TESTS FOR CHANGING MEAN WITH MONOTONIC POWER

TED JUHL AND ZHIJIE XIAO

Abstract. Several widely used tests for a changing mean exhibit nonmonotonic power in finite samples due to "incorrect" estimation of nuisance parameters under the alternative. In this paper, we study the issue of nonmonotonic power in testing for changing mean. We investigate the asymptotic power properties of the tests using a new framework where alternatives are characterized as having "large" changes. The asymptotic analysis provides a theoretical explanation to the power problem. Modified tests that have monotonic power against a wide range of alternatives of structural change are proposed. Instead of estimating the nuisance parameters based on ordinary least squares residuals, the proposed tests use modified estimators based on nonparametric regression residuals. It is shown that tests based on the modified long-run variance estimator provide an improved rate of divergence of the tests under the alternative of a change in mean. Tests for structural breaks based on such an estimator are able to remain consistent while still retaining the same asymptotic distribution under the null hypothesis of constant mean.

1. Introduction

Testing for structural changes in dynamic models is a common practice in empirical time series analysis. Two widely used procedures for structural changes are the CUSUM (Kolmogorov-Smirnoff) test and the QS (or Cramer von-Mises) test. The CUSUM test based on recursive residuals was proposed by Brown, Durbin and Evans (1975). Ploberger and Kramer (1992) studied CUSUM test based on OLS residuals. The QS tests based on the Cramer von-Mises measure were studied by Gardner (1969), MacNeil (1978) and Perron (1991). For other tests used in econometrics, see, e.g., Andrews (1993), Andrews and Ploberger (1994), Perron (1991), Hansen (1992), and Elliott and Müller (2006).

An important issue in tests for structural change is that some tests can exhibit nonmonotonic power (see, e.g., Perron (1991), Vogelsang (1999), Deng and Perron (2007)). That is, as the structural change grows, the power of the test actually decreases. The problem of nonmonotonic power is found in models with a structural break in the mean.

Version: January, 2008. We thank Bruce Hansen, Pierre Perron, a Co-Editor, an Associate Editor, and two anonymous referees for very helpful comments on an earlier version of this paper. Ted Juhl, Department of Economics, University of Kansas, 415 Snow Hall, 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).

Monte Carlo studies indicate that the main reason for nonmonotonic power is that nuisance parameters are poorly estimated under the (global) alternative. For the above mentioned tests, a common feature is that they are based on empirical processes that satisfy functional central limit theorems under the null hypotheses of no structural change. The test statistics converge to functionals of (demeaned) standard Brownian motions after standardization by an estimate of the long-run variance of the time series. Monte Carlo results of Crainiceanu and Vogelsang (2007) show that, in the presence of a changing mean, the value of the data dependent bandwidth will be very large, and nonmonotonic power can be found when such variance estimators are used.

In this paper, we investigate the problem of nonmonotonic power in tests for changing mean. To provide a theoretical explanation for the nonmonotonic power problem, we develop an asymptotic analysis for the tests under the alternative hypothesis with large (diverging) changes. We then propose a modification to the tests that is designed to avoid nonmonotonic power. Since the source of nonmonotonic power is the long-run variance estimator, we want to find a variance estimator that is robust under very general deterministic structures. We propose a nonparametric regression on the original data first, and then construct an estimator of the long-run variance parameter based on the nonparametrically demeaned data. Asymptotic analysis on the proposed estimators and test statistics is provided. We show that the tests using long-run variance estimators based on the nonparametric regression residuals diverge at a slower rate than standard long-run variance estimators, yet they retain correct size under the null hypothesis. We show that tests constructed using such a long-run variance estimator retain their consistency under various diverging alternative hypotheses.

The structure of the paper is as follows. In section 2, we describe the model and existing tests. In section 3, we investigate the problem of nonmonotonic power via asymptotics of large structural changes. The modified tests are proposed and examined in section 4. We illustrate the performance of the tests using a small Monte Carlo experiment in section 5 and illustrate the usefulness of our procedure using a series containing expected U.S. inflation in section 6. Section 7 concludes.

2. Model and Tests

Consider a time series model characterized by

$$y_t = \theta_t + u_t, \tag{2.1}$$

where $\theta_t = \theta\left(\frac{t}{T}\right)$ is the deterministic component of y_t, and u_t is the stochastic component that is potentially serially correlated. Notice that θ_t is allowed to vary over time in the current setup. Our purpose is to test whether or not there is a change (or changes) in the deterministic component θ_t. 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 in time series y_t.

One way for testing structural change in θ_t is to look at the fluctuation in the OLS residuals of (2.1), $\hat{u}_t$, through the following empirical process

$$V_T(r) = \frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} \hat{u}_t. \tag{2.2}$$

Under H_0, the fluctuation in $\hat{u}_t$ is stable and $V_T(r)$ converges weakly to a demeaned Brownian motion with variance $\omega^2 = \sum_{j=-\infty}^{\infty} \gamma(j)$, where $\gamma(j)$ is the autocovariance function of u_t. In the presence of structural breaks in θ, the fluctuation in $\hat{u}_t$ will be larger. Since the limiting process of $V_T(r)$ has variance ω^2, appropriate standardization is needed to remove this nuisance parameter. Let $\hat{\omega}$ be a consistent estimator for ω, we may consider the following standardized empirical process:

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

where $U_T(r)$ converges weakly to a standard demeaned Brownian motion, say $\underline{W}(r)$, under the null hypothesis. Let $g(\cdot)$ be a continuous functional that measures the fluctuation of $U_T(r)$, we can use $g(U_T(r))$ as a test statistic for H_0. By the continuous mapping theorem,

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

When we take $g(\cdot)$ as the classical Kolmogorov-Smirnoff measure, we obtain the CUSUM test

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

which was studied by Ploberger and Kramer (1992). Using the Cramer-von Mises type metric, we can construct the following test

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

The above test is also labeled as QS in the literature (see, for example, Gardner (1969), MacNeill (1978), and Perron (1991)). Under the null hypothesis,

$$CUSUM_T \Rightarrow \sup_{0\le r\le 1}|\underline{W}(r)|, \text{ and } QS_T \Rightarrow \int_0^1 \underline{W}(r)^2 dr.$$

The CUSUM test and the QS test are two of the most widely used testing procedures for structural breaks. In this paper, we focus our discussion on the CUSUM and QS tests, but our analysis in this paper extends naturally to other tests that employ nonparametric estimation of the serial correlation structure through a long-run variance.

The above tests make use of an estimated long-run variance, which is proportional to the spectral density of u_t at zero frequency. This quantity is usually estimated by a nonparametric kernel method in the following form:

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

where $\hat{\gamma}(j)$ is the sample autocovariance function

$$\hat{\gamma}(j) = \frac{1}{T}\sum_{t=j+1}^{T}\hat{u}_t\hat{u}_{t-j}^{\top}, \text{ for} j \ge 0; \text{ and } \hat{\gamma}(j) = \hat{\gamma}(-j) \text{ for} j < 0.$$

$k(x)$ is a lag window kernel function, and b_T is a bandwidth parameter which controls the length of lag truncation, satisfying the property $b_T \to \infty$ and $b_T/T \to 0$. For example, Andrews (1991) suggests the following optimal kernel (quadratic spectral kernel)

$$k(x) = \frac{25}{12\pi^2 x^2}\left(\frac{\sin(6\pi x/5)}{6\pi x/5} - \cos(6\pi x/5)\right). \tag{2.3}$$

Another popular kernel is the Bartlett kernel with the following form:

$$k(x) = 1 - |x|, \text{ for } |x| \le 1. \tag{2.4a}$$

See, e.g., Andrews (1991) for more discussions on other kernels.

In practice, the long-run variance estimate is sensitive to the choice of bandwidth parameter b_T. In many econometric applications, the AR(1)-based plug-in data dependent bandwidth (Andrews (1991)) based on a mean-squared-error criterion is widely used. Let $\hat{\rho}$

be the estimated AR(1) parameter from a univariate autoregression of $\hat{u}_t$ on $\hat{u}_{t-1}$, the data dependent bandwidth parameter b_{opt} is given as

$$b_{opt} = \xi(k) \times (\hat{\alpha}(q)T)^{1/(2q+1)}$$

where $\xi(k)$ is a constant depending on the kernel function and q is the order of the kernel. $\alpha(q)$ is a function of the unknown spectral density and is estimated by a quantity based on $\hat{\rho}$. If the quadratic spectral kernel (2.3) is used, $q = 2$, $\xi(k) = 1.3221$, and

$$\hat{\alpha}(2) = \frac{4\hat{\rho}^2}{(1-\hat{\rho})^4}. \tag{2.5}$$

If the Bartlett kernel (2.4a) is used, $q = 1$, $\xi(k) = 1.1447$, and

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

These data-dependent choices of bandwidth are widely used and are available as a standard option in many software packages.

The use of a data-dependent bandwidth is very important in applied work since long-run variances are sensitive to the bandwidth. Andrews (1991) illustrates the improvement in estimation of an optimal bandwidth via a Monte Carlo experiment. The data-dependent bandwidth turns out to be an important part of the cause for nonmonotonic power in the structural break tests we discuss.

