Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

J Bus Econ Stat. Author manuscript; available in PMC 2016 April 01.

Published in final edited form as:

J Bus Econ Stat. 2015 April 1; 33(2): 296–306. doi:10.1080/07350015.2014.948177.

Inference for local autocorrelations in locally stationary models

Zhibiao Zhao*

Department of Statistics, Penn State University

Abstract

For non-stationary processes, the time-varying correlation structure provides useful insights into the underlying model dynamics. We study estimation and inferences for local autocorrelation process in locally stationary time series. Our constructed simultaneous confidence band can be used to address important hypothesis testing problems, such as whether the local autocorrelation process is indeed time-varying and whether the local autocorrelation is zero. In particular, our result provides an important generalization of the R function acf() to locally stationary Gaussian processes. Simulation studies and two empirical applications are developed. For the global temperature series, we find that the local autocorrelations are time-varying and have a "V" shape during 1910–1960. For the S&P 500 index, we conclude that the returns satisfy the efficient-market hypothesis whereas the magnitudes of returns show significant local autocorrelations.

Keywords

Local autocorrelation; Locally stationary model; Nonparametric regression; Simultaneous confidence band; Testing for zero autocorrelation; Time series

1 Introduction

Practical data often exhibit complicated time-varying patterns that can hardly be captured by simple parametric models, such as the classical autoregressive moving average models. To model non-stationary data, locally stationary models allow the model dynamics to change smoothly in time and thus can provide flexible and tractable alternatives over stationary models. Using spectral representations, Dahlhaus (1997) and Adak (1998) studied locally stationary models; other related works include Priestley (1965) and Ombao et al. (2005). In time domain analysis, Subba Rao (1970) proposed time-varying autoregressive models, Davis et al. (2006) studied piecewise stationary autoregressive models, and Dahlhaus and Subba Rao (2006) considered time-varying autoregressive conditional heteroscedastic (ARCH) models. For locally stationary wavelet processes, see Nason et al. (2000) and Van Bellegem and Von Sachs (2008). See Dahlhaus (2012) for a survey of contributions.

For non-stationary data, the time-varying dependence/correlation structure can provide useful insights into the dynamics of the underlying process, such as whether and how the dependence changes over time. For example, in global temperature studies, it is of interest to

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

The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIDA or the NIH.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

examine whether the dependence of temperature measurements has changed over the past century. In children's growth studies, it is reasonable to believe that the dependence of growth measurements has different patterns during early childhood, middle childhood, and adolescence. In financial markets, we can ask whether the market returns are more correlated in periods with high volatilities. In these applications, one common feature is the potentially time-varying dependence/correlation structure, and traditional tools based on the autocorrelations of stationary processes are of little use.

Motivated by the above discussion, we aim to study autocorrelations of locally stationary processes. For stationary processes, there is a vast literature on autocovariances/autocorrelations; see Wu and Xiao (2011) for a survey. On the other hand, developments on autocovariances/autocorrelations of non-stationary processes are quite limited. One approach is to divide data into blockwise stationary models [Ombao et al. (2001); Davis et al. (2006)], but this approach relies on selecting the number, location, and length of the blocks. For autocovariance estimation of locally stationary processes, some alternative approaches include the local cosine function method in Mallat et al. (1998), the wavelet periodograms method in Nason et al. (2000), and the local kernel autocovariance estimation in Dahlhaus (2012). However, these works did not obtain the related pointwise and simultaneous asymptotic distributions, the essential ingredients in statistical inference.

Our first goal is to quantify and estimate the time-varying correlation structure for a rich class of locally stationary time series introduced in Wu and Zhou (2011). At each lag k, we introduce the local autocorrelation curve, denoted by $\{\rho_k(t)\}_{t\in(0,1)}$, to characterize the local correlation of measurements near time $t \in (0, 1)$. As a time-varying curve, the local autocorrelation curve is a natural generalization of the classical autocorrelation function for stationary processes to locally stationary processes. Using the nonparametric kernel smoothing method, we propose a local sample autocorrelation estimate of $\rho_k(t)$.

Our second goal is to construct a simultaneous confidence band (SCB) for $\{\rho_k(t)\}_{t\in(0,1)}$. Compared to a pointwise confidence interval, an SCB can provide more useful insights into the overall trend of $\rho_k(t)$ on the time interval $t \in (0, 1)$. In the literature, SCB has been constructed for various marginal characteristics, including the marginal density function [Bickel and Rosenblatt (1973)], nonparametric mean function [Eubank and Speckman (1993); Wu and Zhao (2007)], and coefficient functions in varying-coefficient linear models [Fan and Zhang (2000)]. These marginal quantities are less informative when the goal is to study how time series observations are correlated with each other.

Our development on SCB for $\{\rho_k(t)\}_{t\in(0,1)}$ provides an appealing tool for a nonparametric assessment of the time-varying dynamics of the underlying model. The constructed nonparametric SCB can be used to examine whether the local autocorrelation curve is of certain parametric patterns by checking whether the parametric curve is contained within the SCB. We develop easy-to-use rule-of-thumb tests for some practically important hypotheses: (i) whether the local autocorrelations are indeed time-varying or constant, and (ii) whether the local autocorrelations are significant or not. For stationary time series, the R function acf() plots sample autocorrelations along with two horizontal 95% critical lines to examine whether there are significant autocorrelations among observations. By constructing

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

the 95% simultaneous cutoff bands, we provide an important generalization of the latter result to local autocorrelations of locally stationary processes.

The developed methodology for local autocorrelations has a wide range of applications in social and scientific fields, such as biomedical, climatic, environmental, economics, and financial fields, where non-stationary processes occur frequently. In Section 6 we illustrate two applications in the global temperature series and the S&P 500 index, and our methodology yields some interesting findings.

2 Two Motivating Data sets

2.1 Global temperature series

Consider the global monthly temperature anomalies during 1850–2010, relative to the 1961–1990 mean. The continually updated data set is available at

http://cdiac.ornl.gov/ftp/trends/temp/jonescru/global.txt,

and detailed background information can be found at

http://cdiac.ornl.gov/trends/temp/jonescru/jones.html.

There are $n = 161 * 12 = 1932$ observations, with a time series plot in Figure 1.

Figure 1 clearly indicates a time-varying trend, and in climate change studies there is a substantial literature on inferences for the mean trend $\mu(t)$. For example, Woodward and Gray (1993) fitted a linear trend, Rust (2003) argued that a linear trend is inadequate and fitted a quadratic trend, Wu et al. (2001) considered an isotonic regression to examine the change point in the mean trend, and Wu and Zhao (2007) constructed a simultaneous confidence band for the mean trend. The thick solid curve in Figure 1 is the nonparametric smoothing estimate of $\mu(t)$; see Section 6.1 for more details.

While the aforementioned works provided an extensive account of inferences for the mean trend $\mu(t)$, there are two limitations. First, they all assumed that the temperature series has a mean trend plus some stationary errors. As a modeling issue, it is hard to justify that the mean is time-varying while the dependence structure of the errors remains the same over the long time period of 161 years. In fact, our methodology (see Section 6.1) suggests that the error process has a complicated time-varying pattern, and thus the stationarity-based inference may be questionable. Second, these works did not address the equally important question of whether and how the dependence structure of temperature series varies over time. In contrast to the first-order marginal mean trend, the autocovariance/autocorrelation concerns the second-order property of the temperature series and can provide useful insights into the evolving dynamics of the series.

2.2 S&P 500 index returns

In finance, the efficient-market hypothesis (EMH) states that financial markets are efficient: stock price behaves like random walks and all subsequent price changes represent random departures from previous prices. See the review paper Fama (1970). In terms of

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

autocorrelations, EMH asserts that stock returns have zero autocorrelations and thus are not predictable. On the other hand, however, there is empirical evidence [Ding et al. (1993)] that the absolute values or magnitudes of returns are correlated.

The S&P 500 index includes 500 large-cap companies in leading industries of the U.S. economy and is one of the most widely used benchmarks for the overall U.S. stock market performance. Since weekly data are less noisy/volatile than daily data, we consider the weekly S&P 500 index during the 20 years period January 6, 1992–December 27, 2011; see the left plot in Figure 2. Denote the logarithm returns by $X_i = \log(S_i) - \log(S_{i-1})$, where S_i is the index at week i. There are $n = 1041$ observations, plotted in the right plot of Figure 2. It is desirable to develop some test for the EMH that can work under possible time-varying dependence of unknown form.

3 Inferences for Local Autocorrelation Curve

Motivated by the two examples in Section 2, in this section we study the time-varying dependence structure of non-stationary processes. For a random variable Z, write $Z \in \mathcal{L}^q$ if $\|Z\|_q := [\mathbb{E}(|Z|^q)]^{1/q} < \infty$, $q > 0$.

