051	.051
.3	.071	.068	.01	.068	.01	.062	.061	.059	.057	.056	.056	.055	.053
.5	.073	.070	.02	.071	.01	.064	.063	.063	.061	.061	.060	.059	.058
.7	.074	.071	.02	.071	.03	.065	.064	.064	.063	.062	.062	.060	.060
.9	.078	.074	.02	.075	.04	.067	.064	.062	.060	.061	.061	.060	.059
.95	.079	.075	.02	.075	.04	.069	.065	.063	.062	.062	.061	.060	.060

Panel F: QS Kernel, $AR(1)$ Prewhitening, $T = 100$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.055	.052	.01	.053	.04	.050	.049	.048	.049	.049	.049	.046	.045
-.5	.060	.058	.01	.059	.02	.057	.052	.052	.051	.051	.050	.047	.046
-.3	.062	.059	.01	.061	.02	.058	.053	.052	.051	.050	.049	.047	.046
.0	.067	.065	.01	.064	.01	.059	.054	.051	.051	.049	.049	.046	.044
.3	.071	.068	.01	.067	.01	.061	.056	.057	.055	.053	.052	.049	.047
.5	.073	.071	.01	.069	.02	.063	.061	.061	.058	.057	.055	.053	.051
.7	.074	.071	.01	.069	.03	.064	.062	.061	.060	.058	.057	.054	.052
.9	.077	.075	.01	.073	.04	.064	.060	.061	.060	.059	.058	.054	.053
.95	.079	.076	.01	.074	.04	.067	.061	.062	.060	.058	.056	.054	.053

Table XI: Empirical Null Rejection Probabilities in Simple Regression Model
5% Nominal Level, 10,000 Replications

$y_t = \theta_1 + \theta_2 x_t + u_t$, $x_t = 0.5x_{t-1} + \varepsilon_t$, $\varepsilon_t \sim i.i.d.\ N(0,1)$, $x_0 = 0$, $u_t = \rho u_{t-1} + \xi_t$, $\xi_t \sim i.i.d.$ $N(0,1)$, $u_0 = 0$ $H_0 : \theta_1 \leq 0$, $H_1 : \theta_1 > 0$

Panel A: Bartlett Kernel, No Prewhitening, $T = 200$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.043	.037	.03	.048	.38	.042	.045	.046	.046	.047	.046	.045	.046
-.5	.047	.043	.02	.048	.24	.047	.047	.048	.049	.048	.049	.048	.048
-.3	.049	.047	.02	.049	.15	.048	.049	.050	.050	.050	.050	.050	.050
.0	.056	.055	.01	.054	.07	.053	.052	.052	.053	.054	.055	.054	.054
.3	.068	.065	.01	.056	.13	.054	.054	.055	.056	.056	.055	.054	.055
.5	.073	.068	.02	.055	.22	.056	.055	.055	.055	.056	.056	.055	.055
.7	.078	.072	.03	.057	.31	.059	.056	.056	.055	.056	.056	.056	.056
.9	.084	.076	.03	.058	.41	.062	.059	.059	.060	.059	.059	.060	.061
.95	.085	.075	.04	.059	.43	.061	.058	.059	.060	.060	.060	.061	.062

Panel B: Parzen Kernel, No Prewhitening, $T = 200$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.046	.043	.03	.047	.38	.047	.048	.047	.048	.048	.050	.050	.051
-.5	.050	.045	.02	.050	.24	.051	.048	.048	.050	.049	.049	.049	.048
-.3	.051	.048	.02	.050	.15	.051	.050	.050	.050	.051	.050	.049	.049
.0	.057	.054	.02	.054	.07	.054	.052	.052	.053	.053	.054	.053	.053
.3	.065	.062	.03	.056	.13	.056	.053	.055	.054	.053	.053	.054	.054
.5	.069	.063	.03	.054	.22	.057	.055	.054	.050	.052	.053	.054	.053
.7	.074	.065	.04	.054	.31	.059	.054	.054	.052	.052	.053	.052	.051
.9	.078	.070	.06	.057	.41	.060	.057	.057	.056	.056	.055	.056	.056
.95	.079	.068	.06	.058	.43	.059	.057	.057	.056	.056	.056	.056	.056

Panel C: QS Kernel, No Prewhitening, $T = 200$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.047	.043	.01	.050	.38	.046	.048	.049	.052	.050	.049	.048	.048
-.5	.049	.045	.01	.050	.24	.049	.050	.049	.048	.049	.049	.048	.048
-.3	.050	.048	.01	.050	.15	.049	.050	.051	.051	.051	.051	.051	.049
.0	.057	.054	.01	.054	.07	.053	.052	.054	.054	.055	.055	.055	.054
.3	.066	.062	.01	.054	.13	.052	.054	.052	.055	.055	.054	.055	.053
.5	.069	.064	.01	.052	.22	.054	.049	.052	.054	.054	.053	.052	.052
.7	.073	.066	.02	.054	.31	.055	.052	.053	.053	.052	.051	.050	.050
.9	.078	.070	.03	.057	.41	.055	.055	.055	.056	.055	.055	.054	.054
.95	.078	.068	.03	.056	.43	.056	.056	.056	.057	.057	.057	.057	.058