In the next section, we see how the CUSUM and QS tests are affected by changes that are not characterized as "local".

3. Asymptotic Theory of Large Changes

It is well known that the $CUSUM$ and QS tests are consistent, so that power goes to one as the sample size increases *when the size of the break is fixed.* Moreover, local power for the $CUSUM$ test is derived in Ploberger and Kramer (1992). However, monte carlo evidence of Creaniceanu and Vogelsang (2007) shows that if one lets the size of a one time discrete change in mean go to infinity with the sample size, power may not go to one. This effect is referred to as nonmonotonic power since power may increase initially and then decrease as changes in mean are larger in magnitude. Given the possibility of nonmonotonic power, we consider alternatives that are allowed to diverge with T.

In this paper, we consider alternatives of the form

$$\theta\left(\frac{t}{T}\right) = \mu + g\left(\frac{t}{T}\right)\eta.$$

By allowing η to approach infinity, we can determine the effects of large changes in mean on the tests. In particular, we will allow $\eta = O(T^b)$ where $b > 0$. We first show (Theorem 3.1) that the long-run variance estimator is the main source of the nonmonotonic power problem, and then analyze the power property of the tests using long-run variance estimator (Theorem 3.2).

Our first result deals with the tests that do not use a long-run variance estimator. In the case that u_t is a martingale difference sequence, only a variance estimator (instead of the long-run variance estimator) is needed for standardization because there is no serial correlation. Let $\widehat{\sigma}^2$ be the sample variance, and

$$\begin{aligned}
\underline{CUSUM}_T &= \max_{j=1,\ldots,T} \frac{1}{\widehat{\sigma}} \left| \frac{1}{\sqrt{T}} \sum_{t=1}^{j} \hat{u}_t \right|, \\
\underline{QS}_T &= \frac{1}{T} \sum_{t=1}^{T} \left(\frac{1}{\widehat{\sigma}\sqrt{T}} \sum_{j=1}^{t} \hat{u}_j \right)^2 .
\end{aligned}$$

The following theorem states that the CUSUM and QS tests still have power against alternatives that increase with sample size so long as we do not use a long-run variance estimator.

Theorem 3.1. *Suppose that u_t is a martingale difference sequence and that $\theta(t/T) = \mu + g(t/T)\eta$, with $|g(t/T) - g(s/T)| \leq |(t-s)/T|c_1$, where c_1 is a constant. Moreover, suppose that $\eta = O(T^b)$ with $0 \leq b \leq 1$. Then we have*

$$\begin{aligned}
T^{-1/2}\underline{CUSUM}_T &= \frac{\eta \sup_r \int_0^r (g(s) - \bar{g})ds}{\sqrt{\sigma^2 + \eta^2 \int_0^1 (g(s) - \bar{g})^2 ds}} + o_p(1) \\
T^{-1}\underline{QS}_T &= \frac{\eta^2 \int_0^1 \left(\int_0^r [g(s) - \bar{g}]\, ds \right)^2 dr}{\sigma^2 + \eta^2 \int_0^1 (g(s) - \bar{g})^2 ds} + o_p(1).
\end{aligned}$$

Although the above result is based on the case where no long-run variance is used or needed, the result can be used to find the order of the statistics that *are* modified by a long run variance estimator. If u_t is modeled as having a general serial correlation structure that

is unknown, we would modify each test by dividing by a function of the long run variance estimator. For the purpose of asymptotic analysis, we introduce the following assumptions:

Assumption 1. *y_t is generated according to (2.1) where $E(u_t) = 0$ and $\theta\left(\frac{t}{T}\right) = \mu + g\left(\frac{t}{T}\right)\eta$ with $|g(t/T) - g(s/T)| \leq |(t-s)/T|c_1$, where c_1 is a constant and $\eta = O(T^b)$, $0 \leq b \leq 1$.*

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

Assumption 3. *$k(\cdot)$ satisfies Assumption A3 in Jansson (2002).*

Assumption 4. *Let $M = \max\{M_1, M_2\}$, where*

$$M_1 \geq \int |u_r u_s u_{s'} u_{r'}|^{1+\delta} dF(u_r, u_{r'}, u_s, u_{s'})$$
$$M_2 \geq \int |u_r u_s u_{s'} u_{r'}|^{1+\delta} dF(u_r) dF(u_{r'}, u_s, u_{s'}),$$

where $M < \infty$.

Assumption 1 is the same as in Theorem 3.1. 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 Hsiao and Li (2001), Fan and Li (1999) and Li (1999). Assumption 3 was used in Jansson (2002), who shows that this is the class of kernels that should be considered for consistent estimation. This class of kernel contains many conventional kernel functions in the literature. Assumption 4 are moment conditions used along with the dependence conditions to find limiting distributions.

We present the initial result in the following theorem.

Theorem 3.2. *Suppose Assumptions 1 - 4 hold and we use the data dependent bandwidth for the long-run variance, if $\eta = T^b$ with $0 < b < 1/2$, we have*

$$CUSUM_T = O_p\left(T^{\frac{q(1-2b)}{2q+1}}\right),$$
$$QS_T = O_p\left(T^{\frac{2q(1-2b)}{2q+1}}\right),$$

where q is the order of the kernel function. If $\eta = T^b$ with $b \geq 1/2$, we have

$$CUSUM_T = O_p(1),$$
$$QS_T = O_p(1).$$

In particular, if $q = 1$,

$$CUSUM_T = O_p\left(T^{\frac{(1-2b)}{3}}\right),\ QS_T = O_p\left(T^{\frac{(2-4b)}{3}}\right),\ when\ 0 < b < 1/2,$$

and if $q = 2$,

$$CUSUM_T = O_p\left(T^{\frac{(2-4b)}{5}}\right),\ QS_T = O_p\left(T^{\frac{(4-8b)}{5}}\right),\ when\ 0 < b < 1/2.$$

Given the results of the above theorem, we see that the divergence rate of the test statistic becomes smaller and smaller as the magnitude of the change (i.e. b) increases, corresponding to less and less power in the presence of a large change. When $b \geq 1/2$, the tests are bounded in probability. In other words, for divergent alternatives, the tests are no longer consistent. In finite samples, this means we may observe very obvious large changes in the mean of a series that will not be detectable by this type of test. As discussed above, several Monte Carlo studies have shown evidence of nonmonotonic power, where large changes are less detectable than small changes. The theorem provides an asymptotic theory to explain the existence of nonmonotinc power for a wide class of changes in mean that are indexed by η. We find this to be a very undesirable feature for the tests, and suggest a remedy in the next section.

4. Modified Tests

In this section, we consider a modification on the existing tests. We use an estimator of the long-run variance that still converges to the true long-run variance even under parameter change (non-constant θ_t). 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_t assumed to be constant. These residuals will be used to construct the partial sums $T^{-1/2}\sum_{t=1}^{[Tr]}\hat{u}_t$. Next, we construct nonparametric regression

residuals $\tilde{u}_t$ based on a nonparametric regression on (2.1). That is, we estimate $\theta(t/T)$ using nonparametric kernel methods,

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

with

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

and $K(\cdot)$ is a kernel function (different from the lag window $k(\cdot)$ used for the long-run variance), and h is a bandwidth parameter such that $h \to 0$. This estimator of θ_t was originally proposed in Priestley and Chao (1972) and is also related to the estimator of Robinson (1989). It differs from the usual nonparametric estimator in that there is no need to have a kernel density estimator in the denominator since θ is a function of time. Alternatively, one could include an estimate of the density in the denominator which would converge (uniformly) to one.

The second set of residuals is defined as

$$\tilde{u}_t = y_t - \tilde{\theta}(t/T).$$

Based on $\tilde{u}_t$, the sample autocovariance function $\tilde{\gamma}(j)$ is then calculated as

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

and long run variance estimated by

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

The modified tests are then constructed based on $\tilde{\omega}^2$ (based on nonparametrically demeaned data) and $\hat{u}_t$ (based on OLS demeaned data). For example, the modified CUSUM and QS tests are based on the following modified empirical process:

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

In particular, we denote the modified tests by

$$CUSUM_T^* = \sup_{0\leq r\leq 1} |\tilde{U}_T(r)|, \; QS_T^* = \int_0^1 \tilde{U}_T(r)^2 dr.$$

In these modified tests, the partial sum process $T^{-1/2}\sum_{t=1}^{[Tr]}\hat{u}_t$ are constructed based on OLS residuals and will capture the fluctuation in the presence of structural changes. The long run variance estimator $\tilde{\omega}^2$ is constructed based on nonparametric demeaned data and, under the null hypothesis of constant mean, should converge to the true long-run variance. Moreover, we posit that under the alternative hypothesis, the modified estimator of the long-run variance diverges at a slower rate than the original unmodified version. Thus tests based on the modified long-run variance estimator are more likely to exhibit monotonic power.[1] Our proposed method uses a nonparametric kernel estimator for $\theta(t/T)$. Moreover, the widely used data-dependent bandwidth b_{opt} can also be calculated based on $\tilde{u}_t$. We propose a modified data dependent bandwidth using the nonparametric residuals $\tilde{u}_t$ and the plug-in formula as Andrews (1991):

$$\tilde{b}_{opt} = \xi(k) \times (\tilde{\alpha}(q)T)^{1/(2q+1)}$$

where $\tilde{\alpha}(q)$ is now estimated based on $\tilde{u}_t$. Let $\tilde{\rho}$ be the AR(1) coefficient estimate from regressing $\tilde{u}_t$ on $\tilde{u}_{t-1}$, $\tilde{\alpha}(q)$ is estimated based on $\tilde{\rho}$. Corresponding to the quadratic spectral kernel,

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

We explore the performance of the modified long-run variance estimator in the following theorems.

