Published in final edited form as:

J Stat Plan Inference. 2013 February 1; 143(2): 237–250. doi:10.1016/j.jspi.2012.07.012.

Testing for changes in autocovariances of nonparametric time series models

Xiaoye Li and Zhibiao Zhao*

Department of Statistics, Penn State University

Abstract

In the literature on change-point analysis, much attention has been paid to detecting changes in certain marginal characteristics, such as mean, variance, and marginal distribution. For time series data with nonparametric time trend, we study the change-point problem for the autocovariance structure of the unobservable error process. To derive the asymptotic distribution of the cumulative sum test statistic, we develop substantial theory for uniform convergence of weighted partial sums and weighted quadratic forms. Our asymptotic results improve upon existing works in several important aspects. The performance of the test statistic is examined through simulations and an application to interest rates data.

Keywords

Autocovariance; Change-point analysis; Functional central limit theorem; Interest rates; Nonparametric regression; Time series; Weighted partial sum

1 Introduction

In the large literature on change-point analysis, much effort has been devoted to detecting changes in certain marginal characteristics. Some representative works include Pet-titt (1980) for proportion changes in binary data, Horváth (1993) for mean and variance changes in Gaussian observations, Inclán and Tiao (1994) for changes in variance of independent observations, Bai and Perron (1998) for coefficient changes in linear regression models, Lavielle (1999) and Mei (2006) for changes in marginal distribution, Kokoszka and Leipus (2000) and Wied (2011) for variance changes in time series models, Aue et al. (2008b) for mean change in time series with stationary errors, Zhao and Li (2011) for mean changes in time series data with non-stationary errors, and the monograph Csörgő and Horváth (1997) for related references.

While working well for independent observations, marginal information, such as mean, variance, and marginal distribution, is usually less informative for time series data. In fact, one of the main goals of time series analysis is to study how observations are dependent on each other, and marginal information is of little use in such contexts. For such purpose, autocovariance (or autocorrelation) is the most widely used measure of dependence among

© 2012 Elsevier B.V. All rights reserved.

*Corresponding author: Zhibiao Zhao, Department of Statistics, Penn State University, 326 Thomas Building, University Park, PA 16802. zuz13@stat.psu.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

time series observations. In fact, autocorrelation function plays a key role in identifying potential candidate models in the classical ARMA modeling (Brockwell and Davis, 1991). Motivated by the above discussion, we study the change-point problem for autocovariances $\text{cov}(e_i, e_{i+k})$ at a lag $k \geq 0$ of a time series $\{e_i\}_{i=1}^n$. Specifically, we consider testing

$$H_0^{(k)}: \text{cov}(e_1, e_{1+k}) = \cdots = \text{cov}(e_{n-k}, e_n), \quad (1)$$

against the alternative that there exists one J , called the change-point, such that

$$H_A^{(k)}: \text{cov}(e_1, e_{1+k}) = \cdots = \text{cov}(e_J, e_{J+k}) \neq \text{cov}(e_{J+1}, e_{J+1+k}) = \cdots = \text{cov}(e_{n-k}, e_n).$$

Most existing works on the change-point analysis (1) deal with parametric time series and the process of interest is directly observable. For example, Picard (1985) studied linear autoregressive models, and Lee, Ha and Na (2003) and Berkes, Gombay and Horváth (2009) considered general linear processes. The change-point analysis on the autoregressive parameters in Davis, Huang and Yao (1995) and Gombay (2008) can be used to detect autocovariance change-points in linear autoregressive models; also, see Ling (2007) for related work on nonlinear parametric autoregressive models. Galeano and Peña (2007) studied the change-point problem for the variance and correlations of vector autoregressive moving average models. On the one hand, a parametric model can provide a parsimonious description of the underlying model dynamics. On the other hand, as a modeling issue, it is hard to believe that we have better information about the explicit form of the underlying model than about the possibility of change-points in such models. Recently, Aue et al. (2009) and Shao and Zhang (2010) proposed tests for change-points in autocovariances of nonparametric time series models, and Wied, Krämer and Dehling (2012) considered testing for change-points in correlations between two processes. In all the aforementioned works, a restrictive assumption is that the process of interest is directly observable.

In practice, however, the process $\{e_i\}$ of interest is often not observable and the actual observations, denoted by $\{X_i\}$, may be non-stationary and contain unknown functions. Consider the important case of the fixed-design nonparametric regression model:

$$X_i = f(i/n) + e_i, \quad i = 1, \dots, n, \quad (2)$$

where $\{e_i\}$ is a process with $\mathbb{E}(e_i) = 0$, and $f(\cdot)$ is an unknown time trend. Model (2) and its variants have been extensively studied under different focuses, mainly on inferences for the (marginal information) mean trend $f(\cdot)$. For example, Altman (1990) considered estimation of $f(\cdot)$, and Wu and Zhao (2007) studied the change-point problem and simultaneous confidence band construction for $f(\cdot)$. Assuming that $f(\cdot)$ is a polynomial function and $\{e_i\}$ are uncorrelated stationary process, Aue et al. (2008a) studied coefficient changes.

Despite the rich literature on the change-point problem and the importance of autocovariances in time series analysis, the change-point analysis for autocovariances of $\{e_i\}$ in (2) seems unexplored, and our goal is to address the latter problem. Our results improve upon earlier studies in several important aspects. First, with the introduction of the nonparametric time trend $f(\cdot)$, model (2) offers a flexible framework for potential applications in climatic, economic, and financial time series fields. In such applications, the data often exhibit complicated time-varying trend that can hardly be captured by parametric forms. Second, our theory is developed under a general dependence structure covering many linear and nonlinear processes, whereas many previous works (e.g., Lee, Ha and Na, 2003;

Berkes, Gombay and Horváth, 2009) depend heavily on the linear process assumption. Third, the presence of the unknown function $f(\cdot)$ makes the problem significantly more challenging, and we develop substantial technical tools, such as moment inequalities and uniform convergence for weighted partial sums and quadratic forms, which may be of independent interest.

The paper is organized as follows. In Section 2, we introduce the dependence structure along with examples. In Section 3, we first establish a functional central limit theorem for the autocovariance process and then study the cumulative sum test. Section 4 contains some useful results on weighted partial sum and weighted quadratic forms. Section 5 contains simulation studies and an application. Proofs are provided in Section 6.

2 Dependence Structure

We adopt Wu (2007)'s framework for the dependence structure on $\{e_i\}$. Under the null hypothesis of no change-points, we assume that e_i has the causal representation

$$e_i = G(\dots, e_{i-1}, \varepsilon_i), \quad (3)$$

for a measurable function G and independent and identically distributed (IID) innovations ε_i .

Let $\{\varepsilon'_i\}_{i \in \mathbb{Z}}$ be an IID copy of $\{\varepsilon_i\}_{i \in \mathbb{Z}}$. Define the coupling process of $\{e_i\}$:

$$e'_i = G(\dots, \varepsilon'_{-1}, \varepsilon'_0, \varepsilon_1, \dots, \varepsilon_i). \quad (4)$$

For a random variable Z , we write $Z \in \mathcal{L}^q$, $q > 0$, if $\|Z\|_q = [\mathbb{E}(|Z|^q)]^{1/q} < \infty$. Following Shao and Wu (2007), we say that $\{e_i\}$ satisfies the *geometric moment contraction* condition GMC(q, ρ) for some $q > 0$ and $\rho \in (0, 1)$ if $e_i \in \mathcal{L}^q$ and

$$\text{GMC}(q, \rho): \quad \|e_i - e'_i\|_q \leq C\rho^i, \quad i \in \mathbb{N}, \quad (5)$$

for some constant $C < \infty$. If we view (3) as an input-output system with $(\dots, e_{i-1}, \varepsilon_i)$, G , and e_i being, respectively, the input, filter, and output, then (5) asserts that the impact of replacing the distant input $\{e_j\}_{j \rightarrow 0}$ by an IID copy $\{\varepsilon'_j\}_{j \leq 0}$ decays exponentially.

Example 1—Consider the linear process $e_i = \sum_{j=1}^{\infty} a_j \varepsilon_{i-j}$. If $e_i \in \mathcal{L}^q$, $a_1 = \alpha \rho^j$ some $\rho \in (0, 1)$, then (5) holds. In particular, it is well-known that causal ARMA models (all roots of the autoregressive polynomial lie outside of the unit disc) satisfy the latter condition; see Brockwell and Davis (1991).

Example 2—Consider the nonlinear autoregressive model of order $p \in \mathbb{N}$

$$e_i = \mu(e_{i-1}, \dots, e_{i-p}) + \sigma(e_{i-1}, \dots, e_{i-p}) \varepsilon_i, \quad (6)$$

where $\mu(\cdot)$ and $\sigma(\cdot)$ are two measurable functions and $\{\varepsilon_i\}$ are IID innovations. Write $x = (x_1, \dots, x_p)$, $x' = (x'_1, \dots, x'_p)$. if $e_0 \in \mathcal{L}^q$ and there exists $a_1, \dots, a_p$ such that

$$|\mu(x) - \mu(x')| + |s(x) - s(x')| \|\varepsilon_0\|_q \leq \sum_{j=1}^p a_j |x_j - x'_j| \quad \text{and} \quad \sum_{j=1}^p a_j < 1, \quad (7)$$

then the causal representation (5) holds for some $\rho \in (0, 1)$; see Theorem 5.1 in Shao and Wu (2007). Many widely used parametric models are special cases of (6), including linear autoregressive model, threshold autoregressive model, exponential autoregressive model, random coefficient model, and autoregressive conditional heteroscedastic model.

Proposition 1 below is adopted from Zhao, Wei and Lin (2012). It shows that, if $\{e_j\}$ satisfies (5), then its properly transformed process also satisfies (5).

Proposition 1—For $0 < \varsigma < 1$ and $v > 0$, define the collection of functions h

$$\mathcal{H}(\varsigma, v) = \{h: |h(x) - h(x')| \leq c|x - x'|^\varsigma (1 + |x| + |x'|)^v, \quad x, x' \in \mathbb{R}\}, \quad (8)$$

where c is a constant. Suppose $\{e_j\}$ satisfies the condition $\text{GMC}(q, \rho)$ in (5). Then the transformed process $e_i^* = h(e_i)$ satisfies $\text{GMC}(q/(\varsigma + v), \rho^\varsigma)$.

