MONOTONIC POWER IN TESTS FOR CHANGING MEAN

TED JUHL AND ZHIJIE XIAO

Abstract. Many tests for a changing mean in a time series require the estimation of a long run variance. Recent research of Crainiceanu and Vogelsang (2001) suggests that tests involving a long run variance estimate exhibit non-monotonic power. We propose an estimate of long run variances that allows for a time varying mean. In this way, we are able to avoid non-monotonic power while still retaining the same asymptotic distribution under the null hypothesis of constant parameters. A Monte Carlo experiment illustrates the increased power.

1. Introduction

It is common to test the stability of economic relationships in an empirical study. There have been many tests proposed to test for possible structural changes in regression models. Some of these tests include Andrews (1993), Andrews and Ploberger (1994), Brown, Durbin, and Evans (1975), Gardner (1969), MacNeill (1978), and Ploberger and Kramer (1992), to name a few. Recent work by Elliott and Müller (2003) suggests that the asymptotic performance of these tests is very similar.

Hansen (2000) notes that many tests for structural change in regression models are sensitive to a change in the marginal distribution of the regressors. For example, if a regressor has a change in mean, some standard tests are vastly oversized. In light of this evidence, it is important to test the stability of the distribution of any time series used in a regression model. The above tests are valid if the regression model only contains a constant and each one can then be used to test a single series for a constant mean.

Vogelsang (1997, 1999) shows that some tests for a change in mean can exhibit nonmonotonic power when estimating a dynamic model. That is, as the change in

Version: July 15, 2004. Ted Juhl, Department of Economics, University of Kansas, 213 Summerfield, Lawrence, KS 66045; Tel (785)-864-2849; (juhl@ku.edu), Zhijie Xiao, Department of Economics, Boston College, Chestnut Hill, MA, 02467; Tel (617)-552-3670; (zhijie.xiao@bc.edu).

mean grows, the power of the test actually decreases with power going to zero in some cases. Moreover, Crainiceanu and Vogelsang (2001) showed that if one estimates static models but uses a long-run variance, nonmonotonic power is possible. In addition, their work shows that the cause of the nonmonotonic power is a diverging long-run variance estimate. Juhl and Xiao (2004) recently proposed a test for changing mean that avoids nonmonotonic power by using nonparametric estimation in a dynamic model.

In this paper, we propose a modification to tests for a changing mean that is designed to avoid nonmonotonic power. These tests are modified using a long-run variance estimator that converges regardless of the stability of the mean. By "robustifying" the long-run variance estimator, we do not obtain divergent estimators. This allows the tests to retain their power and avoid nonmonotonic power. We illustrate the performance of the tests using a small Monte Carlo experiment.

2. Model and Tests

Consider a simple process governed by

$$y_t = \theta\left(\frac{t}{T}\right) + u_t,$$

where u_t is potentially serially correlated. Notice that θ is allowed to vary over time in the current setup. Our purpose is to test whether or not there is a change in θ. More precisely, we want to test the null hypothesis

$$H_0 : \theta\left(\frac{t}{T}\right) = \bar{\theta} = \text{constant},$$

so that there is no change in mean.

A well-known approach for testing structural change in the parameters is to look at the fluctuation in the OLS residuals, $\hat{u}_t$. Under H_0, the fluctuation in $\hat{u}_t$ is stable and has a well-behaved limiting distribution. However, under H_1, the fluctuation in $\hat{u}_t$ will be larger. Since u_t may be serially correlated, we must estimate a long-run variance for the process. That is, we need to find a variance for the scaled process

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}u_t.$$

We denote the long-run variance of u_t as ω^2 so that

$$\omega^2 = \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\sum_{t'=1}^{T}E\left(u_t u_{t'}\right).$$

Consider the following empirical process (or partial sum process)

$$U_T(r) = \frac{1}{\hat{\omega}\sqrt{T}}\sum_{t=1}^{[Tr]}\hat{u}_t$$

where $\hat{\omega}$ is a consistent estimator for ω. Under the null hypothesis, $U_T(r)$ converges weakly to a demeaned Brownian motion, say $\tilde{W}(r)$. Consider a continuous functional $g(\cdot)$ that measures the fluctaion of $U_T(r)$ so that we can use $g(U_T(r))$ as a test statistic for the stationarity of u_t, and hence the null hypothesis. By the continuous mapping theorem,

$$g(U_T(r)) \Rightarrow g(\tilde{W}(r)).$$

In principle, any metric that measures the fluctuation in $U_T(r)$ is a candidate for the functional $g(\cdot)$. The classical Kolmogorov-Smirnoff or Cramer-von Mises type measures are of particular interest and have been widely used in the literature for testing for structural change. When we take $g(\cdot)$ as the classical Kolmogorov-Smirnoff measure, we obtain the following CUSUM type statistic

$$KS_T = \sup_{0\le r\le 1}|U_T(r)| = \max_{j=1,\ldots,T}\frac{1}{\hat{\omega}}\left|\frac{1}{\sqrt{T}}\sum_{t=1}^{j}\hat{u}_t\right|.$$

Moreover, we may consider $R \subseteq [0,1]$, and a generalized version of the Kolmogorov-Smirnoff test given as

$$GKS_T = \sup_{r\in R}|U_T(r)|.$$

Using the Cramer-von Mises type metric, we can construct the test

$$CVT = \int_0^1 U_T(r)^2 dr = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{\hat{\omega}\sqrt{T}}\sum_{j=1}^{t}\hat{u}_j\right)^2.$$

The above test is also labeled as QS in the literature (see, for example, Gardner (1969), MacNeill (1978), and Perron (1991)). Again, a generalized statistic may be constructed for a suitable chosen weight function, $v(r)$, and $R \subseteq [0,1]$,

$$GCV_T = \int_{r \in R} v(r) U_T(r)^2 dr.$$

In a recent paper, Elliott and Müller (2003) develop a test that has optimal power against a very general class of alternatives for $\theta(t/T)$. The statistic itself is calculated using several very simple steps. First, we calculate the series $\hat{u}_t = y_t - \bar{y}$ as an estimate of u_t. Next, denote the long-run variance of u_t as ω^2 and let the estimator for this variance be calculated using $\hat{u}_t$. Denote the estimator as $\hat{\omega}^2$ and construct the standardized series $\hat{z}_t = \hat{\omega}^{-1}\hat{u}_t$. Then construct a new series denoted by $\hat{w}_t = \bar{r}\hat{w}_{t-1} + \Delta\hat{z}_t$ and $\hat{w}_1 = \hat{z}_1$ with $\bar{r} = 1 - 10/T$. Now regress $\{\hat{w}_t\}$ on $\{\bar{r}^t\}$ and denote the residuals as $\hat{e}_t$. The final statistic is given as

$$EM_T = \bar{r}\sum_{t=1}^{T} \hat{e}_t^2 - \sum_{t=1}^{T} \hat{z}_t^2.$$

As noted by Elliott and Müller (2003), the above statistic is equivalent to the Most Powerful Invariant (MPI) test of Franzini and Harvey (1983) when there is no serial correlation in u_t.