Assumption 5. *$K(\cdot)$ is a bounded continuous density such that $\int_{-\infty}^{\infty}|uK(u)|du < \infty$.*

Assumption 6. *Let $h \to 0$ and $T^{3/10}h^{1/2} \to \infty$.*

Theorem 4.1. *Suppose that Assumptions 1-6 hold and let $\tilde{\omega}^2$ be estimated using the nonparametric residuals, under the null hypothesis of no structural change, we have*

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

This theorem states that tests using the modified long-run variance will have the correct size. Moreover, as a minor corollary to this theorem, it is easy to show that the quantity

[1] Related to the idea of local smoothing, Crainiceanu and Vogelsang (2007) suggest using moving averages of the data as a simple estimate of the mean. However, no further analysis is provided and no proofs or regularity conditions are given.

$\tilde{\alpha}(2)$ will converge to the relevant population quantity $\alpha(2)$, so that we can expect to have good size in finite samples.

The purpose in using modified residuals, $\tilde{u}_t$, in the construction of the long-run variance is to maintain consistency of the tests. The next theorem shows that this intuition is correct.

Theorem 4.2. *Suppose that the long-run variance is estimated using $\tilde{\omega}^2$, assumptions 1-5 hold, $\eta = T^b$, and $h = O(T^{-a})$ where $0 < a < 1/3$. Let q be the order of the kernel function, Then we have the following results:*
Case 1: $0 < b \leq a - 1/(2(2q+1))$.

$$CUSUM_T^* = O_p(T^{\frac{1}{2}+b}),\ QS_T^* = O_p(T^{1+2b}).$$

Case 2: $a - 1/(2(2q+1)) < b \leq a$.

$$CUSUM_T^* = O_p(T^{\frac{2aq+q+a}{2q+1}}),\ QS_T^* = O_p(T^{\frac{2(2aq+q+a)}{2q+1}}).$$

Case 3: $a < b \leq a/2 + 1/2$.

$$CUSUM_T^* = O_p(T^{\frac{4qa-2qb+q+a}{2q+1}}),\ QS_T^* = O_p(T^{\frac{2(4qa-2qb+q+a)}{2q+1}}).$$

Case 4: $1/2 + a/2 < b$.

$$CUSUM_T^* = O_p(T^{\frac{3qa+a}{2q+1}}),\ QS_T^* = O_p(T^{\frac{2(a+3qa)}{2q+1}}).$$

The main point of the theorem is that the modified tests are consistent against all alternatives. This consistency holds for large order of $\eta = O(T^b)$. The conventional $CUSUM_T$ and QS_T tests that use the standard long-run variance estimators are not consistent for $b \geq 1/2$. Moreover, even for values of b where the standard tests remain consistent, the modified tests diverge at a faster rate due to the presence of $h = O(T^{-a})$ where $a > 0$. Therefore, we expect the performance of $CUSUM_T^*$ and QS_T^* to dominate $CUSUM_T$ and QS_T for all classes of (non-local) alternatives. We verify this claim in the Monte Carlo section.

5. Monte Carlo

In this section, we conduct a Monte Carlo experiment to examine the effectiveness of the proposed modification. In particular, we compare the original tests (denoted by $CUSUM_T$ and QS_T) with the proposed tests (denoted by $CUSUM_T^*$ and QS_T^*, using long-run variance estimator based nonparametric regression residuals). The kernel and lag truncation used in the long-run variance estimation are the quadratic spectral kernel and the corresponding data-dependent bandwidth $\tilde{b}_{opt}$.

The proposed new tests require a bandwidth parameter h and a kernel $K(\cdot)$. For the kernel used in the nonparametric regression 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. 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

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

We calculate size adjusted power using $h = 2 \times T^{-1/5}$ with $T = 200$ for different values of ρ, and we report the resulting power graphs for the case $\rho = 0.7$ in Figures 1-2.

The results shown in the graphs are striking. The original tests both exhibit non-monotonic power for this experiment, with power going to near zero for large changes in mean. Although we show only the results for $\rho = 0.7$, other simulations show that the effect is more pronounced as the level of serial correlation in u_t increases. However, as predicted from the asymptotic results in the last section, the modified tests alleviate the problem of nonmonotonic power. Moreover, the size-adjusted power of the new tests in the small η range is never less than the size-adjusted power of the original tests. The figures verify our intuition and asymptotic results which suggest the modified tests have some immunity to monotonic power and inconsistency.

The proposed new tests uses a nonparametric estimation of the regression function which involve a choice of bandwidth parameter h. In our Monte Carlo, we considered a simple choice of $h = cT^{-1/5}$ with $c = 1, 2, 3$. The Monte Carlo results indicate that when $c = 2$ and 3, the proposed tests have reasonably good sampling performance. However, when $c = 1$, the tests overreject as ρ increases. This is because that as the temporal dependence in u_t increases, a larger bandwidth is needed. Like many other inference procedures based on nonparametric estimation, the choice of bandwidth h affects the sampling performance and thus it would be helpful to use some data-dependent information in selecting h. Existing data-dependent bandwidth selection procedures are not directly applicable to our inference problem. For practical purpose, notice that the design of regressors are fixed in our model and what affects the bandwidth selection is the serial correlation in the error term. Based on Monte Carlo experiments for different design of the error correlation, it suggests that

the bandwidth choice $h = 2 \times T^{-1/5}$ provides reasonable sampling performance for a wide range of serial correlation. For this reason, we recommend the use of this bandwidth in the nonparametric regression. More complicated ways on bandwidth selection in our model is possible. In principle, a bandwidth selection that optimize the size and power trade-off would require higher order expansion. Alternatively, bootstrap based inference methods may also help. We wish to investigate these alternatives in future research.

6. Empirical Illustration

In this section, we provide an empirical example using our new modified statistics. The series we analyze is the expected median inflation rate. The data was collected by the University of Michigan Survey Research Center and is available on the Federal Reserve Bank of St. Louis web site. Figure 4 shows the data from January of 1978 to September of 2004. From the graph, it appears that there is an obvious change in mean. As is well known, from the early 1980's forward, inflation has a much lower mean and appears to be more stable. We calculate the standard $CUSUM$ and QS tests and find values of 0.913 and 0.362 respectively. The corresponding critical values of these tests are 1.36 and 0.463 respectively. Thus, we fail to reject the null hypothesis of a constant mean over the entire range of the data. Next, we calculate the $CUSUM^*$ and QS^* tests using the modified long-run variance. The values are 1.55 and 1.06 respectively. We reject the null hypothesis in both cases. This example shows how some standard tests will fail to reject the null hypothesis of a constant mean in situations with obvious instability. However, the new modified tests are able to easily detect such a change.

7. Conclusion

Using an asymptotic analysis of large changes in mean, we have shown that the problem of nonmonotonic power in tests of changing mean can be alleviated through the use of nonparametric residuals. The theoretical results show that we can expect the modified tests to reduce the problem due to the local smoothing of the series which provides a better estimate of the residuals *under the alternative* hypothesis. The properties of our

modification are apparent from the Monte Carlo experiment as well as from our empirical example using expected inflation.

The problem of nonmonotonic power affects tests of changing mean. However, this remains a very serious and pervasive issue. In a regression setting, if there are changes in the mean of the regressors, Hansen (2000) shows that there can be very severe empirical problems. Researchers who ignore changes in mean of the regressors will incorrectly find structural change in the regression parameters rather than correctly attributing the change to the mean of the regressors. Hence, it becomes important to test for changes in mean of each regressor, and our tests may serve an important role in such a task.

Appendix A. Proofs

Let $\check{u}_t = (Th)^{-1}\sum_{s=1} u_s K_{ts}$ and $\check{\theta}_t = (Th)^{-1}\sum_{s=1}(\theta(t/T) - \theta(s/T))K_{ts}$.