In (8), $\mathcal{H}(\varsigma, 0)$ is the class of uniformly Hölder-continuous functions with index ς . If $h(x) = |x|^b$, $b > 1$, then $h \in \mathcal{H}(1, b - 1)$. Proposition 1 provides a way to generate many nonlinear processes satisfying (5). In particular, by Proposition 1 and Examples 1–2, our subsequent results have much wider applications compared to the linear process framework in Lee, Ha and Na (2003) and Berkes, Gombay and Horváth (2009).

Proposition 2 below gives some useful moment inequalities.

Proposition 2—Suppose $\{e_j\}$ satisfies (5) for some $q \geq 2$.

- i. For any $i \neq j$, $\mathbb{E}(e_i e_j) = O(\rho^{|i-j|})$ holds uniformly.
- ii. If $q \geq 4$, then $\mathbb{E}(e_i e_j e_r e_s) = O[\sqrt{\rho}^{(j-i)+(s-r)}]$ for all integers $i \neq j \neq r \neq s$.

3 Autocovariance Process and Cumulative Sum Test

For $\{e_j\}$ in (2), denote by $\gamma_k = \text{cov}(e_t, e_{t+k})$ the autocovariance function under the null. Since $e_1, \dots, e_n$ are unknown, we need to estimate them first in order to study their autocovariances. For each $t \in (0, 1)$, we can estimate $f(t)$ by the Priestley-Chao estimator (Priestley and Chao, 1972):

$$\widehat{f}(t) = \frac{1}{nb} \sum_{i=1}^n X_i K\left(\frac{i/n - t}{b}\right), \quad (9)$$

where $b > 0$ is a bandwidth, and $K(\cdot) \geq 0$ is a kernel function. Based on the residuals e_i , we define the sample autocovariance process:

$$\widehat{\gamma}_k(t) = \frac{\sum_{i=\lfloor nt \rfloor}^{\lfloor nt \rfloor + \lfloor n\delta \rfloor} \widehat{e}_i \widehat{e}_{i+k}}{\lfloor nt \rfloor - \lfloor n\delta \rfloor + 1}, \quad t \in (\delta, 1 - \delta), \quad \text{where} \quad \widehat{e}_i = X_i - \widehat{f}(i/n). \quad (10)$$

Here $\delta > 0$ is a small constant to avoid the boundary issue since the estimate $\widehat{f}(t)$ is usually not consistent at the boundaries.

Assumption 1—(i) $K(\cdot)$ is bounded, symmetric, integrates to one, and has bounded support and bounded derivative; (ii) $\sup_{t \in [0,1]} |f''(t)| < \infty$; (iii) $nb^2/\log^2 n \rightarrow \infty$ and $nb^8 \rightarrow 0$.

Theorem 1 below establishes the convergence of the normalized autocovariance process to a time-shifted Brownian motion, where the shift d is due to the boundary truncation.

Theorem 1—Suppose (5) holds for some $q > 4$ and Assumption 1 holds. Under (3), we have the functional central limit theorem in the Skorokhod space (Billingsley, 1968):

$$\left\{ \sqrt{n}(t-\delta)[\widehat{\gamma}_k(t)-\gamma_k] \right\}_{t \in (\delta, 1-\delta)} \Rightarrow \tau(k)\{IB_{t-\delta}\}_{t \in (\delta, 1-\delta)},$$

where $\{B_t\}$ is a standard Brownian Motion, and

$$\tau^2(k) = \mathbb{E}(\eta_0^2) + 2 \sum_{i=1}^{\infty} \mathbb{E}(\eta_0 \eta_i), \quad \text{with} \quad \eta_i = e_i e_{i+k} - \gamma_k. \quad (11)$$

To address the change-point problem (1), consider the cumulative sum test:

$$T_n^{(k)} = \sup_{t \in (\delta, 1-\delta)} |T_n^{(k)}(t)|, \quad \text{where} \quad T_n^{(k)}(t) = n^{-1/2} \sum_{i=\lfloor n\delta \rfloor}^{\lfloor nt \rfloor} [\widehat{e}_i \widehat{e}_{i+k} - \widehat{\gamma}_k(1-\delta)]. \quad (12)$$

By writing $T_n^{(k)}(t) = n^{-1/2}(\lfloor nt \rfloor - \lfloor n\delta \rfloor + 1)[\widehat{\gamma}_k(t) - \widehat{\gamma}_k(1-\delta)]$, we see that $T_n^{(k)}(t)$ measures the deviation between the two estimates of γ_k : $\widehat{\gamma}_k(t)$ based on the residuals up to time t , and $\widehat{\gamma}_k(1-\delta)$ based on all residuals (excluding the boundaries). Under $H_0^{(k)}$, we expect that such deviation should be small. Thus, $T_n^{(k)}$ measures the overall maximal deviation over $t \in (\delta, 1-\delta)$, with a large value indicating the rejection of $H_0^{(k)}$.

Theorem 2—Assume the same conditions in Theorem 1. Under (3), we have

$$T_n^{(k)} \Rightarrow \tau(k) \sup_{t \in (\delta, 1-\delta)} \left| IB_{t-\delta} - \frac{t-\delta}{1-2\delta} IB_{1-\delta} \right|,$$

where $\tau(k)$ and $\{B_t\}$ are defined as in Theorem 1.

For $\delta = 0.05$, by simulations, the 90%, 95%, and 99% quantiles of $\sup_{t \in (\delta, 1-\delta)} |B_t - (t-\delta)/(1-2\delta) B_1|$ are 1.14, 1.26, and 1.52, respectively. Therefore, at significance level 0.05, we reject $H_0^{(k)}$ if $T_n^{(k)} \geq 1.26\widehat{\tau}(k)$, where $\widehat{\tau}^2(k)$ is a consistent estimator of $\tau^2(k)$.

Remark 1—The causal representation (3) implies the strict stationarity of $\{e_i\}$. Thus, the asymptotic null distribution in Theorem 2 is based on the strict stationarity of $\{e_i\}$, which implies but is stronger than the null hypothesis $H_0^{(k)}$ in (1). Unfortunately, it is unclear how to derive the asymptotic null distribution under $H_0^{(k)}$ instead of the stronger assumption (3). In the absence of strict stationarity, different segments of the data may have completely

different behavior and it is technically challenging to study them. For example, consider the following model: the first half of the data are IID random variables with zero mean and unit variance whereas the second half of the data follow the autoregressive conditional

heteroscedastic model (Engle, 1982) $e_i = \sqrt{0.5 + 0.5e_{i-1}^2} \varepsilon_i$, where ε_i are IID random variables with $\mathbb{E}(\varepsilon_i) = 0$ and $\mathbb{E}(\varepsilon_i^2) = 1$. Then both segments have the same autocovariances $\text{cov}(e_i, e_j) = 1$ for $i = j$ and $\text{cov}(e_i, e_j) = 0$ for $i \neq j$.

Using the well-known Bartlett kernel estimator, we can estimate $\tau^2(k)$ by

$$\widehat{\tau}^2(k) = \widehat{\text{cov}}(\eta_0, \eta_0) + 2 \sum_{j=1}^I \left(1 - \frac{j}{I+1}\right) \widehat{\text{cov}}(\eta_0, \eta_j), \quad (13)$$

where I is a bandwidth parameter, $\widehat{\text{cov}}(\eta_0, \eta_j)$ is the sample version of $\mathbb{E}(\eta_0 \eta_j)$ given by

$$\widehat{\text{cov}}(\eta_0, \eta_j) = \frac{1}{\ell_j} \sum_{i=\lfloor n\delta \rfloor}^{\lfloor n(1-\delta) \rfloor - j} [\widehat{e}_{i+k} - \widehat{\gamma}_k(1-\delta)][\widehat{e}_{i+j+k} - \widehat{\gamma}_k(1-\delta)],$$

and $\ell_j = \lfloor n(1-\delta) \rfloor - \lfloor n\delta \rfloor + 1 - j$

Theorem 3—Suppose (5) holds for some $q \geq 8$ and Assumption 1 holds. Further assume

$$I \rightarrow \infty \quad \text{and} \quad I \left[b^2 + \left(\frac{\log n}{nb} \right)^{1/2} \right] \rightarrow 0. \quad (14)$$

Under (3), for each given $k \geq 0$, we have $\tau \mathcal{F}^2(k) = \tau^2(k) + o_p(1)$.

The proof of Theorem 3, given in Section 6, relies on the decomposition

$\widehat{\tau}^2(k) = \tau_f^2 + R_n + O_p(1)\Omega_n$ in (45), where τ_f^2 is a truncated version of $\tau^2(k)$ in (11) using finite terms, and R_n and Ω_n quantifies the influence from estimating the unknown function $f(\cdot)$. In (14), the condition $I \rightarrow \infty$ is needed to ensure $\tau_f^2 \rightarrow \tau^2(k)$, and the condition $I \{ b^2 + [(nb)^{-1/2} \log n]^{1/2} \} \rightarrow 0$ guarantees the negligibility of R_n and Ω_n (the effect of estimating f).

We now address the selection of the bandwidth I . It is well-known that the optimal choice of b in (9) is $b \asymp n^{-1/5}$. If $I \asymp n^\beta$, then (14) holds for $\beta \in (0, 2/5)$. For long-run variance estimation with Bartlett kernel, the optimal bandwidth is $I = cn^{1/3}$ for some constant c (Andrews, 1991), which satisfies (14) when using $b \asymp n^{-1/5}$. By the proof of Theorem 3, under (5), $\mathbb{E}(\eta_0 \eta_j) = O(\rho^j)$ in (11) decays exponentially fast. Thus, even in the presence of strong dependence, a relatively small truncation lag I can result in a reasonably accurate

convergence of τ_f^2 to τ^2 . For example, if $\rho = 0.7$, the tail $\sum_{j=11}^{\infty} \rho^j \approx 0.066$ is small. Furthermore, a small I is preferred so that $I \{ b^2 + [(nb)^{-1/2} \log n]^{1/2} \} \rightarrow 0$, and thus the effect of estimating f is negligible. In summary, we recommend using a relatively small I ; based on our simulations, taking $c = 1$ in the optimal choice $I = cn^{1/3}$ works quite well.