Table XI: Continued

Panel D: Bartlett Kernel, $AR(1)$ Prewhitening, $T = 200$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.053	.053	.01	.050	.03	.047	.047	.048	.049	.049	.049	.049	.048
-.5	.054	.053	.01	.053	.02	.050	.049	.050	.050	.049	.049	.049	.049
-.3	.055	.055	.01	.054	.01	.051	.050	.051	.051	.050	.050	.049	.050
.0	.058	.058	.01	.057	.01	.054	.052	.053	.053	.055	.054	.053	.053
.3	.060	.060	.01	.060	.01	.053	.054	.054	.056	.054	.053	.053	.054
.5	.060	.060	.01	.059	.01	.053	.054	.054	.054	.055	.055	.054	.054
.7	.061	.060	.01	.059	.02	.056	.053	.054	.054	.054	.053	.052	.053
.9	.064	.064	.02	.059	.03	.055	.055	.055	.054	.055	.055	.055	.055
.95	.064	.063	.02	.060	.03	.055	.054	.054	.057	.057	.057	.057	.057

Panel E: Parzen Kernel, $AR(1)$ Prewhitening, $T = 200$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.054	.052	.01	.051	.03	.048	.048	.047	.048	.049	.049	.049	.049
-.5	.054	.052	.01	.053	.02	.052	.048	.049	.049	.048	.049	.047	.047
-.3	.055	.054	.01	.055	.01	.053	.051	.051	.051	.051	.051	.049	.048
.0	.058	.057	.01	.058	.01	.055	.052	.052	.053	.053	.054	.053	.053
.3	.060	.059	.01	.059	.01	.055	.053	.054	.054	.054	.053	.054	.053
.5	.060	.059	.01	.059	.01	.056	.055	.054	.051	.052	.053	.053	.052
.7	.060	.059	.01	.059	.02	.057	.054	.053	.052	.052	.052	.051	.050
.9	.064	.063	.01	.062	.03	.056	.056	.056	.055	.054	.054	.054	.054
.95	.064	.062	.01	.062	.03	.055	.056	.056	.055	.055	.056	.056	.055

Panel F: QS Kernel, $AR(1)$ Prewhitening, $T = 200$

ρ	$t_{\widehat{b}}$	$t^*_{\widehat{b}}$	$ave(\widehat{b})$	$t^*_{\widetilde{b}}$	$ave(\widetilde{b})$	$t^*_{0.1}$	$t^*_{0.25}$	$t^*_{0.35}$	$t^*_{0.5}$	$t^*_{0.65}$	$t^*_{0.75}$	$t^*_{0.9}$	$t^*_{1.0}$
-.8	.054	.051	.01	.049	.03	.046	.048	.050	.051	.049	.047	.045	.044
-.5	.054	.052	.01	.054	.02	.050	.050	.049	.049	.048	.047	.045	.044
-.3	.055	.054	.01	.054	.01	.050	.051	.052	.050	.049	.048	.047	.044
.0	.058	.057	.01	.058	.01	.053	.053	.054	.053	.053	.053	.052	.050
.3	.060	.059	.01	.059	.01	.053	.053	.053	.054	.054	.053	.052	.050
.5	.060	.059	.01	.059	.01	.054	.050	.052	.052	.051	.051	.049	.048
.7	.061	.059	.01	.059	.02	.055	.051	.053	.051	.051	.050	.048	.047
.9	.064	.063	.01	.059	.03	.053	.054	.053	.054	.054	.053	.051	.050
.95	.064	.062	.01	.060	.03	.055	.055	.054	.055	.055	.054	.054	.053