All of the above tests make use of an estimated long-run variance. A widely used method is the nonparametric method suggested by Andrews (1991). For example, an estimate of ω^2 is

$$\hat{\omega}^2 = \sum_{j=-T+1}^{T-1} k\left(\frac{j}{b_T}\right)\hat{\gamma}(j)$$

where

$$\hat{\gamma}(j) = \begin{cases} \frac{1}{T}\sum_{t=j+1}^{T} \hat{u}_t \hat{u}_{t-j}^\top & \text{for} j \geq 0, \\ \frac{1}{T}\sum_{t=-j+1}^{T} \hat{u}_{t+j} \hat{u}_t^\top & \text{for} j < 0. \end{cases}$$

The function $k(x)$ is a kernel function and b_T is a bandwidth parameter satisfying the property $b_T \to \infty$ and $b_T/T \to 0$. Such an estimate of the long-run variance is consistent and guaranteed to be positive definite under simple regularity conditions.

Andrews (1991) suggests the optimal kernel

$$k(x) = \frac{25}{12\pi^2 x^2}\left(\frac{\sin(6\pi x/5)}{6\pi x/5} - \cos(6\pi x/5)\right).$$

The optimal bandwidth associated with this kernel is based on an AR(1) approximation. Let $\hat{\rho}$ and $\hat{\sigma}^2$ be the estimated AR(1) parameter and variance from a univariate autoregression of $\hat{u}_t$ on $\hat{u}_{t-1}$. The estimated optimal bandwidth parameter b_{opt} is given as

$$b_{opt} = 1.3221 \times (\hat{\alpha}(2)T)^{1/5}$$

with

$$\hat{\alpha}(2) = \frac{4\hat{\rho}_a^2}{(1-\hat{\rho}_a)^4}.$$

The KS_T and CV_T tests are very common and have been analyzed thoroughly (see Ploberger and Kramer (1992), for example). However, Crainiceanu and Vogelsang (2001) recently showed that these tests may exhibit nonmonotonic power when testing for a change in mean. Moreover, they show that power can go to zero as the change in parameters becomes very large. The reason for this unusual power behavior lies in the estimation of the long-run variance parameter. As the change in parameters becomes large, the optimal bandwidth parameter diverges too quickly. When the bandwidth diverges, the long-run variance diverges too fast relative to the rest of the test statistic. Since the long-run variance always appears in the denominator, power suffers and goes to zero in certain cases. The analytical and empirical evidence in Crainiceanu and Vogelsang (2001) shows that the nonmonotonic power is a potentially serious problem.

We propose an estimator of the long-run variance that still converges to the true long-run variance even under a class of alternatives including parameter change (non-constant θ). To this end, we estimate two sets of residuals. First let $\hat{u}_t$ be the OLS residuals from regressing y_t on a constant. These residuals do not allow for any structural change, with θ assumed to be constant. Next, we estimate $\theta(t/T)$ using nonparametric kernel methods. The Nadaraya-Watson estimator for $\theta(t/T)$ is given as

$$\tilde{\theta}(t/T) = \left(\frac{1}{Th}\sum_{s=1}^{T} K_{ts}\right)^{-1} \frac{1}{Th}\sum_{s=1}^{T} K_{ts} y_s,$$

where

$$K_{ts} = K\left(\frac{t-s}{Th}\right),$$

$K(\cdot)$ is a kernel function (different from the kernel function used for the long-run variance), and h is a bandwidth parameter such that $h \to 0$. The second set of residuals is defined as

$$\begin{aligned}\tilde{u}_t &= y_t - \tilde{y}_t \\ &= y_t - \tilde{\theta}(t/T).\end{aligned}$$

The KS_T, CV_T, and EM_T statistics all use OLS residuals and an estimate of the long-run variance. We propose modifying each of these statistics by using $\tilde{u}_t$ in the estimator for the the long-run variance. In this way, the long-run variance estimator will not diverge, and power will not be lost.

We provide the conditions under which the long-run variance estimator is consistent in the next section. These conditions will include a large class of situations under the alternative hypothesis, as this is precisely what is needed to avoid nonmonotonic power.

3. Asymptotic Results

We provide the conditions under which the modified long-run variance estimator is consistent in this section.

Assumption 1. *y_t is generated according to*

$$y_t = \theta\left(\frac{t}{T}\right) + u_t$$

where $E(u_t) = 0$ and $\theta(\cdot)$ is first order Lipshitz continuous.

Assumption 2. *u_t is fourth order stationary and absolutely regular with mixing coefficients $\beta(j)$ satisfying $\sum_{j=1}^{\infty} j^2 \beta(j)^{\frac{\delta}{1+\delta}} \leq \infty$ for some $\delta > 0$.*

Assumption 3. *$k(\cdot) \in \mathcal{K}_2$ where $\mathcal{K}_2$ is the set of kernels defined in Andrews (1991).*

Assumption 4. *$K(\cdot)$ is a bounded continuous density such that $\int_{-\infty}^{\infty} |uK(u)|du < \infty$.*

Assumption 5. *Let*

$$\begin{aligned} M_1 &= \max\{M_{11}, M_{12}\} \\ M_{11} &\geq \int |u_r u_s u_{s'} u_{r'}|^{1+\delta} dF(u_r, u_{r'}, u_s, u_{s'}) \\ M_{12} &\geq \int |u_r u_s u_{s'} u_{r'}|^{1+\delta} dF(u_r) dF(u_{r'}, u_s, u_{s'}), \end{aligned}$$

where $M_1 < \infty$.

Assumption 1 allows for the parameters to be time-varying. Typically, some sort of smoothness conditions are required in nonparametric estimation. In our case, a first order Lipshitz condition is included in the assumption.

Assumption 2 limits the amount of dependence in the data. The mixing concept of absolute regularity is common and includes ARMA processes under certain restrictions of the density of innovations. This type of mixing condition is employed in several nonparametric estimation papers including recent articles by Fan and Li (1999) and Li (1999).

Assumption 3 ensures that the long-run variance estimator produces positive semi-definite matrices. This class of kernel contains the quadratic spectral kernel presented in the last section. The second kernel $K(\cdot)$ is used for the nonparametric estimation of the mean of the series and satisfies the conditions in Assumption 4.

Assumption 5 are moment conditions. We list them in current form to highlight the weakest forms necessary. There are higher order moment conditions that are simpler to verify but we choose to list them as shown.