Remark 2—In (13), we can also use other kernels; see, e.g. Andrews (1991) for several other choices. For the class of flat-top kernels developed by Politis and Romano in a series of papers, Politis (2003) proposed some data-driven bandwidth choice.

Now we study the power of the proposed test statistic. Consider the fixed alternative

$$H_a: e_i = \begin{cases} G_1(\dots, \varepsilon_{i-1}, \varepsilon_i), & i=1, \dots, \lfloor nt_0 \rfloor, \\ G_2(\dots, \varepsilon_{i-1}, \varepsilon_i), & i=\lfloor nt_0 \rfloor+1, \dots, n, \end{cases} \quad (15)$$

for two measurable functions G_1 and G_2 and a change-point $t_0 \in (0, 1)$. Under H_a , $\{e_i\}$ consists of two stationary processes of the form (3).

Theorem 4—Suppose both G_1 and G_2 satisfy (5) for some $q > 4$ and Assumption 1 holds. For the two stationary processes in (15), denote their autocovariances by $\gamma_{k,1}$ and $\gamma_{k,2}$, respectively. Further assume $t_0 \in (\delta, 1 - \delta)$, $\gamma_{k,1} \neq \gamma_{k,2}$, and that the bandwidth I in (13) satisfies $I/n \rightarrow 0$. Then $\mathbb{P}\{T_n^{(k)} > \widehat{\tau}(k)\} \rightarrow 1$ as $n \rightarrow \infty$.

By Theorem 4, the proposed test statistic has asymptotic power one under the fixed alternative (15). Our simulation studies in Section 5.1 further demonstrate that the power increases very fast as the before change-point model moves further away from the after change-point model [i.e., as λ in (23) increases]. We point out that, it is quite challenging to study the asymptotic power under local alternatives and it deserves future research.

The null hypothesis (1) tests the constancy of γ_k for each given $k \geq 0$. Alternatively, we can also test the joint constancy of $\gamma_0, \gamma_1, \dots, \gamma_T$ for a lag T . Such joint constancy test can address whether there is a change-point in the overall autocovariance structure but does not tell which specific autocovariance has a change-point. By contrast, testing (1) is more informative and it tells us whether there is a change-point in the autocovariance at a particular lag k . For example, it is possible that we have a change-point in the variance (γ_0) but not in the autocovariance at lag one (γ_1). Furthermore, it is technically more challenging to test the joint constancy due to the complicated correlations among the sample autocovariance functions. For example, Berkes, Gombay and Horváth (2009) also considered only the test for the constancy of individual γ_k .

4 Weighted Partial Sum and Quadratic Forms

The presence of the unknown function $f(\cdot)$ makes it substantially more challenging to prove Theorem 1 and Theorem 3. Intuitively, the residual $e/\mathcal{I}_j$ in (10) has two components: one is the underlying noise e_j and the other is the estimation error from estimating $f(\cdot)$ by $\widehat{f}(\cdot)$. To control the second component, three technical issues arise. First, we need to study certain weighted partial sum of $\{e_j\}$. Second, since the estimation error also contains $\{e_j\}$, we must also deal with some weighted quadratic form in order to study $\gamma\mathcal{I}_k(t)$ in (10). Third, to study the process $\{\gamma\mathcal{I}_k(t)\}_t$, it is necessary to obtain uniform convergence of the latter weighted partial sum process and quadratic form. In this section we present some useful results in these directions, which may be of independent interest.

Proposition 3—Suppose (5) holds for some $q \geq 2$. Then, for any sequence $\{a_i\}_{i=1}^n$,

$$\max_{1 \leq k \leq n} \left| \sum_{i=1}^k a_i e_i \right| = O_p \left\{ \log n \left[\max_{1 \leq i \leq n} |a_i| \sum_{i=1}^n |a_i| \right]^{1/2} \right\}. \quad (16)$$

The bound in (16) is optimal up to a logarithm factor. For example, if $a_i \equiv 1$ and $\{e_i\}$ are IID, then $\max_{1 \leq k \leq n} \left| \sum_{i=1}^k e_i \right| = O_p(\sqrt{n})$ via Kolmogorov's maximal inequality. Similarly, the bounds in Proposition 4 below are also optimal up to a logarithm factor.

Proposition 4—Suppose (5) holds for some $q \geq 2$. Then, for any double array $\{\theta_{i,j}\}_{i,j=1}^n$,

$$(i) \quad \max_{1 \leq k \leq n} \left| \sum_{i=1}^k \sum_{j=1}^n \theta_{i,j} e_j \right| = O_p(\sqrt{\chi_n n \log n}), \quad (17)$$

$$\text{where } \chi_n = \max_{1 \leq i, j \leq n} |\theta_{i,j}| \max_{1 \leq i \leq n} \sum_{j=1}^n |\theta_{i,j}|; \quad (18)$$

$$(ii) \quad \max_{1 \leq i \leq n} \left| \sum_{j=1}^n \theta_{i,j} e_j \right| = O_p \left\{ \sqrt{\Theta_n \log n} + L_n n^{1/\min(q,4)} \log n \right\}, \quad (19)$$

$$\text{where } \Theta_n = \max_{1 \leq i \leq n} \sum_{j=1}^n \theta_{i,j}^2, \quad L_n = \max_{1 \leq i \leq n} \left(|\theta_{i,n}| + \sum_{j=1}^{n-1} |\theta_{i,j} - \theta_{i,j+1}| \right). \quad (20)$$

Proposition 5 below presents a uniform bound for a weighted quadratic form of $\{e_j\}$.

Proposition 5—Suppose (5) holds for some $q \geq 4$. Let $\{\theta_{i,j}\}_{i,j=1}^n$ be any double array. Assume that there exists N_n such that

$$\max_{1 \leq i \leq n} |\{1 \leq j \leq n: \theta_{i,j} \neq 0\}| \leq N_n, \quad \max_{1 \leq j \leq n} |\{1 \leq i \leq n: \theta_{i,j} \neq 0\}| \leq N_n. \quad (21)$$

Here $|\mathcal{S}|$ denotes the cardinality of a finite set $\mathcal{S}$. Then

$$\max_{1 \leq k \leq n} \left| \sum_{i=1}^k \sum_{j=1}^n \theta_{i,j} e_i e_j \right| = O_p \left(\sqrt{n N_n \log n} \max_{1 \leq i, j \leq n} |\theta_{i,j}| \right). \quad (22)$$

5 Numerical Results

Throughout our subsequent analysis we use significance level 0.05 and $\delta = 0.05$ in (10).

5.1 Simulation Studies

In this section we examine the empirical size and power of the test based on 1000 realizations. To reduce the boundary effect, in (9) we use the local linear smoothing method implemented by the command `locpoly` in R package `KernSmooth`, with the optimal plug-in

bandwidth selection using the command `dpill`. In (2), we let $f(t) = \cos(2\pi t)$ and sample size $n = 600$. For e_i , we consider two models below.

First, we consider the linear autoregressive model:

$$\text{Model 1: } e_i = \begin{cases} 0.2e_{i-1} + 0.2\varepsilon_i, & i \leq 250; \\ (0.2 + \lambda)e_{i-1} + 0.2\varepsilon_i, & i \geq 251, \end{cases} \quad (23)$$

where ε_i are IID $N(0,1)$ random variables. For $i \leq 250$, the autocovariance function of $\{e_i\}$ is $\gamma_k = 0.2^{2+k}/(1 - 0.2^2)$, $k \geq 0$; for $i \geq 251$, the autocovariance function is $\gamma_k = 0.2^2(0.2 + \lambda)^k/[1 - (0.2 + \lambda)^2]$, $k \geq 0$. Thus, $\lambda = 0$ corresponds to the null hypothesis of no change-point, whereas $\lambda \neq 0$ indicates a change-point at time 251. By the discussion following Theorem 3, the optimal I is of order $n^{1/3} \approx 8.4$; three choices $I = 8, 10, 12$ are examined to assess the effect of I . We consider testing change-points in $\gamma_0, \gamma_1, \gamma_2$, and report the empirical size ($\lambda = 0.0$), empirical power and size-corrected power ($\lambda = 0.2, 0.4, 0.6$) in Table 1. We see that the empirical size is close to the nominal level 0.05, and the empirical power increases with λ . Also, the three choices of I give similar performance, with $I = 8$ slightly outperforming $I = 10, 12$. Interestingly, it is the easiest to detect change-points in γ_1 whereas it is the most difficult to detect change-points in the variance γ_0 . Note that, for $\lambda = 0.2$, the differences between before and after change-points autocovariances are 0.006, 0.011 and 0.006 for γ_0, γ_1 and γ_2 , respectively. The largest power of γ_1 is due to its most visible difference. One possible explanation for the relatively better power of γ_2 over γ_0 is: although the differences for γ_2 and γ_0 are the same, the variances of the estimates γ_2/n_2 and γ_0/n_0 could be very different, especially under the alternative hypothesis. For different models, it may be easier to detect change-points in autocovariances at different lags.

Next, we consider the nonlinear model:

$$\text{Model 2: } e_i = \begin{cases} \lambda e_{i-1} + 0.2\varepsilon_i, & i \leq 250; \\ 0.2\sqrt{1 + e_{i-1}^2}\varepsilon_i, & i \geq 251, \end{cases} \quad (24)$$

where ε_i are IID $N(0,1)$ random variables. For $i \leq 250$, we have a linear autoregressive model with the autocovariance function $\gamma_k = 0.2^2\lambda^k/(1 - \lambda^2)$, $k \geq 0$; for $i \geq 251$, we have the autoregressive conditional heteroscedastic model (Engle, 1982), which is widely used in financial econometrics. For $i \geq 251$, we can show that $\gamma_0 = 0.2^2/(1 - 0.2^2)$, and by the martingale difference property $\gamma_k = 0$, $k \geq 1$. Therefore, for $\lambda = 0$, although the model dynamics are quite different before and after the change-point, the autocovariance structures are almost the same (except that the variance changes from $0.2^2 = 0.04$ to $0.2^2/(1 - 0.2^2) \approx 0.042$). As λ increases, the discrepancy between the autocovariances increases. The empirical power in Table 2 shows the same pattern as in Table 1.