3.1 Local autocorrelation curve

We consider the non-stationary time series in Wu and Zhou (2011):

$$X_i = g(t_i, \xi_i), \quad \text{where} \quad t_i = i/n, \quad \xi_i = (\ldots, \varepsilon_{i-1}, \varepsilon_i), \quad i = 1, \ldots, n, \quad (1)$$

for i.i.d. innovations $\{\varepsilon_i\}$ and a time-varying transformation $g(t_i, \cdot)$. Assume

$$\sup_t \|g(t, \varepsilon_0)\|_2 < \infty \quad \text{and} \quad \limsup_{|t-s| \to 0} \|g(t, \xi_0) - g(s, \xi_0)\|_2 = 0. \quad (2)$$

Then $X_i = g(t_i, \xi_i) = g(t, \xi_i) + o_p(1)$ for $t_i \approx t$. That is, in the local time window $t_i \approx t$, $\{X_i\}_i$ can be approximated by the stationary process $\{g(t, \xi_i)\}_i$. Following Dahlhaus (1997) and Wu and Zhou (2011), we say that $\{X_i\}_i$ is locally stationary. Therefore, (1) allows the dependence structure to vary over time, and meanwhile it imposes a local stationarity. In time series analysis, autocorrelation functions measure the strength of dependence in a process. The autocorrelation function $\rho(\cdot, \cdot)$ for $\{X_i\}$ at lag $k \geq 0$ is

$$\rho(i, k) = \frac{\gamma(i, i+k)}{\sqrt{\gamma(i, i)\gamma(i+k, i+k)}}, \quad \text{where} \quad \gamma(i, i+k) = \operatorname{cov}(X_i, X_{i+k}). \quad (3)$$

By stationarity of ξ_i, $\gamma(i, i + k) = \operatorname{cov}\{g(t_i, \xi_0), g(t_{i+k}, \xi_k)\}$. Let $t \in (0, 1)$ be any fixed time point. Consider $i = \lfloor nt \rfloor$, with $\lfloor x \rfloor$ being the integer part of x, so that $t_i = \lfloor nt \rfloor / n \to t$ as $n \to \infty$. Let k be fixed. Using $g(t_i, \xi_0)g(t_{i+k}, \xi_k) - g(t, \xi_0)g(t, \xi_k) = \{g(t_i, \xi_0) - g(t, \xi_0)\}g(t_{i+k}, \xi_k) + \{g(t_{i+k}, \xi_k) - g(t, \xi_k)\}g(t, \xi_0)$ and the inequality $\mathbb{E}|UV| \leq \|U\|_2\|V\|_2$, we obtain the following result:

Proposition 1—Suppose (2) holds. Then for any fixed $k \geq 0$ and $t \in (0, 1)$,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

$$\lim_{n\to\infty}\rho(\lfloor nt\rfloor, k)=\frac{\gamma_k(t)}{\gamma_0(t)}\overset{\text{def}}{=}\rho_k(t), \quad where \quad \gamma_k(t)=\operatorname{cov}\{g(t,\xi_0), g(t,\xi_k)\}. \quad (4)$$

If $g(t, \xi_i)$ does not depend on t, then $\rho_k(t)$ is a constant and reduces to the usual autocorrelation function for stationary processes. In our locally stationary case, the time-varying curve $\rho_k(t)$ measures the correlation of observations in a neighborhood of t. We call $\{\rho_k(t)\}_{t\in[0,1]}$ the local autocorrelation curve.

Example 1—Consider the fixed-design nonparametric regression model: $X_i = \mu(t_i) + \sigma(t_i)e_i$, $i = 1, \ldots, n$, where $\{e_i \in \mathcal{L}^2\}_{i\in\mathbb{Z}}$ is a centered stationary process, $\mu(t)$ and $\sigma(t)$ are continuous functions on [0, 1] representing the time-varying trend and scale, respectively. Then $\rho_k(t) = \operatorname{corr}(e_0, e_k)$ is a constant, although the variance changes over time.

Example 2—The classical linear process is $e_i=\sum_{j=0}^{\infty}\alpha_j\varepsilon_{i-j}$, where $\{\varepsilon_i \in \mathcal{L}^2\}_{i\in\mathbb{Z}}$ are i.i.d. random variables with $\mathbb{E}(\varepsilon_i) = 0$, and $\{a_j\}_{j\geq 0}$ are constants satisfying $\sum_{j=0}^{\infty}\alpha_j^2<\infty$. For a natural extension, we consider the linear process with time-varying coefficients:

$$e_i(t)=\sum_{j=0}^{\infty}\alpha_j(t)\varepsilon_{i-j}, \quad i\in\mathbb{Z}, \quad t\in[0,1], \quad (5)$$

where $a_j(t)$, $j \geq 0$, are continuous functions on [0, 1] satisfying $\sup_{t\in[0,1]}\sum_{j=0}^{\infty}\alpha_j^2(t)<\infty$. Let $X_i = e_i(t_i)$ be discrete samples. Then $\rho_k(t)=\sum_{j=0}^{\infty}\alpha_j(t)\alpha_{j+k}(t)/\sum_{j=0}^{\infty}\alpha_j^2(t)$. If $a_j(t) = 0$ for $j \geq q + 1$, then (5) is the time-varying coefficient moving average of order q. Consider the time-varying autoregressive model: $e_i(t) = a(t)e_{i-1}(t) + \sigma(t)\varepsilon_i$ for some continuous functions $a(t)$ and $\sigma(t)$ on [0, 1]. Note that this is different from the time-varying autoregressive model $e_i = a(i/n)e_i + \sigma(i/n)e_i$ in, e.g., Subba Rao (1970), which does not admit the representation (5). If $\sup_{t\in[0,1]}|a(t)| < 1$, then (5) holds with $agr;_j(t) = \sigma(t)a^j(t)$ and $\rho_k(t) = a^k(t)$.

Example 3—Consider the threshold autoregressive model $e_i=\alpha|e_{i-1}|+\sqrt{1-\alpha^2}\varepsilon_i$, where $\{\varepsilon_i\}$ are i.i.d. $N(0, 1)$ random variables, and $a \in (-1, 1)$ controls the dependence. By Andel et al. (1984), the stationary distribution has the skew-normal density $f(u)=2\phi(u)\Phi(\alpha u/\sqrt{1-\alpha^2})$, where φ and Φ are the standard normal density and distribution functions, respectively. We consider the extension to the time-varying coefficient case:

$$e_i(t)=\alpha(t)|e_{i-1}(t)|+\sqrt{1-\alpha^2(t)}\varepsilon_i, \quad i\in\mathbb{Z}, \quad t\in[0,1], \quad (6)$$

where $a(t)$ is a continuous function and satisfies $\sup_t|a(t)| < 1$. Let $X_i = e_i(t_i)$ be discrete samples. Using the skew-normal density and $\operatorname{cov}\{e_i(t), e_{i-1}(t)\} = a(t)\operatorname{cov}\{|e_{i-1}(t)|, e_{i-1}(t)\}$, we can show

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

$$\rho_1(t)=\frac{2\alpha(t)}{\pi-2\alpha^2(t)}\left[\alpha(t)\left\{\sqrt{1-\alpha^2(t)}-1\right\}+\arctan\left\{\frac{\alpha(t)}{\sqrt{1-\alpha^2(t)}}\right\}\right]. \quad (7)$$

Figure 3 shows that $\{\rho_1(t)\}_t$ has quite different patterns for different choices of $\alpha(t)$.

Example 4—In econometrics, the widely used GARCH(1,1) model is $e_i = \sigma_i\varepsilon_i$, $\sigma_i^2=\beta_1+\beta_2 e_{i-1}^2+\alpha\sigma_{i-1}^2$, where $\beta_1 > 0$, $\beta_2 \geq 0$, $\alpha \geq 0$, and $\{\varepsilon_i\}$ are i.i.d. noises with zero mean and unit variance. As in Examples 2–3, the corresponding time-varying coefficient GARCH(1,1) is $e_i(t) = \sigma_i(t)\varepsilon_i$, $\sigma_i^2(t)=\beta_1(t)+\beta_2(t)e_{i-1}^2(t)+\alpha(t)\sigma_{i-1}^2(t)$, $i \in \mathbb{Z}$, $t \in [0, 1]$, where $\beta_1(t) > 0$ and $\beta_2(t) \geq 0$, $\alpha(t) \geq 0$ are continuous functions on $[0, 1]$ satisfying $\inf_t \beta_1(t) > 0$ and $\sup_t[\beta_2(t) + \alpha(t)] < 1$. As a special case, if $\alpha(t) = 0$, then the latter model reduces to the time-varying coefficient ARCH(1) model in Dahlhaus and Subba Rao (2006). Let $X_i = e_i(t_i)$ be discrete samples. Then $\rho_k(t) = 0$ for $k \geq 1$.

3.2 Local sample autocorrelation process

Throughout this paper, our theory and methods rely on the assumption $\mu(t) := \mathbb{E}[g(t, \xi_0)] = 0$ for all t so that $\mathbb{E}(X_i) = 0$. In practice, we can subtract a nonparametric estimate of $\mu(t_i)$ from X_i; see Section 6.1. Due to technical difficulty, it is non-trivial to theoretically quantify the impact from pre-centering the data. From our simulation study in Section 5, our methods still yield reasonable performance when we pre-center the data using local linear regression. Based on our empirical studies, we recommend:

- **i.** For time series data (e.g., global temperature series) with a clear time trend, use the local linear regression to remove the trend; see Section 6.1.
- **ii.** For daily or weekly financial returns data, the returns usually oscillate around zero and are negligible relative to the volatility, and thus we can assume that the returns are already centered. For absolute returns, we can use the local linear regression to remove the trend. For squared returns, the time trend is usually negligible (for a 1% return, the squared return is 0.0001) relative to the strength of the autocorrelation and the potential error from fitting a complicated local linear regression, and thus we recommend simply subtracting the sample mean squared returns.