Recall that

$$\omega^2 = \lim_{T \to \infty} \frac{1}{T} \sum_{s=1}^{T} \sum_{t=1}^{T} E\left(u_s u_t^{\top}\right)$$

and let

$$\tilde{\omega}^2 = \sum_{j=-T+1}^{T-1} k\left(\frac{j}{b_T}\right)\tilde{\gamma}(j)$$

where

$$\tilde{\gamma}(j) = \begin{cases} \frac{1}{T}\sum_{t=j+1}^{T} \tilde{u}_t \tilde{u}_{t-j}^{\top} & \text{for} j \geq 0, \\ \frac{1}{T}\sum_{t=-j+1}^{T} \tilde{u}_{t+j} \tilde{u}_t^{\top} & \text{for} j < 0. \end{cases}$$

We present the main result in the following theorem.

Theorem 3.1. *Suppose that* $b_T \to \infty$*,* $h \to 0$*,* $b_T h^2 \to 0$*,* $T^{-1} b_T h^{-3/2} \to 0$*,* $T^{-1/2} b_T h^{-1/2} \to 0$*, and that Assumptions 1 - 5 hold. Then*

$$\tilde{\omega}_T^2 \xrightarrow{p} \omega^2$$

The result states that consistency depends on two different bandwidth parameters. The bandwidth h must go to zero but there are limitations as to the rate of convergence, and that b_T must go to infinity, but with restrictions on the speed of divergence.

In practice, the long-run variance estimate is sensitive to the choice of bandwidth parameter b_T. Therefore, Andrews (1991) suggests the use of the data dependent bandwidth, b_{opt} given earlier. This choice of bandwidth is widely used and is available as a standard option in many software packages. However, as noted by Crainiceanu and Vogelsang (2001), if there is a change in mean, b_{opt} diverges at a fast rate causing a nonmonotonic power problem. We propose a modified bandwidth using the nonparametric residuals $\tilde{u}_t$ so that ,

$$b_{opt}^* = 1.3221 \times (\tilde{\alpha}(2)T)^{1/5}$$

with

$$\tilde{\alpha}(2) = \frac{4\tilde{\rho}^2}{(1-\tilde{\rho})^4}.$$

where $\tilde{\rho}$ is the estimate AR(1) coefficient from regression $\tilde{u}_t$ on $\tilde{u}_{t-1}$. The asymptotic behavior of $\tilde{\alpha}(2)$ is given in the following theorem.

Theorem 3.2.

$$\tilde{\alpha}(2) \xrightarrow{p} \tilde{\alpha}(2) = \frac{4\rho^2}{(1-\rho)^4}$$

where ρ is the first (population) autocorrelation for u_t.

Since u_t is stationary, $|\rho| < 1$ which means that $\alpha(2)$ is bounded. The theorem shows that under a class of mean changes, the modified bandwidth will still diverge at rate $T^{1/5}$. Hence, the long-run variance estimator using $\tilde{u}_t$ and the bandwidth b^*_{opt} converges to the true long-run variance of u_t under the class of changes in mean governed by Assumption 1.

[Perhaps a discussion of unit root and data dependent bandwidth???]

4. Monte Carlo

We provide a small Monte Carlo experiment in this section. The test statistics of interest are the $CUSUM_T$ (Kolmogorov-Smirnoff), QS_T (Cramer von-Mises), and the optimal statistic proposed by Elliott and Müller, which we denote simply by EM_T. Each statistic is also estimated using the "robust" long-run variance estimator, $\tilde{\omega}^2$, which is calculated with the data-dependent bandwidth b^*_{opt}. The modified statistics are denoted $CUSUM^*_T$, QS^*_T, and EM^*_T respectively.

The modified statistics all require a bandwidth parameter h and a kernel $K(\cdot)$. For the kernel used in the nonparametric estimation of $\theta(t/T)$, we use the Epanechnikov kernel given by

$$K(x) = \begin{cases} \frac{3}{4}(1-x^2) & \text{if } |x| \leq 1 \\ 0 & \text{if } |x| > 1. \end{cases}$$

For comparison purposes, we use three choices of bandwidth, $h = cT^{-1/5}$, with $c = 1, 2, 3$.

First, size is examined by generating the following process

$$y_t = \theta + u_t$$

where $\theta = 1$ and $u_t = \rho u_{t-1} + \epsilon_t$ with ϵ_t iid $N(0,1)$. We consider sample sizes of $T = 100, 200, 300$ and rho takes the values $0.0, 0.5$, and 0.7. The nominal size is $\%5$, the number of replications is 2000, and the percentage of rejections appears in Table 1. There are several interesting points to note. First, each of the original statistics have reasonable size properties as none of the original tests are grossly oversized in any case. For 100 observations, the EM_T test is more undersized as ρ increases. As expected, size performance improves as T increases. The behavior of the modified tests depends on the choice of bandwidth h used in the nonparametric estimation of the mean, $\theta(t/T)$. For the bandwidth $h = 1 \times T^{-1/5}$, the modified tests are oversized, especially with larger values of ρ. However, for $h = 2 \times T^{-1/5}$, size is much better, especially for $T = 200$ and $T = 300$. Of course, the goal of the modified tests is not to improve the already reasonable size of the tests. We are trying to maintain good size and avoid power problems, the issue we address next.

Power comparisons are made using the following process.

$$y_t = \theta(t/T) + u_t$$

$$u_t = \rho u_{t-1} + \epsilon_t$$

where ϵ_t is again $N(0,1)$ iid. However, now we let $\theta(t/T)$ be governed by

$$\theta(t/T) = \begin{cases} 1 & \text{for } T = 1, \ldots, 0.5T \\ 1+\delta & \text{for } t = 0.5T+1, \ldots T. \end{cases}$$

We calculate size adjusted power using $h = 2 \times T^{-1/5}$ with $T = 200$ for $\rho = 0.5$, and 0.7. The resulting power graphs appear in Figures 1-6.

Note that the original tests all exhibit nonmonotonic power for this experiment. The effect is more pronounced as the level of serial correlation in u_t increases. However, as predicted from the asymptotic results in the lasts section, the modified tests alleviate the problem of nonmonotonic power. Moreover, the size-adjusted power of the new tests in the small δ range is never less than the size-adjusted power of the original tests.

5. Conclusion

Crainiceahu (2001), Juhl and Xiao (2004), and Vogelsang (1997, 1999) have shown that nonmonotonic power can be a serious problem when testing time series models

models for a change in mean. The discussion in Vogelsang (1999) gets to the heart of the matter; whenever there is a change in mean, the autoregressive parameters are biased towards one.