5.2 An Application to Interest Rates

Interest rates play an important role in virtually every aspect of financial markets, including consumer spending, housing markets, banks' borrowing costs, inflation, and stock market performance. Here we consider market yields on U.S. treasury securities at one-year constant maturity. The data, available at <http://www.federalreserve.gov/releases/h15/data.htm>, consist of weekly (Friday) rates during January 5, 1962–June 10, 2011, with a total of $n = 2579$ observations. From the plot in the left panel of Figure 1, we see dramatic interest rates hike in the early 1980s with a peak as high as 17% in 1981. Lai (2008) argued that the observed unit-root behavior of interest rates may be due to structural breaks. Here we shall examine whether there are change-points in autocovariances.

The left plot of Figure 1 shows a complicated time-varying trend that can hardly be described using any parametric form. Thus, we use the nonparametric model (2). The right panel of Figure 1 shows the residuals $e/\hat{\sigma}_i$ in (10) after removing the nonparametric time trend $\hat{f}$. We consider testing the null hypothesis $H_0^{(k)}$ in (1) for $k = 0, 1, 2$. Since $n^{1/3} \approx 13.7$, we use four choices $I = 12, 15, 18, 20$ in (13), and all lead to the same conclusion that we can reject $H_0^{(0)}, H_0^{(1)}, H_0^{(2)}$ at the significance level 0.05. For example, with $I = 15$, the corresponding p -values for $H_0^{(0)}, H_0^{(1)}, H_0^{(2)}$ are 0.013, 0.017, 0.023, respectively, indicating strong evidence to reject the null hypothesis.

With $I = 15$, Figure 2 plots the cumulative sum process $|T_n^{(k)}(t)|/\hat{\tau}(k)$ for testing the constancy of γ_0 ($k = 0$, solid curve), γ_1 ($k = 1$, dotted curve), and γ_2 ($k = 2$, dashed curve), where $T_n^{(k)}(t)$ is defined in (12). It is in December 1981 when the cumulative sum processes exceed the critical level 1.26 (corresponding to the significance level 0.05), and we conclude that the autocovariances $\gamma_0, \gamma_1, \gamma_2$ have changed since January 1982. Also, the sharp increase in $|T_n^{(k)}(t)|/\hat{\tau}(k)$ in the early 1980s roughly corresponds to the “interest rates hike” period (also see the unusual activities of the residuals plot in Figure 1).

During the 1970s, U.S. economy experienced rising unemployment and accelerating inflation. According to U.S. Bureau of Labor Statistics (<http://www.bls.gov/data>), the Consumer Price Index doubled from 37.9 in January 1970 to 76.9 in December 1979, compared to a mere 25% increase during the 1960s. At a meeting in October 1979, federal policymakers approved new policy procedures aiming to bring down the inflation by increasing the interest rates. As a result, during the next three years (1980–1982), the interest rates increased sharply and averaged 13%. Such abrupt changes likely resulted in the change-point of autocovariances in December 1981 as detected by our analysis above.

Acknowledgments

We are grateful to two anonymous referees for their insightful comments that have significantly improved the paper. Zhao's research was partially supported by a NIDA grant P50-DA10075-15. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIDA or the NIH.

References

- Abramowitz, M.; Stegun, IA. Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables. Dover; 1964.
- Altman NS. Kernel smoothing of data with correlated errors. J Amer Statist Assoc. 1990; 85:749–759.
- Andrews DWK. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. Econometrica. 1991; 59:817–858.
- Aue A, Hörmann S, Horváth L, Reimherr M. Break detection in the covariance structure of multivariate time series models. Ann Statist. 2009; 37:4046–4087.
- Aue A, Horváth L, Hušková M, Kokoszka P. Testing for changes in polynomial regression. Bernoulli. 2008a; 14:637–660.
- Aue A, Horváth L, Kokoszka P, Steinebach J. Monitoring shifts in mean: asymptotic normality of stopping times. Test. 2008b; 17:515–530.
- Bai J, Perron P. Estimating and testing linear models with multiple structural changes. Econometrica. 1998; 66:47–78.
- Berkes I, Gombay E, Horváth L. Testing for changes in the covariance structure of linear processes. J Statist Plann Inference. 2009; 139:2044–2063.
- Billingsley, P. Convergence of Probability Measures. New York: Wiley; 1968.
- Brockwell, PJ.; Davis, R. Time Series: Theory and Methods. 2. Springer-Verlag; New York: 1991.

- Davis RA, Huang DW, Yao YC. Testing for a change in the parameter values and order of an autoregressive model. *Ann Statist.* 1995; 23:282–304.
- Galeano P, Peña D. Covariance changes detection in multivariate time series. *J Statist Plann Inference.* 2007; 137:194–211.
- Gombay E. Change detection in autoregressive time series. *J Multivariate Anal.* 2008; 99:451–464.
- Horváth L. The maximum likelihood method for testing changes in the parameters of normal observations. *Ann Statist.* 1993; 21:671–680.
- Inclán C, Tiao GC. Use of cumulative sums of squares for retrospective detection of changes of variance. *J Amer Statist Assoc.* 1994; 89:913–923.
- Kokoszka P, Leipus R. Change-point estimation in ARCH models. *Bernoulli.* 2000; 6:513–539.
- Kufner, A.; John, O.; Fucik, S. *Function Spaces.* Noordhoff; Leyden: 1977.
- Lai KS. The puzzling unit root in the real interest rate and its inconsistency with intertemporal consumption behavior. *J Int Money Financ.* 2008; 27:140–155.
- Lavielle M. Detection of multiple changes in a sequence of dependent variables. *Stoch Proc Appl.* 1999; 83:79–102.
- Lee S, Ha J, Na O. The CUSUM test for parameter change in time series models. *Scand J Statist.* 2003; 30:781–796.
- Ling S. Testing for change points in time series models and limiting theorems for NED sequences. *Ann Statist.* 2007; 35:1213–1237.
- Mei Y. Sequential change-point detection when unknown parameters are present in the pre-change distribution. *Ann Statist.* 2006; 34:92–122.
- Picard D. Testing and estimating change-points in time series. *Adv Appl Probab.* 1985; 17:841–867.
- Politis DN. Adaptive bandwidth choice. *J Nonparametr Statist.* 2003; 15:517–533.
- Priestley MB, Chao MT. Non-parametric function fitting. *J R Statist Soc B.* 1972; 34:385–392.
- Shao X, Wu WB. Asymptotic spectral theory for nonlinear time series. *Ann Statist.* 2007; 35:1773–1801.
- Shao X, Zhang X. Testing for change points in time series. *J Amer Statist Assoc.* 2010; 105:1228–1240.
- Wied D, Arnold M, Bissantz N, Ziggel D. A new fluctuation test for constant variances with applications to finance. *Metrika.* 2012 Forthcoming.
- Wied D, Krämer W, Dehling H. Testing for a change in correlation at an unknown point in time using an extended functional delta method. *Econometric Theory.* 2012; 28:570–589.
- Wu WB. Strong invariance principles for dependent random variables. *Ann Probab.* 2007; 35:2294–2320.
- Wu WB, Zhao Z. Inference of trends in time series. *J R Statist Soc B.* 2007; 69:391–410.
- Zhao Z, Li X. Inference for modulated stationary processes. To appear, *Bernoulli.* 2012
- Zhao Z, Wei Y, Lin D. Asymptotics of nonparametric L-1 regression models with dependent data. 2012 Manuscript.

6 Appendix: Technical Proofs

6.1 Proof of Propositions 1–2

Proof of Proposition 1

For completeness, we include the proof from Zhao, Wei and Lin (2012). Let $q^* = q/(\zeta + \nu)$, $p_1 = \nu/\zeta + 1$, and $p_2 = \zeta/\nu + 1$ so that $\zeta q^* p_1 = q$, $\nu q^* p_2 = q$, and $1/p_1 + 1/p_2 = 1$. By (8) and Hölder's inequality [p. 67, Kufner, John and Fucik (1977)] $\mathbb{E}|Z_1 Z_2| \leq \|Z_1\|_{p_1} \|Z_2\|_{p_2}$ we have

$$\begin{aligned}
\|h(e'_i) - h(e_i)\|_{q^*}^{q^*} &\leq O(1) \mathbb{E}[|e'_i - e_i|^{q^*} (1 + |e_i| + |e'_i|)^{q^*}] \\
&\leq O(1) \{\mathbb{E}[|e_i - e'_i|^{q^* \cdot p_1}]\}^{1/p_1} \{\mathbb{E}[(1 + |e_i| + |e'_i|)^{q^* \cdot p_2}]\}^{1/p_2} \\
&= O(1) \|e_i - e'_i\|_q^{q/p_1} \|e_0\|_q^{q/p_2} = O(\rho^{iq/p_1}).
\end{aligned}$$

The above expression gives $\|h(e'_i) - h(e_i)\|_{q^*} \leq O(1) [\rho^{q/(p_1 q^*)}]^i = O(1) (\rho^S)^i$.

Proof of Proposition 2

Let $\{\varepsilon'_i\}_{i \in \mathbb{Z}}$ be an IID copy of $\{\varepsilon_i\}_{i \in \mathbb{Z}}$. Define the coupling process

$$e_{\ell, k} = G(\dots, \varepsilon'_{k-1}, \varepsilon'_k, \varepsilon_{k+1}, \dots, \varepsilon_\ell), \quad \ell > k.$$

By condition (5), we have $\|e_\ell - e_{\ell, k}\|_q \leq C\rho^{\ell-k}$. Recall that $q \geq 2$. Thus, by Hölder's inequality, $\|e_\ell - e_{\ell, k}\|_2 \leq \|e_\ell - e_{\ell, k}\|_q = O(\rho^{\ell-k})$.

(i) For $i < j$, e_i is independent of $e_{j,i}$, which implies $\mathbb{E}(e_i e_{j,i}) = \mathbb{E}(e_i) \mathbb{E}(e_{j,i}) = 0$. Thus, by the Cauchy-Schwarz inequality, $|\mathbb{E}(e_i e_j)| = |\mathbb{E}(e_i (e_j - e_{j,i}))| \leq \|e_i\|_2 \|e_j - e_{j,i}\|_2 = O(\rho^{j-i})$.