Given $X_1, \ldots, X_n$ from (1), we discuss estimation of $\rho_k(t)$ at a given $t \in (0, 1)$. If $\{X_i\}$ is stationary, then $\rho_k(t)$ is a constant and can be estimated by the sample correlation $(n^{-1}\sum_{i=1}^{n-k} X_i X_{i+k})/(n^{-1}\sum_{i=1}^{n} X_i^2)$. For (1), in light of the local stationarity, we propose estimating $\rho_k(t)$ based on local data with $t_i \approx t$. Specifically, we propose the following nonparametric kernel smoothing estimate:

$$\hat{\rho}_k(t)=\frac{\hat{\gamma}_k(t)}{\hat{\gamma}_0(t)}, \quad \text{where} \quad \hat{\gamma}_k(t)=\frac{1}{nb}\sum_{i=1}^{n} X_i X_{i+k} K\left(\frac{t_i-t}{b}\right), \quad (8)$$

for a kernel function $K(\cdot)$ and a bandwidth $b > 0$. Since $\hat{\rho_k}(t)$ is the sample autocorrelation based on local data, we call $\{\hat{\rho_k}(t)\}_{t \in (0,1)}$ the local sample autocorrelation process.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

In (1), the pair (g, ξ_i) determines the probabilistic structure of $\{X_i\}$. To study asymptotic properties, we impose the following dependence condition:

Assumption 1—For $p, q > 0$, we write $(g, \xi_i) \in SLC(p)\text{-}GMC(q)$ if they satisfy $\mathcal{L}^p$ stochastic Lipschitz continuity (*SLC*) condition and $\mathcal{L}^q$ geometric moment contraction (*GMC*) condition in the sense: (*i*) $\sup_{s \neq t} \|g(t, \xi_0) - g(s, \xi_0)\|_p/|t - s| < \infty$; (ii) for an i.i.d. copy ε_0' of ε_0 and $\xi_i' = (\xi_{-1}, \varepsilon_0', \varepsilon_1, \ldots, \varepsilon_i)$, there exists a constant $\lambda \in (0, 1)$ such that

$$\sup_{t\in[0,1]} \|g(t,\xi_0)\|_q < \infty, \quad \omega_i = \sup_{t\in[0,1]} \|g(t,\xi_i) - g(t,\xi_i')\|_q = O(\lambda^i), \quad i \geq 0. \tag{9}$$

Assumption 1 (i) imposes a stochastic ($\mathcal{L}^p$ norm) Lipschitz continuity condition on g as inferences would be impossible if data jump randomly without any continuity. In (9), $\sup_{t \in [0,1]} \|g(t, \xi_0)\|_q < \infty$ imposes the finite q-th moment assumption, and ω_i measures the effect of replacing ε_0 with an i.i.d. copy ε_0', which in turn imposes explicit conditions on the strength of the dependence. Assumption 1 is satisfied for many linear and nonlinear locally stationary time series; see Wu and Zhou (2011).

Assumption 2—(i) $K(\cdot)$ has support $[-1, 1]$, is symmetric and continuously differentiable, and $\int_{\mathbb{R}} K(u)du = 1$; throughout, $\psi_K = \int_{\mathbb{R}} u^2K(u)du/2$, $\phi_K = \int_{\mathbb{R}} K^2(u)du$. (ii) $nb^9 \log n \to 0$ and $nb^2/(\log n)^8 \to \infty$. (iii) $\gamma_k(t)$ and $\gamma_0(t)$ are four times continuously differentiable.

Theorem 1—Let $(g, \xi_i) \in SLC(4)\text{-}GMC(8)$ (see Assumption 1) and Assumption 2 hold. Define $\beta_i(t) = [g(t, \xi_i)g(t, \xi_{i+k}) - \rho_k(t)g^2(t, \xi_i)]/\gamma_0(t)$,

$$\Gamma_k^2(t) = \mathrm{var}\{\beta_0(t)\} + 2\sum_{r=1}^{\infty} \mathrm{cov}\{\beta_0(t), \beta_r(t)\}. \tag{10}$$

Further assume $\inf_t \Gamma_k(t) > 0$ and $\inf_t \gamma_0(t) > 0$. Then, for each fixed $t \in (0, 1)$,

$$\left(\frac{nb}{\varphi_K}\right)^{1/2} \frac{\hat{\rho}_k(t) - \rho_k(t) - b^2\psi_K\lambda(t)}{\Gamma_k(t)} \Rightarrow N(0,1), \quad \lambda(t) = \frac{\gamma_k''(t) - \rho_k(t)\gamma_0''(t)}{\gamma_0(t)}.$$

It is worth pointing out that the fourth differentiability on $\gamma_k(t)$ in Assumption 2 (iii) can be weakened to twice differentiability under a stronger bandwidth condition.

3.3 Simultaneous confidence band

Theorem 1 can be used to construct an asymptotic pointwise confidence interval for $\rho_k(t)$ at each $t \in [b, 1 - b]$. However, in practice, it is desirable to construct a simultaneous confidence band (SCB). We say that $[l(t), u(t)]$ is an asymptotic $(1 - \alpha)$ SCB for $\rho_k(t)$ if

$$\lim_{n\to\infty} \mathbb{P}\{l(t) \leq \rho_k(t) \leq u(t), \quad \text{for all } t \in (0,1)\} = 1 - \alpha. \tag{11}$$

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

In practice, a typical choice is the $1 - \alpha = 95\%$ SCB.

Theorem 2—Assume the same conditions in Theorem 1. Then, for all $z \in \mathbb{R}$,

$$\mathbb{P}\left\{\left(\frac{nb}{\varphi_K}\right)^{1/2} \max_{t\in[b,1-b]} \frac{|\hat{\rho}_k(t)-\rho_k(t)-b^2\psi_K\lambda(t)|}{\Gamma_k(t)} \le \sqrt{-2\log b}+\frac{C_K+z}{\sqrt{-2\log b}}\right\} \to e^{-2e^{-z}},$$

where $\Gamma_k(t)$ and $\lambda(t)$ are defined in Theorem 1 and $C_K = 0.5 \log[\int_{\mathbb{R}} |K'(u)|^2 du/(4\pi^2\phi_K)]$.

The limiting distribution in (12) is the extreme value distribution. In Section 3.4, we show that the bias $b^2\psi_K\lambda(t)$ can be reduced to a negligible high-order term by using a higher-order kernel. Therefore, let $\hat{\Gamma}_k(t)$ be a consistent estimate of $\Gamma_k(t)$ (see Section 4.1), based on Theorem 2, an asymptotic $(1 - \alpha)$ SCB for $\{\rho_k(t)\}_{t\in[b,1-b]}$ is

$$\hat{\rho}_k(t) \pm C(b,\alpha)\hat{\Gamma}_k(t)\left(\frac{\varphi_K}{nb}\right)^{1/2}, \quad C(b,\alpha)=\sqrt{-2\log b}+\frac{C_K-\log\log\frac{1}{\sqrt{1-\alpha}}}{\sqrt{-2\log b}}. \quad (12)$$

The factor $C(b, \alpha)$ reflects the simultaneousness of the confidence band. By contrast, if we replace $C(b, \alpha)$ in (12) by the $(1 - \alpha/2)$ standard normal quantile, by Theorem 1, then we have a pointwise $(1 - \alpha)$ confidence band. In practice, SCB can depict the overall trend and variability of the process $\rho_k(t)$; see Section 4.2–4.3 for applications of SCB.

3.4 Bandwidth selection and bias correction

(Bandwidth selection)—The estimator $\hat{\gamma}_k(t)$ in (8) is the Priestley and Chao (1972) nonparametric kernel smoothing estimator of the mean trend $\text{cov}(X_i, X_{i+k})$ in the fixed-design regression $X_iX_{i+k}=\text{cov}(X_i, X_{i+k})+e_i^*$ with $e_i^*=X_iX_{i+k}-\text{cov}(X_i, X_{i+k})$. For nonparametric regression of independent data, Ruppert et al. (1995) proposed an automatic bandwidth selector, which is well-known to undersmooth correlated data. Let b^* be the automatic bandwidth implemented using the function dpill in the R package KernSmooth. Based on our simulations in Section 5, $b = 1.5b^*$ works quite well. To avoid choosing b for each k, we choose b based on $k = 0$ and use the same b for all k.

(Bias correction)—In Theorems 1–2, due to the unknown derivatives $\gamma_k''(t)$ and $\gamma_0''(t)$, it is difficult to estimate the second-order bias $b^2\psi_K\lambda(t)$, and we adopt a higher-order kernel to remove it. We use the kernel $K(u)=2\phi(u)-\phi(u/\sqrt{2})/\sqrt{2}$ with the Gaussian kernel $\phi(u) = (2\pi)^{-1/2} \exp(-u^2/2)$. Then, in Theorems 1–2, simple calculation shows $\psi_K = 0$, $\phi_K = 0.41$, $C_K = -1.98$, and thus the bias $b^2\psi_K\lambda(t) = 0$. We point out that using this higher-order kernel does not completely remove the bias, instead there are still negligible higher-order bias terms of the order $O(b^4)$.

4 Gaussianity-based Rule-of-thumb Inferences

In this section we assume that $\{g(t,\xi_i)\}_{i\in\mathbb{Z}}$ is a Gaussian process for each t and develop some easy-to-use inferences, which we term by the rule-of-thumb inferences.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

While it may seem restrictive at first, the Gaussian assumption is quite reasonable here based on two facts: (i) By the Wold's decomposition, any zero-mean covariance-stationary process can be represented by a linear process with uncorrelated innovations (assuming no deterministic component); and (ii) Gaussian process and linear processes with independent innovations share the equivalent second-order moment properties. Thus, in light of the local stationarity of the locally stationary process, in order to study our (second-order) local autocorrelations it is reasonable to work under the locally stationary Gaussian framework, provided that we ignore the gap between the uncorrelation and independence in the aforementioned (i)–(ii). In fact, in the literature on locally stationary processes, the Gaussian assumption is frequently imposed to facilitate inferences [Nason et al. (2000); Van Bellegem and Von Sachs (2008)]. Furthermore, simulation studies in Section 5 show that the rule-of-thumb inferences work reasonably well even for non-Gaussian processes.

4.1 Rule-of-thumb estimation of $\Gamma_k^2(t)$ and rule-of-thumb SCB

To implement the SCB in Section 3.3, we need to estimate the long-run variance $\Gamma_k^2(t)$ in (10). For stationary processes, long-run variance estimation involves the selection of a block length parameter and has the optimal convergence rate $O(n^{1/3})$ [Lahiri (2003)]. For locally stationary processes, this problem becomes even more challenging as we can only use local data, and as a result we need to choose both a block length parameter and a local bandwidth, and the convergence rate of the estimator can be quite slow. To attenuate these issues, we propose a rule-of-thumb estimator based on the Gaussian assumption.