Lemma A.1. *Under Assumptions 1 - 4,*

$$D = T^{-1}\sum_{t=j+1}^{T} \check{u}_t \check{u}_{t-j} = O_p(T^{-1}h^{-1}).$$

Proof:

$$\begin{aligned} E|D| &\leq \frac{1}{T}\sum_{t=j+1}^{T} E|\check{u}_t\check{u}_{t-j}| \\ &\leq \frac{1}{T}\sum_{t=j+1}^{T} \sqrt{E(\check{u}_t^2)E(\check{u}_{t-j}^2)} \end{aligned}$$

$$\begin{aligned} E(\check{u}_t^2) &= \frac{1}{T^2h^2}\sum_{s=1}^{T} E(u_s^2)K_{ts}^2 + \frac{1}{T^2h^2}\sum\sum_{s\neq s'} E(u_s u_{s'})K_{ts}K_{ts'} \\ &\leq O(T^{-1}h^{-1}) + \frac{2}{T^2h^2}\sum\sum_{s<s'} \beta(s'-s)^{\frac{\delta}{1+\delta}} M^{\frac{1}{1+\delta}} K(0)K_{ts} \\ &= O(T^{-1}h^{-1}) \end{aligned}$$

where $E|u_s u_{s'}| < M^{\frac{1}{1+\delta}}\beta(s'-s)^{\frac{\delta}{1+\delta}}$ comes from an application of Lemma A.5. Then we have

$$\begin{aligned} E|D| &\leq \frac{1}{T}\sum_{t=j+1}^{T} \sqrt{O(T^{-2}h^{-2})} \\ &= O(T^{-1}h^{-1}) \end{aligned}$$

Lemma A.2. *Under Assumptions 1 - 4,*

$$F = \frac{1}{T}\sum_{t=j+1}^{T} \check{u}_t\check{\theta}_t = O_p(T^{-1/2}h^{1/2}\eta).$$

Proof:

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

We have

$$E(\check{u}_t\check{u}_{t'}) = \frac{1}{T^2h^2}\sum_{s=1}^{T}\sum_{s'=1}^{T} K_{ts}K_{t's'}E(u_su_{s'}) \leq \frac{1}{T^2h^2}\sum_{s=1}^{T}\sum_{s'=1}^{T} K_{ts}K_{t's'}|E(u_su_{s'})|$$
$$\sim \frac{1}{T^2h^2}\sum_{s'=s+1}^{T} K_{t's'}\sum_{s=1}^{T-1} K(0)\beta(s'-s)^{\frac{\delta}{1+\delta}}M^{\frac{1}{1+\delta}} = O(T^{-1}h^{-1}),$$

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

$$\begin{aligned}\check{\theta}_t &= \frac{1}{Th}\sum_{s=1}^{T}(\theta(t/T)-\theta(s/T))K_{ts}\\ &\sim \frac{\eta}{h}\int_0^1 (g(u)-g(v))K\left(\frac{u-v}{h}\right)dv \leq \frac{\eta}{h}\int_0^1 |u-v|c_1K\left(\frac{u-v}{h}\right)dv\\ &= \frac{\eta}{h}\int_{-u/h}^{(1-u)/h} |wh|c_1K(w)\,dw\; h = O(\eta h).\end{aligned}$$

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

Lemma A.3. *Under Assumptions 1 - 4,*

$$G = \frac{1}{T}\sum_{t=1}^{T} u_t\check{\theta}_t = O_p(T^{-1/2}\eta h).$$

Proof:

$$\begin{aligned}E(G^2) &= \frac{1}{T^2}\sum_{t=1}^{T} E(u_t^2)\check{\theta}_t^2\\ &\quad + \frac{1}{T^2}\sum\sum_{t\neq t'}^{T} E(u_tu_{t'})\check{\theta}_t\check{\theta}_{t'}\\ &= G_1 + G_2\end{aligned}$$

Since $\check{\theta}_t = O(\eta h)$, $G_1 = O(T^{-1}\eta^2h^2)$. Then

$$G_2 \leq \frac{1}{T^2}\sum\sum_{t\neq t'}^{T} |E(u_tu_{t'})|O(\eta^2h^2) \leq \frac{1}{T^2}\sum\sum_{t\neq t'}^{T} \beta(t-t')^{\frac{\delta}{1+\delta}}M^{\frac{1}{1+\delta}}O(\eta^2h^2) = O(T^{-1}\eta^2h^2),$$

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

Lemma A.4. *Under Assumptions 1 - 4,*

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

Proof:

$$
\begin{aligned}
E(H^2) &\leq \frac{1}{T^2}\sum_{t=1}^{T} E(u_t^2 \check{u}_{t-j}^2) \\
&\quad + \frac{1}{T^2}\sum_{t\neq t'}^{T}\sum^{T} E(u_t u_{t'} \check{u}_{t-j}\check{u}_{t'-j}) \\
&= \frac{1}{T^4 h^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{s'=1}^{T} E(u_t^2 u_s u_{s'}) K_{t-j,s} K_{t-j,s'} \\
&\quad + \frac{1}{T^4 h^2}\sum_{t\neq t'}^{T}\sum^{T}\sum_{s=1}^{T}\sum_{s'=1}^{T} E(u_t u_{t'} u_s u_{s'}) K_{t-j,s} K_{t'-j,s'} \\
&= H_1 + H_2.
\end{aligned}
$$

Consider H_2. Without loss of generality, suppose that $t < t' < s < s'$. There are several subcases:

$$
\begin{aligned}
&a)\ t' - t \geq \max\{s - t', s' - s\} \\
&b)\ s - t' \geq \max\{t' - t, s' - s\} \\
&c)\ s' - s \geq \max\{t' - t, s - t'\}
\end{aligned}
$$

For case a), we have

$$
\begin{aligned}
H_2 &\leq \frac{1}{T^4 h^2}\sum_{t=1}^{T-3}\sum_{t'=t+1}^{T-2}\sum_{s=t'+1}^{T-1}\sum_{s'=s+1}^{T} E|u_t u_{t'} u_s u_{s'}| K_{t-j,s} K_{t'-j,s'} \\
&\leq \frac{1}{T^4 h^2}\sum_{t=1}^{T-3}\sum_{t'=t+1}^{T-2}\sum_{s=t'+1}^{T-1}\sum_{s'=s+1}^{T} \beta(t'-t)^{\frac{\delta}{1+\delta}} M^{\frac{1}{1+\delta}} K_{t-j,s} K_{t'-j,s'} \\
&\leq \frac{1}{T^4 h^2}\sum_{t=1}^{T-3}\sum_{t'=t+1}^{T-2} (t'-t)\beta(t'-t)^{\frac{\delta}{1+\delta}} \sum_{s=1}^{T} K(0)^2 \\
&= O(T^{-2}h^{-2})
\end{aligned}
$$

where the second inequality comes from Lemma A.5 and the last line from Assumption 2. The other cases are similar and the proof of H_1 is similar so that $H = O_p(T^{-1}h^{-1})$.

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:

$$\hat{u}_t = y_t - \bar{y} = u_t - \bar{u} + \theta(t/T) - \bar{\theta}_T$$

where $\bar{\theta}_T = \frac{1}{T}\sum_{t=1}^{T} \theta(t/T)$. The partial sums are such that

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} \hat{u}_t \Rightarrow W(r) - \bar{W} + \sqrt{T}\eta \int_0^r (g(s) - \bar{g}) ds + o(\sqrt{T}\eta)$$

In addition, we have

$$\frac{1}{T}\sum_{t=1}^{T} \hat{u}_t^2 = \frac{1}{T}\sum_{t=1}^{T} (u_t - \bar{u})^2 + \frac{2}{T}\sum_{t=1}^{T} (\theta(t/T) - \bar{\theta}_T)(u_t - \bar{u}) + \frac{1}{T}\sum_{t=1}^{T} (\theta(t/T) - \bar{\theta}_T)^2$$
$$= \sigma^2 + o_p(1) + H_1 + H_2$$

We have

$$H_1 = \frac{\eta}{\sqrt{T}} \frac{2}{\sqrt{T}} \sum_{t=1}^{T} (g(t/T) - \bar{g}_T)(u_t - \bar{u})$$
$$\xrightarrow{d} \frac{\eta}{\sqrt{T}} X$$

where the variable X has a normal distribution with variance $4\int_0^1 (g(s) - \bar{g})^2 ds$, and we have $H_2 \to \eta^2 \int_0^1 (g(s) - \bar{g})^2 ds$ so that

$$\frac{1}{T}\sum_{t=1}^{T} \hat{u}_t^2 \xrightarrow{d} \sigma^2 + \frac{\eta}{\sqrt{T}} X + \eta^2 \int_0^1 (g(s) - \bar{g})^2 ds.$$

The result follows from the Continuous Mapping Theorem.

Proof of Theorem 3.2: Let $\hat{\gamma}(j) = T^{-1}\sum_{t=j+1}^{T} u_t u_{t-j}$. We find the order of

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

in terms of b_T and then we will find the order of b_T to complete the proof.

We can write

$$\hat{\gamma}(j) = \frac{1}{T}\sum_{t=j+1}^{T}(u_t - \bar{u}_T)(u_{t-j} - \bar{u}_T) + \frac{1}{T}\sum_{t=j+1}^{T}(\theta(t/T) - \bar{\theta}_T)(u_{t-j} - \bar{u}_T)$$
$$+ \frac{1}{T}\sum_{t=j+1}^{T}(u_t - \bar{u}_T)(\theta((t-j)/T - \bar{\theta}_T) + \frac{1}{T}\sum_{t=j+1}^{T}(\theta(t/T) - \bar{\theta}_T)(\theta((t-j)/T) - \bar{\theta}_T)$$

The middle two terms are of order $\frac{\eta}{\sqrt{T}}$ from the proof of Theorem 3.1. The last term is the dominant term and is of order η^2.