(ii) First, we prove that, for any $i < j < r < s$, the following assertions hold:

$$(iia) \quad \mathbb{E}(e_i e_j e_r e_s) = O(\sqrt{\rho}^{(j-i)+(s-r)});$$

$$(iib) \quad \mathbb{E}(e_i^2 e_j e_r) = O(\rho^{r-j}), \mathbb{E}(e_i e_j e_r^2) = O(\rho^{j-i}), \mathbb{E}(e_i e_j^2 e_r) = O(\sqrt{\rho}^{r-i});$$

$$(iic) \quad \mathbb{E}(e_i e_j^3) = O(\rho^{j-i}), \mathbb{E}(e_i^3 e_j) = O(\rho^{j-i}).$$

By assumption, for $q \geq 4$, $\|e_\ell - e_{\ell, k}\|_4 \leq \|e_\ell - e_{\ell, k}\|_q \leq C\rho^{\ell-k}$.

(iia) By Hölder's inequality,

$$|\mathbb{E}(e_i (e_j - e_{j,i}) e_r e_s)| \leq \|e_i\|_4 \times \|e_j - e_{j,i}\|_4 \times \|e_r\|_4 \times \|e_s\|_4 = O(\rho^{j-i}).$$

Similarly, $|\mathbb{E}(e_i e_j (e_r - e_{r,i}) e_s)| = O(\rho^{r-i})$ and $|\mathbb{E}(e_i e_j e_r (e_s - e_{s,i}))| = O(\rho^{s-i})$.

The independence between e_i and $e_{j,i} e_{r,i} e_{s,i}$ implies $\mathbb{E}(e_i e_{j,i} e_{r,i} e_{s,i}) = \mathbb{E}(e_i) \mathbb{E}(e_{j,i} e_{r,i} e_{s,i}) = 0$. Thus,

$$|\mathbb{E}(e_i e_j e_r e_s)| = |\mathbb{E}([e_i (e_j - e_{j,i}) e_r e_s] + [e_i e_{j,i} (e_r - e_{r,i}) e_s] + [e_i e_{j,i} e_{r,i} (e_s - e_{s,i})])| = O(\rho^{j-i}).$$

The same argument also shows $|\mathbb{E}(e_i e_j e_r e_s)| = |\mathbb{E}(e_i e_j e_r (e_s - e_{s,i}))| = O(\rho^{s-i})$.

Combining the two bounds together and using the inequality $\min(a, b) \leq \sqrt{ab}$ for all $a, b \geq 0$, we get

$$|\mathbb{E}(e_i e_j e_r e_s)| = O(1) \min(\rho^{j-i}, \rho^{s-i}) \leq O(1) \sqrt{\rho}^{(j-i)+(s-i)}. \quad (25)$$

(iib) As in (iia), by Hölder's inequality, we have

$$|\mathbb{E}(e_i^2 e_j e_r)| = |\mathbb{E}[e_i^2 e_j(e_r - e_{r,j})]| \leq \|e_i^2\|_2 \times \|e_j\|_4 \times \|e_r - e_{r,j}\|_4 = O(\rho^{r-j}).$$

For $\mathbb{E}(e_i e_j e_r^2)$, using $|\mathbb{E}[e_i e_{j,i}(e_r^2 - e_{r,i}^2)]| = |\mathbb{E}[e_i e_{j,i}(e_r - e_{r,i})(e_r + e_{r,i})]| = O(\rho^{r-i})$, we have

$$|\mathbb{E}(e_i e_j e_r^2)| \leq |\mathbb{E}[e_i(e_j - e_{j,i})e_r^2]| + |\mathbb{E}[e_i e_{j,i}(e_r^2 - e_{r,i}^2)]| = O(\rho^{j-i} + \rho^{r-i}) = O(\rho^{j-i}).$$

Similarly, $\mathbb{E}(e_i e_j^2 e_r) = \mathbb{E}[e_i(e_j^2 e_r - e_{j,i}^2 e_{r,i})] = O(\rho^{j-i})$ and

$\mathbb{E}(e_i e_j^2 e_r) = \mathbb{E}[e_i e_j^2(e_r - e_{r,j})] = O(\rho^{r-j})$. Thus, as in (25),

$$|\mathbb{E}(e_i e_j^2 e_r)| = O(1) \min(\rho^{j-i}, \rho^{r-j}) = O(\sqrt{\rho^{r-i}}).$$

(iic) We skip the proof since the argument above also applies here.

Notice that $\rho \leq \sqrt{\rho}$. We conclude the proof by noting that (iia)–(iic) imply

$\mathbb{E}(e_i e_j e_r e_s) = O[\sqrt{\rho}^{(j-i)+(s-r)}]$ for all $i < j < r < s$. For example, if $i < j < r < s$, (iia) gives the latter bound; if $i = j < r < s$, then (iib) gives $\mathbb{E}(e_i e_i e_r e_s) = \mathbb{E}(e_i^2 e_r e_s) = O(\rho^{s-r}) = O[\sqrt{\rho}^{(i-i)+(s-r)}]$; if $i = j < r = s$, then the Cauchy-Schwarz inequality gives

$$\mathbb{E}(e_i e_i e_r e_r) \leq \|e_i^2\|_2 \|e_r^2\|_2 = O(1) = O[\sqrt{\rho}^{(i-i)+(r-r)}];$$

if $i = j = r < s$, then (iic) gives

$$\mathbb{E}(e_i e_i e_i e_s) = \mathbb{E}(e_i^3 e_s) = O(\rho^{s-i}) = O[\sqrt{\rho}^{(i-i)+(s-i)}].$$

6.2 Proof of Propositions 3–5

First, we state some elementary inequalities. Let $a_i, i \in \mathbb{N}$, be any sequence of real numbers. Then for any $\rho \in (0, 1)$ and integers $1 \leq \ell_1 < \ell_2$,

$$\sum_{i=\ell_1}^{\ell_2} |a_i| \rho^{|j-i|} \leq \max_{\ell_1 \leq i \leq \ell_2} |a_i| \sum_{i=\ell_1}^{\ell_2} \rho^{|j-i|} \leq \frac{1+\rho}{1-\rho} \max_{\ell_1 \leq i \leq \ell_2} |a_i|, \quad \text{for all } j. \quad (26)$$

Here the last inequality follows from $\sum_{i=\ell_1}^{\ell_2} \rho^{|j-i|} \leq \sum_{k=-\infty}^{\infty} \rho^{|k|} = (1+\rho)/(1-\rho)$. By (26),

$$\sum_{i=\ell_1}^{\ell_2} \sum_{j=\ell_1}^{\ell_2} |a_i a_j| \rho^{|j-i|} = \sum_{i=\ell_1}^{\ell_2} \left(|a_i| \sum_{j=\ell_1}^{\ell_2} |a_j| \rho^{|j-i|} \right) \leq \frac{1+\rho}{1-\rho} \max_{\ell_1 \leq i \leq \ell_2} |a_i| \sum_{i=\ell_1}^{\ell_2} |a_i|. \quad (27)$$

Lemma 1

Let $N_1, \dots, N_n$ be any normal random variables satisfying $\mathbb{E}(N_i) = 0$ and $\mathbb{E}(N_i^2) \leq s_n^2$ for some $s_n > 0$. Then $\max_{1 \leq i \leq n} |N_i| = O_p(s_n \sqrt{\log n})$.

Proof—Let N be a standard normal random variable. The proof is completed via

$$\mathbb{P}\left(\max_{1 \leq i \leq n} |N_i| \geq 2s_n \sqrt{\log n}\right) \leq \sum_{i=1}^n \mathbb{P}(|N_i| \geq 2s_n \sqrt{\log n}) \leq n\mathbb{P}(|N| \geq 2\sqrt{\log n}) \rightarrow 0.$$

Here the last convergence follows from $\mathbb{P}(N > z) = [1 + \alpha(1)] \frac{\varphi(z)}{z}$ as $z \rightarrow \infty$ [see formula 7.1.13 in Abramowitz and Stegun (1964)] and $\varphi(2\sqrt{\log n}) = O(n^{-2})$, with $\varphi(z)$ being the standard normal density.

Lemma 2 below presents a useful maximal inequality from Wu (2007).

Lemma 2

(Proposition 1 in Wu (2007)). Let d be a positive integer, and $Z_i \in \mathcal{L}^2$, $1 \leq i \leq 2^d$, be any random variables. Write $S_k = \sum_{i=1}^k Z_i$. Then

$$\left\| \max_{1 \leq k \leq 2^d} |S_k| \right\|_2 \leq \sum_{r=0}^d \left[\sum_{m=1}^{2^{d-r}} \left\| \sum_{i=(m-1)2^r+1}^{m2^r} Z_i \right\|_2^2 \right]^{1/2}.$$

Proof of Proposition 3

Let $H_k = \sum_{i=1}^k a_i e_i$, write $d = \lceil \log_2 n \rceil$, where $\lceil x \rceil$ is the smallest integer that is not less than x . Then $n \leq 2^d \leq 2n$. For convenience, write $a_i = 0$ for $n+1 \leq i \leq 2^d$ so that $H_j = H_n$ for $n+1 \leq j \leq 2^d$. By Proposition 2 (i), $|\mathbb{E}(e_i e_j)| = O(\rho^{j-i})$ for all i, j . Thus, by (27), for any $1 \leq \ell_1 < \ell_2 \leq 2^d$,

$$\begin{aligned} \left\| \sum_{i=\ell_1}^{\ell_2} a_i e_i \right\|_2^2 &\leq \sum_{i=\ell_1}^{\ell_2} \sum_{j=\ell_1}^{\ell_2} |a_i a_j| |\mathbb{E}(e_i e_j)| \\ &= O(1) \sum_{i=\ell_1}^{\ell_2} \sum_{j=\ell_1}^{\ell_2} |a_i a_j| \rho^{j-i} = O(1) \max_{1 \leq i \leq n} |a_i| \sum_{i=\ell_1}^{\ell_2} |a_i|. \end{aligned} \quad (28)$$

Therefore, by (28) and Lemma 2,

$$\begin{aligned} \left\| \max_{1 \leq k \leq n} |H_k| \right\|_2 &= \left\| \max_{1 \leq k \leq 2^d} |H_k| \right\|_2 \leq \sum_{r=0}^d \left[\sum_{m=1}^{2^{d-r}} \left\| \sum_{i=(m-1)2^r+1}^{m2^r} a_i e_i \right\|_2^2 \right]^{1/2} \\ &= O(1) \sum_{r=0}^d \left[\max_{1 \leq i \leq n} |a_i| \sum_{m=1}^{2^{d-r}} \sum_{i=(m-1)2^r+1}^{m2^r} |a_i| \right]^{1/2} \\ &= O \left\{ \log n \left[\max_{1 \leq i \leq n} |a_i| \sum_{i=1}^n |a_i| \right]^{1/2} \right\}, \end{aligned}$$

completing the proof.