Proposition 2—Suppose $\{g(t,\xi_i)\}_{i\in\mathbb{Z}}$ is a Gaussian process. Then, for $\Gamma_k^2(t)$ in (10),

$$\Gamma_k^2(t)=\sum_{r=1}^{\infty}\left[2\rho_k(t)\rho_r(t)-\rho_{k-r}(t)-\rho_{k+r}(t)\right]^2 \quad (13)$$

Recall $\hat{\rho_k}(t)$ in (8). Proposition 2 suggests the estimator of $\Gamma_k^2(t)$ (L is a truncation lag)

$$\hat{\Gamma}_k^2(t)=\sum_{r=1}^{L}\left[2\hat{\rho}_k(t)\hat{\rho}_r(t)-\hat{\rho}_{k-r}(t)-\hat{\rho}_{k+r}(t)\right]^2. \quad (14)$$

The estimator $\hat{\Gamma}_k^2(t)$ only requires L and does not involve any extra estimation. However, it is non-trivial to choose L. For stationary processes, the optimal $L \propto n^{1/3}$. In our locally stationary case, the optimal bandwidth $b \propto n^{-1/5}$ and the local sample size $\propto n^{4/5}$, thus we propose $L = \varrho(n^{4/5})^{1/3} = \varrho n^{4/15}$ for some ϱ. By our simulations, $\varrho = 2$ performs well.

Substituting the rule-of-thumb estimator $\hat{\Gamma}_k^2(t)$ in (14) into (12), we obtain the rule-of-thumb SCB. With the given kernel in Section 3.4, the $1-\alpha = 95\%$ SCB in (12) becomes

$$\hat{\rho}_k(t) \pm \hat{\Gamma}_k(t)\frac{0.64}{\sqrt{nb}}\left(\sqrt{-2\log b}+\frac{1.68}{\sqrt{-2\log b}}\right). \quad (15)$$

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

4.2 Test for constant or zero local autocorrelations at a given lag

Nonparametric SCB for $\{\rho_k(t)\}$ offers a reference to which other parametric autocorrelations can compare. By examining whether the parametric autocorrelation is contained within the 95% SCB (15), we can decide whether to reject a parametric null hypothesis at significance level 5%. This and next sections address some important hypotheses.

Consider testing whether $\{\rho_k(t)\}$ at a given lag k is constant or time-varying, i.e. $H_0 : \rho_k(t) = \rho_k$, $t \in [0, 1]$, for some constant ρ_k. Under H_0, we estimate ρ_k by

$$\hat{\rho}_k = \frac{n^{-1}\sum_{i=1}^{n-k} X_i X_{i+k}}{n^{-1}\sum_{i=1}^{n} X_i^2}, \quad (16)$$

and check whether $\hat{\rho}_k$ is contained within the SCB.

Another interesting question is to test $H_0 : \rho_k(t) = 0$, $t \in [0, 1]$, at a given lag k. We can examine whether the constant horizontal line $\rho_k(t) = 0$ is contained within the SCB.

4.3 Test for local uncorrelation: An extension of R function acf()

In R, the function acf() computes the sample autocorrelations along with two horizontal dashed blue lines $\pm 1.96/\sqrt{n}$, which are the 95% acceptance region for the null hypothesis of zero autocorrelations based on the assumption of n i.i.d. normal random variables. Any significant departure from the two lines would indicate significant autocorrelations. We shall develop a counterpart result for local autocorrelations of locally stationary processes.

As in R, we assume that $\{g(t, \xi_i)\}_{i\in\mathbb{Z}}$ is a Gaussian process for each t. We are interested in the null hypothesis of zero local autocorrelations at all lags (or local uncorrelation):

$$H_0 : \rho_k(t) = 0, \quad t \in [0, 1], \quad \text{for all } k = 1, 2, \ldots [\text{note that } \rho_0(t) = 1]. \quad (17)$$

Under H_0, by Proposition 2, $\Gamma_k^2(t) = 1$. By (12), the simultaneous acceptance bands are

$$\pm C(b, \alpha)\left(\frac{\varphi_K}{nb}\right)^{1/2}, \quad C(b, \alpha) = \sqrt{-2\log b} + \frac{C_K - \log\log\frac{1}{\sqrt{1-\alpha}}}{\sqrt{-2\log b}}, \quad t \in [b, 1-b]. \quad (18)$$

The bands (18) are similar to the cutoff lines $\pm 1.96/\sqrt{n}$ produced by acf() for stationary processes. In our locally stationary case, nb/φ_K reflects the effective local sample size, and $C(b, \alpha)$ is from the simultaneousness of the bands. In practice, we use $1 - \alpha = 95\%$ and the kernel $K(\cdot)$ given in Section 3.4, then (18) becomes

$$\pm \frac{0.64}{\sqrt{nb}}\left(\sqrt{-2\log b} + \frac{1.68}{\sqrt{-2\log b}}\right), \quad t \in [b, 1-b]. \quad (19)$$

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

The bands (19) are model-free, easy to implement, and can serve as a rule-of-thumb criterion for testing for zero autocorrelations of locally stationary Gaussian processes. Any significant departure from the bands would suggest rejection of the null hypothesis.

Remark 1—We can apply the above method to test for local uncorrelation of residuals from parametric models. Consider $X_i = \mu(\theta, U_i) + e_i$, where U_i is the covariate, $\mu(\theta, \cdot)$ is a parametric function with unknown parameter vector θ, and $\{e_i\}$ is a locally stationary process. For example, $\mu(\theta, U_i) = \theta^T U_i$ corresponds to the linear model. Let $\hat{\theta}$ be a $\sqrt{n}$-consistent estimate of θ. Suppose each component of the partial derivative $\partial\mu(\theta, U_i)/\partial\theta$ can be bounded by some random variable $L(U_i) \in \mathcal{L}^2$ uniformly in i and a neighborhood of θ. Since $\hat{\theta}$ has a faster parametric convergence rate ($\sqrt{n}$) than the nonparametric rate ($\sqrt{nb}$) of the local sample autocorrelation, it can be shown that the impact from estimating θ is negligible and thus the method can be used to test for local uncorrelation of $\{\hat{e}_i\}$.

Remark 2—To test for uncorrelation of stationary process, the cutoff lines $\pm 1.96/\sqrt{n}$ work for a given lag, and we can use some portmanteau test, such as the Ljung–Box test, to combine autocorrelations at multiple lags. Our SCB is simultaneous with respect to the time t but for a given lag k, and it seems quite challenging to extend the Ljung–Box test to our local autocorrelation setting. This will serve as a direction for future research.

5 Simulation studies

Consider $X_i = e_i(t_i)$, $i = 1, \ldots, n$, where $\{e_i(t_i)\}_i$ are discrete samples from a process $\{e_i(t)\}_i$. Let $\{\varepsilon_i\}$ be i.i.d. standard normal variables. Consider three models for $\{e_i(t)\}$:

$$\text{Model 1:} \quad e_i(t) = \varepsilon_i + 3\theta t\varepsilon_{i-1} - \cos(\pi t)\varepsilon_{i-2}, \quad (20)$$

$$\text{Model 2:} \quad e_i(t) = 0.6\{(1-\theta) + \theta\sin(2\pi t)\}e_{i-1}(t) + \varepsilon_i, \quad (21)$$

$$\text{Model 3:} \quad e_i(t) = 0.9\theta t|e_{i-1}(t)| + \sqrt{1-(0.9\theta t)^2}\varepsilon_i. \quad (22)$$

Here the parameter $\theta \in [0, 1]$ controls the strength of the dependence and the shape of local autocorrelations. Models 1–2 are special cases of Example 2 and Model 3 is a special case of Example 3. Due to nonlinearity, Model 3 is a non-Gaussian process. We use Model 3 to examine the impact from pre-centering the data. Specifically, we consider two approaches: (i) subtract the theoretical mean $0.9\theta t_i\sqrt{2/\pi}$ from X_i; and (ii) subtract the estimated mean using the local linear regression (24) in Section 6.1 below.

For each choice of $\theta = 0.0, 0.2, 0.5, 0.8, 1.0$, we examine the empirical coverage probabilities of the constructed SCB's for $\{\rho_1(t)\}_{t\in[0.1,0.9]}$ and $\{\rho_2(t)\}_{t\in[0.1,0.9]}$ and the power of the test based on SCB. For the empirical coverage probabilities, two nominal levels 90% and 95% are considered. For SCB based tests in Section 4, we consider the null hypotheses:

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

$$H_0^k: \rho_k(t), t \in [0.1, 0.9], \text{ is a constant function}, \quad k=1, 2, \quad (23)$$

for each of the three models with significance level 5%. To examine the effect of the bandwidth b, following the discussion in Section 3.4, we use three choices $b = b^*$, $1.5b^*$, $2b^*$, where b^* (the second column of Table 1) is the average of 1000 bandwidths based on the automatic bandwidth selector implemented using the R command dpill. In all cases, we use sample size $n = 500$ and compute the coverage probabilities and power based on 1000 realizations. As discussed in Section 4.1, we use $L = 10 \approx 2n^{4/15}$ in (14).

The results are summarized in Table 1. Overall, the empirical coverage probabilities are close to the two nominal levels 90% and 95%, except that Model 3 with $\theta = 1.0$ has a slight undercoverage due to the strong dependence and nonlinearity. Furthermore, the bandwidth choice $b = 1.5b^*$ slightly outperforms the other choices $b = b^*$, $2b^*$. For the power study, as the model moves away from the null hypotheses H_0^1 and H_0^2, the power of the SCB based test generally increases. Also, the proposed test for H_0^1 has better power than H_0^2. This phenomenon is due to the relatively larger magnitude of $|\rho_1(t)|$ than that of $|\rho_2(t)|$. In general, it is difficult to test the shape of a curve when the magnitude of the curve is too small. For example, for Model 2, $\rho_2(t) = [0.6\{(1 - \theta) + \theta \sin(2\pi t)\}]^2 \leq 0.36$, which explains the much lower power in testing for H_0^2 than H_0^1. For Model 3, the methods have comparable performance when we subtract either the nonparametrically estimated mean or the theoretical mean.

