Statistica Sinica Preprint No: SS-2022-0244	
Title	Time-Varying Correlation for Noncentered Nonstationary Time Series: Simultaneous Inference and Visualization
Manuscript ID	SS-2022-0244
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202022.0244
Complete List of Authors	Ting Zhang and Yu Shao
Corresponding Authors	Ting Zhang
E-mails	tingzhang@uga.edu

Statistica Sinica

TIME-VARYING CORRELATION FOR NONCENTERED NONSTATIONARY TIME SERIES: SIMULTANEOUS INFERENCE AND VISUALIZATION

Ting Zhang* and Yu Shao†

University of Georgia and Boston University†*

Abstract: We consider simultaneous inference of the time-varying correlation, as a function of time, between two nonstationary time series, when their trend functions are unknown. Unlike the stationary setting, where the effect of precentering using the sample mean is trivially negligible, in the nonstationary setting, it is difficult to quantify the effect of precentering using nonparametric trend function estimators. This is mainly because the trend estimators are time-varying across different time points, which makes it difficult to quantify their cumulative interaction with the error process in a time series setting. We propose using a centering scheme that, instead of aligning with the time point at which the data are observed, aligns with the time point at which the local correlation estimation is performed. We show that the proposed centering scheme leads to simultaneous confidence bands with a solid theoretical guarantee for the time-varying correlation between two nonstationary time series when their trend functions are unknown. Lastly, we demonstrate the proposed method using numerical examples, including a real-data analysis.

Key words and phrases: kernel smoothing, local linear estimation, noncentered data, simultaneous confidence band.

1. Introduction

The correlation coefficient is a popular metric for quantifying the dependence between two variables. In a time series setting, we can use the correlation between two observed time series to understand their relationship or co-movement over time, or the correlation between the time series and its lagged version to study the underlying dependence structure. The latter is often referred to as the autocorrelation; see Wu and Xiao (2012). In addition, we can use the correlation between a time series and the lagged version of another time series to understand the lagged effect of one on the other, referred to as the Granger causality in time series analysis. The problem of estimating the correlation and autocorrelation has been studied extensively for stationary time series; see, for example, Anderson (1971), Hannan (1976), Hall and Heyde (1980), Priestley (1981), Brockwell and Davis (1991), Phillips and Solo (1992), Hosking (1996), Wu and Min (2005), Wu (2009), Wu and Xiao (2012), and the references therein. In the aforemen-

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0244

tioned results, the underlying process is mostly assumed to be stationary, which means the correlation coefficient is a constant that does not change over time, facilitating the estimation and statistical inference.

However, in nonstationary time series applications, it is generally expected that certain aspects of the observed data evolve over time, making it better to consider time-varying correlations as a function of time. For this, Mallat et al. (1998) consider a covariance estimation using a local cosine basis approximation for locally stationary processes. Dahlhaus (2012) considers a data tapering method for the covariance estimation of locally stationary processes using kernel functions. Fu et al. (2014) estimate the time-varying covariance between two locally stationary biological processes, and provide an asymptotic analysis of the resulting estimation bias and variance. Choi and Shin (2021) consider a nonparametric estimation of the time-varying correlation coefficient, and establish its asymptotic normality when the joint error process is strong mixing and stationary, except for a scale difference. Most of the aforementioned results focus on the estimation or pointwise inference of the time-varying covariance at a given time point. Few works explore the difficult task of developing a simultaneous inference procedure for the time-varying correlation as a function of time.

In an important work, Zhao (2015) provides a solution to this problem by constructing simultaneous confidence bands for local autocorrelations of locally stationary time series. However, the theory and methods rely on the assumption that the mean trend function of the underlying process is known to be uniformly zero, which the author argues is satisfied, in general, in daily or weekly data on financial returns. For data with potentially nonzero trend functions, Zhao (2015) proposes first precentering the data using parametric or nonparametric trend estimators, and then applying their methods to the residual process. However, it is nontrivial to quantify the effect of such a precentering procedure on the subsequent correlation inference procedure, and was left as an open problem; see the discussion in Section 3.2 of Zhao (2015). In Section 2, we discuss in detail why the effect of precentering is trivially negligible in the stationary case, but suddenly becomes difficult to understand in the nonstationary case. We suggest that this is largely because of the time-varying nature of the trend function, which makes it difficult to quantify its cumulative interaction with the error process. In Section 3, we propose the locally homogenized centering method, which alleviates the problems of traditional centering schemes, and leads to a simultaneous inference of time-varying correlations with a solid theoretical guarantee when the underlying trend functions are unknown. In addition to allowing unknown trend functions, we consider the more general setting

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0244

in which the time-varying correlation between two time series can depend on each other in a nontrivial way. In particular, when one time series is a lagged version of the other, this reduces to the autocorrelation setting considered in Zhao (2015). Additionally, Zhao (2015) requires a geometric moment contraction condition, under which the dependence decays geometrically quickly, whereas we allow processes with an algebraic decay; see the discussion in Section 3.3. Numerical examples, including Monte Carlo simulations and a real-data analysis, are provided in Section 4 to illustrate the proposed method.

