CAE Working Paper #01-14

Spectral Density Bandwith Choice: Source of Nonmonotonic Power for Tests of a Mean Shift in a Time Series

by

Ciprian Crainiceanu
and
Timothy Vogelsang

February 2001

.

Spectral Density Bandwidth Choice: Source of Nonmonotonic Power for Tests of a Mean Shift in a Time Series

Ciprian M. Crainiceanu* Timothy J. Vogelsang†

February, 2001

Abstract

Data dependent bandwidth choices for zero frequency spectral density estimators of a time series are shown to be an important source of nonmonotonic power when testing for a shift in mean. It is shown that if the spectral density is estimated under the null hypothesis of a stable mean using a data dependent bandwidth (with or without prewhitening), non-monotonic power appears naturally for some popular tests including the CUSUM test. On the other hand, under some fixed bandwidth choices, power is monotonic. Empirical examples and simulations illustrate these power properties. Theoretical explanations for the power results are provided.

Keywords: Spectral Density Estimator, HAC Estimator, Serial Correlation, CUSUM Test, Change Point

*Department of Statistics, Cornell University, Malott Hall, NY 14853-4201, USA. E-mail: cmc59@cornell.edu

†Corresponding Author: Departments of Economics and Statistics, Cornell University, Uris Hall, NY 14853-7601, USA. E-mail:tjv2@cornell.edu

1 Introduction

In general, a statistical analysis of time series data requires some stationarity assumptions about the underlying stochastic process. Testing these assumptions is important for assessing the quality of an estimation procedure or the goodness-of-fit of the model used. Testing for the structural stability of the mean of a time series is one example. A reasonable property for any shift in mean test to satisfy is that for a given sample size, the larger the shift in mean the higher the probability of detecting the shift. In other words, one would like to see a finite sample power function that is monotonically increasing in the shift magnitude. Recent research by Perron (1991) and Vogelsang (1999) has shown that some well known mean shift tests, configured to allow for serial correlation in the data, can have nonmonotonic power functions. And, in some cases, for large shifts in mean power can drop to zero.

To illustrate these potential power problems consider the three time series plotted in Figure 1. These series are logarithms of monthly high bond prices for bonds issued by the Argentinean, Brazilian and Chilean governments from January 1927 to December 1936. Each series spans 10 years and has 120 observations. The data was collected from various issues of the Commercial and Financial Chronicle. All three time series exhibit obvious mean shifts at about the middle of the interval considered. Therefore any reasonable test for shift in mean should easily reject the null of a stable mean. Table 1 reports the values for two mean shift tests: the well known CUSUM test and a related test we label QS. The formal definitions of the tests are given in Section 3. Below each statistic are p-values based on asymptotic critical values under the null of a stable mean. Because prices of assets like bond prices are well known to be serially correlated over time, the CUSUM and QS statistics require the computation of a consistent estimator of the variance of the sample mean of each series. A heteroskedasticity and autocorrelation consistent (HAC) estimator of this variance was used. These estimators are equivalent to zero frequency spectral density estimators. HAC estimators require the choice of a truncation lag, or bandwidth. Three different choices for the bandwidth were used: i) fixed bandwidth (FB) which is dependent only on the sample size, ii) data dependent bandwidth (DDB) and iii) data dependent bandwidth with prewhitening (DDB-PW). See the note to Table 1 for details.

Table 1: **Empirical Example: Tests of the Null of a Stable Mean**

	Argentina	Brazil	Chile
CUSUM (FB)	1.42 (0.033)	2.37 (0.00003)	2.41 (0.00001)
QS (FB)	0.57 (0.026)	1.95 (0.00001)	1.96 (0.00001)
CUSUM (DDB)	0.75 (0.603)	0.99 (0.266)	1.05 (0.207)
QS (DDB)	0.16 (0.352)	0.34 (0.105)	0.37 (0.087)
CUSUM (DDB-PW)	0.44 (0.987)	0.61 (0.832)	0.53 (0.930)
QS (DDB-PW)	0.05 (0.876)	0.13 (0.453)	0.10 (0.583)

Notes: Tests are right tail tests of the null hypothesis of a stable mean. Values in parentheses are p-values based on asymptotic null distributions. The CUSUM and QS statistics are defined in Section 3 by (3) and (4) respectively. The HAC estimator is given by (2). The formulas for the bandwidths are given in Section 4: FB is given by (6), DDB is given by (7)-(9), and DDB-PW is given by (11)-(13).

Some surprising and interesting patterns appear in Table 1. When a fixed bandwidth is used, the null hypothesis of a stable mean is easily rejected for Brazil and Chile using both statistics. The null can be rejected at the 5% level for Argentina. On the other hand, when a data dependent bandwidth is used, with or without prewhitening, the null hypothesis is not rejected in nearly all cases. Only when the QS statistic is used, but without prewhitening, can the null be rejected at the 10% level for Brazil and Chile. This empirical example suggests that power of the CUSUM and QS tests is very sensitive to the choice of bandwidth even when there is a large and obvious mean shift in the data. The fact that the mean shifts are detected with a fixed bandwidth but not with a data dependent bandwith is an unfortunate situation because data dependent bandwidths are usually recommended over fixed bandwidths in practice.

A number of interesting questions arise from this example:

1. Are the power properties suggested by the example general properties of the CUSUM and QS tests?

2. What is the source of low power when data dependent bandwidths are used?

3. Can differences in power with respect to bandwidth choice be explained theoretically?

In this paper we provide answers to these questions. We show that for the CUSUM and QS tests, fixed bandwidths can lead to nice power functions whereas data dependent bandwidths can lead to nonmonotonic power. In fact, power often drops to zero for large mean shifts under data dependent bandwidths. We show that, provided a HAC estimator is used, these power patterns occur whether serial correlation in the time series is strong or weak. Power is good under a FB provided the bandwidth is relatively small whereas low power under a DDB occurs because of the tendency of data dependent bandwidths to choose very large bandwidths when the mean shift is big. In constrast, when a DDB-PW is used, low power results because of the bad behavior of the prewhitened estimator itself in conjuction with the tendency of a DDB-PW to choose small bandwidths. As the theoretical analysis shows, a key factor in explaining the power properties is that the HAC estimator is constructed under the null hypothesis of a stable mean and is thus constructed under a misspecified model when there is a mean shift.

The rest of the paper is organized as follows. In section 2 the null model and the basic assumptions are introduced. In section 3 the statistics for testing for a shift in mean are defined. In section 4 we provide finite sample evidence on power using simulations. In section 5 theoretical results are provided that explain the patterns of the power function with respect to bandwidth. The theoretical and practical implications of these results are assessed. In section 6 the conclusions of the paper are presented. Proofs are given in a mathematical appendix.

2 Data Generating Process and Assumptions

We consider the following data generating process (DGP) for a univariate time series y_t:

$$y_t = \mu + u_t, \tag{1}$$

where u_t is a stationary mean zero error process. Define the partial sums of u_t as $S_t = \sum_{j=1}^{t} u_j$. Let $W(r)$ denote the standard Wiener process defined on [0,1], let $\Rightarrow$ denote weak convergence and let $[x]$ denote the integer part of x.

Assumption 1 *u_t is second order stationary with $\sigma^2 = \lim_{T\to\infty} E[\frac{1}{T}\left(\sum_{t=1}^{T} u_t\right)^2]$ and $T^{-\frac{1}{2}} S_{[rT]} \Rightarrow \sigma W(r)$.*

Note that under stationarity, σ^2 is proportional to the spectral density of u_t at frequency zero and is given by

$$\sigma^2 = \sum_{j=-\infty}^{\infty} \gamma_j,$$

where $\gamma_j = cov(u_t, u_{t-j})$ is the autocovariance function of u_t. There are a variety of regularity conditions under which *Assumption 1* (a functional limit theorem) holds. See, for example, Herrndorf (1984) and Phillips (1987).

The parameter σ^2 needs to be estimated in order to test the null hypothesis that the mean of y_t is stable. Typical estimators of σ^2 include the class of non-parametric spectral density estimators given by

$$\widehat{\sigma}^2 = \sum_{j=-(T-1)}^{T-1} K\left(\frac{j}{s(T)}\right)\widehat{\gamma}_j, \tag{2}$$

where $K(.)$ is a kernel function, $\widehat{\gamma}(j) = T^{-1}\sum_{t=j+1}^{T} \widehat{u}_t\widehat{u}_{t-j}$ and $\widehat{u}_t = y_t - \bar{y}$ are the OLS residuals from (1), and $s(T)$ is the bandwidth or the truncation lag. In order to ensure that $\widehat{\sigma}^2$ is a consistent estimator of σ^2, basic requirements are $s(T)/T \to 0$ and $s(T) \to \infty$ as $T \to \infty$. See Priestley (1981) for details.