6 Applications

Now we apply the methodology to the two examples in Section 2.

6.1 Global temperature series in Section 2.1

We study local autocorrelations for the global monthly temperature anomalies in Section 2.1. Denote by $\mu(t)$ the time-varying mean trend. Due to the clear time trend, we need to center the data first. Consider the local linear estimate of $\mu(t)$ with Gaussian kernel $K(\cdot)$:

$$\left(\hat{\mu}(t), \hat{\mu}'(t)\right) = \underset{(\mu,\mu')}{\operatorname{argmin}} \sum_{i=1}^{n} \{X_i - \mu - \mu'(t_i - t)\}^2 K\left(\frac{t_i - t}{\tau}\right). \quad (24)$$

To choose the bandwidth τ, we follow the discussions in Sections 3.4 and 5 and let $\tau = 1.5\tau^*$, where τ^* is the automatic bandwidth selected by the R function dpill assuming independence. The thick solid curve in Figure 1 is the estimated trend $\hat{\mu}(t)$. We then apply our methods to the centered data $\tilde{X}_i = X_i - \hat{\mu}(i/n)$.

By the method in Section 3.4, we use bandwidth $b = 0.033$ in (8) for all k. The left plot of Figure 4 plots $\hat{\rho}_k(t)$ at lags $k = 1, \ldots, 4$. The four local sample autocorrelation processes exhibit similar but complicated time-varying patterns. Interestingly, there is a clear "V" shape during 1910–1960. To test for the null hypothesis of zero local autocorrelations, we plot 95% simultaneous cutoff bands in (19) as the two horizontal lines. Clearly, the null

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

hypothesis is rejected, and the measurements are significantly positively correlated. To examine whether the local autocorrelations are indeed time-varying, in the middle and right plots of Figure 4 we plot $\hat{\rho_1}(t)$ and $\hat{\rho_2}(t)$ along with their 95% SCB, where $L = 15 \approx 2n^{4/15}$ in (14). The horizontal lines, computed using (16), are not entirely contained within the SCB, thus we conclude that the local autocorrelations are indeed time-varying. Our findings suggest that the temperature series have an intrinsically complicated structure and should not be modeled by a mean trend plus a stationary noise as in the existing literature. A future research direction is to explore specific time-varying coefficient models, such as the time-varying autoregressive models in Example 2.

6.2 S&P 500 index returns in Section 2.2

By the discussion in Section 2.2, we apply our local autocorrelations methodology to examine two important hypotheses: (i) the EMH that stock returns have zero autocorrelations; and (ii) the hypothesis that the absolute values or magnitudes of returns are correlated. Furthermore, numerous empirical studies have shown that squared returns are positively correlated, a phenomenon that can be very well explained by the widely used ARCH/GARCH models. We shall also examine this phenomenon using our methods.

Figure 5 plots the local autocorrelations $\hat{\rho_k}(t)$ at lags $k = 1, \ldots, 4$, along with the 95% simultaneous cutoff bands (horizontal long-dashed lines) in (19), for the returns X_i (left plot), the absolute returns $|X_i|$ (middle plot), and squared returns X_i^2 (right plot). When calculating $\hat{\rho_k}(t)$, we adopt the following rule: (i) do not center the returns; (ii) use local linear regression to center $|X_i|$, i.e., $|X_i| - \hat{\mu}(t_i)$; and (iii) subtract the sample mean squared returns to center X_i^2, i.e., $X_i^2 - n^{-1}\sum_{i=1}^{n} X_i^2$. The rationale is to avoid local linear regression (which always introduces estimation errors) when the mean trend is weak relative to the strength of autocorrelations. See Section 3.2 for detailed explanations.

From Figure 5, we make the following observations:

First, in the left plot, since all the curves $\hat{\rho_1}(t), \ldots, \hat{\rho_4}(t)$ are contained within the cutoff bands, we have no evidence to reject the null hypothesis of zero local autocorrelations. Thus, we conclude that the returns are uncorrelated, which confirms the EMH that it should be impossible to predict stock returns.

Second, in the middle plot, since $\hat{\rho_1}(t)$ is not entirely contained within the bands, we reject the null hypothesis of zero local autocorrelations at the 5% significance level. Thus, although the returns are uncorrelated, the absolute returns are significantly correlated at lag $k = 1$. Our conclusion is consistent with those in Fama (1970) and Ding et al. (1993). It is worth pointing out that, the absolute values of returns are important in risk management. For example, a risk-aversion investor may want to stay on the sidelines when he/she expects large magnitudes of returns, equally likely to be positive or negative. Interestingly, comparing the left plot of Figure 2 and the middle plot of Figure 5, we see that the peaks and bottoms of the stock market correspond to the bottoms and peaks of the local autocorrelation process $\hat{\rho_1}(t)$ of the absolute returns. For example, the stock market peaked in 2007 when the local autocorrelation bottomed; the subsequent financial crisis during 2008 and early 2009

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

corresponds to a dramatic increase in $\hat{\rho}_1(t)$ with a peak at approximately the time when the market bottomed.

Third, the right plot suggests significant and mostly positive local autocorrelations for the squared returns, which is consistent with the ARCH/GARCH modeling in existing literature. A future research direction is to investigate specific time-varying coefficient ARCH/GARCH models (see Example 4) with nonparametrically estimated coefficients and data-driven order selection.

Acknowledgments

We are grateful to an associate editor and a referee for their constructive comments. Zhao's research was supported by a NSF grant DMS-1309213 and a NIDA grant P50-DA10075-15.

References

Adak S. Time-dependent spectral analysis of nonstationary time series. Journal of the American Statistical Association. 1998; 93:1488–1501.

Andel J, Netuka I, Svara K. On threshold autoregressive processes. Kyber-netika. 1984; 20:89–106.

Bickel PJ, Rosenblatt M. On some global measures of the deviations of density function estimates. Annals of Statistics. 1973; 1:1071–1095.

Dahlhaus R. Fitting time series models to nonstationary processes. Annals of Statistics. 1997; 25:1–37.

Dahlhaus, R. Locally stationary processes. In: Subba Rao, T.; Subba Rao, S.; Rao, CR., editors. Handbook of Statistics, Time Series Analysis: Methods and Applications. Vol. 30. 2012. p. 351-413.

Dahlhaus R, Subba Rao S. Statistical inference for time-varying ARCH processes. Annals of Statistics. 2006; 34:1075–1114.

Davis RA, Lee T, Rodriguez-Yam G. Structural break estimation for non-stationary time series models. Journal of the American Statistical Association. 2006; 101:223–239.

Ding Z, Granger CWJ, Engle RF. A long memory property of stock market returns and a new model. Journal of Empirical Finance. 1993; 1:83–106.

Eubank RL, Speckman PL. Confidence bands in nonparametric regression. Journal of the American Statistical Association. 1993; 88:1287–1301.

Fama E. Efficient capital markets: A review of theory and empirical work. Journal of Finance. 1970; 25:383–417.

Fan J, Zhang W. Simultaneous confidence bands and hypothesis testing in varying-coefficient models. Scandinavian Journal of Statistics. 2000; 27:715–731.

Isserlis L. On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. Biometrika. 1918; 12:134–139.

Lahiri, SN. Resampling Methods for Dependent Data. New York: Springer-Verlag; 2003.

Mallat S, Papanicolaou G, Zhang Z. Adaptive covariance estimation of locally stationary processes. Annals of Statistics. 1998; 26:1–47.

Nason GP, Von Sachs R, Kroisandt G. Wavelet processes and adaptive estimation of the evolutionary wavelet spectrum. Journal of the Royal Statistical Society, Series B. 2000; 62:271–292.

Ombao H, Raz J, Von Sachs R, Malow BA. Automatic statistical analysis of bivariate nonstationary time series. Journal of the American Statistical Association. 2001; 96:543–560.

Ombao H, Von Sachs R, Guo W. SLEX analysis of multivariate nonstationary time series. Journal of the American Statistical Association. 2005; 100:519–531.

Priestley MB, Chao MT. Non-parametric function fitting. Journal of the Royal Statistical Society, Series B. 1972; 34:385–392.

Ruppert D, Sheather SJ, Wand MP. An effective bandwidth selector for local least squares regression. Journal of the American Statistical Association. 1995; 90:1257–1270.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Rust BW. Signal from noise in global warming. Computing Science and Statistics. 2003; 35:263–277.

Subba Rao S. On multiple regression models with nonstationary correlated errors. Biometrika. 2004; 91:645–659.

Subba Rao T. The fitting of nonstationary time series models with time-dependent parameters. Journal of the Royal Statistical Society, Series B. 1970; 32:312–322.

Van Bellegem S, Von Sachs R. Locally adaptive estimation of evolutionary wavelet spectra. Annals of Statistics. 2008; 36:1879–1924.

Woodward WA, Gray HL. Global warming and the problem of testing for trend in time-series data. Journal of Climate. 1993; 6:953–962.

Wu WB, Woodroofe M, Mentz G. Isotonic regression: another look at the change point problem. Biometrika. 2001; 88:793–804.

Wu WB, Xiao H. Covariance matrix estimation in time series. Handbook of Statistics. 2011; 30

Wu WB, Zhao Z. Inference of trends in time series. Journal of the Royal Statistical Society, Series B. 2007; 69:391–410.