Proof of Proposition 4

First, we prove (17). Define $W_i = \sum_{j=1}^n \theta_{i,j} e_j$. As in (28),

$$\|W_i\|_2^2 = O(1) \max_{1 \leq j \leq n} |\theta_{i,j}| \sum_{j=1}^n |\theta_{i,j}| \leq O(1) \max_{1 \leq i, j \leq n} |\theta_{i,j}| \max_{1 \leq i \leq n} \sum_{j=1}^n |\theta_{i,j}| = O(\chi_n). \quad (29)$$

Applying the triangle inequality, for any $\ell_1 < \ell_2$, we can obtain

$$\left\| \sum_{i=\ell_1}^{\ell_2} W_i \right\|_2 \leq \sum_{i=\ell_1}^{\ell_2} \|W_i\|_2 = O[(\ell_2 - \ell_1) \sqrt{\chi_n}]. \quad (30)$$

Then (17) follows by applying (30) and Lemma 2 to $\left\| \max_{1 \leq k \leq n} \left| \sum_{i=1}^k W_i \right| \right\|_2$.

Next, we prove (19). By Corollary 4 in Wu (2007), there exists a standard Brownian Motion $\{B_t\}$ such that

$$\max_{1 \leq j \leq n} \left| \sum_{i=1}^j e_i - sIB_j \right| = O_{a.s.} \{n^{1/\min(q,4)} \log(n)\}. \quad (31)$$

Here $s^2 = \mathbb{E}(e_0^2) + 2 \sum_{i=1}^{\infty} \mathbb{E}(e_0 e_i)$ is the long-run variance of $\{e_i\}$. Write $S_j = \sum_{i=1}^j e_i$ so that $e_j = S_j - S_{j-1}$. Then by (31) and the summation-by-part formula,

$$\begin{aligned} \sum_{j=1}^n \theta_{i,j} e_j &= \sum_{j=1}^n \theta_{i,j} (S_j - S_{j-1}) \\ &= \sum_{j=1}^{n-1} (\theta_{i,j} - \theta_{i,j+1}) S_j + \theta_{i,n} S_n \\ &= \sum_{j=1}^{n-1} (\theta_{i,j} - \theta_{i,j+1}) sIB_j + \theta_{i,n} sIB_n + O_{a.s.} [L_n n^{1/\min(q,4)} \log(n)] \\ &= s \sum_{j=1}^n \theta_{i,j} (IB_j - IB_{j-1}) + O_{a.s.} [L_n n^{1/\min(q,4)} \log(n)]. \end{aligned} \quad (32)$$

Since $\{IB_j - IB_{j-1}\}_{j=1}^n$ are IID standard normals, $\sum_{j=1}^n \theta_{i,j} (IB_j - IB_{j-1})$, $i = 1, \dots, n$, are normally distributed, and the variances $\sum_{j=1}^n \theta_{i,j}^2 \leq \Theta_n$. Therefore, by Lemma 1, $\max_{1 \leq i \leq n} \left| \sum_{j=1}^n \theta_{i,j} (IB_j - IB_{j-1}) \right| = O_p[(\Theta_n \log n)^{1/2}]$. Finally, (20) follows from (32).

Proof of Proposition 5

Recall N_n in (21). Suppose we can prove that, uniformly for all $\ell_1 < \ell_2$,

$$\left\| \sum_{i=\ell_1}^{\ell_2} U_i \right\|_2^2 = O(1) N_n (\ell_2 - \ell_1) \max_{1 \leq i, j \leq n} |\theta_{i,j}|^2, \quad \text{where} \quad U_i = e_i \sum_{j=1}^n \theta_{i,j} e_j. \quad (33)$$

Then the result easily follows by applying Lemma 2 to $\left\| \max_{1 \leq k \leq n} \left| \sum_{i=1}^k U_i \right| \right\|_2$.

In order to prove (33), we observe that

$$\left\| \sum_{i=\ell_1}^{\ell_2} U_i \right\|_2^2 = \left\| \sum_{i=\ell_1}^{\ell_2} \sum_{j=1}^n \theta_{i,j} e_i e_j \right\|_2^2 \leq \sum_{i=\ell_1}^{\ell_2} \sum_{i_2=\ell_1}^{\ell_2} \sum_{j_1=1}^n \sum_{j_2=1}^n |\theta_{i_1,j_1} \theta_{i_2,j_2}| \times |\mathbb{E}(e_{i_1} e_{i_2} e_{j_1} e_{j_2})|. \quad (34)$$

By considering the 24 different orderings along with Proposition 2 (ii), we can get

$$|\mathbb{E}(e_{i_1} e_{i_2} e_{j_1} e_{j_2})| = O\left(\sqrt{\rho}^{|i_1-i_2|+|j_1-j_2|} + \sqrt{\rho}^{|i_1-j_1|+|i_2-j_2|} + \sqrt{\rho}^{|i_1-j_2|+|i_2-j_1|}\right), \quad (35)$$

for all i_1, j_1, i_2, j_2 . We consider the three terms separately below.

For convenience, write $\theta_* = \max_{i,j} |\theta_{i,j}|$. Applying (26) and (27), we have

$$\sum_{j_1, j_2=1}^n |\theta_{i_1, j_1} \theta_{i_2, j_2}| \sqrt{\rho}^{|j_1-j_2|} \leq \theta_* \frac{1+\sqrt{\rho}}{1-\sqrt{\rho}} \sum_{j_1=1}^n |\theta_{i_1, j_1}| \leq \frac{1+\sqrt{\rho}}{1-\sqrt{\rho}} N_n \theta_*^2. \quad (36)$$

Thus, for the first term in (35), by (36) and (27), we have

$$\sum_{i_1, i_2=\ell_1}^{\ell_2} \sum_{j_1, j_2=1}^n |\theta_{i_1, j_1} \theta_{i_2, j_2}| \sqrt{\rho}^{|i_1-i_2|+|j_1-j_2|} = O(N_n \theta_*^2) \sum_{i_1, i_2=\ell_1}^{\ell_2} \sqrt{\rho}^{|i_1-i_2|} = O[N_n (\ell_2 - \ell_1) \theta_*^2].$$

Now we consider the second term in (35). Observe that, by (26),

$$\sum_{i_1=\ell_1}^{\ell_2} \sum_{j_1=1}^n |\theta_{i_1, j_1}| \sqrt{\rho}^{|i_1-j_1|} \leq \theta_* \sum_{i_1=\ell_1}^{\ell_2} \sum_{j_1=1}^n \sqrt{\rho}^{|i_1-j_1|} = O[(\ell_2 - \ell_1) \theta_*]. \quad (37)$$

On the other hand, using the transformation $k = i_1 - j_1$, we can easily see

$$\sum_{i_1=\ell_1}^{\ell_2} \sum_{j_1=1}^n |\theta_{i_1, j_1}| \sqrt{\rho}^{|i_1-j_1|} = \sum_k \left(\sqrt{\rho}^{|k|} \sum_{i_1} |\theta_{i_1, i_1+k}| \right) \leq N_n \theta_* \sum_k \sqrt{\rho}^{|k|} = O(N_n \theta_*). \quad (38)$$

Combining the two bounds in (37) and (38), we get

$$\sum_{i_1=\ell_1}^{\ell_2} \sum_{j_1=1}^n |\theta_{i_1, j_1}| \sqrt{\rho}^{|i_1-j_1|} = O(1) \min\{(\ell_2 - \ell_1) \theta_*, N_n \theta_*\} = O(1) \theta_* \sqrt{N_n (\ell_2 - \ell_1)}.$$

Therefore, we have

$$\sum_{i_1, i_2=\ell_1}^{\ell_2} \sum_{j_1, j_2=1}^n |\theta_{i_1, j_1} \theta_{i_2, j_2}| \sqrt{\rho}^{|i_1-j_1|+|i_2-j_2|} = \left(\sum_{i_1, j_1} |\theta_{i_1, j_1}| \sqrt{\rho}^{|i_1-j_1|} \right)^2 = O[N_n (\ell_2 - \ell_1) \theta_*^2].$$

Similarly, the bound $O[N_n(\ell_2 - \ell_1)\theta_*^2]$ holds for the third term in (35). So (33) holds.

6.3 Proof of Theorems 1–4

Lemma 3

Suppose the assumptions in Theorem 1 hold. Then

$$\sup_{[n\delta] \leq m \leq [n(1-\delta)]} |D_m| = o_p(\sqrt{n}), \quad \text{where} \quad D_m = \sum_{i=[n\delta]}^m \widetilde{e_i e_{i+k}} - \sum_{i=[n\delta]}^m e_i e_{i+k}.$$

Proof—It is easy to see the decomposition $e\mathcal{N}_j = e_j + f(i/n) - f(j/n) = e_j + \Delta_j - \zeta_j$ where

$$\Delta_i = f(i/n) - \frac{1}{nb} \sum_{j=1}^n f(j/n) K\left(\frac{i-j}{nb}\right) \quad \text{and} \quad \zeta_i = \frac{1}{nb} \sum_{j=1}^n e_j K\left(\frac{i-j}{nb}\right). \quad (39)$$

Thus, we can write

$$D_m = \sum_{i=[n\delta]}^m (e_i \Delta_{i+k} + \Delta_i e_{i+k} + \Delta_i \Delta_{i+k} + \zeta_i \zeta_{i+k} - e_i \zeta_{i+k} - \zeta_i e_{i+k} - \zeta_i \Delta_{i+k} - \Delta_i \zeta_{i+k}). \quad (40)$$

It suffices to show that each of the eight terms above is uniformly bounded by $o_p(\sqrt{n})$.