Now $b_T^{-1}\sum_{j=-T+1}^{T-1}|k(j/b_T)| \to \int_{-\infty}^{\infty}|k(x)|dx$, which implies that

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

Now we find the order of the optimal bandwidth, which is a function of $\hat{\rho} - 1$

$$\hat{\gamma}_1 - \hat{\gamma}_0 = \gamma_1 - \gamma_0 + \frac{1}{T}\sum_{t=2}^{T}(u_{t-1} - \bar{u}_T)(\theta(t/T) - \bar{\theta}_T) + \frac{1}{T}\sum_{t=2}^{T}(u_t - \bar{u}_T)(\theta((t-1)/T) - \bar{\theta}_T)$$
$$+ \frac{1}{T}\sum_{t=2}^{T}(\theta(t/T) - \bar{\theta}_T)(\theta((t-1)/T) - \bar{\theta}_T) - \frac{1}{T}\sum_{t=1}^{T}(u_t - \bar{u}_T)(\theta(t/T) - \bar{\theta}_T)$$
$$- \frac{1}{T}\sum_{t=1}^{T}(u_t - \bar{u}_T)(\theta(t/T) - \bar{\theta}_T) - \frac{1}{T}\sum_{t=1}^{T}(\theta(t/T) - \bar{\theta}_T)(\theta(t/T) - \bar{\theta}_T) + o_p(1)$$

Define

$$J_1 = \frac{1}{T}\sum_{t=2}^{T}(u_{t-1} - \bar{u}_T)(\theta(t/T) - \bar{\theta}_T) - \frac{1}{T}\sum_{t=1}^{T}(u_t - \bar{u}_T)(\theta(t/T) - \bar{\theta}_T)$$
$$J_2 = \frac{1}{T}\sum_{t=2}^{T}(u_t - \bar{u}_T)(\theta((t-1)/T) - \bar{\theta}_T) - \frac{1}{T}\sum_{t=1}^{T}(u_t - \bar{u}_T)(\theta(t/T) - \bar{\theta}_T)$$
$$J_3 = \frac{1}{T}\sum_{t=2}^{T}(\theta(t/T) - \bar{\theta}_T)(\theta((t-1)/T) - \bar{\theta}_T) - \frac{1}{T}\sum_{t=1}^{T}(\theta(t/T) - \bar{\theta}_T)(\theta(t/T) - \bar{\theta}_T)$$

so that

$$\hat{\gamma}_1 - \hat{\gamma}_0 = \gamma_1 - \gamma_0 + J_1 + J_2 + J_3 + o_p(1).$$

For J_2 we have

$$
\begin{aligned}
E|J_2| &\leq \frac{\eta}{T} E\left|(u_1-\bar{u})\right| \left|(g(1/T)-\bar{g}_T)\right| + \frac{\eta}{T}\sum_{t=2}^{T} E\left|(u_t-\bar{u}_T)\right| \left|(g((t-1)/T)-g(t/T))\right| \\
&\sim \frac{\eta}{T}\sum_{t=2}^{T} E\left|(u_t-\bar{u}_T)\right| \frac{1}{T} \\
&= O\left(\frac{\eta}{T}\right)
\end{aligned}
$$

The analysis of J_1 is similar. Then

$$
\begin{aligned}
|J_3| &\leq \frac{\eta^2}{T}(g(1)-\bar{g}_T)^2 + \frac{\eta^2}{T}\sum_{t=2}^{T} \left|g((t-1)/T)-\bar{g}_T\right| \left|g((t-1)/T)-g(t/T)\right| \\
&\leq \frac{\eta^2}{T}\sum_{t=2}^{T} \left|g((t-1)/T)-\bar{g}_T\right| \frac{1}{T} \\
&= O\left(\frac{\eta^2}{T}\right).
\end{aligned}
$$

The order of $\hat{\gamma}_0$ is $O_p(\eta^2)$ from the proof of Theorem 3.1.

We consider three cases for the behavior of η.

In particular, in the case when $q = 2$,

Case 1: $\frac{\eta^2}{T} \to \infty$. Hence, the dominant term in $\hat{\gamma}_1 - \hat{\gamma}_0$ is η^2/T so that $\hat{\rho} - 1 = O_p(\frac{1}{T})$ and then

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

so that

$$
\begin{aligned}
b_{opt} &= O_p((\hat{\alpha}(2)T)^{1/5}) \\
&= O_p(T)
\end{aligned}
$$

Setting $b_T = b_{opt}$ gives $\hat{\omega}^2 = O_p(T\eta^2)$ so that

$$
\begin{aligned}
CUSUM_T &= O_p\left(\frac{T^{1/2}\eta}{T^{1/2}\eta}\right) \\
QS_T &= O_p\left(\frac{T\eta^2}{T\eta^2}\right)
\end{aligned}
$$

Case 2: $\frac{\eta^2}{T} \to 0$. The dominating term in $\hat{\gamma}_1 - \hat{\gamma}_0$ is $\gamma_1 - \gamma_0$ so that $\hat{\rho} - 1 = O_p(\eta^{-2})$. and then

$$\hat{\alpha}(2) = \frac{4\hat{\rho}^2}{(1-\hat{\rho})^4} = O_p(\eta^8)$$

so that

$$\begin{aligned} b_{opt} &= O_p((\hat{\alpha}(2)T)^{1/5}) \\ &= O_p(\eta^{8/5}T^{1/5}) \end{aligned}$$

where $\eta = T^b$. Setting $b_T = b_{opt}$ gives $\hat{\omega}^2 = O_p(\eta^{8/5}T^{1/5}\eta^2)$ so that

$$\begin{aligned} CUSUM_T &= O_p\left(\frac{T^{1/2}\eta}{T^{(1+8b)/10}\eta}\right) \\ QS_T &= O_p\left(\frac{T\eta^2}{T^{(1+8b)/5}\eta^2}\right) \end{aligned}$$

Case 3: $\eta = T^{1/2}$. This case is similar to Case 1.

In general, notice that the component that is determined by the plug-in procedure can be written as $\hat{\alpha}(q) = g(\hat{\rho})/(1-\hat{\rho})^{2q}$, where $g(\hat{\rho}) = O_p(1)$. when $\frac{\eta^2}{T} \to \infty$, Hence, the dominant term in $\hat{\gamma}_1 - \hat{\gamma}_0$ is η^2/T so that $\hat{\rho} - 1 = O_p(\frac{1}{T})$ and then

$$\hat{\alpha}(q) = g(\hat{\rho})\frac{1}{(1-\hat{\rho})^{2q}} = O_p(T^{2q})$$

so that

$$b_{opt} = O_p((\hat{\alpha}(q)T)^{1/(2q+1)}) = O_p(T).$$

Setting $b_T = b_{opt}$ gives $\hat{\omega}^2 = O_p(T\eta^2)$ so that

$$CUSUM_T = O_p(1), \; QS_T = O_p(1).$$

Case 2: $\frac{\eta^2}{T} \to 0$. The dominating term in $\hat{\gamma}_1 - \hat{\gamma}_0$ is $\gamma_1 - \gamma_0$ so that $\hat{\rho} - 1 = O_p(\eta^{-2})$. and then

$$\hat{\alpha}(q) = g(\hat{\rho})\frac{1}{(1-\hat{\rho})^{2q}} = O_p(\eta^{4q})$$

so that

$$b_{opt} = O_p((\hat{\alpha}(q)T)^{1/5}) = O_p(\eta^{4q/(2q+1)}T^{1/(2q+1)})$$

where $\eta = T^b$. Setting $b_T = b_{opt}$ gives $\hat{\omega}^2 = O_p(\eta^{4q/(2q+1)}T^{1/(2q+1)}\eta^2)$ so that

$$CUSUM_T = O_p\left(\frac{T^{1/2}\eta}{\eta^{2q/(2q+1)}T^{1/(4q+2)}\eta}\right) = O_p\left(T^{q(1-2b)/(2q+1)}\right)$$
$$QS_T = O_p\left(\frac{T\eta^2}{\eta^{4q/(2q+1)}T^{1/(2q+1)}\eta^2}\right) = O_p\left(T^{2q(1-2b)/(2q+1)}\right)$$

Case 3: $\eta = T^{1/2}$. This case is similar to Case 1.