3 Statistics Used for Testing for a Shift in Mean

We focus on two statistics for testing the null hypothesis that the mean of y_t is constant. Both statistics were originally motivated in part by a search for tests that have power for detecting fairly general forms of mean shifts without having to estimate a particular alternative model. These tests are similar to Lagrange multiplier tests in that the model only needs to be estimated under the null hypothesis that the mean is stable.

The first statistic is the well known CUSUM statistic:

$$CUSUM = \frac{1}{\widehat{\sigma}} \sup_{1\leq j\leq T} \left| \frac{1}{\sqrt{T}} \sum_{t=1}^{j} \widehat{u}_t \right|. \tag{3}$$

This version of the CUSUM test was proposed and analyzed by Ploberger and Kramer (1992). The CUSUM test was originally proposed by Brown, Durbin and Evans (1975) and was constructed using recursive residuals (rather than OLS residuals as in (3)). The second statistic, which is similar in spirit to the CUSUM statistic, was proposed by Gardner (1969), extended by MacNeill (1978), further extended by Perron (1991) and is defined as:

$$QS = \frac{1}{T^2\widehat{\sigma}^2} \sum_{j=1}^{T} \left(\sum_{t=1}^{j} \widehat{u}_t \right)^2 . \tag{4}$$

If $\widehat{\sigma}^2$ is a consistent estimator for σ^2, then under model (1), the null hypothesis of a stable mean,

$$CUSUM \Rightarrow \sup_{r\in[0,1]} |B(r)| , \qquad QS \Rightarrow \int_0^1 B^2(r)dr,$$

where $B(r) = W(r) - rW(1)$ is the standard Brownian bridge on [0,1]. Both tests are right tail tests and asymptotic critical values can be found in Ploberger and Kramer (1992) for CUSUM and MacNeill (1978) for QS. We reproduce the critical values in Table 2 .

Table 2: **Asymptotic critical values for CUSUM and QS**

Statistic	**Quantiles**		
	0.9	0.95	0.99
$CUSUM$	1.22	1.36	1.63
QS	0.35	0.46	0.74

4 Simulations

In this section we examine finite sample power of the CUSUM and QS tests for detecting a single shift in mean at an unknown date. Let T_b denote the date of a mean shift, and define the dummy variable $DU_t = 1_{(t>T_b)}$ where $1_{t>T_b} = 1$ if $t > T_b$ and 0 otherwise. Consider the DGP:

$$y_t = \mu + \delta DU_t + u_t, \tag{5}$$

where $u_t = \rho u_{t-1} + \epsilon_t$, ϵ_t are i.i.d. distributed $N(0, 1)$ and δ is the magnitude of the shift.

As we shall soon illustrate, the power of the CUSUM and QS tests for detecting the alternative model (5) depends critically on how the bandwidth, $s(T)$, is chosen when constructing $\widehat{\sigma}^2$. Choices for $s(T)$ that minimize the mean square error (MSE) of $\widehat{\sigma}^2$ depend on the kernel $K(x)$. For concreteness, we focus on the Bartlett kernel defined as $K(x) = 1 - |x|$ for $|x| \leq 1$ and $K(x) = 0$ for $|x| > 1$. Similar results are obtained for other popular kernels.

When the Bartlett kernel is used, $s(T)$ needs to increase at rate $T^{1/3}$ for the MSE of $\widehat{\sigma}^2$ to be minimized. As is well known in the spectral density and HAC literature, this rate result essentially places no restriction on the choice of $s(T)$ for a given sample because any choice of $s(T)$ can be justified as satisfying the rate $T^{1/3}$ by setting $s(T) = cT^{1/3}$ and suitably choosing the constant c. For the sake of illustration we take $c = 1$ and denote this choice of $s(T)$ by

$$s_0(T) = T^{1/3}. \tag{6}$$

We refer to this choice of $s(T)$ as the "fixed bandwidth".

Because of the arbitrary nature of the choice of $s(T)$, there has emerged a literature on data dependent choices of the bandwidth. Here we follow Andrews (1991) and consider the AR(1) plug-in choice of $s(T)$:

$$\widehat{s}(T) = 1.1447(\widehat{\alpha}(1)T)^{1/3}, \tag{7}$$

$$\widehat{\alpha}(1) = \frac{4\widehat{\rho}^2}{(1-\widehat{\rho}^2)^2}, \tag{8}$$

$$\widehat{\rho} = \frac{\sum_{t=2}^{T} \widehat{u}_t \widehat{u}_{t-1}}{\sum_{t=2}^{T} \widehat{u}_{t-1}^2}. \tag{9}$$

Andrews (1991) showed that $\widehat{s}(T)$ given by (7) minimizes the approximate MSE of $\widehat{\sigma}^2$ given the AR(1) structure of u_t.

Andrews and Monahan (1992) showed that prewhitening can improve the performance of $\widehat{\sigma}^2$. Therefore we also consider AR(1) prewhitened estimation of $\widehat{\sigma}^2$. Let $\widehat{\sigma}_{PW}^2$ denote the prewhitened estimator:

$$\widehat{\sigma}_{PW}^2 = \frac{\widehat{\sigma}_\epsilon^2}{(1-\widehat{\rho})^2}, \tag{10}$$

where $\widehat{\rho}$ is given by equation (9) and

$$\widehat{\sigma}_\epsilon^2 = \sum_{j=-(T-1)}^{(T-1)} K\left(j/\widehat{s}_{PW}(T)\right) \widehat{\gamma}_j^\epsilon,$$

where $\widehat{\gamma}_j^{\epsilon} = T^{-1} \sum_{t=j+1}^{T} \widehat{\epsilon}_t \widehat{\epsilon}_{t-j}$ and $\widehat{\epsilon}_t = \widehat{u}_t - \widehat{\rho} \widehat{u}_{t-1}$. As in the case without prewhitening, $\widehat{s}_{PW}(T)$ is chosen using the AR(1) plug-in method except that $\widehat{\alpha}(1)$ is calculated using $\widehat{\epsilon}_t$ rather than $\widehat{u}_t$:

$$\widehat{s}_{PW}(T) = 1.1447(\widehat{\alpha}_{PW}(1)T)^{1/3}, \tag{11}$$

$$\widehat{\alpha}_{PW}(1) = \frac{4\widehat{\rho}_{\epsilon}^2}{(1-\widehat{\rho}_{\epsilon}^2)^2}, \tag{12}$$

$$\widehat{\rho}_{\epsilon} = \frac{\sum_{t=2}^{T} \widehat{\epsilon}_t \widehat{\epsilon}_{t-1}}{\sum_{t=2}^{T} \widehat{\epsilon}_{t-1}^2}. \tag{13}$$

For the simulations we generated data according to (3) for $T = 50$, 100, 200, 500, $\rho = 0.1$, 0.4, 0.7, 0.9 and $\delta = 0$, 1, 2, ..., 20. In all cases 2000 replications were used. We calculated two types of finite sample power functions. The first is power using empirical null 5% critical values taken from the $\delta = 0$ simulations. In this case, empirical null rejection probabilities are exactly 0.05 for both CUSUM and QS. We refer to this case as size corrected power. The second is power using the asymptotic 5% critical values (the 95% percentiles from Table 2). In this case empirical null rejection probabilities will differ from 0.05. We refer to this as uncorrected power. Uncorrected power is relevant in practice since asymptotic critical values will typically be used. Size corrected power, while not feasible in practice, allows theoretical power comparisons while holding size equivalent across tests.

The finite sample power results are given in Figures 2-9. We only report size adjusted power as patterns are similar with unadjusted power. We only report results for $T = 100$. Similar results were obtained for other values of T and are available upon request. Regardless of the value of ρ, a clear pattern emerges from the figures. When the fixed bandwidth is used, power is monotonically increasing in δ and reaches 1 for large mean shifts. On the other hand, with the data dependent bandwidth, power is nonmonotonic in δ and drops to zero for large δ. Prewhitening does not improve matters. These patterns hold for all the values of ρ.

Clearly, the choice of $s(T)$ matters. But why do DDB choices for $s(T)$ generate nonmonotonic power whereas the FB choice does not? The first step for answering this question is to examine the behavior of $\widehat{s}(T)$ as δ increases. In Table 3 we report averages of $\widehat{s}(T)$ across the 2000 replications for the power simulations for $\rho = 0.7$ and $T = 100$. As δ increases, $\widehat{s}(T)$ grows using a DDB without prewhitening. This contrasts with the FB case where $s_0(T)$ remains fixed at 5 for all values of δ.