In this paper, we use nonparametric residuals to estimate a data-dependent bandwidth for the long-run variance parameter. By using nonparametric residuals, we are able to avoid the bias for the data-dependent bandwidth, and our long-run variance estimator does not diverge. This method avoids nonmonotonic power. Monte Carlo evidence indicates that power is very good using our procedure, and the modifications are able to control the power problems associated with long-run variance estimation.

Appendix A

Lemma A.1.

$$D = \frac{1}{T}\sum_{t=j+1}^{T} \tilde{u}_t \tilde{u}_{t-j} = O_p(T^{-1}h^{-3/2})$$

where

$$\tilde{u}_t = \frac{1}{Th}\sum_{s=1} u_s K_{ts}.$$

Proof:

$$E(D)^2 = E\left[\frac{1}{T}\sum_{t=j+1}^{T}\left(\frac{1}{Th}\sum_{s=1}^{T} u_s K_{ts}\right)\left(\frac{1}{Th}\sum_{r=1}^{T} u_r K_{t-j,r}\right)\right]^2$$

$$= \frac{1}{T^6 h^4}\sum_{t=j+1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T}\sum_{t'=j+1}^{T}\sum_{s'=1}^{T}\sum_{r'=1}^{T} E(u_s u_r u_{s'} u_{r'}) K_{ts} K_{t-j,r} K_{t's'} K_{t'-j,r'}$$

The dominant term occurs when $s \neq r \neq s' \neq r'$. Suppose that $s < s' < r < r'$ and denote this case 1. We consider the following subcases.

Case 1a: $s' - s \geq \max\{r - s', r' - r\}$

Case 1b: $r - s' \geq \max\{s' - s, r' - r\}$

Case 1c: $r' - r \geq \max\{s' - s, r - s'\}$

For case 1a,

$$
\begin{aligned}
E(D^2) &\leq \frac{1}{T^6h^4}\sum_{s=1}^{T-5}\sum_{s'=s+1}^{T-4}\sum_{r=s'+1}^{T-3}\sum_{r'=r+1}^{T-2}\sum_{t=r'+1}^{T-1}\sum_{t'=t+1}^{T} K_{ts}K_{t-j,r}K_{t's'}K_{t'-j,r'}|E(u_su_ru_{s'}u_{r'})| \\
&\leq \frac{1}{T^6h^4}\sum_{s=1}^{T-5}\sum_{s'=s+1}^{T-4}\sum_{r=s'+1}^{T-3}\sum_{r'=r+1}^{T-2}\sum_{t=r'+1}^{T-1}\sum_{t'=t+1}^{T} K_{ts}K_{t-j,r}K_{t's'}K_{t'-j,r'}M_2^{\frac{1}{1+\delta}}\beta(s'-s)^{\frac{\delta}{1+\delta}} \\
&\leq \frac{1}{T^6h^4}\sum_{s=1}^{T-5}\sum_{s'=s+1}^{T-4}(s'-s)M_2^{\frac{1}{1+\delta}}\beta(s'-s)^{\frac{\delta}{1+\delta}}\sum_{r=1}^{T-3}\sum_{r'=1}^{T-2}\sum_{t=1}^{T} K_{t-j,r}K(0)^3 \\
&= O\left(\frac{1}{T^2h^3}\right).
\end{aligned}
$$

where the first inequality comes from Lemma $A.5$.

For case 1b,

$$
\begin{aligned}
E(D^2) &\leq \frac{1}{T^6h^4}\sum_{s=1}^{T-5}\sum_{s'=s+1}^{T-4}\sum_{r=s'+1}^{T-3}\sum_{r'=r+1}^{T-2}\sum_{t=r'+1}^{T-1}\sum_{t'=t+1}^{T} K_{ts}K_{t-j,r}K_{t's'}K_{t'-j,r'}M_2^{\frac{1}{1+\delta}}\beta(r-s')^{\frac{\delta}{1+\delta}} \\
&\leq \frac{1}{T^6h^4}\sum_{s=1}^{T-5}\sum_{s'=s+1}^{T-4}\sum_{r=s'+1}^{T-3}(r-s')M_2^{\frac{1}{1+\delta}}\beta(r-s')^{\frac{\delta}{1+\delta}}\sum_{r'=1}^{T-2}\sum_{t'=1}^{T} K(0)^3 \\
&= O\left(\frac{1}{T^2h^3}\right).
\end{aligned}
$$

For case 1c,

$$
\begin{aligned}
E(D^2) &= \frac{1}{T^6h^4}\sum_{t'=6}^{T}\sum_{t=5}^{t'-1}\sum_{r'=4}^{t-1}\sum_{r=3}^{r'-1}\sum_{s'=2}^{r-1}\sum_{s=1}^{s'-1} K_{ts}K_{t-j,r}K_{t's'}K_{t'-j,r'}|E(u_su_ru_{s'}u_{r'})| \\
&\leq \frac{1}{T^6h^4}\sum_{t'=6}^{T}\sum_{t=5}^{t'-1}\sum_{r'=4}^{t-1}\sum_{r=3}^{r'-1}\sum_{s'=2}^{r-1}\sum_{s=1}^{s'-1} K_{ts}K_{t-j,r}K_{t's'}K_{t'-j,r'}M_2^{\frac{1}{1+\delta}}\beta(r'-r)^{\frac{\delta}{1+\delta}} \\
&\leq \frac{1}{T^6h^4}\sum_{t'=6}^{T}\sum_{t=5}^{t'-1}\sum_{r'=4}^{t-1}\sum_{r=3}^{r'-1}(r'-r)M_2^{\frac{1}{1+\delta}}\beta(r'-r)^{\frac{\delta}{1+\delta}}\sum_{s'=2}^{r-1} K_{t's'}K(0)^3 \\
&= O\left(\frac{1}{T^2h^3}\right).
\end{aligned}
$$

The same arguments apply for different orderings of $\{s, s', r, r'\}$.

Lemma A.2.

$$F = \frac{1}{T}\sum_{t=j+1}^{T} \tilde{u}_t \left[\frac{1}{Th}\sum_{s=1}^{T} \left(\theta(s/T) - \theta(t/T)\right) K_{ts} \right] = O_p(T^{-1/2}h^{1/2})$$

Proof:

$$E(F^2) = \frac{1}{T^2}\sum_{t=1}^{T} E(\tilde{u}_t^2) \left[\frac{1}{Th}\sum_{s=1}^{T}(\theta(t/T) - \theta(s/T))K_{ts} \right]^2$$
$$+ \frac{1}{T^2}\sum^{T}\sum_{t\neq t'}^{T} E(\tilde{u}_t\tilde{u}_{t'}) \left[\frac{1}{Th}\sum_{s=1}^{T}(\theta(t/T) - \theta(s/T))_{ts} \right] \left[\frac{1}{Th}\sum_{s'=1}^{T}(\theta(t'/T) - \theta(s'/T)).$$