Proof of Theorem 4.1:

$$\begin{aligned}\tilde{\gamma}(j) &= \frac{1}{T}\sum_{t=j+1}^{T}(u_t - \check{u}_t)(u_{t-j} - \check{u}_{t-j}) \\ &= \frac{1}{T}\sum_{t=j+1}^{T} u_t u_{t-j} - \frac{1}{T}\sum_{t=j+1}^{T} u_t \check{u}_{t-j} \\ &\quad - \frac{1}{T}\sum_{t=j+1}^{T} \check{u}_t u_{t-j} + \frac{1}{T}\sum_{t=j+1}^{T} \check{u}_t \check{u}_{t-j} \\ &= \frac{1}{T}\sum_{t=j+1}^{T} u_t u_{t-j} + O_p(T^{-1/2}h^{-1/2}) + O_p(T^{-1}h^{-1})\end{aligned}$$

from Lemmas A.1 and A.4. Then

$$\sum_{j=-T+1}^{T-1} k\left(\frac{j}{b_T}\right)\left(\tilde{\gamma}(j) - \frac{1}{T}\sum_{t=j+1}^{T} u_t u_{t-j}\right) = O_p(b_T T^{-1/2}h^{-1/2} + b_T T^{-1}h^{-1})$$

We find the order of $\tilde{b}_{opt}$ using the residuals from the nonparametric regression. Note that under the null, $y_t - \tilde{y}_t = u_t - \check{u}_t$ as defined in the beginning of the Appendix. We have

$$\begin{aligned}\tilde{\gamma}_1 - \tilde{\gamma}_0 &= \frac{1}{T}\sum_{t=2}^{2} u_t u_{t-1} - \frac{1}{T}\sum_{t=2}^{T} u_t \check{u}_{t-1} - \frac{1}{T}\sum_{t=2}^{T} \check{u}_t u_{t-1} + \frac{1}{T}\sum_{t=2}^{T} \check{u}_t \check{u}_{t-1} \\ &\quad - \frac{1}{T}\sum_{t=1}^{T} u_t^2 + \frac{2}{T}\sum_{t=2}^{T} u_t \check{u}_t - \frac{1}{T}\sum_{t=2}^{T} \check{u}_t \check{u}_t \\ &= \gamma_1 - \gamma_0 + O_p(T^{-1/2}h^{-1/2} + T^{-1}h^{-1})\end{aligned}$$

and similarly,

$$\tilde{\gamma}_0 = \gamma_0 + O_p(T^{-1/2}h^{-1/2} + T^{-1}h^{-1})$$

so that $\tilde{b}_{opt} = O(T^{1/5})$ under the null hypothesis. Replacing b_T with $\tilde{b}_{opt}$, we have

$$\begin{aligned}\sum_{j=-T+1}^{T-1} k\left(\frac{j}{b_T}\right)\left(\tilde{\gamma}(j) - \frac{1}{T}\sum_{t=j+1}^{T} u_t u_{t-j}\right) &= O_p(T^{1/5}T^{-1/2}h^{-1/2} + T^{1/5}T^{-1}h^{-1}) \\ &= o_p(1)\end{aligned}$$

by the assumptions on the bandwidth so that $\tilde{\omega}^2$ is consistent.

Proof of Theorem 4.2:

$$\begin{aligned}\tilde{\gamma}(j) &= \frac{1}{T}\sum_{t=j+1}^{T}(u_t - \check{u}_t + \check{\theta}_t)(u_{t-j} - \check{u}_{t-j} + \check{\theta}_{t-j}) \\ &= \frac{1}{T}\sum_{t=j+1}^{T} u_t u_{t-j} - \frac{1}{T}\sum_{t=j+1}^{T} u_t \check{u}_{t-j} + \frac{1}{T}\sum_{t=j+1}^{T} u_t \check{\theta}_{t-j} \\ &\quad - \frac{1}{T}\sum_{t=j+1}^{T} \check{u}_t u_{t-j} + \frac{1}{T}\sum_{t=j+1}^{T} \check{u}_t \check{u}_{t-j} - \frac{1}{T}\sum_{t=j+1}^{T} \check{u}_t \check{\theta}_{t-j} \\ &\quad + \frac{1}{T}\sum_{t=j+1}^{T} \check{\theta}_t u_{t-j} - \frac{1}{T}\sum_{t=j+1}^{T} \check{\theta}_t \check{u}_{t-j} + \frac{1}{T}\sum_{t=j+1}^{T} \check{\theta}_t \check{\theta}_{t-j} \\ &= \gamma_j + O_p(T^{-1/2}h^{1/2}\eta + T^{-1}h^1\eta + T^{-1}h^{-1} + T^{-1/2}h^{-1/2}) + O(h^2\eta^2)\end{aligned}$$

from Lemmas A.1-A.4 and the fact that $\check{\theta}_t = O(h\eta)$. Since $h = T^{-a}$ and $0 < a < 1/3$, the dominant term is either a constant or $\eta^2 h^2$ depending on whether $\eta^2 h^2$ diverges. Therefore, the maximum order is

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

We find the order of $\tilde{b}_{opt}$ using $\tilde{\gamma}_j$ under the alternative hypothesis. We have

$$\begin{aligned}\tilde{\gamma}_1 - \tilde{\gamma}_0 &= \frac{1}{T}\sum_{t=2}^{T}(u_t - \check{u}_t + \check{\theta}_t)(u_{t-1} - \check{u}_{t-1} + \check{\theta}_{t-1}) - \frac{1}{T}\sum_{t=1}^{T}(u_t - \check{u}_t + \check{\theta}_t)^2 \\ &= \frac{1}{T}\sum_{t=2}^{T} u_t u_{t-1} - \frac{1}{T}\sum_{t=1}^{T} u_t^2 - \frac{1}{T}\sum_{t=2}^{T} u_t \check{u}_{t-1} + \frac{1}{T}\sum_{t=1}^{T} u_t \check{u}_t + \frac{1}{T}\sum_{t=2}^{T} u_t \check{\theta}_{t-1} - \frac{1}{T}\sum_{t=1}^{T} u_t \check{\theta}_t \\ &\quad - \frac{1}{T}\sum_{t=2}^{T} \check{u}_t u_{t-1} + \frac{1}{T}\sum_{t=1}^{T} \check{u}_t u_t + \frac{1}{T}\sum_{t=2}^{T} \check{u}_t \check{u}_{t-1} - \frac{1}{T}\sum_{t=1}^{T} \check{u}_t^2 - \frac{1}{T}\sum_{t=2}^{T} \check{u}_t \check{\theta}_{t-1} + \frac{1}{T}\sum_{t=1}^{T} \check{u}_t \check{\theta}_t \\ &\quad + \frac{1}{T}\sum_{t=2}^{T} \check{\theta}_t u_{t-1} - \frac{1}{T}\sum_{t=1}^{T} \check{\theta}_t u_t - \frac{1}{T}\sum_{t=2}^{T} \check{\theta}_t \check{u}_{t-1} + \frac{1}{T}\sum_{t=1}^{T} \check{\theta}_t \check{u}_t + \frac{1}{T}\sum_{t=2}^{T} \check{\theta}_t \check{\theta}_{t-1} - \frac{1}{T}\sum_{t=1}^{T} \check{\theta}_t \check{\theta}_t \\ &= L_1 - L_2 + L_3 - L_4 + L_5 - L_6 + L_7 - L_8 + L_9\end{aligned}$$

where the L terms are defined as

$$L_1 = \frac{1}{T}\sum_{t=2}^{T} u_t u_{t-1} - \frac{1}{T}\sum_{t=1}^{T} u_t^2$$
$$L_2 = \frac{1}{T}\sum_{t=2}^{T} u_t \check{u}_{t-1} - \frac{1}{T}\sum_{t=1}^{T} u_t \check{u}_t$$
$$L_3 = \frac{1}{T}\sum_{t=2}^{T} u_t \check{\theta}_{t-1} - \frac{1}{T}\sum_{t=1}^{T} u_t \check{\theta}_t$$
$$L_4 = \frac{1}{T}\sum_{t=2}^{T} \check{u}_t u_{t-1} - \frac{1}{T}\sum_{t=1}^{T} \check{u}_t u_t$$
$$L_5 = \frac{1}{T}\sum_{t=2}^{T} \check{u}_t \check{u}_{t-1} - \frac{1}{T}\sum_{t=1}^{T} \check{u}_t^2$$
$$L_6 = \frac{1}{T}\sum_{t=2}^{T} \check{u}_t \check{\theta}_{t-1} - \frac{1}{T}\sum_{t=1}^{T} \check{u}_t \check{\theta}_t$$
$$L_7 = \frac{1}{T}\sum_{t=2}^{T} \check{\theta}_t u_{t-1} - \frac{1}{T}\sum_{t=1}^{T} \check{\theta}_t u_t$$
$$L_8 = \frac{1}{T}\sum_{t=2}^{T} \check{\theta}_t \check{u}_{t-1} - \frac{1}{T}\sum_{t=1}^{T} \check{\theta}_t \check{u}_t$$
$$L_9 = \frac{1}{T}\sum_{t=2}^{T} \check{\theta}_t \check{\theta}_{t-1} - \frac{1}{T}\sum_{t=1}^{T} \check{\theta}_t \check{\theta}_t$$

We find the convergence of each term in turn. First, $L_1 = \gamma_1 - \gamma_0 + o_p(1)$. The terms L_2 and L_4 are $O_p(T^{-1}h^{-1})$ by Lemma A.4. L_5 is $O_p(T^{-1}h^{-1})$ by Lemma A.1. We can write

$$\begin{aligned}
L_3 &= \frac{1}{T}\sum_{t=2}^{T} u_t \check{\theta}_{t-1} - \frac{1}{T}\sum_{t=1}^{T} u_t \check{\theta}_t \\
&= \frac{1}{T}\sum_{t=2}^{T} u_t \left(\check{\theta}_{t-1} - \check{\theta}_t\right) - \frac{1}{T} u_1 \check{\theta}_1 \\
&\leq \frac{1}{T}\sum_{t=2}^{T} |u_t| |\left(\check{\theta}_{t-1} - \check{\theta}_t\right)| + \frac{1}{T}|u_1||\check{\theta}_1| \\
&\leq \frac{1}{T}\sum_{t=2}^{T} |u_t| |\frac{\eta}{T}| + \frac{1}{T}|u_1||\check{\theta}_1| \\
&= O_p\left(\frac{\eta}{T}\right) + O_p\left(\frac{\eta h}{T}\right)
\end{aligned}$$