Clearly, it is large values of $\widehat{s}(T)$ that is crippling power for large δ. But this is not the story for DDB-PW because $\widehat{s}_{PW}(T)$ is small and roughly constant for all values of δ. Whereas large $\widehat{s}(T)$ hurts power when there is no prewhitening, small $\widehat{s}_{PW}(T)$ hurts power when prewhitening is used. It is not obvious why this relationship between power and $s(T)$ should occur, and we now turn to a theoretical analysis for further insight.

Table 3: **Behavior of $s(T)$ in finite samples for T $= 100$ and $\phi = 0.7$.**

δ	$s_0(T)$	$\widehat{s}(T)$	$\widehat{s}_{PW}(T)$
0	5.000	9.900	1.143
1	5.000	10.958	1.172
2	5.000	13.748	1.304
3	5.000	17.598	1.473
4	5.000	21.948	1.595
5	5.000	26.418	1.652
6	5.000	30.767	1.661
7	5.000	34.596	1.636
8	5.000	38.596	1.597
9	5.000	41.975	1.548
10	5.000	44.988	1.496
11	5.000	47.657	1.443
12	5.000	50.011	1.389
13	5.000	52.082	1.338
14	5.000	53.904	1.290
15	5.000	55.506	1.245
16	5.000	56.918	1.204
17	5.000	58.165	1.164
18	5.000	59.267	1.129
19	5.000	60.246	1.095
20	5.000	61.116	1.063

Notes: The numbers denote the average value of $s(T)$ across 2000 replications. $s_0(T) = T^{1/3}$ is the fixed bandwidth, $\widehat{s}(T)$ is the data dependent bandwidth given by (7), and $\widehat{s}_{PW}(T)$ is the data dependent bandwidth under prewhitening.

5 Theoretical Explanations

In this section we provide theoretical explanations for the finite sample power patterns shown in section 4. We use an asymptotic analysis using model (5), i.e. a fixed alternative. We do not consider a local asymptotic analysis where δ is modeled as local to zero because $\widehat{\sigma}^2$ is asymptotically invariant to δ in this case and power would not depend on $\widehat{\sigma}^2$.

5.1 Data dependent bandwidth without prewhitening

The first step is to explain why $\widehat{s}(T)$ increases as δ grows. Recall that $\widehat{s}(T)$ depends on $\widehat{\rho}$ through $\widehat{\alpha}(1)$. The behavior of $\widehat{\rho}$ under model (5) follows from Perron (1990) as:

$$\lim_{T\to\infty} \widehat{\rho} = \phi(\delta), \tag{14}$$

where $\phi(\delta) = (\gamma_1 + f(\delta,\lambda)) \,/\, (\gamma_0 + f(\delta,\lambda))$ and $f(\delta,\lambda) = \lambda(1-\lambda)\delta^2$. Using equations (8) and (14) it follows that for the Bartlett kernel

$$\lim_{T\to\infty} \widehat{\alpha}(1) = C_1(\delta), \tag{15}$$

where

$$C_1(\delta) = \frac{4\phi^2(\delta)}{(1-\phi^2(\delta))^2}.$$

Note that as δ increases, $\phi(\delta)$ approaches one (in which case $\widehat{\rho}$ is severely biased) and $C_1(\delta)$ becomes large. In fact, $1-\phi(\delta) = O(\delta^{-2})$, and it follows that $C_1(\delta) = O(\delta^4)$. Therefore, as δ increases, $\widehat{\alpha}(1)$ becomes large and this leads to a large $\widehat{s}(T)$.

Using equations (14) and (15) we can approximate the behavior of $\widehat{s}(T)$ as

$$\widehat{s}(T) \approx 1.1447\,(C_1(\delta)T)^{1/3}. \tag{16}$$

In Table 4 we provide values of $\widehat{s}(T)$ based on the approximation (16) for $T = 50, 100, 200, 500$ for the error model $u_t = 0.7u_{t-1} + \epsilon_t$, $\epsilon_t \sim i.i.d.\ N(0,1)$. It is obvious from the table that $\widehat{s}(T)$ increases very quickly as δ increases. The patterns in Table 4 match the patterns in the finite sample simulations. We can conclude from this analysis that $\widehat{s}(T)$ will be large when δ is large because $\widehat{\rho}$ is biased towards one and $\widehat{\alpha}(1)$ becomes very large.

To explain why power falls as δ increases we need to understand how the CUSUM and QS tests behave when $\widehat{s}(T)$ is large. To obtain a useful approximation for this case, we examine the behavior of the tests under the assumption that $s(T) = \vartheta T$ where ϑ is a constant. For example, from Table 3 we see that for a DDB with $\delta = 10$, $mean(\widehat{s}(T))$= 44.988. Thus it is sensible to approximate the behavior of the tests for this case by setting ϑ= 0.44988.

Table 4: **Predicted Data Dependent Bandwidth, $\widehat{s}(T)$, Without Prewhitening.**

δ	$T = 50$	$T = 100$	$T = 200$	$T = 500$
1	9.1225	11.4936	14.4810	19.6538
2	11.4573	14.4353	18.1874	24.6841
3	14.8692	18.7340	23.6034	32.0348
4	19.0583	24.0120	30.2532	41.0599
5	23.8383	30.0344	37.8410	51.3581
6	29.0924	36.6542	46.1813	62.6777
7	34.7436	43.7742	55.1521	74.8529
8	40.7384	51.3271	64.6682	87.7682
9	47.0376	59.2637	74.6675	101.3394
10	53.6115	67.5463	85.1030	115.5025
11	60.4368	76.1456	95.9374	130.2071
12	67.4945	85.0377	107.1408	145.4124
13	74.7689	94.2029	118.6882	161.0847
14	82.2469	103.6246	130.5588	177.1955
15	89.9172	113.2885	142.7346	193.7206
16	97.7699	123.1823	155.2000	210.6388
17	105.7965	133.2952	167.9415	227.9316
18	113.9893	143.6176	180.9468	245.5826
19	122.3416	154.1408	194.2052	263.5770
20	130.8471	164.8571	207.7069	281.9016

Note: Values in the table correspond to equation (16) evaluated for errors given by $u_t = 0.7u_{t-1} + \epsilon_t$ with ϵ_t i.i.d. $N(0, 1)$.

The following theorem describes the behavior of the tests under the assumptions that $s(T) = \vartheta T$ and the model is given by (5):

Theorem 1 *Suppose that y_t is generated by (5) and Assumption 1 holds. If $s(T) = \vartheta T$ where $0 < \vartheta < 1$ and the Bartlett kernel is used to construct $\widehat{\sigma}^2$ then*

$$\lim_{T\to\infty} CUSUM(T) = \frac{\sup_{r\in(0,1)} |g(r,\lambda)|}{\sqrt{\frac{2}{\vartheta}\left(\int_0^1 g^2(r,\lambda)dr - \int_0^{1-\vartheta} g(r+\vartheta,\lambda)g(r,\lambda)dr\right)}}$$

$$\lim_{T\to\infty} QS(T) = \frac{\int_0^1 g^2(r,\lambda)dr}{\frac{2}{\vartheta}\left(\int_0^1 g^2(r,\lambda)dr - \int_0^{1-\vartheta} g(r+\vartheta,\lambda)g(r,\lambda)dr\right)},$$

where $g(r,\lambda) = (r-\lambda)1_{(r>\lambda)} - r(1-\lambda)$, and $1_{(r>\lambda)} = 1$ for $r > \lambda$ and 0 otherwise.

Notice from the Theorem that the limits do not depend on δ. However, the limits do depend on ϑ, and ϑ implicitly depends on δ through $\widehat{s}(T)$.