Wu WB, Zhou Z. Gaussian approximations for non-stationary multiple time series. Statistica Sinica. 2011; 21:1397–1413.

Appendix: Proofs

Theorem 3

Recall t_i and ξ_i in (1). Define $\eta_i = h(t_i, \xi_i)$ for a transformation $h(\cdot, \cdot)$ such that $\mathbb{E}(\eta_i) = 0$. Let $H_n(t)=\sum_{i=1}^{n} K\{(t_i-t)/b\}\eta_i$. Suppose $h \in$ SLC(2)-GMC(4) (see Assumption 1), $K(\cdot)$ satisfies Assumption 2 (i), $b \to 0$ and $nb^2/(\log n)^8 \to \infty$. Further assume $\sum^2(t):=\mathrm{var}\{h(t,\xi_0)\}+2\sum_{r=1}^{\infty}\mathrm{cov}\{h(t,\xi_0),h(t,\xi_r)\}>0$ for all $t \in [0, 1]$. Then

- **i.** *[Asymptotic normality] For each fixed $t \in (0, 1)$, $(nb)^{-1/2}H_n(t)/\Sigma(t) \Rightarrow N(0, \phi_K)$.*
- **ii.** *[Maximum deviation] Let C_K be defined in Theorem 2. For all $z \in \mathbb{R}$, we have*

$$\mathbb{P}\left\{\frac{1}{\sqrt{nb\varphi_K}}\max_{t\in[b,1-b]}\frac{|H_n(t)|}{\sum(t)} \le \sqrt{-2\log b}+\frac{C_K+z}{\sqrt{-2\log b}}\right\} \to e^{-2e^{-z}}.$$

Proof

Let $S_m=\sum_{i=1}^{m}\eta_i$. By Wu and Zhou (2011), under the condition $h \in$ SLC(2)-GMC(4) (recall that this condition imposes a finite fourth moment; see Assumption 1), there exist i.i.d. standard normal variables $Z_1, \ldots, Z_n$ such that

$$\max_{1\le m\le n}|S_m-\tilde{S}_m|=O_p[n^{1/4}(\log n)^{3/2}], \quad \text{where} \quad \tilde{S}_m=\sum_{i=1}^{m}\sum(t_i)Z_i. \quad (25)$$

Write $S_0 = 0$ and $K_i = K\{(t_i - t)/b\}$. By the summation by parts, $H_n(t)=\sum_{i=1}^{n}K_i(S_i-S_{i-1})=\sum_{i=1}^{n-1}(K_i-K_{i+1})S_i+K_nS_n$. By (25) and the summation by parts again,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

$$H_n(t)=\sum_{i=1}^{n-1}(K_i-K_{i+1})\tilde{S}_i+K_n\tilde{S}_n+O_p(\chi_n)=\sum_{i=1}^{n}K_i\sum(t_i)Z_i+O_p(\chi_n), \quad (26)$$

where $\chi_n = \Delta_n n^{1/4}(\log n)^{3/2}$ and $\Delta_n=\sum_{i=1}^{n-1}|K_i-K_{i+1}|+|K_n|$. Note that $K_i - K_{i+1} = O[(nb)^{-1}]$, and that there are $O(nb)$ indices i such that $K_i - K_{i+1} \neq 0$. Thus, $\Delta_n = O(1)$ uniformly in $t \in [0, 1]$, and consequently $\chi_n=O[n^{1/4}(\log n)^{3/2}]=o(\sqrt{nb/\log n})$. By (26), it suffices to establish the corresponding results for the Gaussian process $\sum_{i=1}^{n}K_i\sum(t_i)Z_i$.

(i) It easily follows since $Z_1, \ldots, Z_n$ are i.i.d. N(0,1). (ii) It follows from the maximum deviation of the Gaussian process $\sum_{i=1}^{n}K_i\sum(t_i)Z_i$ [Lemma 2 in Wu and Zhao (2007)].

Proof of Proposition 2

By Isserlis (1918), for any normal vector (Z_1, Z_2, Z_3, Z_4) with zero mean, $\mathbb{E}(Z_1Z_2Z_3Z_4) = \mathbb{E}(Z_1Z_2)\ \mathbb{E}(Z_3Z_4)+ \mathbb{E}(Z_1Z_3)\ \mathbb{E}(Z_2Z_4)+ \mathbb{E}(Z_1Z_4)\ \mathbb{E}(Z_2Z_3)$. Using this identity and after some careful calculations, we can show the result.

Proof of Theorems 1–2

Recall $\mathbb{E}[g(t, \xi_0)] = 0$. For $\gamma_k(t)$ in (4), $\mathbb{E}[g(t_i, \xi_i)g(t_i, \xi_{i+k})] = \gamma_k(t_i)$. Let $u_{i,t} = (t_i - t)/b$. Note the decomposition $\hat{\gamma_k}(t) = V_k(t) + U_k(t) + R_k(t)$, where

$$V_k(t)=\frac{1}{nb}\sum_{i=1}^{n}\gamma_k(t_i)K(u_{i,t}), \quad (27)$$

$$\begin{aligned} U_k(t)&=\tfrac{1}{nb}\sum_{i=1}^{n}K(u_{i,t})\,\{g(t_i,\xi_i)g(t_i,\xi_{i+k})-\mathbb{E}[g(t_i,\xi_i)g(t_i,\xi_{i+k})]\},\\ R_k(t)&=\tfrac{1}{nb}\sum_{i=1}^{n}g(t_i,\xi_i)[g(t_{i+k},\xi_{i+k})-g(t_i,\xi_{i+k})]K(u_{i,t}). \end{aligned} \quad (28)$$

By elementary calculus, $(nb)^{-1}\sum_{i=1}^{n}u_{i,t}^{r}K(u_{i,t})=\int_{\mathbb{R}}u^{r}K(u)du+O[(nb)^{-1}]$, $r = 0, 1, 2, 3$, uniformly on $t \in [b, 1 - b]$. The symmetry of $K(\cdot)$ implies $\int_{\mathbb{R}} uK(u)du = \int_{\mathbb{R}}u^3K(u)du = 0$. Thus, by Taylor's expansion $\gamma_k(t_i)=\sum_{r=0}^{3}[\partial^{r}\gamma_k(t)/\partial t]b^{r}u_{i,t}^{r}/r!+O(b^4)$ for $|t_i - t| \le b$,

$$V_k(t)=\gamma_k(t)+b^2\psi_K\gamma_k''(t)+O[(nb)^{-1}+b^4], \quad \text{uniformly on } [b,1-b]. \quad (29)$$

By Assumption 1 (i) and the Schwarz inequality, $\mathbb{E}|g(t_i, \xi_i)[g(t_{i+k}, \xi_{i+k}) - g(t_i, \xi_{i+k})]| \le \|g(t_i, \xi_i)\|_2 \times \|g(t_{i+k}, \xi_{i+k}) - g(t_i, \xi_{i+k})\|_2 = O(t_{i+k} - t_i) = O(1/n)$. Thus,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

$$\mathbb{E}\left[\sup_t|R_k(t)|\right] \le \frac{\sup_u|K(u)|}{nb}\sum_{i=1}^{n}\mathbb{E}|g(t_i,\xi_i)[g(t_{i+k},\xi_{i+k})-g(t_i,\xi_{i+k})]|=\frac{O(1)}{nb}. \quad (30)$$

By (29) and (30), we have the uniform approximation on $t \in [b, 1 - b]$:

$$\hat{\gamma}_k(t)-\gamma_k(t)=b^2\psi_K\gamma_k''(t)+U_k(t)+O_p(\varepsilon_n) \quad \text{with} \quad \varepsilon_n=(nb)^{-1}+b^4. \quad (31)$$

Recall $\lambda(t)=[\gamma_k''(t)-\rho_k(t)\gamma_0''(t)]/\gamma_0(t)$ and $\gamma_k(t) = \gamma_0(t)\,\rho_k(t)$. Thus, by (31),

$$\begin{aligned}\hat{\rho}_k(t)-\rho_k(t)&=\frac{[\hat{\gamma}_k(t)-\gamma_k(t)]-[\hat{\gamma}_0(t)-\gamma_0(t)]\rho_k(t)}{\hat{\gamma}_0(t)}\\ &=\frac{\gamma_0(t)}{\hat{\gamma}_0(t)}b^2\psi_K\lambda(t)+\frac{\gamma_0(t)}{\hat{\gamma}_0(t)}\frac{U_k(t)-\rho_k(t)U_0(t)}{\gamma_0(t)}+\frac{O_p(\varepsilon_n)}{\hat{\gamma}_0(t)}. \quad (32)\end{aligned}$$

Let $W_n(t) = \{U_k(t) - \rho_k(t)U_0(t)\}/\gamma_0(t)$. We can rewrite

$$W_n(t)=\frac{1}{nb}\sum_{i=1}^{n}K(u_{i,t})[d_i-\mathbb{E}(d_i)], \quad d_i=\frac{g(t_i,\xi_i)g(t_i,\xi_{i+k})-\rho_k(t)g^2(t_i,\xi_i)}{\gamma_0(t)}. \quad (33)$$

Since $g \in$ SLC(4)-GMC(8), it can be shown that $g(t_i, \xi_i)g(t_i, \xi_{i+k}) - \mathbb{E}[g(t_i, \xi_i)g(t_i, \xi_{i+k})] \in$ SLC(2)-GMC(4). For example, for ξ_i' in Assumption 1 (ii), using $\|UV\|_4 \le \|U\|_8\|V\|_8$, we obtain $\|g(t_i,\xi_i)g(t_i,\xi_{i+k})-g(t_i,\xi_i')g(t_i,\xi_{i+k}')\|_4 \le \|g(t_i,\xi_i)[g(t_i,\xi_{i+k})-g(t_i,\xi_{i+k}')]\|_4 + \|g(t_i,\xi_{i+k}')[g(t_i,\xi_i)-g(t_i,\xi_i')]\|_4 = O(\lambda^i)$. By the maximum deviation in Theorem 3 (ii), $\max_{t\in[b,1-b]}|U_0(t)| = O_p[(nb/\log n)^{-1/2}]$. Thus, by (31), $\max_{t\in[b,1-b]}|\hat{\gamma_0}(t) - \gamma_0(t)| = O_p[b^2 + (nb/\log n)^{-1/2}]$. Note that $d_i \in$ SLC(2)-GMC(4). In view of (32), Theorems 1–2 follow from the asymptotic normality and maximum deviation of $W_n(t)$ via Theorem 3.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