The order of L_7 is the same as L_3. Similarly, we have

$$
\begin{aligned}
L_6 &= \frac{1}{T}\sum_{t=2}^{T}\check{u}_t\check{\theta}_{t-1} - \frac{1}{T}\sum_{t=1}^{T}\check{u}_t\check{\theta}_t \\
&= \frac{1}{T}\sum_{t=2}^{T}\check{u}_t\left(\check{\theta}_{t-1}-\check{\theta}_t\right) - \frac{1}{T}\check{u}_1\check{\theta}_1 \\
&\leq \frac{1}{T}\sum_{t=2}^{T}|\check{u}_t||\left(\check{\theta}_{t-1}-\check{\theta}_t\right)| + \frac{1}{T}|\check{u}_1||\check{\theta}_1| \\
&\leq \frac{1}{T}\sum_{t=2}^{T}|u_t||\frac{\eta}{T}| + \frac{1}{T}|u_1||\check{\theta}_1| \\
&= O_p\left(T^{-1/2}h^{-1/2}\right)\frac{\eta}{T} + O_p\left(T^{-1/2}h^{-1/2}\right)\frac{\eta h}{T}.
\end{aligned}
$$

The order of L_8 is identical to L_6. Finally,

$$
\begin{aligned}
L_9 &= \frac{1}{T}\sum_{t=2}^{T}\check{\theta}_t\check{\theta}_{t-1} - \frac{1}{T}\sum_{t=1}^{T}\check{\theta}_t\check{\theta}_t \\
&= \frac{1}{T}\sum_{t=2}^{T}\check{\theta}_t(\check{\theta}_{t-1}-\check{\theta}_t) - \frac{1}{T}\check{\theta}_1^2 \\
&\leq \frac{1}{T}\sum_{t=2}^{T}|\check{\theta}_t||(\check{\theta}_{t-1}-\check{\theta}_t)| + \frac{1}{T}\check{\theta}_1^2 \\
&\leq \frac{1}{T}\sum_{t=2}^{T}|\check{\theta}_t|\frac{\eta}{T} + \frac{1}{T}\check{\theta}_1^2 \\
&= O\left(\frac{\eta^2 h}{T}\right) + O\left(\frac{\eta^2 h^2}{T}\right)
\end{aligned}
$$

Since $h = T^{-a}$ with $a < 1/3$, the dominant term of $\tilde{\gamma}_0$ is $\eta^2 h^2$ or a constant. The dominant terms of $\tilde{\gamma}_1 - \tilde{\gamma}_0$ are η/T or $\eta^2 h/T$ or a constant. Recall that $\tilde{\rho} - 1 = (\tilde{\gamma}_1 - \tilde{\gamma}_0)/\tilde{\gamma}_0$. Since $\eta = T^b$ and $h = T^{-a}$, we have the following cases:

Case 1: $0 < b \leq a - 1/(2(2q+1))$, then η/T dominates $\eta^2 h/T$ but $\eta/T \to 0$. Hence $\tilde{\rho} - 1 = O_p(1)$ and $\tilde{b}_{opt} = O_p(T^{1/(2q+1)})$. Then

$$\begin{aligned}\tilde{\omega}^2 &= \sum_{j=-T+1}^{T-1} k\left(\frac{j}{\tilde{b}_{opt}}\right)\left(\gamma_j + O_p(\eta^2 h^2)\right) \\ &= \omega^2 + O_p(T^{1/(2q+1)}\eta^2 h^2) \\ &= \omega^2 + O_p(1)\end{aligned}$$

thus

$$CUSUM_T^* = O_p(T^{\frac{1}{2}+b}),\ QS_T^* = O_p(T^{1+2b}).$$

Case 2: $a - 1/(2(2q+1)) < b \leq a$. Then η/T dominates $\eta^2 h/T$, and $\tilde{b}_{opt} = O_p(T^{1/(2q+1)})$. But now

$$\begin{aligned}\tilde{\omega}^2 &= \sum_{j=-T+1}^{T-1} k\left(\frac{j}{\tilde{b}_{opt}}\right)\left(\gamma_j + O_p(\eta^2 h^2)\right) \\ &= O_p(T^{1/(2q+1)}\eta^2 h^2)\end{aligned}$$

where $T^{1/(2q+1)}\eta^2 h^2 \to \infty$.

$$CUSUM_T^* = O_p(T^{\frac{2aq+q+a}{2q+1}}),\ QS_T^* = O_p(T^{\frac{2(2aq+q+a)}{2q+1}}).$$

Case 3: $a < b \leq a/2 + 1/2$. Then $\eta^2 h/T$ dominates η/T and $\tilde{\gamma}_0 = O_p(T^{2(b-a)})$, $\tilde{\gamma}_1 - \tilde{\gamma}_0 = O_p(1)$. Then $\tilde{\rho} - 1 = O_p(\eta^{-2}h^{-2})$ so that $\tilde{b}_{opt} = O_p(T^{(1+4q(b-a))/(2q+1)})$. This implies that the order of $\tilde{\omega}^2$ is $O_p(T^{(1+4q(b-a))/(2q+1)}\eta^2 h^2)$. Thus

$$CUSUM_T^* = O_p(T^{\frac{1+2a}{2}} T^{-(1+4q(b-a))/(4q+2)}) = O_p(T^{\frac{4qa-2qb+q+a}{2q+1}}),$$

and

$$QS_T^* = O_p(T^{\frac{2(4qa-2qb+q+a)}{2q+1}}).$$

Case 4: $1/2 + a/2 < b$. Again $\tilde{\gamma}_0 = O_p(T^{2(b-a)})$, but

$$\tilde{\gamma}_1 - \tilde{\gamma}_0 = O_p(T^{2b-a-1})$$

Thus $\tilde{b}_{opt} = O_p(T^{(1+2q(1-a))/(2q+1)})$, and the order of $\tilde{\omega}^2$ is $O_p(T^{(1+2q(1-a))/(2q+1)}\eta^2 h^2)$. Thus

$$CUSUM_T^* = O_p(T^{\frac{1+2a}{2}} T^{-(1+2q(1-a))/(4q+2)}) = O_p(T^{\frac{3qa+a}{2q+1}}),$$

and

$$QS_T^* = O_p(T^{\frac{2(a+3qa)}{2q+1}}).$$

In the Special case with a quadratic kernel, we have

Case 1: $0 < b \leq a - 1/10$. $\tilde{b}_{opt} = O_p(T^{1/5})$. Then

$$\tilde{\omega}^2 = \omega^2 + O_p(T^{1/5}\eta^2 h^2) = \omega^2 + O_p(1)$$

where the second to last line comes from the fact that $\sum_{j=-T+1}^{T-1} k\left(\frac{j}{\tilde{b}_{opt}}\right) = O_p(\tilde{b}_{opt}) = O_p(T^{1/5})$ and the last line is because

$$\begin{aligned} T^{1/5}\eta^2 h^2 &= T^{1/5}T^{2b}T^{-2a} \\ &\leq T^{1/5}T^{2a-1/5}T^{-2a} \\ &= 1. \end{aligned}$$

Case 2: $0 < a - 1/10 < b \leq a$. Then

$$\tilde{\omega}^2 = \sum_{j=-T+1}^{T-1} k\left(\frac{j}{\tilde{b}_{opt}}\right)\left(\gamma_j + O_p(\eta^2 h^2)\right) = O_p(T^{1/5}\eta^2 h^2)$$

where $T^{1/5}\eta^2 h^2 \to \infty$.

Case 3: $a < b \leq a/2 + 1/2$. Then $\eta^2 h/T$ dominates η/T but

$$\begin{aligned} \eta^2 h/T &= T^{2b}T^{-a}/T \\ &= \frac{T^{2b}}{T^{1+a}} \\ &\to 0 \end{aligned}$$

since $2b < 1 + a$. Then $\tilde{\rho} - 1 = O_p(\eta^{-2}h^{-2})$ so that $\tilde{\alpha}(2) = O_p(\eta^8 h^8)$ and $\tilde{b}_{opt} = O_p(\eta^{8/5}h^{8/5}T^{1/5})$. This implies that the order of $\tilde{\omega}^2$ is $O_p(T^{1/5}\eta^{8/5}h^{8/5}\eta^2 h^2)$.

Case 4: $1/2 + a/2 < b$. Now the dominant term in the numerator of $\tilde{\rho} - 1$ is $\eta^2 h/T = \frac{T^{2b}}{T^{1+a}} \to \infty$. So $\tilde{\rho} - 1 = O_p(T^{-1}h^{-1})$ and $\tilde{b}_{opt} = O_p(Th^{4/5})$ so that $\tilde{\omega}^2 = O_p(Th^{4/5}\eta^2 h^2)$.