Given the complexity of the dependence of the limits on ϑ we focus on the concrete case of $\lambda = 0.5$ which matches the DGP of the simulations in section 4. With $\lambda = 0.5$ it is easy to show that

$$\lim_{T\to\infty} CUSUM(T) = \begin{cases} \frac{1}{2\sqrt{\vartheta(1-\vartheta)}} & if\ \vartheta \le \frac{1}{2} \\ \frac{\sqrt{6\vartheta}}{2\sqrt{1-2(1-\vartheta)^3}} & if\ \vartheta > \frac{1}{2} \end{cases}, \tag{17}$$

$$\lim_{T\to\infty} QS(T) = \begin{cases} \frac{1}{12\vartheta(1-\vartheta)} & if\ \vartheta \le \frac{1}{2} \\ \frac{\vartheta}{2-4(1-\vartheta)^3} & if\ \vartheta > \frac{1}{2} \end{cases}. \tag{18}$$

In Figure 10 we plot these limits for $0 < \vartheta < 1$. The solid lines in the figure plot the limiting functions whereas the dashed lines plot the 5% asymptotic critical values. The reason for nonmonotonic power is revealed by the plots. Notice that for small values of ϑ, the limits of the CUSUM and QS statistics are above the critical values. Thus we expect the statistics to reject the null. On the other hand, as ϑ increases, the limits of CUSUM and QS drop and are often below the critical value. Thus fewer rejections are obtained as ϑ grows and power drops.

We summarize these results as follows. Because σ^2 is estimated under the null hypothesis of a stable mean, when the true model has a mean shift, $\widehat{\rho}$ is biased towards 1 (see Perron (1990)). As

δ increases, $\widehat{\rho}$ is more biased towards one and $\widehat{\alpha}(1)$ grows. $\widehat{s}(T)$ becomes very large and CUSUM and QS decrease in value. Fewer rejections occur, power falls and nonmonotonic power results.

Finally, the reason that power is monotonic for the fixed bandwidth is also shown in Figure 10. For $T = 100$ the fixed bandwidth is $s_0(T) = 5$ which corresponds to $\vartheta = 0.05$. We see in the figures that very small values of ϑ are associated with large values of CUSUM and QS and rejections are frequent.

5.2 Data dependent bandwidth choice with prewhitening

In this section we provide a theoretical explanation for why power can be nonmonotonic when prewhitening is used. Because the behavior $\widehat{s}_{PW}(T)$ is different from $\widehat{s}(T)$, a different analysis from that in Section 5.1 is required. We begin with a theorem that describes the asymptotic behavior of $\widehat{s}_{PW}(T)$ under AR(1) prewhitening.

Theorem 2 *Let y_t be generated by (5), let $\widehat{\rho}_\epsilon$ be given by (13), and let $\widehat{\alpha}_{PW}(1)$ be given by (12). If Assumption 1 holds, then*

(i) $\lim_{T\to\infty} \widehat{\rho}_\epsilon = \phi_\epsilon(\delta)$, *where*

$$\phi_\epsilon(\delta) = \frac{(\gamma_1 + f(\delta,\lambda))(\gamma_1^2 - \gamma_0\gamma_1 + f(\delta,\lambda)(2\gamma_1 - \gamma_0 - \gamma_2))}{(\gamma_0 - \gamma_1)(\gamma_1 + \gamma_0 + 2f(\delta,\lambda))(\gamma_0 + f(\delta,\lambda))}.$$

(ii) $\lim_{T\to\infty} \widehat{\alpha}_{PW}(1) = C_2(\delta)$, *where*

$$C_2(\delta) = 4\phi_\epsilon^2(\delta)/(1 - \phi_\epsilon^2(\delta))^2.$$

Using the results of Theorem 2, we can approximate the behavior of $\widehat{s}_{PW}(T)$ by

$$\widehat{s}_{PW}(T) \approx 1.1447\,(C_2(\delta)T)^{1/3}. \tag{19}$$

The formulas for $\phi_\epsilon(\delta)$ and $C_2(\delta)$ are complicated as Theorem 2 shows. But, with some algebraic manipulation, it is easy to show that for large δ,

$$\phi_\epsilon(\delta) \approx \frac{2\gamma_1 - \gamma_0 - \gamma_2}{2(\gamma_0 - \gamma_1)}.$$

Unlike in the case without prewhitening, $\widehat{\rho}_\epsilon$ is not systematically biased towards one as δ increases. For example, with $u_t = \rho u_{t-1} + \epsilon_t$ and $\epsilon_t \sim$ *i.i.d.* $N(0,1)$, we have $\phi_\epsilon(\delta) \approx \frac{1}{2}(\rho - 1)$ for large δ.

When $\rho = 0.7$, we have $\phi_\epsilon(\delta) \approx -0.15$ for large δ and $\widehat{\alpha}(1) \approx 0.0942$ which gives $\widehat{s}_{PW}(T) = 2.417$. Therefore, as δ increases, $\widehat{s}(T)$ does not increase and remains relatively small. To illustrate this further, we used (19) to approximate the sampling behavior of $\widehat{s}_{PW}(T)$ for $T = 50$, 100, 200, 500 for the error model $u_t = 0.7u_{t-1} + \epsilon_t$, $\epsilon_t \sim i.i.d.\ N(0,1)$. The results are given in Table 5. We see that regardless the value of δ, $\widehat{s}_{PW}(T)$ is small, and $\widehat{s}_{PW}(T)$ decreases slightly as δ increases. These analytic results match the finite sample simulations where $\widehat{s}_{PW}(T)$ was small (see Table 3).

Table 5: **Predicted Data Dependent Bandwidth, $\widehat{s}_{PW}(T)$, with AR(1) Prewhitening.**

δ	$T = 50$	$T = 100$	$T = 200$	$T = 500$
1	3.0013	3.7814	4.7643	6.4662
2	2.7843	3.5080	4.4198	5.9986
3	2.5626	3.2287	4.0679	5.5210
4	2.3915	3.0131	3.7963	5.1523
5	2.2709	2.8611	3.6048	4.8925
6	2.1871	2.7556	3.4719	4.7120
7	2.1283	2.6814	3.3784	4.5852
8	2.0859	2.6281	3.3112	4.4940
9	2.0548	2.5889	3.2618	4.4269
10	2.0314	2.5593	3.2246	4.3764
11	2.0133	2.5366	3.1960	4.3376
12	1.9992	2.5189	3.1736	4.3072
13	1.9880	2.5047	3.1558	4.2830
14	1.9789	2.4933	3.1414	4.2635
15	1.9715	2.4840	3.1296	4.2475
16	1.9654	2.4762	3.1199	4.2343
17	1.9602	2.4697	3.1117	4.2232
18	1.9559	2.4643	3.1048	4.2139
19	1.9522	2.4596	3.0989	4.2059
20	1.9490	2.4556	3.0939	4.1990

Notes: Values in the table correspond to $\widehat{s}_{PW}(T)$ given by (19) for errors given by $u_t = 0.7u_{t-1} + \epsilon_t$ with ϵ_t i.i.d. $N(0,1)$.

Given that $\widehat{s}_{PW}(T)$ is small under AR(1) prewhitening, we now turn to the limiting behavior of the CUSUM and QS for small $\widehat{s}_{PW}(T)$. A useful approximation can be obtained by assuming that $\widehat{s}_{PW}(T)$ remains fixed at a small value as T increases. Suppose $s_{PW}(T) = k$ where k is a constant (small). The following theorem provides the limiting behavior of CUSUM and QS under prewhitening.

Theorem 3 *Suppose that y_t is generated by (5) and Assumption 1 holds. If $s_{PW}(T) = k$, where k is a constant, and the Bartlett kernel is used with AR(1) prewhitening, then*

$$\lim_{T\to\infty} T^{-1/2} CUSUM(T) = \frac{|\delta|\lambda(1-\lambda)}{\sigma_k}, \qquad \lim_{T\to\infty} T^{-1} QS(T) = \frac{\delta^2\lambda^2(1-\lambda)^2}{3\sigma_k^2},$$

where

$$\sigma_k^2 = \frac{1}{(1-\phi(\delta))^2}\left(\psi_0 + 2\sum_{t=1}^{k-1}\left(1-\frac{t}{k}\right)\psi_t\right),$$

$$\psi_k = f(\delta,\lambda)(1-\phi(\delta))^2 + \gamma_k(1+\phi^2(\delta)) - \phi(\delta)(\gamma_{k-1}+\gamma_{k+1}).$$

It is important to note from the theorem that the statistics must be scaled by $T^{-1/2}$ and T^{-1} to obtain non-degenerate limits. This shows that the tests are consistent under prewhitening and one might be tempted to conclude that power is high in finite samples. Of course, for a fixed δ, as $T \to \infty$, power will eventually reach one. The relevant question here is, what happens to power for a fixed T as δ increases!