We have

$$\begin{aligned}
E(\tilde{u}_t\tilde{u}_{t'}) &= \frac{1}{T^2h^2}\sum_{s=1}^{T}\sum_{s'=1}^{T} K_{ts}K_{t's'}E(u_s u_{s'}) \\
&\leq \frac{1}{T^2h^2}\sum_{s=1}^{T}\sum_{s'=1}^{T} K_{ts}K_{t's'}|E(u_s u_{s'})| \\
&\sim \frac{1}{T^2h^2}\sum_{s'=s+1}^{T} K_{t's'}\sum_{s=1}^{T-1} K(0)\beta(s'-s) \\
&= O(T^{-1}h^{-1}),
\end{aligned}$$

where the $\beta(s'-s)$ terms come from the application of Lemma A.5. Next,

$$\begin{aligned}
\frac{1}{Th}\sum_{s=1}^{T}(\theta(t/T) - \theta(s/T))K_{ts} &\sim \frac{1}{h}\int_0^1 (\theta(u) - \theta(v))K\left(\frac{u-v}{h}\right) dv \\
&\leq \frac{1}{h}\int_0^1 |u-v|K\left(\frac{u-v}{h}\right) dv \\
&= \frac{1}{h}\int_{-u/h}^{(1-u)/h} |wh|K\left(w\right) dw\ h \\
&= O(h).
\end{aligned}$$

These results imply that $E(F^2) = O(T^{-1}h)$ so that $F = O_p(T^{-1/2}h^{1/2})$.

Lemma A.3.

$$G = \frac{1}{T}\sum_{t=1}^{T} u_t \left[\frac{1}{Th}\sum_{s=1}^{T} \left(\theta(s/T) - \theta(t/T)\right) K_{ts} \right] = O_p(T^{-1/2}h)$$

Proof:

$$E(G^2) = \frac{1}{T^2}\sum_{t=1}^{T} E(u_t^2)\left[\frac{1}{Th}\sum_{s=1}^{T}(\theta(t/T)-\theta(s/T))K_{ts}\right]^2$$
$$+\frac{1}{T^2}\sum_{t\neq t'}^{T}\sum^{T} E(u_t u_{t'})\left[\frac{1}{Th}\sum_{s=1}^{T}(\theta(t/T)-\theta(s/T))K_{ts}\right]\left[\frac{1}{Th}\sum_{s'=1}^{T}(\theta(t'/T)-\theta(s'/T\right.$$
$$= G_1 + G_2$$

Since $\frac{1}{Th}\sum_{s=1}^{T}(\theta(t/T)-\theta(s/T))K_{ts} = O(h)$, $G_1 = O(T^{-1}h^2)$. Then

$$\begin{aligned} G_2 &\leq \frac{1}{T^2}\sum_{t\neq t'}^{T}\sum^{T} |E(u_t u_{t'})| O(h^2) \\ &\leq \frac{1}{T^2}\sum_{t\neq t'}^{T}\sum^{T} \beta(t-t')O(h^2) \\ &= O(T^{-1}h^2), \end{aligned}$$

making $E(G^2) = O(T^{-1}h^2)$ and $G = O_p(T^{-1/2}h)$.

Lemma A.4.

$$H = \frac{1}{T}\sum_{t=j+1}^{T} u_t \tilde{u}_{t-j} = O_p(T^{-1/2}h^{-1/2})$$

Proof:

$$\begin{aligned} E(H^2) &= \frac{1}{T^2}\sum_{t=1}^{T} E(u_t^2\tilde{u}_{t-j}^2) + \frac{1}{T^2}\sum_{t\neq t'}^{T}\sum^{T} u_t u_{t'}\tilde{u}_{t-j}\tilde{u}_{t'-j} \\ &= H_1 + H_2 \end{aligned}$$

Then

$$H_1 = \frac{1}{T^3h^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{s'=1}^{T} K_{ts}K_{ts'}E(u_s u_{s'} u_t^2)$$

which is $O(T^{-1}h^{-1})$ and

$$H_2 = \frac{1}{T^4h^2}\sum_{t\neq t'}^{T}\sum^{T}\sum_{s=1}^{T}\sum_{s'=1}^{T} K_{ts}K_{t's'}E(u_t u_{t'} u_s u_{s'})$$

which is $O(T^{-2}h^{-1})$. H_1 is the dominant term making $H = O_p(T^{-1/2}h^{-1/2})$.

Lemma A.5. *(Lemma 1, Yoshihara (1976)) Let* $x_{t_1}, x_{t_2}, \ldots, x_{t_k}$ *(with* $t_1 < t_2 < \cdots < t_k$*) be absolutely regular random vectors with mixing coefficients* β*. Let* $h(x_{t_1}, x_{t_2}, \ldots, x_{t_k})$ *be a Borel measurable function and let there be a* $\delta > 0$ *such that*

$$P = \max\{P_1, P_2\} < \infty$$

where

$$P_1 = \int |h(x_{t_1}, x_{t_2}, \ldots, x_{t_k})|^{1+\delta} \, dF(x_{t_1}, x_{t_2}, \ldots, x_{t_k})$$
$$P_2 = \int |h(x_{t_1}, x_{t_2}, \ldots, x_{t_k})|^{1+\delta} \, dF(x_{t_1}, \ldots, x_{t_j}) dF(x_{t_{j+1}}, \ldots, x_{t_k}).$$

Then

$$\Big| \int h(x_{t_1}, x_{t_2}, \ldots, x_{t_k}) dF(x_{t_1}, x_{t_2}, \ldots, x_{t_k}) - h(x_{t_1}, x_{t_2}, \ldots, x_{t_k}) dF(x_{t_1}, \ldots, x_{t_j}) dF(x_{t_{j+1}}, \ldots, x_{t_k}) \Big| \leq 4P^{\frac{1}{1+\delta}} \beta_\tau^{\frac{\delta}{1+\delta}}$$

for all $\tau = t_{j+1} - t_j$*.*

Proof of Theorem 3.1: Let

$$\check{\gamma}(j) = \frac{1}{T} \sum_{t=j+1}^{T} u_t u_{t-j}.$$

From Parzen (1957) and Andrews (1991), we know that

$$\sum_{j=-T+1}^{T+1} k\left(\frac{j}{b_T}\right) \check{\gamma}(j) \xrightarrow{p} \omega^2.$$

The proof is completed by showing that

$$\sum_{j=-T+1}^{T-1} k\left(\frac{j}{b_T}\right) (\tilde{\gamma}(j) - \check{\gamma}(j)) = o_p(1).$$