The proof is completed by $\hat{\omega}^2$ from Theorem 3.2 with $\tilde{\omega}^2$.

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 (2007): "Nonmonotonic Power for Tests of a Mean Shift in a Time Series," *Journal of Statistical Computation and Simulation*, pp. 457–476.

Deng, A., and P. Perron (2007): "A Non-local Perspective on the Power Properties of the CUSUM and CUSUM of Squares Tests for Structural Change," Boston University Department of Economics.

Elliott, G., and U. K. Muller (2006): "Efficient Tests for General Persistent Time Variation in Regression Coefficients," *Review of Economic Studies*, 73, 907–940.

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.

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

Hansen, B. E. (1992): "Tests for Parameter Instability in Regressions with I(1) Processes," *Journal of Business and Economic Statistics*, 10, 321–335.

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

Hsiao, C., and Q. Li (2001): "A Consistent Test for Conditional Heteroskedasticity," *Econometric Theory*, 17, 188–221.

Jansson, M. (2002): "Consistent Covariance Matrix Estimation for Linear Processes," *Econometric Theory*, 18, 1449–1460.

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. (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.

Priestley, M., and M. Chao (1972): "Non-parametric Function Fitting," *Journal of the Royal Statistical Society, Series B*, 34, 386–92.

Robinson, P. M. (1989): "Nonparametric Estimation of Time-Varying Parameters," in *Statistical Analysis and Forecasting of Economic Structural Change*, ed. by P. Hackl. Springer Verlag: New York.

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

Yoshihara, K. (1976): "Limiting Behavior of U-statistics for Stationary, Absolutely Regular Processes," *Zeitschrift fur Wahrscheinlichkeitstheorie und verwandte Gebiete*, 35, 237–252.

Table 1: Size

c	T	ρ	CUSUM$_T$	QS$_T$	CUSUM$^*_T$	QS$^*_T$
	100	0.0	0.029	0.050	0.050	0.063
	100	0.5	0.014	0.052	0.092	0.121
	100	0.7	0.005	0.045	0.153	0.187
	200	0.0	0.039	0.055	0.048	0.062
1	200	0.5	0.040	0.066	0.094	0.110
	200	0.7	0.022	0.069	0.123	0.156
	300	0.0	0.036	0.051	0.046	0.059
	300	0.5	0.042	0.065	0.080	0.097
	300	0.7	0.030	0.055	0.097	0.115
	100	0.0	0.029	0.050	0.037	0.057
	100	0.5	0.014	0.052	0.051	0.087
	100	0.7	0.005	0.045	0.055	0.105
	200	0.0	0.039	0.055	0.044	0.060
2	200	0.5	0.040	0.066	0.066	0.087
	200	0.7	0.022	0.069	0.074	0.109
	300	0.0	0.036	0.051	0.040	0.055
	300	0.5	0.042	0.065	0.061	0.081
	300	0.7	0.030	0.055	0.058	0.080
	100	0.0	0.029	0.050	0.031	0.053
	100	0.5	0.014	0.052	0.029	0.067
	100	0.7	0.005	0.045	0.023	0.077
	200	0.0	0.039	0.055	0.041	0.058
3	200	0.5	0.040	0.066	0.049	0.077
	200	0.7	0.022	0.069	0.050	0.086
	300	0.0	0.036	0.051	0.039	0.052
	300	0.5	0.042	0.065	0.052	0.073
	300	0.7	0.030	0.055	0.043	0.071

FIGURE 1

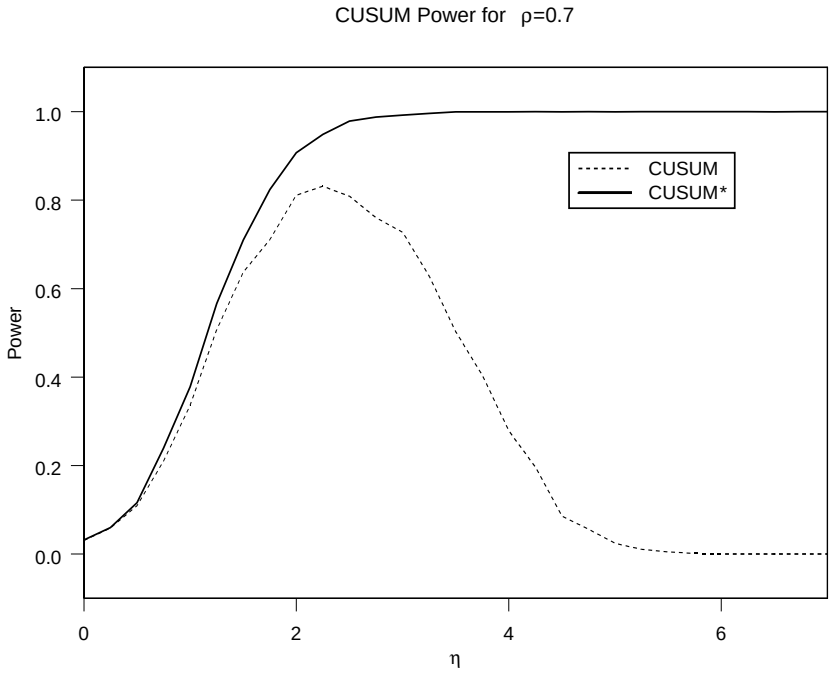


FIGURE 2

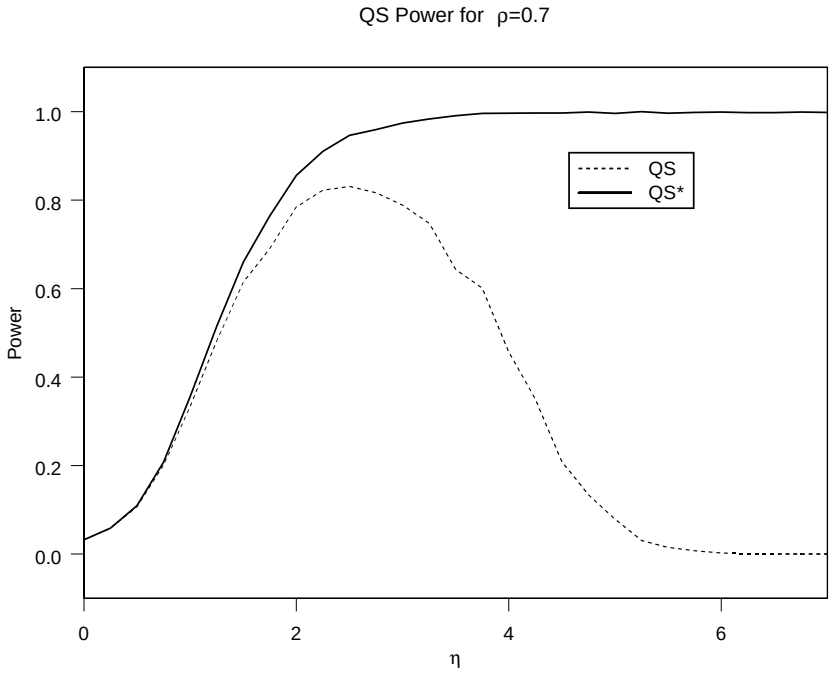
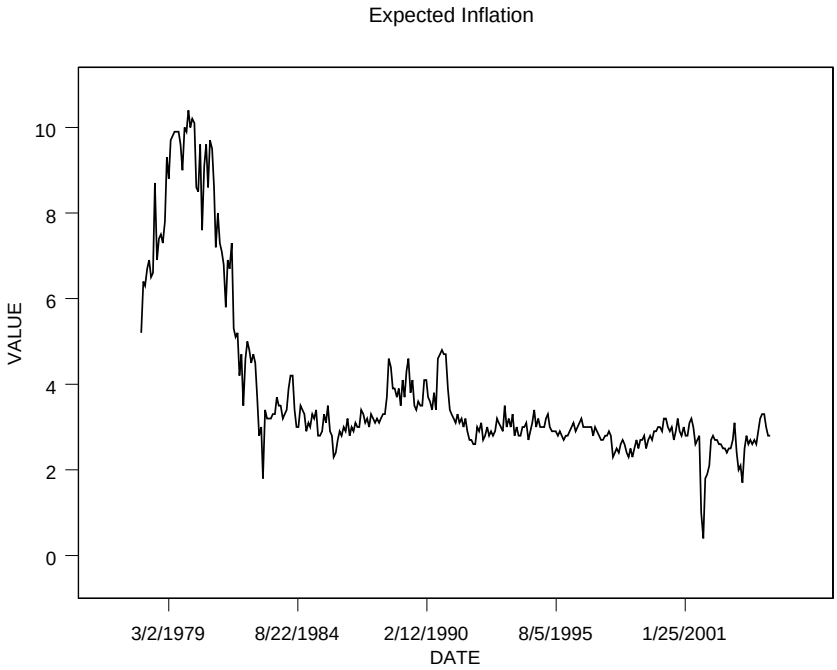


FIGURE 3


Expected Inflation
VALUE
10
8
6
4
2
0
3/2/1979
8/22/1984
2/12/1990
8/5/1995
1/25/2001
DATE