For concreteness, consider the case of $\lambda = 0.5$ and $k = 1$. Then, the limits in Theorem 3 simplify to

$$\lim_{T\to\infty} T^{-1/2} CUSUM(T) = \frac{|\delta|(1-\phi(\delta))}{4\sqrt{\frac{\delta^2}{4}(1-\phi(\delta))^2 + (1+\phi^2(\delta))\gamma_0 - 2\phi(\delta)\gamma_1}},$$

$$\lim_{T\to\infty} T^{-1} QS(T) = \frac{\delta^2(1-\phi(\delta))^2}{48\left(\frac{\delta^2}{4}(1-\phi(\delta))^2 + (1+\phi^2(\delta))\gamma_0 - 2\phi(\delta)\gamma_1\right)}.$$

Using these limits, we can approximate the behavior of CUSUM and QS through the approximations

$$CUSUM(T) \approx T^{1/2}\frac{|\delta|(1-\phi(\delta))}{4\sqrt{\frac{\delta^2}{4}(1-\phi(\delta))^2 + (1+\phi^2(\delta))\gamma_0 - 2\phi(\delta)\gamma_1}}, \tag{20}$$

$$QS(T) \approx T \frac{\delta^2(1-\phi(\delta))^2}{48\left(\frac{\delta^2}{4}(1-\phi(\delta))^2 + (1+\phi^2(\delta))\gamma_0 - 2\phi(\delta)\gamma_1\right)}. \tag{21}$$

Recall from the previous section that as δ increases, $\phi(\delta)$ approaches one. As $\phi(\delta)$ approaches one, $1 - \phi(\delta)$ approaches zero very quickly, and $\delta(1 - \phi(\delta))$ approaches zero. The denominators are bounded as δ increases. Because the limits of the scaled CUSUM and QS tests are proportional to $\delta(1 - \phi(\delta)) = o(1)$ (for large δ), the statistics decrease in value as δ increases for fixed T.

In Figure 11, we plot the approximate limits given by equations (20) and (21) for $T = 100$ and $u_t = 0.7u_{t-1} + \epsilon_t$ where ϵ_t are i.i.d. $N(0, 1)$. Note that $k = 1$ is appropriate for this case as $\widehat{s}(T)$ was approximately 1 in the finite sample simulations reported in Table 3. In the figure we see that as δ increases, the statistics decrease in value and are always below the 5% critical values. Thus, we should expect few rejections and low power for large δ and nonmonotonic power.

The ultimate reason why nonmonotonic power is obtained with prewhitening is that $\widehat{\rho}$ is biased towards 1 and $\widehat{\sigma}^2_{PW}$ is proportional to $(1 - \widehat{\rho})^2$ thus reducing the values of CUSUM and QS as δ grows. The bandwidth does not play a central role here since it remains small as δ increases. In fact, the AR(1) prewhitened estimator with $s(T) = 1$ is equivalent to an AR(1) parametric estimator of σ^2. Our results, therefore, suggest that parametric estimators of σ^2 will also generate nonmonotonic power for CUSUM and QS. Indeed, in unreported simulations (comparable to those used for Figures 2-9) we found that an AR(1) parametric estimator of σ^2 leads to nonmonotonic power.

6 Conclusions

In this paper we have shown that use of a data dependent bandwidth when estimating the spectral density at frequency zero of a time series can result in the CUSUM and QS tests for a stable mean to have nonmonotonic power. Power can drop to zero for large one-time mean shifts. Prewhitening does not improve the situation. On the other hand, deterministic bandwidth rules (fixed bandwidth rules) can result in monotonic power provided the bandwidth is small.

This is an unfortunate situation. Fixed bandwidths are arbitrary in finite samples. In fact, any bandwidth choice can be justified by some fixed bandwidth rule. It was for this reason that data dependent bandwidths were developed in the first place. Our theoretical analysis suggests that small bandwidths can result in monotonic power, whereas large bandwidths can result in nonmonotonic power. But, the analysis does not and cannot provide universal small sample recommendations as to when a bandwidth is too big (if nonmonotonic power is to be avoided).

The underlying problem with the CUSUM and QS tests is that they are constructed using OLS residuals from a model estimated under the null hypothesis of a stable mean. While this approach is attractive because no alternative model needs to be specified, it can lead to tests with poor power because zero frequency spectral density estimates required for the tests are estimated using a misspecified model when the mean is unstable. When there is a single shift in mean, our analysis shows that the spectral density estimators are poorly behaved, and power is crippled.

Our analysis suggests that tests with nice power properties require the specification of an alternative model. For example, for the alternative of a single shift in mean at an unknown time, Vogelsang (1998,1999) suggested some tests that have monotonic power and stable size. Of course, it may not be obvious how to model an unstable mean (hence the appeal of the CUSUM and QS tests). But, at the end of the day, if the stable mean null is rejected, some alternative needs to be specified. Therefore, practitioners should seriously consider alternative models and use tests designed to detect them. Null based tests simply do not always deliver the goods as they can have negligible power for mean shifts that would be obvious from eye-balling a plot of the data.

Appendix

We begin by establishing some basic results. If $\widehat{u}_t$ are the residuals of the regression of y_t on a constant then

$$\widehat{\rho} = \frac{\sum_{t=2}^{T} \widehat{u}_t \widehat{u}_{t-1}}{\sum_{t=2}^{T} \widehat{u}_{t-1}^2},$$

where

$$\widehat{u}_t = \begin{cases} u_t - \bar{u} - \delta(1-\lambda) & for\ t \leq [\lambda T] \\ u_t - \bar{u} + \delta\lambda & for\ t > [\lambda T] \end{cases},$$

and $\bar{u} = T^{-1} \sum_{t=1}^{T} u_t$. We show that

$$\lim_{T\to\infty} \widehat{\gamma}_i = \gamma_i + f(\delta, \lambda), \tag{22}$$

for $i \geq 0$. Denote $\widehat{v}_t = u_t - \bar{u}$. Then, we can write,

$$\widehat{\gamma}_0 = T^{-1} \sum_{t=1}^{T} \widehat{u}_t^2 = T^{-1} \sum_{t=1}^{T} \widehat{v}_t^2 + f(\delta, \lambda) - 2\delta T^{-1} \sum_{t=1}^{[\lambda T]} \widehat{v}_t.$$

Because $p\lim T^{-1}\sum_{t=1}^{T}\widehat{v}_t^2=\gamma_0$ and $p\lim T^{-1}\sum_{t=1}^{[\lambda T]}\widehat{v}_t=0$, we obtain result (22) for $i=0$. Additional algebra gives

$$\widehat{\gamma}_1=T^{-1}\sum_{t=2}^{T}\widehat{u}_t\widehat{u}_{t-1}=T^{-1}\sum_{t=2}^{T}\widehat{v}_t\widehat{v}_{t-1}+f(\delta,\lambda)-2\delta T^{-1}\sum_{t=1}^{[\lambda T]}\widehat{v}_t+h(\delta,\lambda,T),$$

where $h(\delta,\lambda,T)=T^{-1}\delta[(1-\lambda)\widehat{v}_1+\widehat{v}_{[\lambda T]}-\widehat{v}_{[\lambda T]+1}-\lambda\widehat{v}_T]-T^{-1}\delta^2(1-\lambda-\lambda^2)$. Since $p\lim T^{-1}\sum_{t=2}^{T}\widehat{v}_t\widehat{v}_{t-1}=\gamma_1$, $p\lim 2\delta T^{-1}\sum_{t=1}^{[\lambda T]}\widehat{v}_t=0$ and $p\lim h(\delta,\lambda,T)=0$ we obtain result (22) for $i=1$. Similar arguments can be used to establish (22) for $i\geq 2$.