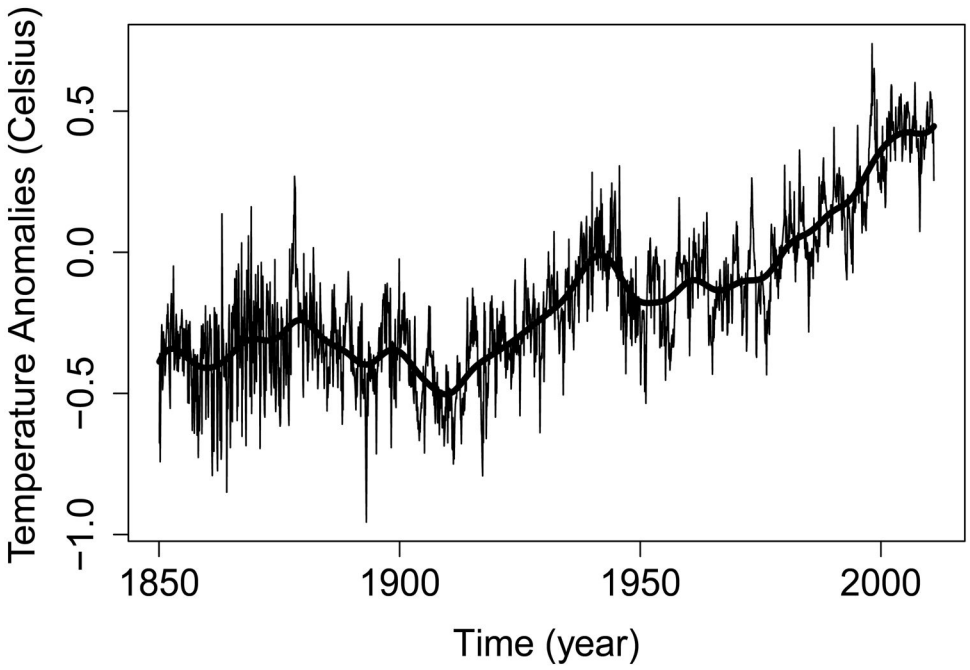

Figure 1.
The global monthly temperature anomalies in Celsius during 1850–2010, and the local linear estimate of the mean trend (thick solid curve).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

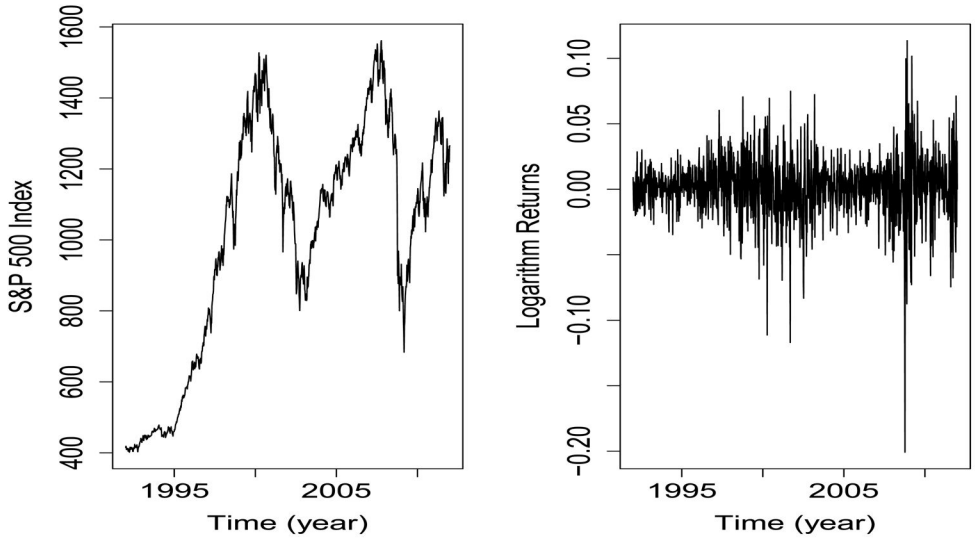

Figure 2.
Weekly S&P 500 index (left plot) and logarithm returns (right plot) during 1992–2011.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

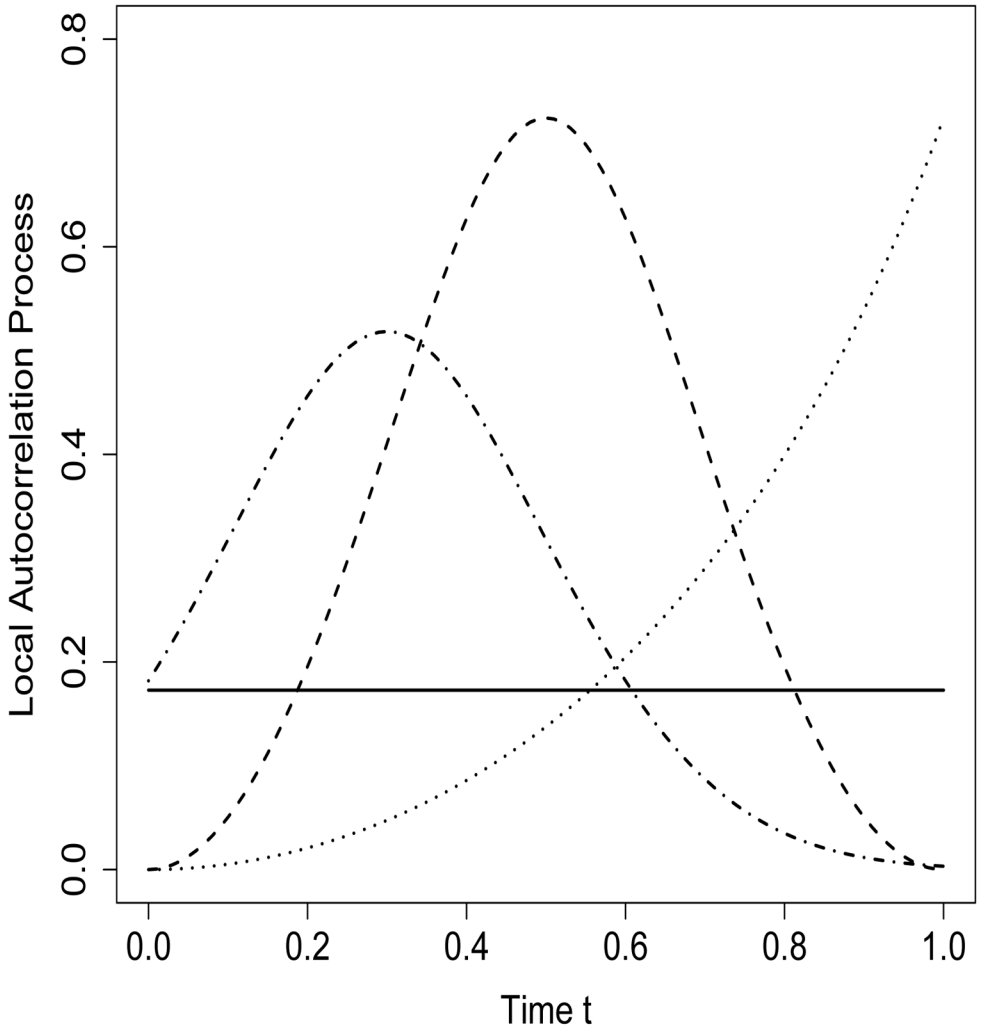

Figure 3.
Plot of the local autocorrelation curve $\{\rho_1(t)\}_{t\in[0,1]}$ in (7). The solid, dotted, dashed, and dot-dashed curves correspond to $\alpha(t) = 0.5$, $0.9t$, $0.9\sin(\pi t)$, and $2\varphi(t - 0.3)$, respectively.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

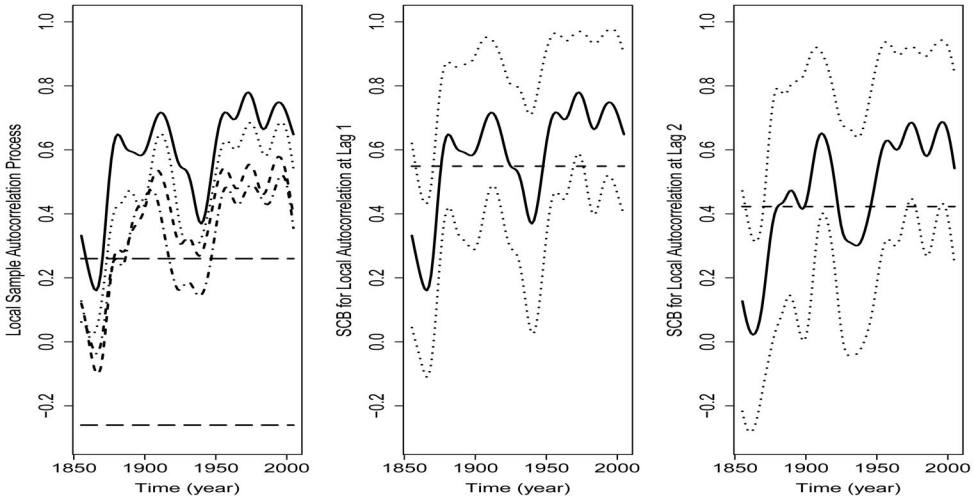

Figure 4.