Note that

$$\tilde{\gamma}(j) = \frac{1}{T} \sum_{t=j+1}^{T} \tilde{v}_t \tilde{u}_{t-j},$$

where

$$\begin{aligned}
\tilde{u}_t &= y_t - \tilde{y}_t \\
&= \theta\left(\frac{t}{T}\right) + u_t - \left(\frac{1}{Th}\sum_{s=1}^{T} K_{ts}\right)^{-1}\left(\frac{1}{Th}\sum_{s=1}^{T}\left[\theta\left(\frac{s}{T}\right) + u_s\right]\right) \\
&= u_t - \left(\frac{1}{Th}\sum_{s=1}^{T} K_{ts}\right)^{-1}\left(\frac{1}{Th}\sum_{s=1}^{T}\left[\theta\left(\frac{s}{T}\right) - \theta\left(\frac{t}{T}\right)\right] K_{ts}\right) \\
&\quad + \left(\frac{1}{Th}\sum_{s=1}^{T} K_{ts}\right)^{-1}\left(\frac{1}{Th}\sum_{s=1}^{T} u_s K_{ts}\right) \\
&= u_t + J_{1t} + J_{2t}
\end{aligned}$$

We can write

$$\begin{aligned}
\tilde{\gamma}(j) = {} & \frac{1}{T}\sum_{t=j+1}^{T} u_t u_{t-j} + \frac{1}{T}\sum_{t=j+1}^{T} u_t J_{1,t-j} + \frac{1}{T}\sum_{t=j+1}^{T} u_t J_{2,t-j} \\
& + \frac{1}{T}\sum_{t=j+1}^{T} J_{1t} u_{t-j} + \frac{1}{T}\sum_{t=j+1}^{T} J_{1t} J_{2,t-j} + \frac{1}{T}\sum_{t=j+1}^{T} J_{1t} J_{1,t-j} \\
& + \frac{1}{T}\sum_{t=j+1}^{T} J_{2t} u_{t-j} + \frac{1}{T}\sum_{t=j+1}^{T} J_{2t} J_{1,t-j} + \frac{1}{T}\sum_{t=j+1}^{T} J_{2t} J_{2,t-j}
\end{aligned}$$

The terms $\frac{1}{T}\sum_{t=j+1}^{T} u_t J_{1,t-j}$ and $\frac{1}{T}\sum_{t=j+1}^{T} J_{1t} u_{t-j}$ are $O_p(T^{-1/2}h)$ by Lemma A.3, $\frac{1}{T}\sum_{t=j+1}^{T} u_t J_{2,t-j}$ and $\frac{1}{T}\sum_{t=j+1}^{T} J_{2t} u_{t-j}$ are $O_p(T^{-1/2}h^{-1/2})$ by Lemma A.4, $\frac{1}{T}\sum_{t=j+1}^{T} J_{1t}J_{2,}$ and $\frac{1}{T}\sum_{t=j+1}^{T} J_{2t}J_{1,t-j}$ are $O_p(T^{-1/2}h^{1/2})$ by Lemma A.2, and $\frac{1}{T}\sum_{t=j+1}^{T} J_{2t}J_{2,t-j}$ is $O_p(T^{-1}h^{-3/2})$ by Lemma A.1. $\frac{1}{T}\sum_{t=j+1}^{T} J_{1t}J_{1,t-j}$ is $O(h^2)$.

Now

$$\frac{1}{b_T}\sum_{j=-T+1}^{T-1}\left|k\left(\frac{j}{b_T}\right)\right| \to \int_{-\infty}^{\infty}|k(x)|dx$$

which implies that

$$\sum_{j=-T+1}^{T-1} k\left(\frac{j}{b_T}\right)\left(\tilde{\gamma}(j) - \check{\gamma}(j)\right) = O_p(b_T h^2 + T^{-1}b_T h^{-3/2} + T^{-1/2}b_T h^{-1/2}) = o_p(1).$$

Proof of Theorem 3.2:

$$\begin{aligned}\tilde{\rho} &= \frac{\tilde{\gamma}(1)}{\tilde{\gamma}(0)} \\ &= \frac{\gamma(1)}{\gamma(0)} + o_p(1)\end{aligned}$$

from the proof of Theorem 3.1.

References

Andrews, D. W. K. (1991): "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817–858.

Andrews, D. W. K. (1993): "Tests for Parameter Instability and Structural Change With Unknown Change Point," *Econometrica*, 61, 821–856.

Andrews, D. W. K., and W. Ploberger (1994): "Optimal Tests When a Nuisance Parameter is Present Only Under the Alternative," *Econometrica*, 62, 1383–1414.

Brown, R. L., J. Durbin, and J. M. Evans (1975): "Techniques for Testing the Constancy of Regression Relationships over Time," *Journal of the Royal Statstical Society, Series B*, 37, 149–172.

Crainiceanu, C. M., and T. J. Vogelsang (2001): "Spectral Density Bandwidth Choice: Source of Nonmonotonic Power for Tests of a Mean Shift in a Time Series," Cornell University.

Elliott, G., and U. K. Muller (2003): "Optimally Testing General Breaking Processes in Linear Time Series Models," UC San Diego.

Fan, Y., and Q. Li (1999): "Central Limit Theorem for Degenerate U-Statistics of Absolutely Regular Processes with Applications to Model Specification Testing," *Journal of Nonparametric Statistics*, 10, 245–271.

Franzini, L., and A. Harvey (1983): "Testing for Deterministic Trend and Seasonal Components in Times Series," *Biometrika*, 70, 673–682.

Gardner, L. A. (1969): "On Detecting Changes in the Mean of Normal Variates," *Annals of Mathematical Statistics*, 10, 116–126.

Hansen, B. E. (2000): "Testing for Structural Change in Conditional Models," *Journal of Econometrics*, 97, 93–115.

Juhl, T., and Z. Xiao (2004): "A Nonparametric Test for Changing Trends," *Journal of Econometrics*, forthcoming.

Li, Q. (1999): "Consistent Model Specification Tests for Time Series Econometric Models," *Journal of Econometrics*, 92, 101–148.

MacNeill, I. B. (1978): "Properties of Sequences of Partial Sums of Polynomial Regression Residuals with Applications to Tests for Change of Regression at Unknown Times," *The Annals of Statistics*, 6, 1422–433.

Perron, P. (1990): "Testing for a Unit Root in a Time Series with a Changing Mean," *Journal of Business and Economic Statistics*, 8, 153–162.

——— (1991): "A Test for Changes in a Polynomial Trend Function for a Dynamic Time Series," Princeton University Department of Economics.

Ploberger, W., and W. Kramer (1992): "The CUSUM Test with OLS Residuals," *Econometrica*, 60, 271–286.

Vogelsang, T. J. (1997): "Wald-Type Tests for Detecting Breaks in the Trend Function of a Dynamic Time Series," *Econometric Theory*, 13, 818–849.