Proof of Theorem 1: For the case of $\vartheta=1$, Kiefer and Vogelsang (2000) showed that $\widehat{\sigma}^2=2T^{-2}\sum_{t=1}^{T}\widehat{S}_t^2$ where $\widehat{S}_t=\sum_{j=1}^{t}\widehat{v}_j$. It is straightforward to extend their results to show that for $s(T)=\vartheta T$:

$$\widehat{\sigma}^2=\frac{2}{\vartheta}\left[T^{-2}\sum_{t=1}^{T}\widehat{S}_t^2-T^{-2}\sum_{t=1}^{T-[\vartheta T]}\widehat{S}_{t+[\vartheta T]}\widehat{S}_t\right].$$

Given this expression for $\widehat{\sigma}^2$, it follows that CUSUM and QS can be written in terms of $T^{-1}\widehat{S}_t$:

$$CUSUM=\frac{\sup_{1\leq t\leq T}\left|T^{-1}\widehat{S}_t\right|}{\left(\frac{2}{\vartheta}\left[T^{-1}\sum_{t=1}^{T}T^{-2}\widehat{S}_t^2-T^{-1}\sum_{t=1}^{T-[\vartheta T]}T^{-1}\widehat{S}_{t+[\vartheta T]}T^{-1}\widehat{S}_t\right]\right)^{1/2}},\tag{23}$$

$$QS=\frac{T^{-1}\sum_{t=1}^{T}T^{-2}\widehat{S}_t^2}{\frac{2}{\vartheta}\left[T^{-1}\sum_{t=1}^{T}T^{-2}\widehat{S}_t^2-T^{-1}\sum_{t=1}^{T-[\vartheta T]}T^{-1}\widehat{S}_{t+[\vartheta T]}T^{-1}\widehat{S}_t\right]}.\tag{24}$$

Using (23) and (24), the proof is completed by deriving the limiting behavior of $T^{-1}\widehat{S}_{[rT]}$. Under model (5) we have

$$\begin{aligned}T^{-1}\widehat{S}_{[rT]}&=T^{-1}\sum_{t=1}^{[rT]}[\delta(DU_t-(1-\lambda))+u_t-\bar{u}]\\&=\delta T^{-1}\sum_{t=1}^{[rT]}(DU_t-(1-\lambda))+o_p(1)\\&\to\delta\int_0^r\left[1_{(r>\lambda)}-(1-\lambda)\right]dr=\delta g(r,\lambda).\end{aligned}$$

Proof of Theorem 2: Direct calculation gives

$$\widehat{\rho}_\epsilon = \frac{\sum_{t=3}^T(\widehat{u}_t - \widehat{\rho}\widehat{u}_{t-1})(\widehat{u}_{t-1} - \widehat{\rho}\widehat{u}_{t-2})}{\sum_{t=2}^T(\widehat{u}_t - \widehat{\rho}\widehat{u}_{t-1})^2} = \frac{T^{-1}\sum_{t=3}^T(\widehat{u}_t\widehat{u}_{t-1} - \widehat{\rho}\widehat{u}_{t-1}^2 - \widehat{\rho}\widehat{u}_t\widehat{u}_{t-2} + \widehat{\rho}^2\widehat{u}_{t-1}\widehat{u}_{t-2})}{T^{-1}\sum_{t=2}^T(\widehat{u}_t^2 - 2\widehat{\rho}\widehat{u}_t\widehat{u}_{t-1} + \widehat{\rho}^2\widehat{u}_{t-1}^2)}.$$

Letting $T \to \infty$ and applying (22), part (i) of the theorem directly follows. Part (ii) of the theorem follows directly from part (i).

Proof of Theorem 3: Let $\widehat{\gamma}_i^\epsilon = T^{-1}\sum_{t=i+2}^T \widehat{\epsilon}_t\widehat{\epsilon}_{t-i}$. Because $\widehat{\epsilon}_t = \widehat{u}_t - \widehat{\rho}\widehat{u}_{t-1}$ it follows that

$$\widehat{\gamma}_i^\epsilon = T^{-1}\sum_{t=i+1}^T \widehat{u}_t\widehat{u}_{t-i} - \widehat{\rho}T^{-1}\sum_{t=i+1}^T \widehat{u}_{t-1}\widehat{u}_{t-i} - \widehat{\rho}T^{-1}\sum_{t=i+2}^T \widehat{u}_t\widehat{u}_{t-i-1} + \widehat{\rho}^2T^{-1}\sum_{t=i+2}^T \widehat{u}_{t-1}\widehat{u}_{t-i-1}.$$

Hence $\lim_{T\to\infty}\widehat{\gamma}_i^\epsilon = \psi_i$ using (22). For a data dependent bandwidth with AR(1) prewhitening

$$\widehat{\sigma}_{PW}^2 = \frac{\widehat{\gamma}_0^\epsilon + 2\sum_{t=1}^{k-1}\left(1 - \frac{t}{k}\right)\widehat{\gamma}_t^\epsilon}{(1-\widehat{\rho})^2},$$

where k is a fixed lag. Using the limiting result for $\widehat{\gamma}_i^\epsilon$ and (14) gives

$$\lim_{T\to\infty}\widehat{\sigma}_{PW}^2 = \frac{\psi_0 + 2\sum_{t=1}^{k-1}\left(1 - \frac{t}{k}\right)\psi_t}{(1-\phi^2(\delta))}. \tag{25}$$

Using simple algebra

$$T^{-1/2}CUSUM = \frac{\sup_{1\le t\le T}|T^{-1}\widehat{S}_t|}{\widehat{\sigma}} \qquad and \qquad T^{-1}QS = \frac{T^{-1}\sum_{t=1}^T T^{-2}\widehat{S}_t^2}{\widehat{\sigma}^2}.$$

The limits of the numerators follow from the proof of Theorem 1. The limits of the denominators follow from (25). The representations follow from the facts that

$$\sup_{r\in(0,1)}|g(r,\lambda)| = |\lambda(1-\lambda)|, \qquad \int_0^1 g^2(r,\lambda) = \frac{1}{3}\lambda^2(1-\lambda)^2.$$

References

[1] Andrews, D.W.K., *Heteroskedasticity and autocorrelation consistent covariance matrix estimation*. Econometrica 59 (1991), 817-858.

[2] Andrews, D.W.K. and Monahan, J., *An Improved Heteroskedasticity and autocorrelation consistent covariance matrix estimator*, Econometrica, 60, (1992), 953-966.

[3] Brown, R.L., Durbin, J. and Evans, J.M., *Techniques for testing the constancy of regression relationship over time*. Journal of the Royal Statistical Society B 37 (1975), 149-163.

[4] Commercial and Financial Chronicle, 1928-1940. William B. Dana Company, New York (various issues).

[5] Gardner, L.A., *On detecting changes in the mean of normal variates.* Annals of Mathematical Statistics 10 (1969), 116-126.

[6] Herrndorf, N., *A functional central limit theorem for weakly dependent sequences of random variables.* Annals of Probability 12 (1984), 141-153.

[7] Kiefer, N.M. and Vogelsang, T.J., *A new approach to the asymptotics of heteroskedasticity-autocorrelation robust testing.* Cornell Department of Economics. Mimeo (2000).

[8] MacNeill, I.B., *Properties of sequences of partial sums of polynomial regression residuals with applications to test for change of regression at unknown times.* Annals of Statistics 6 (1978), 422-433.

[9] Perron, P., *Testing for a unit root in a time series with a changing mean.* Journal of Business and Economic Statistics 8 (1990), 153-162.

[10] Perron, P., *A test for changes in a polynomial trend function for a dynamic time series.* Mimeo. Princeton University Department of Economics. 1991.

[11] Phillips, P.C.B., *Time series with a unit root.* Econometrica 55 (1987), 277-301.

[12] Ploberger, W. and Kramer, W., *The CUSUM Test with OLS Residuals,* Econometrica, 60 (1992), 271-285.

[13] Priestley, M.B., *Spectral analysis and time series*, Vol. 1, Academic Press, New York. 1981.

[14] Vogelsang, T. J., *Testing for a shift in mean without having to estimate serial correlation parameters.* Journal of Business and Economic Statistics 16 (1998), 73-80.

[15] Vogelsang, T. J., *Sources of nonmonotonic power when testing for a shift in mean of a dynamic time series.* Journal of Econometrics 88 (1999), 283-299.