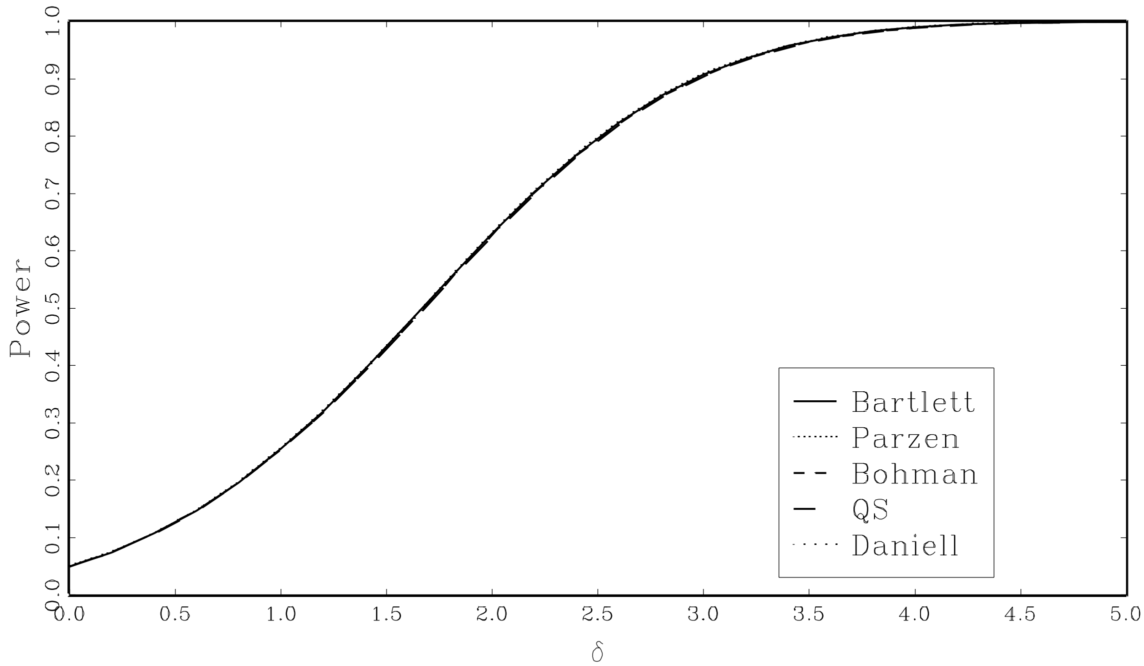

Figure 1: Local Asymptotic Power of t_b^*, b = 0.02

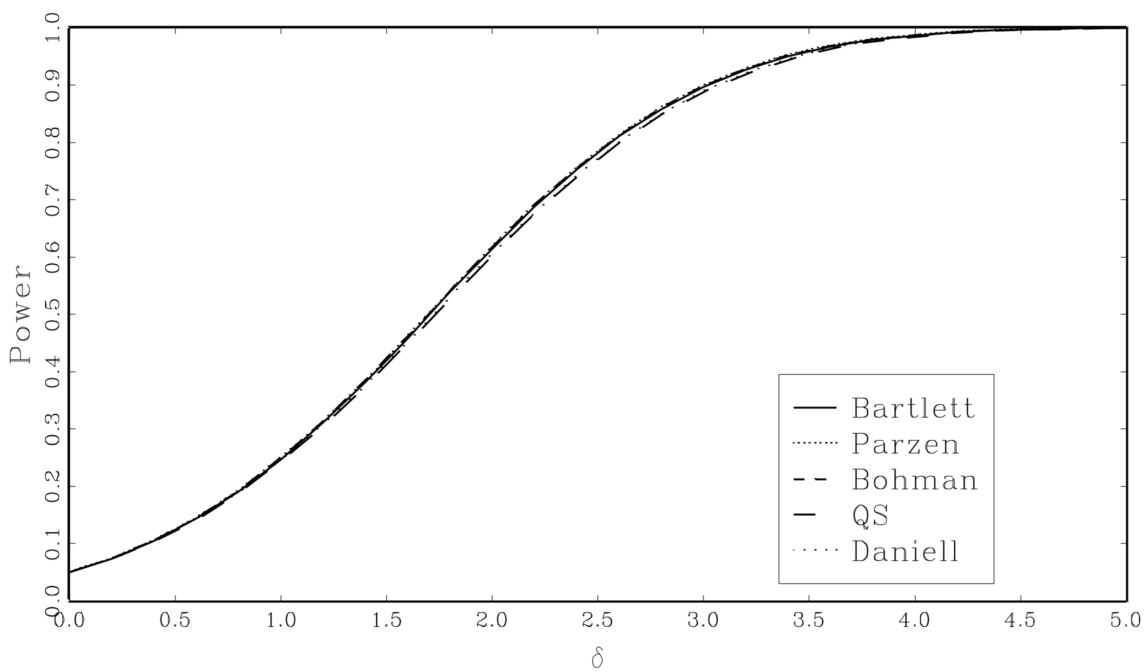

Figure 2: Local Asymptotic Power of t_b^*, b = 0.06