By Lemma 4 below, $\Delta_i = O[b^2 + (nb)^{-1}]$ uniformly over $[n\delta] \leq i \leq [n(1-\delta)]$. Thus, $\sum_{i=[n\delta]}^m \Delta_i \Delta_{i+k} = O\{n[b^2 + (nb)^{-1}]^2\} = o(\sqrt{n})$ under Assumption 1 (iii).

For $\sum_{i=[n\delta]}^m e_i \Delta_{i+k}$, by Proposition 3, we have under Assumption 1 (iii),

$$\sum_{i=[n\delta]}^m e_i \Delta_{i+k} = O_p\{[b^2 + (nb)^{-1}] \sqrt{n} \log n\} = o_p(\sqrt{n}),$$

uniformly over m . The same bound also holds for $\sum_{i=[n\delta]}^m \Delta_i e_{i+k}$.

For $\sum_{i=[n\delta]}^m \zeta_i \Delta_{i+k}$, we shall apply (17) in Proposition 4 by writing

$$\sum_{i=[n\delta]}^m \zeta_i \Delta_{i+k} = \frac{1}{nb} \sum_{i=[n\delta]}^m \sum_{j=1}^n \theta_{i,j} e_j, \quad \text{where} \quad \theta_{i,j} = \Delta_{i+k} K\left(\frac{i-j}{nb}\right).$$

For χ_n in (18), using $\theta_{i,j} = O[b^2 + (nb)^{-1}] K\{(i-j)/(nb)\}$ and $K\{(i-j)/(nb)\} = 0$ only when $|i-j| = O(nb)$, it is easy to see $\chi_n = O\{[b^2 + (nb)^{-1}]^2 nb\}$. Thus, by (17) in Proposition

4, $\sum_{i=\lfloor n\delta \rfloor}^m \zeta_i \Delta_{i+k} = O_p\{[b^2 + (nb)^{-1}] \sqrt{nb} \log n / (nb)\} = o_p(\sqrt{n})$, uniformly over m . The same bound also holds for $\sum_{i=\lfloor n\delta \rfloor}^m \Delta_i \zeta_{i+k}$.

Similarly, by Proposition 5, we can show that both $\sum_{i=\lfloor n\delta \rfloor}^m e_i \zeta_{i+k}$ and $\sum_{i=\lfloor n\delta \rfloor}^m \zeta_i e_{i+k}$ have the uniform bound $O_p[\sqrt{n^2 b \log n / (nb)}] = o_p(\sqrt{n})$. For $\sum_{i=\lfloor n\delta \rfloor}^m \zeta_i \zeta_{i+k}$, by (28), $\|\zeta_i\|_2^2 = O[(nb)^{-1}]$, uniformly over i . Thus, by the triangle inequality,

$$\mathbb{E} \left[\max_{\lfloor n\delta \rfloor \leq m \leq \lfloor n(1-\delta) \rfloor} \left| \sum_{i=\lfloor n\delta \rfloor}^m \zeta_i \zeta_{i+k} \right| \right] \leq \sum_{i=\lfloor n\delta \rfloor}^{\lfloor n(1-\delta) \rfloor} \mathbb{E}(|\zeta_i \zeta_{i+k}|) = O(1/b) = o(\sqrt{n}).$$

Here we have used the Cauchy-Schwarz inequality $\mathbb{E}(|\zeta_i \zeta_{i+k}|) \leq \|\zeta_i\|_2 \|\zeta_{i+k}\|_2 = O[(nb)^{-1}]$. This completes the proof.

Lemma 4

Let Δ_j be defined in (39). Assume $b \rightarrow 0$, $nb \rightarrow \infty$, and that Assumption 1 (i)–(ii) hold. Then $\Delta_j = O[b^2 + (nb)^{-1}]$, uniformly over $\lfloor n\delta \rfloor \leq j \leq \lfloor n(1-\delta) \rfloor$.

Proof—Without loss of generality, we assume that $K(\cdot)$ has support $[-1, 1]$ so that we only need to consider $|j - i|/n \leq b$ in Δ_j . By Taylor's expansion $f(j/n) = f(i/n) + f'(i/n)(j - i)/n + O(b^2)$, we write $\Delta_j = f(i/n)(1 - I_0) - f'(i/n)I_1 + O(b^2)I_0$, where

$$I_0 = \frac{1}{nb} \sum_{j=1}^n K\left(\frac{j-i}{nb}\right), \quad I_1 = \frac{1}{nb} \sum_{j=1}^n \frac{j-i}{n} K\left(\frac{j-i}{nb}\right).$$

The result then follows if we can prove $I_0 = 1 + O[(nb)^{-1}]$ and $I_1 = O(n^{-1})$. We only prove the first assertion since the second one can be similarly treated. Observe that

$$\begin{aligned} \left| I_0 - \frac{1}{nb} \int_0^n K\left(\frac{x-i}{nb}\right) dx \right| &\leq \frac{1}{nb} \sum_{j=1}^n \int_{j-1}^j \left| K\left(\frac{j-i}{nb}\right) - K\left(\frac{x-i}{nb}\right) \right| dx \\ &\leq \frac{\sup_u |K'(u)|}{nb} \sum_{|j-i| \leq nb+1} \int_{j-1}^j \frac{1}{nb} dx = \frac{O(1)}{nb}. \end{aligned} \quad (41)$$

Here, the second inequality in (41) follows from two facts: (1°) when $|j - i| > nb + 1$, we have $|j - i|/(nb) > 1$ and $|(x - i)/(nb)| > 1$ for all $x \in [j - 1, j]$, and consequently $K\{(j - i)/(nb)\} = 0$ and $K\{(x - i)/(nb)\} = 0$; (2°) $|K(u) - K(u')| \leq |u - u'| \sup_u |K'(u)|$. By the change-of-variable $(x - i)/(nb) = u$, it is easy to see that

$$\frac{1}{nb} \int_0^n K\left(\frac{x-i}{nb}\right) dx = \int_{-i/(nb)}^{(n-i)/(nb)} K(u) du = \int_{-1}^1 K(u) du = 1, \quad (42)$$

where the second equality in (42) follows since, for $\lfloor n\delta \rfloor \leq i \leq \lfloor n(1-\delta) \rfloor$ with a fixed constant $\delta \in (0, 1)$, we have $-i/(nb) < -1$ and $(n-i)/(nb) > 1$ for sufficiently small $b \rightarrow 0$. By (41) and (42), we obtain $I_0 = 1 + O[(nb)^{-1}]$.

Proof of Theorems 1–2

Let $\eta_i = e_i e_{i+k} - \gamma_k$ be defined in (11). Note that $\mathbb{E}(\eta_i) = 0$ and η_i is of the form (3). As in (4), consider the coupling process $\eta'_i = e'_i e'_{i+k} - \gamma_k$, where e'_i is defined in (4). By (5) and the inequality $\|XY\|_{q/2} \leq \|X\|_q \times \|Y\|_q$ for $X, Y \in \mathcal{L}^q$,

$$\|\eta'_i - \eta_i\|_{q/2} \leq \|e'_i(e'_{i+k} - e_{i+k})\|_{q/2} + \|e_{i+k}(e'_i - e_i)\|_{q/2} = O(\rho^i). \quad (43)$$

Thus $\{\eta_i\}$ satisfies GMC($q/2, \rho$). For $q > 4$, $n^{-1/2+1/\min(q/2, 4)} \log n = o(1)$. By a similar strong approximation in (31) for $\sum_{i=\lfloor n\delta \rfloor}^{\lfloor nt \rfloor} \eta_i$, $\{n^{-1/2} \sum_{i=\lfloor n\delta \rfloor}^{\lfloor nt \rfloor} \eta_i\}_t \Rightarrow \tau(k)\{IB_{t-\delta}\}_t$, where $\tau^2(k)$, defined in (11), is the long-run variance of $\{\eta_i\}$. By Lemma 3, Theorem 1 follows. Finally, Theorem 2 follows from Theorem 1 and the continuous mapping theorem.

Proof of Theorem 3

Let Δ_i and ζ_i be defined as in the proof of Lemma 3. By the proof of Lemma 3, $\max_i |\Delta_i| = O(b^2 + (nb)^{-1})$. For ζ_i , by (19) in Proposition 4, $\max_i |\zeta_i| = O_p((nb)^{-1} n^{1/4} \log n + [\log n / (nb)]^{1/2})$. Thus, under the condition $nb^2/(\log^2 n) \rightarrow \infty$,

$$\widehat{e}_i = e_i + \Delta_i + \zeta_i = e_i + O_p(\varepsilon_n), \quad \text{where} \quad \varepsilon_n = b^2 + \left(\frac{\log n}{nb}\right)^{1/2}, \quad (44)$$

uniformly over i . Also, by Theorem 1, $\widehat{\gamma}_k(1-\delta) = \gamma_k + O_p(1/\sqrt{n}) = \gamma_k + O_p(\varepsilon_n)$. Recall $\eta_i = e_i e_{i+k} - \gamma_k$ in (11). Then $e_i \eta_i e_{i+k} - \gamma_k(1-\delta) = \eta_i + O_p(\varepsilon_n) v_i$, where $v_i = 1 + |e_i| + |e_{i+k}|$. Therefore,

$$\begin{aligned} \text{cov}(\widehat{\eta}_0, \widehat{\eta}_j) &= \frac{1}{\ell_j} \sum_{i=\lfloor n\delta \rfloor}^{\lfloor n(1-\delta) \rfloor - j} [\eta_i + O_p(\varepsilon_n) v_i][\eta_{i+j} + O_p(\varepsilon_n) v_{i+j}] \\ &= \frac{1}{\ell_j} \sum_{i=\lfloor n\delta \rfloor}^{\lfloor n(1-\delta) \rfloor - j} \eta_i \eta_{i+j} + \frac{O_p(\varepsilon_n)}{\ell_j} \sum_{i=\lfloor n\delta \rfloor}^{\lfloor n(1-\delta) \rfloor - j} (|v_i \eta_{i+j}| + |\eta_i v_{i+j}| + |v_i v_{i+j}|). \end{aligned}$$

Since $I = o(n)$, we have $j_j \asymp n(1-2\delta)$. Consequently, we have the decomposition

$$\widehat{\tau}^2(k) = \tau_I^2 + R_n + O_p(1)\Omega_n, \quad (45)$$

where

$$\begin{aligned} \tau_I^2 &= \mathbb{E}(\eta_0^2) + 2 \sum_{j=1}^I \left(1 - \frac{j}{I+1}\right) \mathbb{E}(\eta_0 \eta_j), \\ R_n &= \Gamma_0 + 2 \sum_{j=1}^I \left(1 - \frac{j}{I+1}\right) \Gamma_j, \quad \text{with} \quad \Gamma_j = \frac{1}{\ell_j} \sum_{i=\lfloor n\delta \rfloor}^{\lfloor n(1-\delta) \rfloor - j} [\eta_i \eta_{i+j} - \mathbb{E}(\eta_0 \eta_j)], \\ \Omega_n &= \frac{\varepsilon_n}{n} \sum_{j=0}^I \sum_{i=\lfloor n\delta \rfloor}^{\lfloor n(1-\delta) \rfloor - j} (|v_i \eta_{i+j}| + |\eta_i v_{i+j}| + |v_i v_{i+j}|). \end{aligned}$$

In what follows we prove: (i) $\tau_I^2 \rightarrow \tau^2(k)$; (ii) $R_n = o_p(1)$; and (iii) $\Omega_n = o_p(1)$.