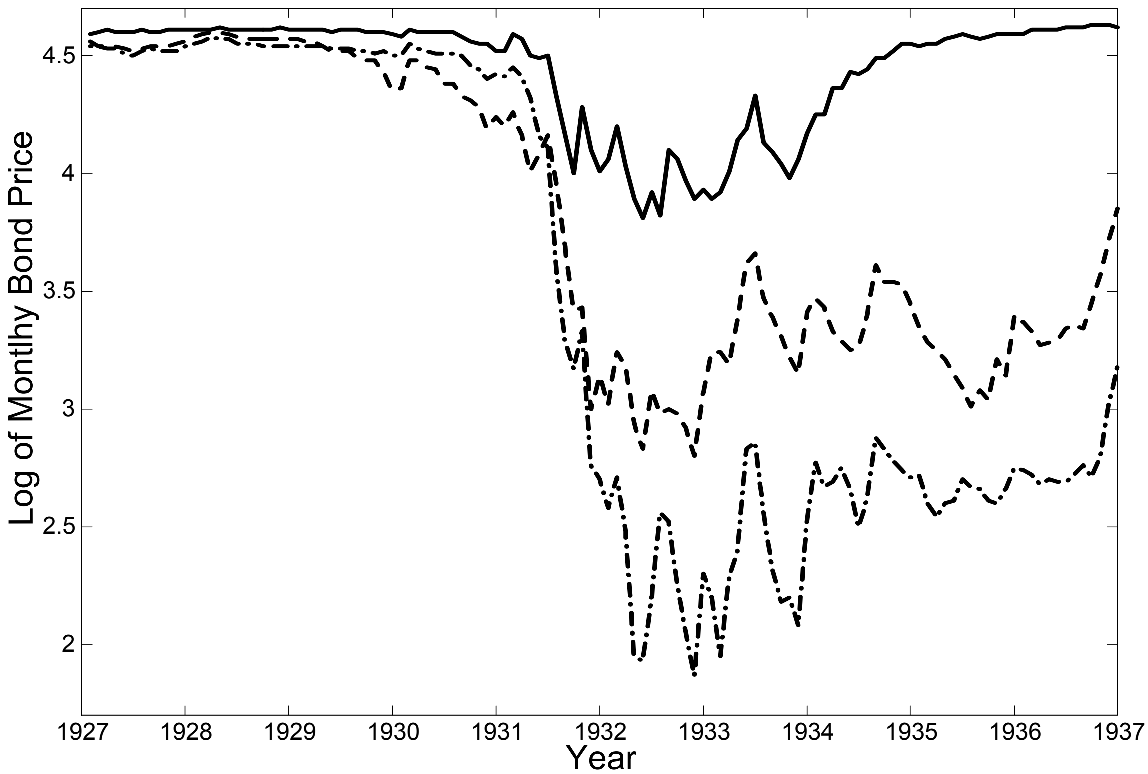

Figure 1: Bond prices for Argentina '—', Brazil '– –' and Chile '–·–'

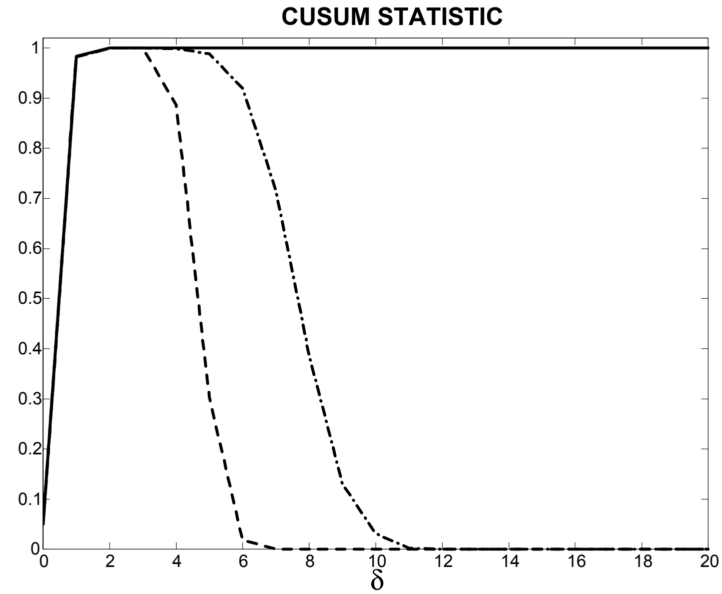

Figure 2: Power functions for FB '—', DDB '– –', DDB-PW '– · –'; $T = 100, \rho = 0.1$.

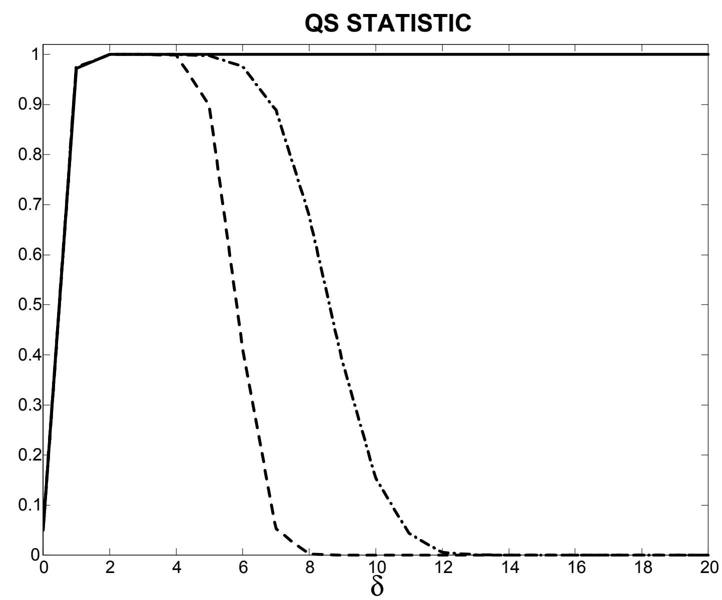

Figure 3: Power functions for FB '—', DDB '– –', DDB-PW '– · –'; $T = 100, \rho = 0.1$.

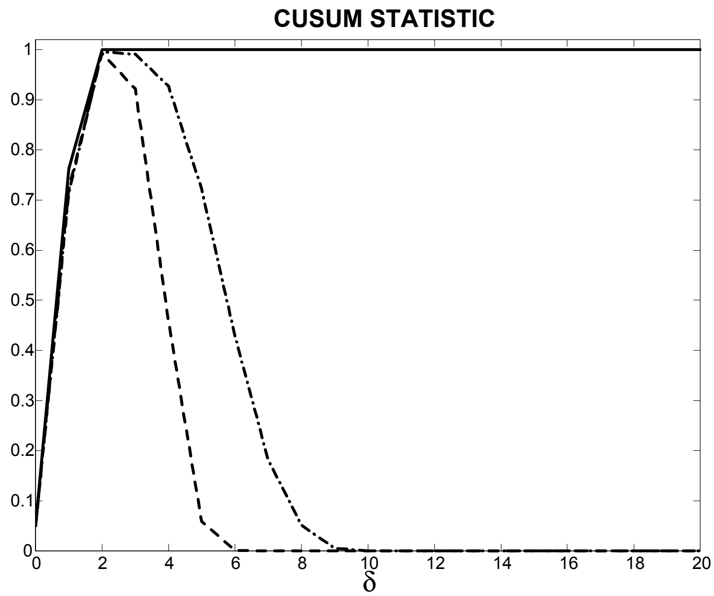

Figure 4: Power functions for FB '—', DDB '– –', DDB-PW '– · –'; $T = 100, \rho = 0.4$.

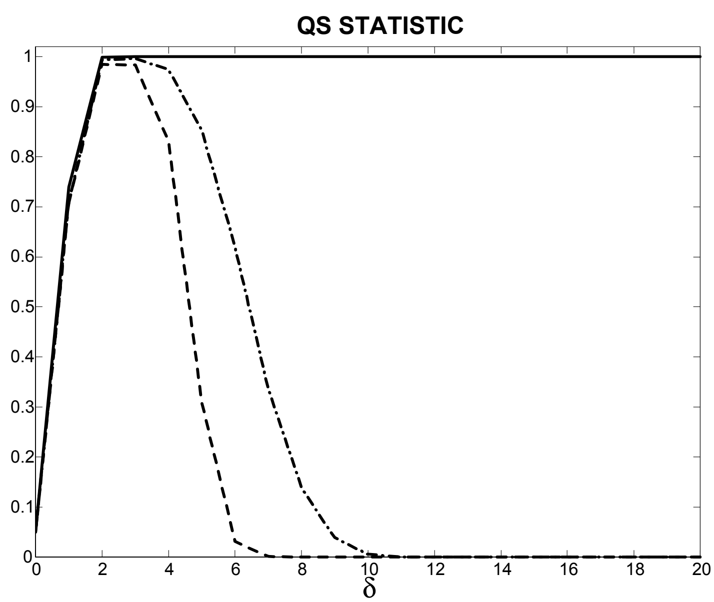

Figure 5: Power functions for FB '—', DDB '– –', DDB-PW '– · –'; $T = 100, \rho = 0.4$.

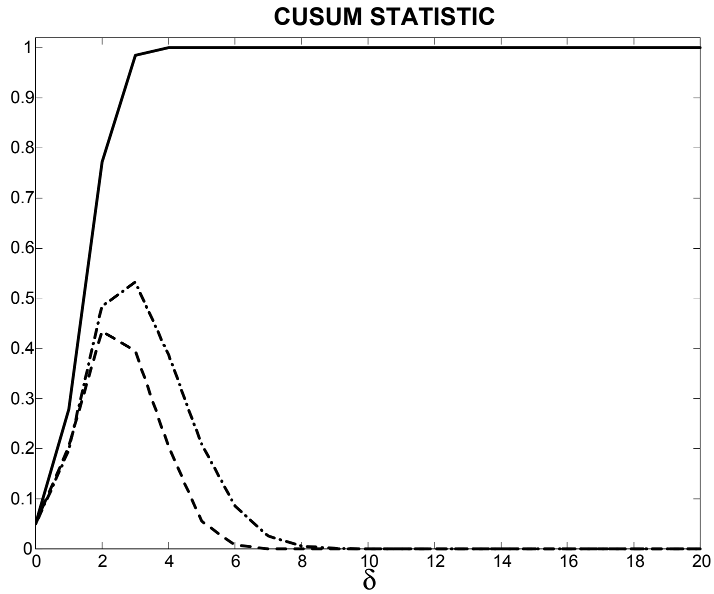

Figure 6: Power functions for FB '—', DDB '– –', DDB-PW '– · –'; $T = 100, \rho = 0.7$.