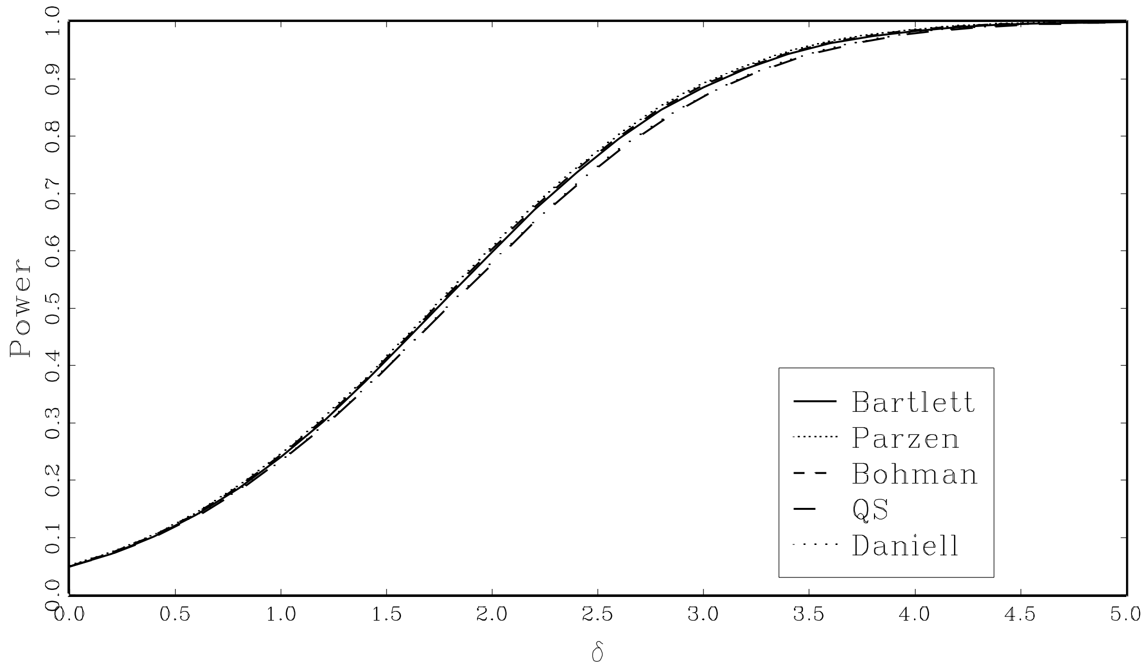

Figure 3: Local Asymptotic Power of t_b^*, b = 0.1

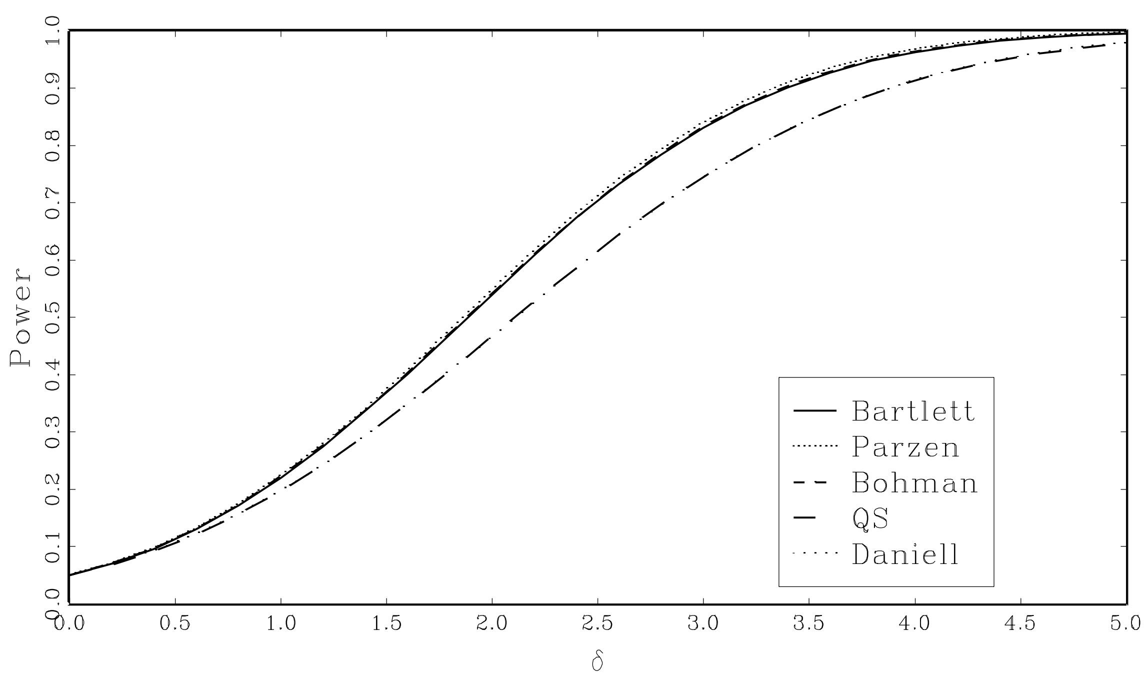

Figure 4: Local Asymptotic Power of t_b^*, b = 0.3

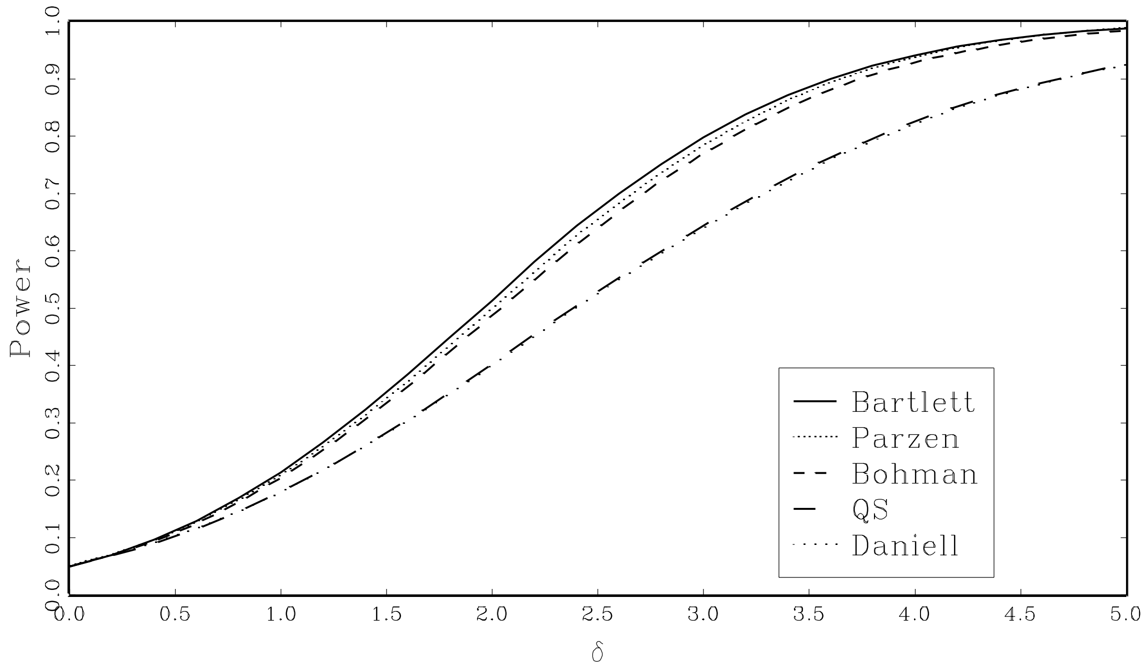