——— (1999): "Sources of Nonmonotonic Power When Testing for Shift in Mean of a Dynamic Time Series," *Journal of Econometrics*, 88, 283–299.

Table 1.
Size

c	T	ρ	CUSUM_T	QS_T	EM_T	CUSUM_T^*	QS_T^*	EM_T^*
	100	0.0	0.029	0.050	0.038	0.050	0.063	0.071
	100	0.5	0.014	0.052	0.013	0.092	0.121	0.139
	100	0.7	0.005	0.045	0.002	0.153	0.187	0.214
	200	0.0	0.039	0.055	0.042	0.048	0.062	0.059
1	200	0.5	0.040	0.066	0.036	0.094	0.110	0.117
	200	0.7	0.022	0.069	0.017	0.123	0.156	0.167
	300	0.0	0.036	0.051	0.045	0.046	0.059	0.053
	300	0.5	0.042	0.065	0.052	0.080	0.097	0.112
	300	0.7	0.030	0.055	0.027	0.097	0.115	0.148
	100	0.0	0.029	0.050	0.038	0.037	0.057	0.051
	100	0.5	0.014	0.052	0.013	0.051	0.087	0.053
	100	0.7	0.005	0.045	0.002	0.055	0.105	0.054
	200	0.0	0.039	0.055	0.042	0.044	0.060	0.047
2	200	0.5	0.040	0.066	0.036	0.066	0.087	0.065
	200	0.7	0.022	0.069	0.017	0.074	0.109	0.071
	300	0.0	0.036	0.051	0.045	0.040	0.055	0.047
	300	0.5	0.042	0.065	0.052	0.061	0.081	0.075
	300	0.7	0.030	0.055	0.027	0.058	0.080	0.073
	100	0.0	0.029	0.050	0.038	0.031	0.053	0.045
	100	0.5	0.014	0.052	0.013	0.029	0.067	0.031
	100	0.7	0.005	0.045	0.002	0.023	0.077	0.015
	200	0.0	0.039	0.055	0.042	0.041	0.058	0.044
3	200	0.5	0.040	0.066	0.036	0.049	0.077	0.049
	200	0.7	0.022	0.069	0.017	0.050	0.086	0.040
	300	0.0	0.036	0.051	0.045	0.039	0.052	0.045
	300	0.5	0.042	0.065	0.052	0.052	0.073	0.064
	300	0.7	0.030	0.055	0.027	0.043	0.071	0.044