Left plot: local sample autocorrelation process $\hat{\rho}_k(t)$ in (8) with $b = 0.033$ for $k = 1$ (solid curve), $k = 2$ (dotted curve), $k = 3$ (dashed curve), $k = 4$ (dot-dashed curve), along with the rule-of-thumb 95% simultaneous cutoff bands (horizontal long-dashed lines) in (19) for testing for zero local autocorrelations. Middle plot and right plot: solid curve is the local sample autocorrelation process at lag 1 (middle plot) and lag 2 (right plot), the two dotted curves are the corresponding rule-of-thumb 95% SCB, and the horizontal long-dashed line, computed using (16), is the sample autocorrelation assuming stationarity.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

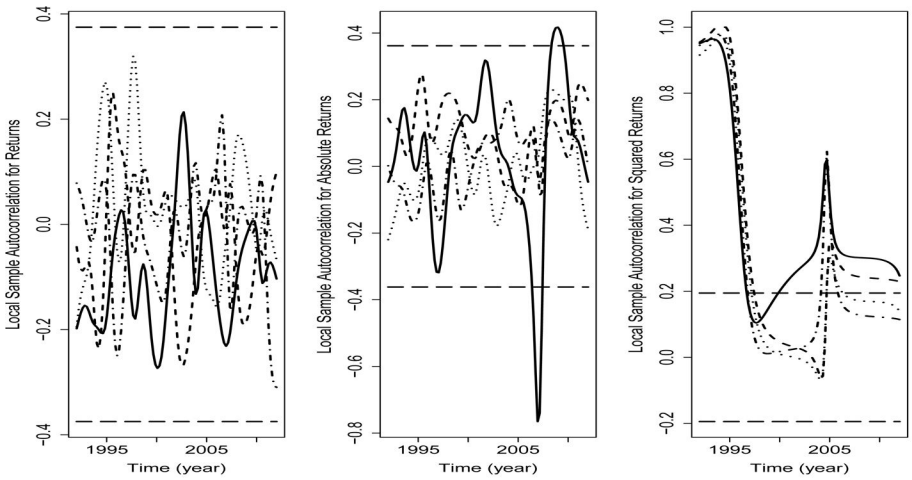

Figure 5.

Left plot: local sample autocorrelation processes $\hat{\rho_k}(t)$ of logarithm returns for $k = 1$ (solid curve), $k = 2$ (dotted curve), $k = 3$ (dashed curve), $k = 4$ (dot-dashed curve), along with the rule-of-thumb 95% simultaneous cutoff bands (horizontal long-dashed lines) in (19) for testing for zero local autocorrelations. Middle plot and right plot: same as in the left plot but for centered absolute returns (middle plot) and squared returns (right plot).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

J Bus Econ Stat. Author manuscript; available in PMC 2016 April 01.

Table 1

Empirical coverage probability and power for Models 1–3 in (20)–(22). Columns 90% and 95% are the empirical coverage probabilities when the nominal levels are 90% and 95%, respectively. Columns H_0^1 and H_0^2 are the empirical power for testing the two null hypotheses in (23) with significance level 5%. In (14), we use $L = 10 \approx 2n^{4/15}$.

Table 1 (a): Model 1 in (20)

		$b = b^*$						$b = 1.5b^*$						$b = 2b^*$					
		$\rho_1(\cdot)$			$\rho_2(\cdot)$			$\rho_1(\cdot)$			$\rho_2(\cdot)$			$\rho_1(\cdot)$			$\rho_2(\cdot)$		
θ	b^*	90%	95%	H_0^1	90%	95%	H_0^2	90%	95%	H_0^1	90%	95%	H_0^2	90%	95%	H_0^1	90%	95%	H_0^2
0.0	0.065	91.1	95.0	4.5	91.9	96.3	100	90.1	95.4	4.7	90.1	94.7	100	88.6	95.5	5.4	88.8	94.4	100
0.2	0.059	89.9	94.3	77.3	91.5	95.6	100	90.8	95.0	89.1	90.9	95.9	100	90.4	95.3	94.9	89.5	95.5	100
0.5	0.051	90.7	95.8	100	89.1	93.2	99.8	90.8	95.7	100	89.9	94.8	100	89.2	94.9	100	89.6	94.0	100
0.8	0.056	92.7	97.3	100	89.7	93.6	98.6	90.6	96.4	100	88.7	94.8	99.7	89.7	95.4	100	87.1	94.4	99.8
1.0	0.061	92.1	96.0	99.9	89.1	93.9	98.2	90.5	95.5	100	89.0	94.2	99.7	87.6	94.4	100	88.0	94.9	100

Table 1 (b): Model 2 in (21)

		$b = b^*$						$b = 1.5b^*$						$b = 2b^*$					
		$\rho_1(\cdot)$			$\rho_2(\cdot)$			$\rho_1(\cdot)$			$\rho_2(\cdot)$			$\rho_1(\cdot)$			$\rho_2(\cdot)$		
θ	b^*	90%	95%	H_0^1	90%	95%	H_0^2	90%	95%	H_0^1	90%	95%	H_0^2	90%	95%	H_0^1	90%	95%	H_0^2
0.0	0.043	90.8	94.9	2.7	86.4	91.4	6.2	90.8	95.3	2.2	89.1	94.2	3.7	91.9	95.6	1.4	89.4	95.1	2.3
0.2	0.058	92.1	97.2	11.2	88.7	94.1	10.4	91.2	96.0	11.6	89.8	94.6	10.3	91.2	96.1	13.8	90.9	95.2	10.9
0.5	0.064	92.2	97.3	79.1	91.2	95.7	21.8	90.7	96.1	89.2	89.6	95.9	28.5	85.6	94.4	93.7	87.2	93.5	33.0
0.8	0.062	91.9	95.8	99.9	90.6	95.4	14.8	90.9	95.4	100	89.0	94.4	15.7	81.6	91.4	100	80.9	89.6	13.6
1.0	0.051	92.2	96.1	100	88.7	93.3	14.9	88.9	95.7	100	87.6	93.4	16.4	83.6	91.3	100	80.5	89.5	12.6

Table 1 (c): Model 3 in (22), centered using the theoretical mean

		$b = b^*$						$b = 1.5b^*$						$b = 2b^*$					
		$\rho_1(\cdot)$			$\rho_2(\cdot)$			$\rho_1(\cdot)$			$\rho_2(\cdot)$			$\rho_1(\cdot)$			$\rho_2(\cdot)$		
θ	b^*	90%	95%	H_0^1	90%	95%	H_0^2	90%	95%	H_0^1	90%	95%	H_0^2	90%	95%	H_0^1	90%	95%	H_0^2
0.0	0.094	92.8	95.9	1.5	93.4	97.0	0.8	90.4	95.6	1.4	90.6	96.1	0.8	89.0	95.1	1.2	88.9	95.2	0.9

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 1 (c): Model 3 in (22), centered using the theoretical mean

		$b = b^*$						$b = 1.5b^*$						$b = 2b^*$					
		$\rho_1(\cdot)$			$\rho_2(\cdot)$			$\rho_1(\cdot)$			$\rho_2(\cdot)$			$\rho_1(\cdot)$			$\rho_2(\cdot)$	
θ	b^*	90%	95%	H_0^1	90%	95%	H_0^2	90%	95%	H_0^1	90%	95%	H_0^2	90%	95%	H_0^1	90%	95%	H_0^2
0.2	0.094	93.3	96.8	1.1	94.2	97.2	0.5	90.1	95.9	1.5	91.2	96.6	0.4	87.9	94.5	1.1	89.5	95.9	0.6
0.5	0.091	92.4	96.9	2.3	93.1	97.3	1.4	90.9	95.8	2.6	91.3	96.2	0.9	88.2	94.9	4.3	89.1	95.1	0.9
0.8	0.084	90.3	95.3	21.7	92.5	96.3	3.5	89.1	94.8	31.3	90.3	95.3	4.0	86.0	93.7	37.9	87.4	94.7	3.7
1.0	0.077	85.7	91.8	71.0	89.7	94.5	22.4	84.9	91.5	83.0	85.7	92.7	30.3	76.1	87.6	88.2	79.8	89.0	35.0

Table 1 (d): Model 3 in (22), centered using local linear regression

		$b = b^*$						$b = 1.5b^*$						$b = 2b^*$					
		$\rho_1(\cdot)$			$\rho_2(\cdot)$			$\rho_1(\cdot)$			$\rho_2(\cdot)$			$\rho_1(\cdot)$			$\rho_2(\cdot)$	
θ	b^*	90%	95%	H_0^1	90%	95%	H_0^2	90%	95%	H_0^1	90%	95%	H_0^2	90%	95%	H_0^1	90%	95%	H_0^2
0.0	0.094	92.0	96.4	0.5	93.9	97.6	0.9	90.1	94.6	1.0	91.5	96.1	1.0	88.1	93.7	1.1	88.5	95.5	1.0
0.2	0.094	90.8	96.4	1.4	92.9	96.5	0.9	88.1	93.7	1.3	90.9	95.2	0.9	88.2	93.6	1.3	88.7	94.4	0.6
0.5	0.089	90.8	96.7	2.7	92.6	96.6	1.6	88.7	95.3	3.2	89.9	95.9	1.3	85.9	93.7	3.6	88.0	95.2	1.1
0.8	0.086	87.7	93.2	22.0	90.2	94.7	3.1	83.9	91.3	31.8	86.5	92.7	3.7	79.4	87.9	37.7	82.0	90.4	4.8
1.0	0.082	83.0	90.4	72.4	81.5	89.0	19.6	75.8	85.1	82.1	75.3	84.7	25.3	65.8	76.8	86.4	63.2	75.1	30.0