Figure 5: Local Asymptotic Power of t_b^*, b = 0.5

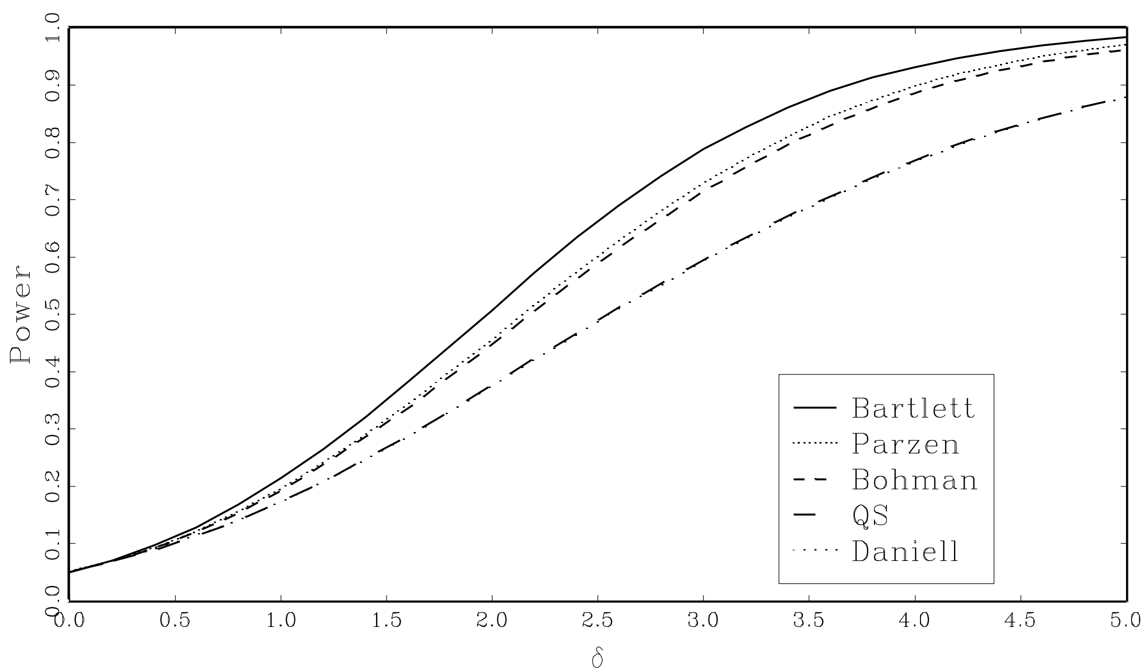

Figure 6: Local Asymptotic Power of t_b^*, b = 0.7

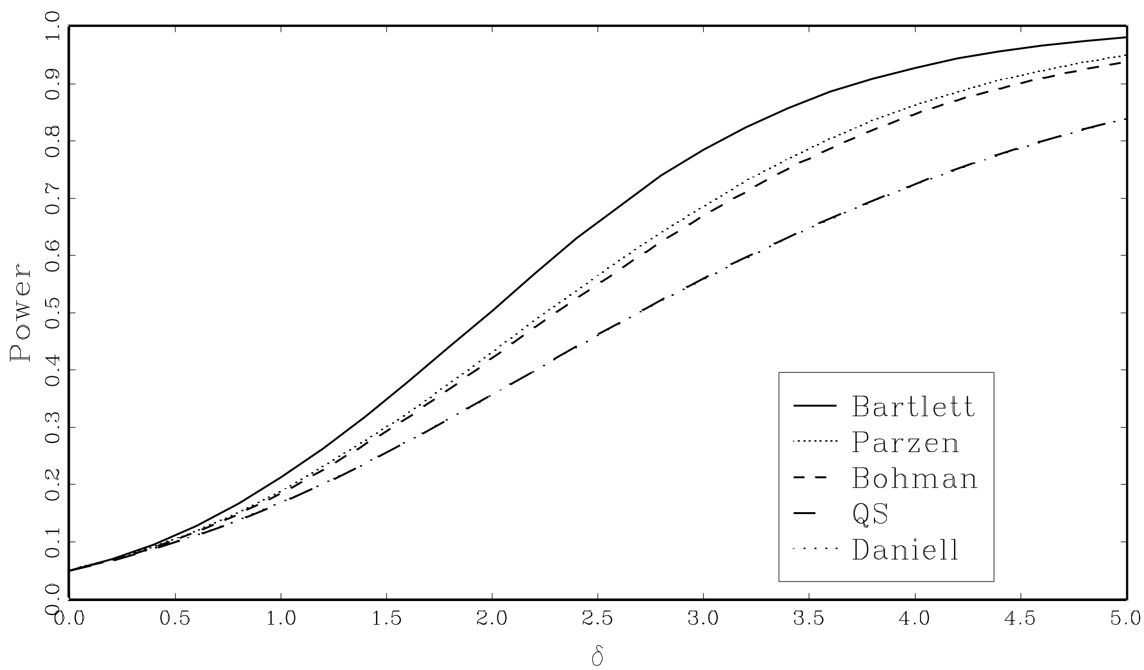