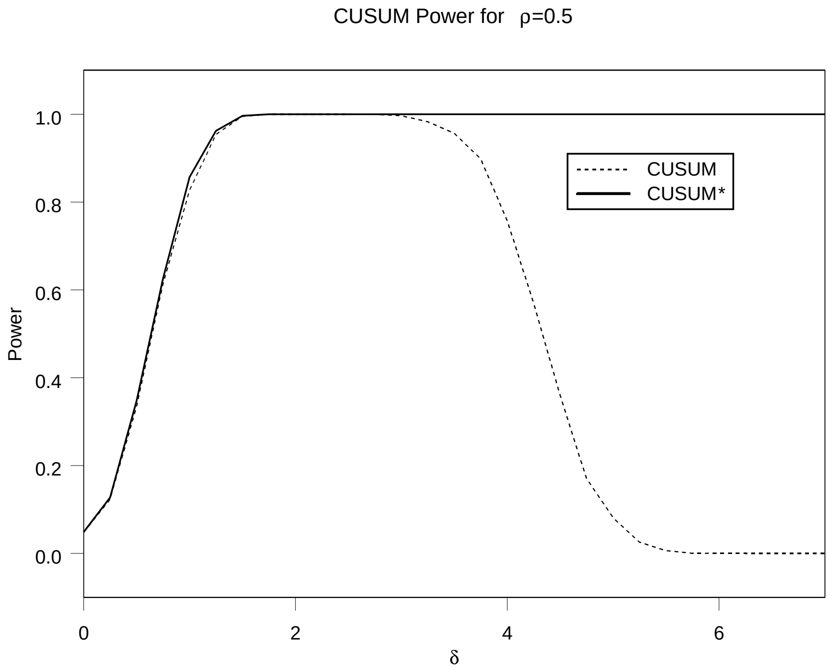

FIGURE 1

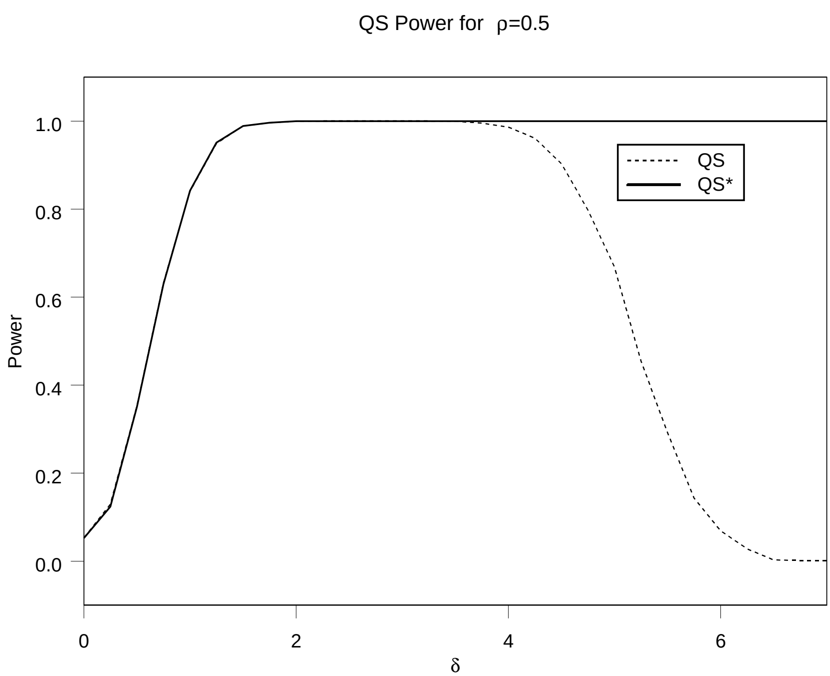

FIGURE 2

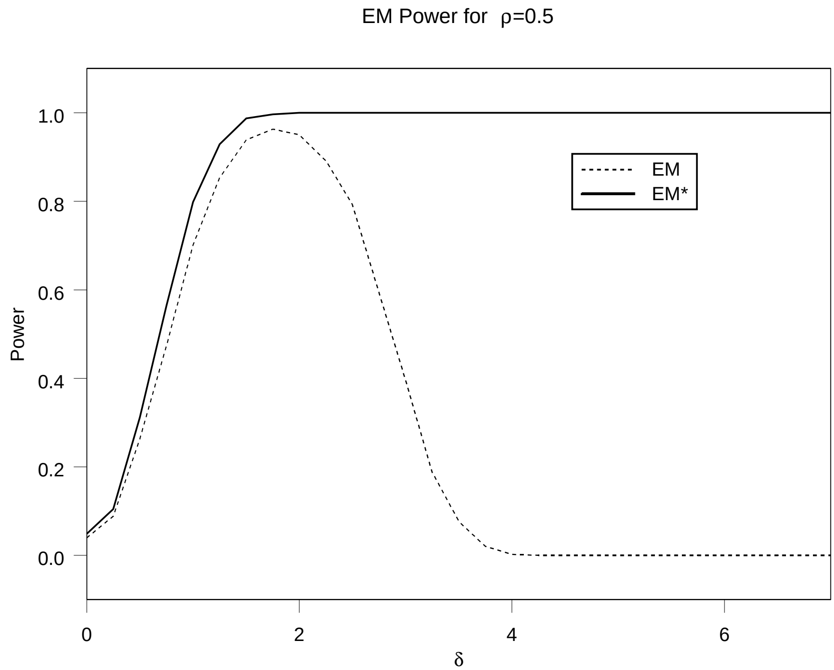

FIGURE 3

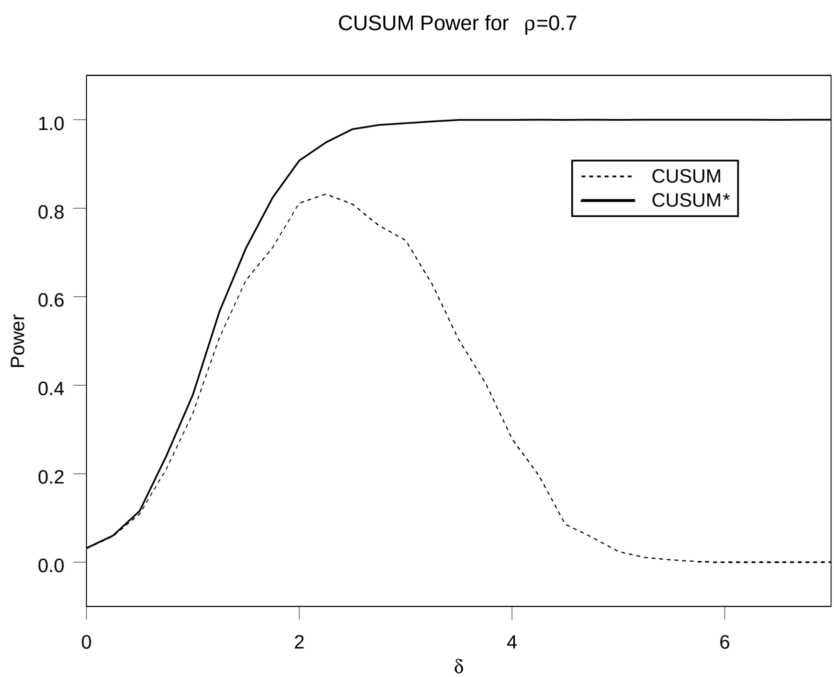

FIGURE 4

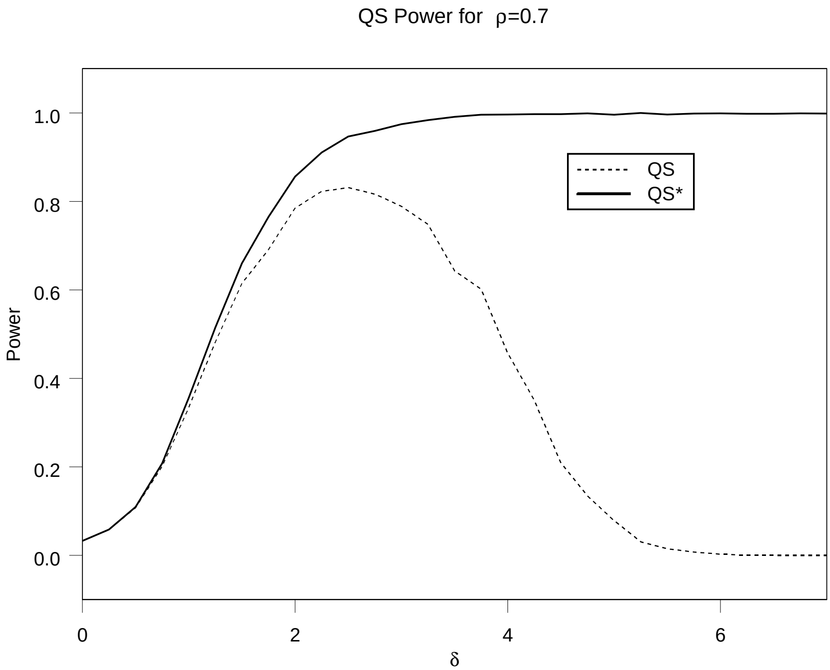

FIGURE 5

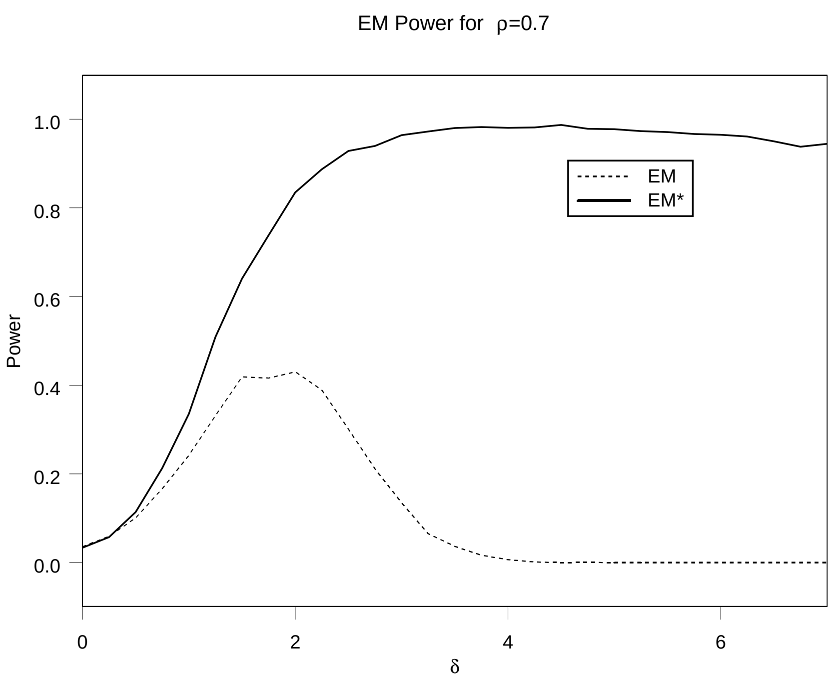

FIGURE 6