- i. $\tau_I^2 \rightarrow \tau^2(k)$. By the proof of Theorem 1, $\{\eta_j\}$ satisfies GMC(4, ρ) in (5). By Proposition 2, $\mathbb{E}(\eta_0 \eta_j)$ decays exponentially fast. Thus, by the Dominated Convergence Theorem, $\tau_I^2 \rightarrow \tau^2(k)$ as $I \rightarrow \infty$.
- ii. $R_n = o_p(1)$. Note that (28) is a general result: as long as $\{e_j\}$ satisfies the condition GMC(q , ρ) for some $q \geq 2$, then (28) holds. For the summands $\{\tilde{e}_j := \eta_j \eta_{i+j} - \mathbb{E}(\eta_0 \eta_j)\}$ in Γ_j by the assumption GMC(8, ρ) and the same argument in (43), we can show that the new process $\{\tilde{e}_j\}$ satisfies GMC(2, ρ). Applying (28), we obtain $\|\Gamma_j\|_2^2 = O(1/\ell_j) = O(1/n)$, uniformly in j . Thus, by the triangle inequality, $\|R_n\|_2 \leq \|\Gamma_0\|_2 + 2 \sum_{j=1}^I \|\Gamma_j\|_2 = O(I/\sqrt{n}) \rightarrow 0$.
- iii. $\Omega_n = o_p(1)$. Under the assumption $\mathbb{E}(|e_j|^3) < \infty$, there exists a constant $c < \infty$ such that $\mathbb{E}|v_i \eta_{i+j}| + \mathbb{E}|\eta_i v_{i+j}| + \mathbb{E}|v_i v_{i+j}| < c$. Therefore, $\mathbb{E}(\Omega_n) = O(e_n I) = o(1)$.

Proof of Theorem 4

It suffices to show that, with probability approaching one, $T_n^{(k)}/\sqrt{n} > c$ for some constant $c > 0$ and $\widehat{\tau}(k)/\sqrt{n} \rightarrow 0$.

Recall the decomposition of D_m in (40). Since Δ_i depends only on t_i , Lemma 4 still holds. In the right hand side of (40), by breaking the summation into two summations according to the two processes and applying the same argument in Lemma 3 to the two summations separately, we see that Lemma 3 still holds. Thus, by Lemma 3 and (10),

$$\widehat{\gamma}_k(1-\delta) = \frac{\sum_{i=\lfloor n\delta \rfloor}^{\lfloor n(1-\delta) \rfloor} \widetilde{e_i e_{i+k}}}{\lfloor n(1-\delta) \rfloor - \lfloor n\delta \rfloor + 1} = \frac{\sum_{i=\lfloor n\delta \rfloor}^{\lfloor n(1-\delta) \rfloor} e_i e_{i+k}}{\lfloor n(1-\delta) \rfloor - \lfloor n\delta \rfloor + 1} + o_p(n^{-1/2}). \quad (46)$$

$$\begin{aligned} & \sum_{i=\lfloor n\delta \rfloor}^{\lfloor n(1-\delta) \rfloor} e_i e_{i+k} \\ &= \sum_{i=\lfloor n\delta \rfloor}^{\lfloor nt_0 \rfloor} (e_i e_{i+k} - \gamma_{k,1}) \\ &+ \sum_{i=\lfloor nt_0 \rfloor}^{\lfloor n(1-\delta) \rfloor} (e_i e_{i+k} - \gamma_{k,2}) \\ &+ \{\lfloor nt_0 \rfloor - \lfloor n\delta \rfloor + 1\} \gamma_{k,1} \\ &+ \{\lfloor n(1-\delta) \rfloor - \lfloor nt_0 \rfloor\} \gamma_{k,2} = O_p(\sqrt{n}) + \{\lfloor nt_0 \rfloor - \lfloor n\delta \rfloor + 1\} \gamma_{k,1} \\ &+ \{\lfloor n(1-\delta) \rfloor - \lfloor nt_0 \rfloor\} \gamma_{k,2} \end{aligned}$$

By the proof of Theorems 1–2, (46),

. Thus, by

$$\widehat{\gamma}_k(1-\delta) \rightarrow \frac{(t_0 - \delta) \gamma_{k,1} + (1 - \delta - t_0) \gamma_{k,2}}{1 - 2\delta}, \quad \text{in probability.} \quad (47)$$

Similarly, $\gamma_k(t_0) \rightarrow \gamma_{k,1}$ in probability. Therefore, $\widehat{\gamma}_k(t_0) - \gamma_k(1-\delta) \rightarrow (\gamma_{k,1} - \gamma_{k,2})(1 - \delta - t_0)/(1 - 2\delta)$ in probability. Recall $T_n^{(k)}(t)$ in (12). Then

$$\frac{T_n^{(k)}}{\sqrt{n}} \geq \frac{|T_n^{(k)}(t_0)|}{\sqrt{n}} = \frac{\lfloor nt_0 \rfloor - \lfloor n\delta \rfloor + 1}{n} |\widehat{\gamma}_k(t_0) - \widehat{\gamma}_k(1-\delta)| \rightarrow c, \quad \text{in probability,}$$

where $c = |\gamma_{k,1} - \gamma_{k,2}|(1 - \delta - t_0)(t_0 - \delta)/(1 - 2\delta) > 0$.

It remains to show $\widehat{\tau}(k) = o_p(\sqrt{n})$. By (44) and (47), $e\mathcal{N}_i e\mathcal{N}_{i+k} - \gamma\mathcal{N}_k(1 - \delta) = e_i e_{i+k} + O_p(\varepsilon_n)(|e_i| + |e_{i+k}|) + O_p(1)$, uniformly over i . Then it can be easily seen that

$$\widehat{\tau}^2(k) \leq |\widehat{\text{cov}}(\eta_0, \eta_0)| + 2 \sum_{j=1}^I |\widehat{\text{cov}}(\eta_0, \eta_j)| = O_p(I). \quad (48)$$

Since $I = o(n)$, we conclude that $\widehat{\tau}(k)/\sqrt{n} \rightarrow 0$ in probability, completing the proof.

Figure 1.

Left: Plot of the weekly rates during January 5, 1962–June 10, 2011. Right: Plot of the residuals $e\mathcal{N}_i$ in (10) after removing the nonparametric time trend f .

Figure 2.

Plot of the cumulative sum process $|T_n^{(k)}(t)|/\widehat{\tau}(k)$ ($\delta=0.05$) ($\delta=0.05$) in (12) for $k=0$ (solid curve), $k=1$ (dotted curve), and $k=2$ (dashed curve). The horizontal line is the critical value 1.26 at the significance level 0.05. Here $\tau\mathcal{R}^2(k)$ is the estimator in (13) with $I=15$.

\$watermark-text

\$watermark-text

\$watermark-text

Table 1
Empirical power and size-corrected power (in parenthesis) for Model 1 in (23) ($\lambda=0.0$ corresponds to the size)

		λ			
k		0.0	0.2	0.4	0.6
$I=8$	0	0.053	0.113(0.102)	0.347(0.327)	0.697(0.676)
	1	0.048	0.437(0.449)	0.946(0.949)	1.000(1.000)
	2	0.055	0.170(0.161)	0.681(0.665)	0.988(0.986)
$I=10$	0	0.045	0.095(0.113)	0.355(0.395)	0.702(0.736)
	1	0.050	0.430(0.431)	0.955(0.955)	0.998(0.998)
	2	0.048	0.151(0.164)	0.695(0.715)	0.977(0.979)
$I=12$	0	0.045	0.078(0.079)	0.300(0.317)	0.694(0.714)
	1	0.046	0.409(0.417)	0.931(0.935)	0.999(1.000)
	2	0.045	0.137(0.159)	0.648(0.685)	0.963(0.973)

Table 2

Empirical power and size-corrected power (in parenthesis) for Model 2 in (24) ($\lambda = 0.0$ corresponds to the size)

		λ				
k		0.0	0.2	0.4	0.6	0.8
$I = 8$	0	0.055	0.044(0.041)	0.079(0.069)	0.408(0.393)	0.813(0.798)
	1	0.040	0.423(0.461)	0.967(0.970)	0.999(1.000)	1.000(1.000)
	2	0.044	0.053(0.059)	0.230(0.247)	0.810(0.823)	0.993(0.993)
$I = 10$	0	0.056	0.048(0.041)	0.088(0.074)	0.359(0.322)	0.762(0.744)
	1	0.039	0.437(0.469)	0.947(0.955)	1.000(1.000)	1.000(1.000)
	2	0.038	0.053(0.067)	0.189(0.234)	0.791(0.825)	0.992(0.996)
$I = 12$	0	0.056	0.039(0.038)	0.086(0.081)	0.341(0.334)	0.749(0.743)
	1	0.042	0.406(0.436)	0.945(0.960)	1.000(1.000)	1.000(1.000)
	2	0.048	0.055(0.055)	0.193(0.198)	0.783(0.785)	0.991(0.991)