Figure 7: Local Asymptotic Power of t_b^*, b = 0.9

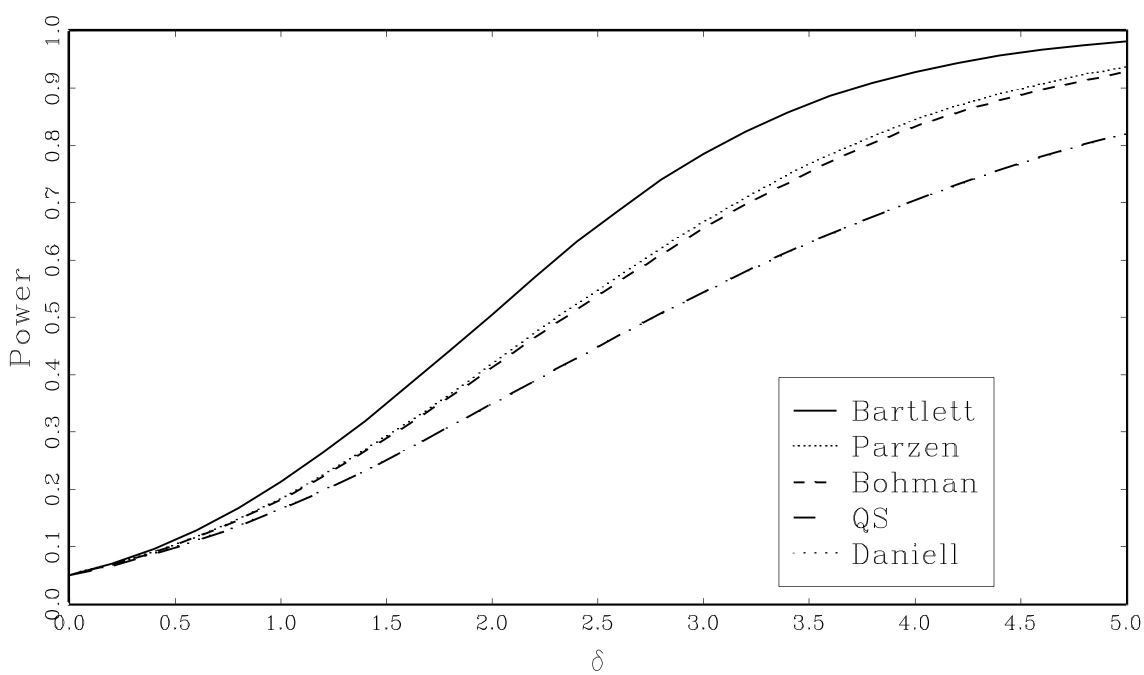

Figure 8: Local Asymptotic Power of t_b^*, b = 1.0

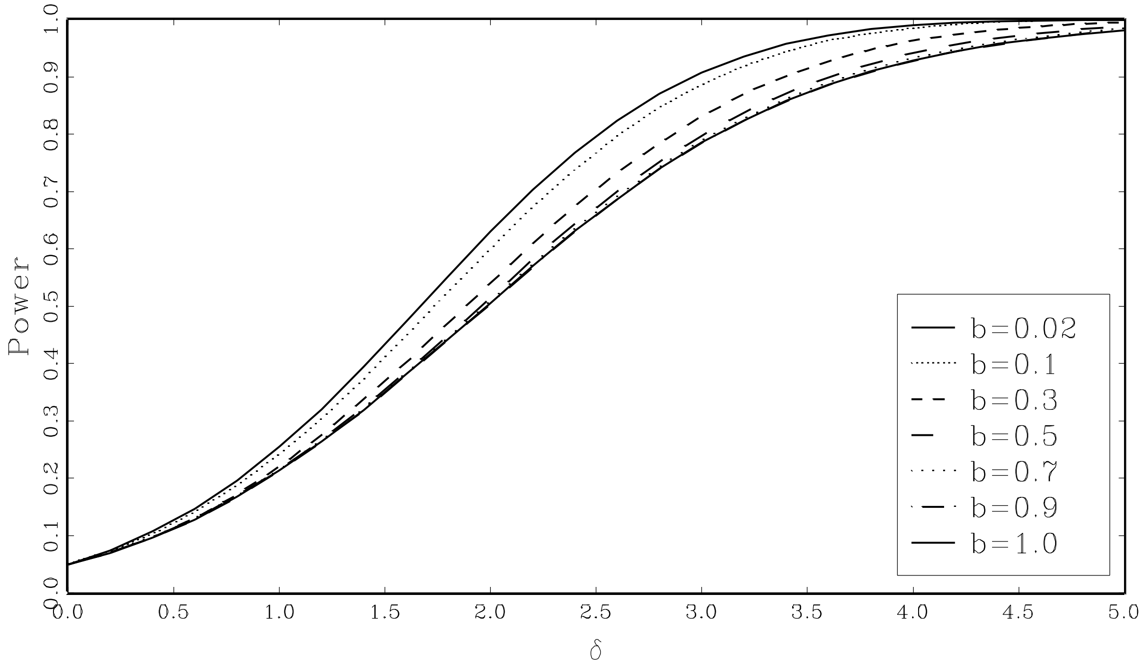

Figure 9: Local Asymptotic Power of t_b^*, Bartlett Kernel