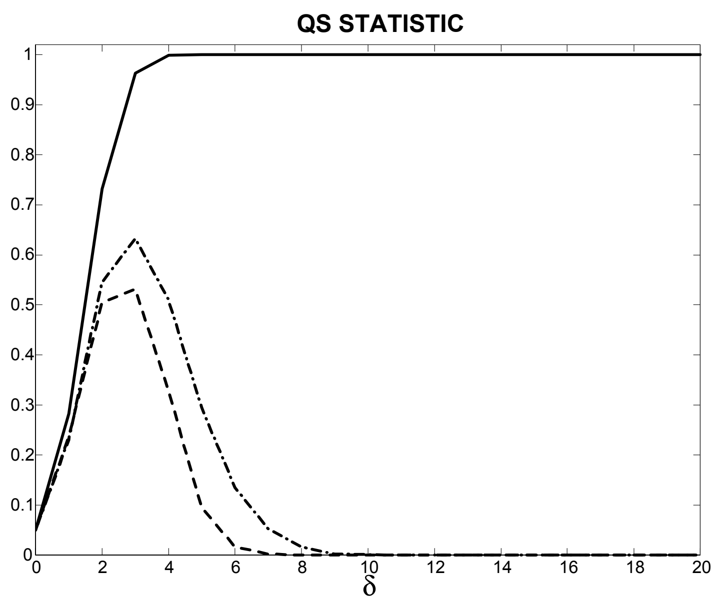

Figure 7: Power functions for FB '—', DDB '– –', DDB-PW '– · –'; $T = 100, \rho = 0.7$.

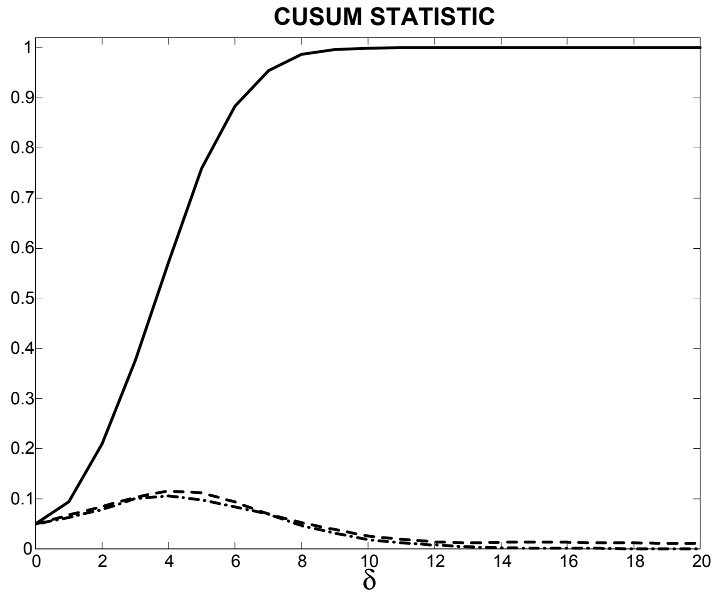

Figure 8: Power functions for FB '—', DDB '– –', DDB-PW '– · –'; $T = 100, \rho = 0.9$.

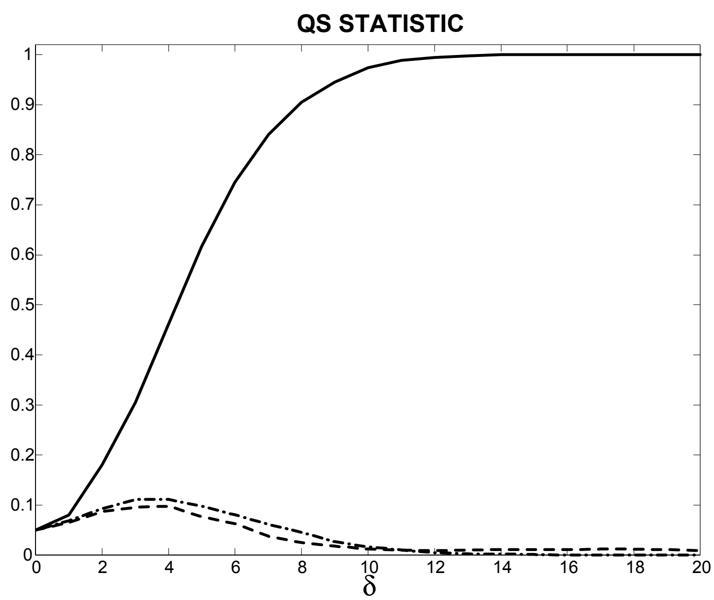

Figure 9: Power functions for FB '—', DDB '– –', DDB-PW '– · –'; $T = 100, \rho = 0.9$.

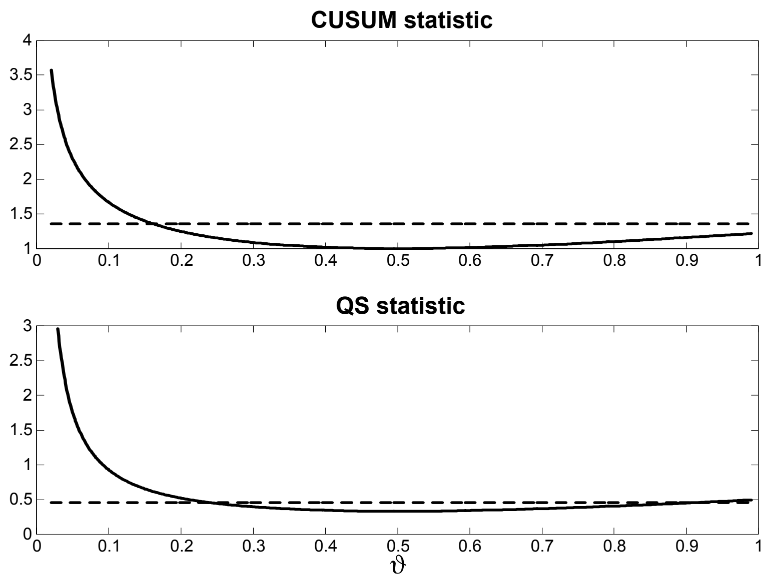

Figure 10: Asymptotic Limits Given by (17) and (18).

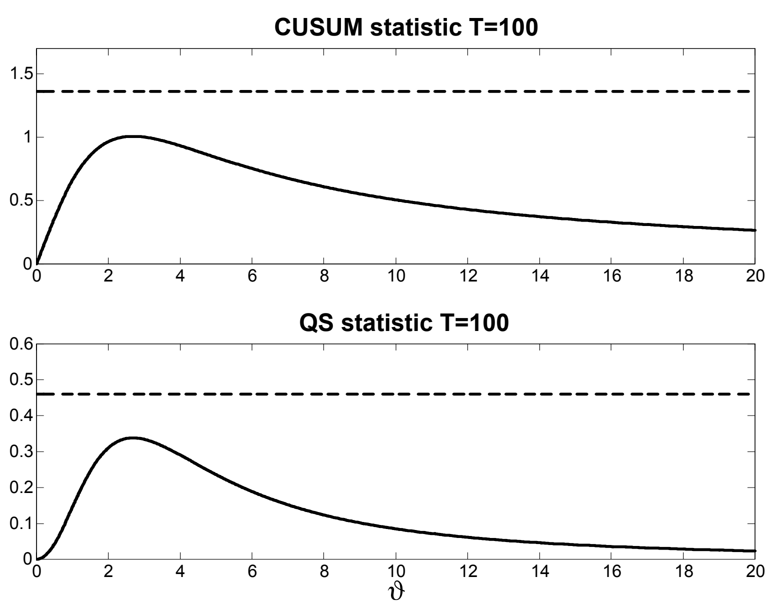

Figure 11: Approximate Asymptotic Limits Given by (20) and (21).