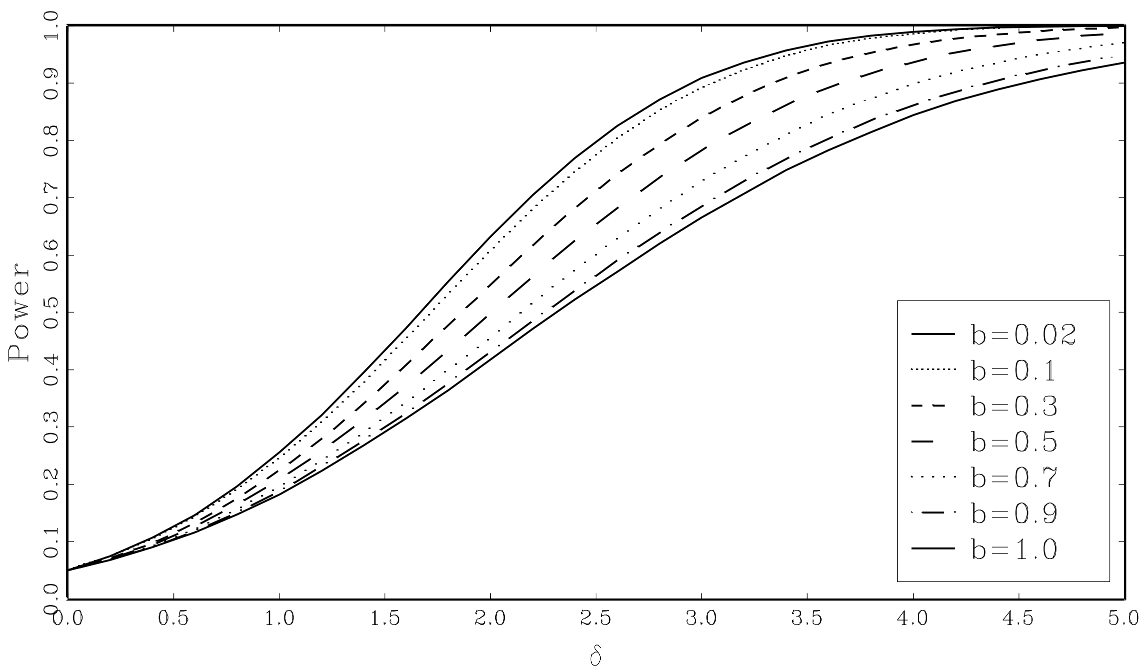

Figure 10: Local Asymptotic Power of t_b^*, Parzen Kernel

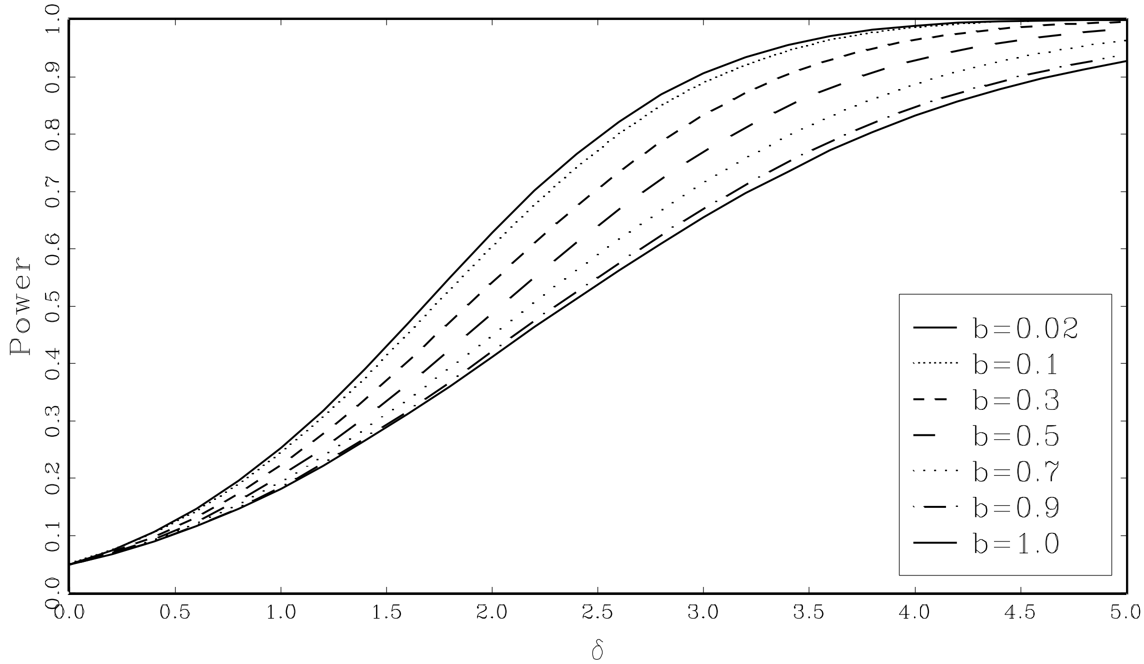

Figure 11: Local Asymptotic Power of t_b^*, Bohman Kernel

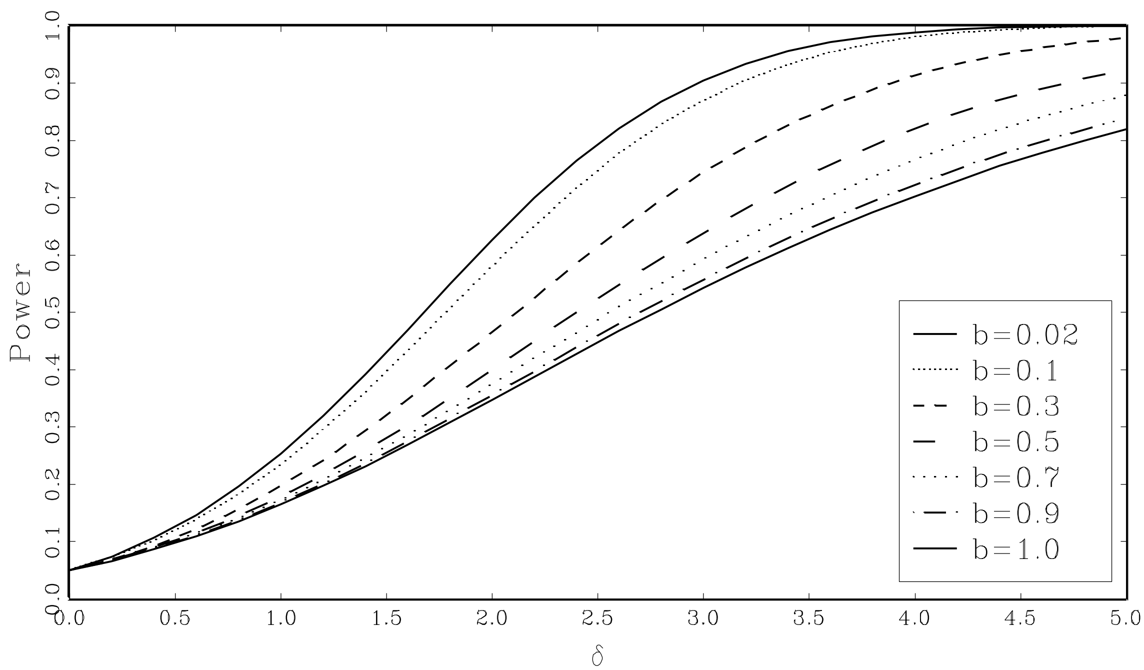

Figure 12: Local Asymptotic Power of t_b^*, Daniell Kernel

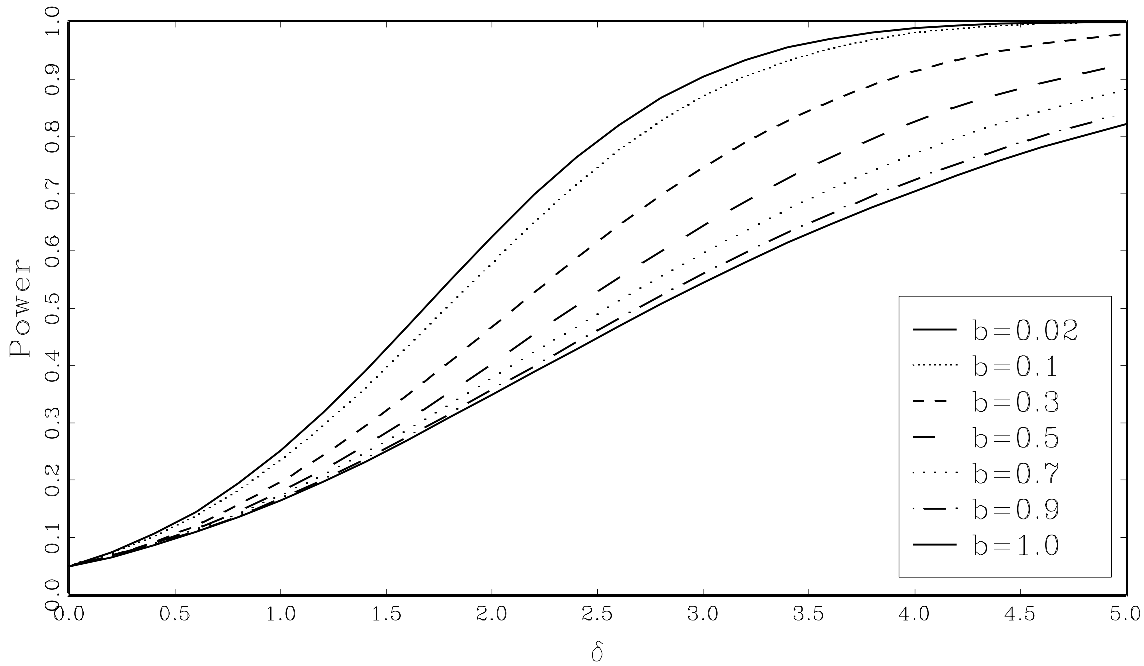

Figure 13: Local Asymptotic Power of t_b^*, QS Kernel