2. Precentering: A Natural Approach, and Its Problem

We first review the stationary case, for which the effect of precentering using the sample mean is trivially negligible. To illustrate, suppose we observe stationary time series X_i and Y_i, for $i = 1, \ldots, n$. Then, assuming that the stationary means $\mu_x = E(X_1)$ and $\mu_y = E(Y_1)$ are known, we can estimate the covariance by the oracle

$$\tilde{\gamma}_n = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_x)(Y_i - \mu_y).$$

When the true means μ_x and μ_y are unknown, we can plug in the sample means $\bar{X}_n = n^{-1}\sum_{i=1}^{n} X_i$ and $\bar{Y}_n = n^{-1}\sum_{i=1}^{n} Y_i$, yielding the covariance estimator

$$\hat{\gamma}_n = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)(Y_i - \bar{Y}_n).$$

Now, we can quantify the effect of using the sample mean to replace the true mean by the difference

$$\begin{aligned}\hat{\gamma}_n - \tilde{\gamma}_n &= (\bar{X}_n - \mu_x)(\bar{Y}_n - \mu_y) - \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_x)(\bar{Y}_n - \mu_y) - \frac{1}{n}\sum_{i=1}^{n}(\bar{X}_n - \mu_x)(Y_i - \mu_y) \\ &= -(\bar{X}_n - \mu_x)(\bar{Y}_n - \mu_y).\end{aligned}$$

Therefore, if the sample means $\bar{X}_n - \mu_x = O_p(n^{-1/2})$ and $\bar{Y}_n - \mu_y = O_p(n^{-1/2})$ have the usual parametric rate (Zhang, 2018), then the difference

$$\hat{\gamma}_n - \tilde{\gamma}_n = O_p(n^{-1}),$$

which is typically of a negligible order for covariance inference.

In the nonstationary case, however, parameters such as the mean or covariance do not necessarily stay as constants, and are often treated as unknown functions of time. For this, a prominent approach is to consider the scaling device under which

$$E(X_i) = \mu_x(i/n), \quad E(Y_i) = \mu_y(i/n), \quad \mathrm{cov}(X_i, Y_i) = \gamma(i/n),$$

for some functions $\mu_x(t)$, $\mu_y(t)$, and $\gamma(t)$, with $t \in [0, 1]$. Note that the scaling device itself does not impose any additional assumptions on the underlying dynamics, but it can work well with certain smoothness conditions to provide an asymptotic justification for nonparametric smoothing estimators; see, for example, Robinson (1989, 1991), Dahlhaus (1996, 1997), Cai (2007), Zhou and Wu (2010), Zhang and Wu (2011), and Zhang (2013). If the true mean functions $\mu_x(\cdot)$ and $\mu_y(\cdot)$ are known, we can follow Zhao (2015), and estimate the covariance as a function of time by

$$\tilde{\gamma}_n(t) = \frac{1}{nb_n}\sum_{i=1}^{n}\{X_i - \mu_x(i/n)\}\{Y_i - \mu_y(i/n)\}K\left(\frac{i/n - t}{b_n}\right),$$

where $K(\cdot)$ is a kernel function and $b_n > 0$ is a bandwidth. When the true mean functions $\mu_x(\cdot)$ and $\mu_y(\cdot)$ are unknown, a natural approach is to plug in their nonparametric estimators $\hat{\mu}_x(\cdot)$ and $\hat{\mu}_y(\cdot)$, respectively, as in the Nadaraya (1964) and Watson (1964) estimators, the Priestley and Chao (1972) estimator, and the local linear estimator of Fan and Gijbels (1996). Doing so leads to the nonparametric covariance estimator

$$\breve{\gamma}_n(t) = \frac{1}{nb_n}\sum_{i=1}^{n}\{X_i - \hat{\mu}_x(i/n)\}\{Y_i - \hat{\mu}_y(i/n)\}K\left(\frac{i/n - t}{b_n}\right). \tag{2.1}$$

This is the same as using the precentered data $X_i - \hat{\mu}_x(i/n)$ and $Y_i - \hat{\mu}_y(i/n)$ to compute the covariance if the trend function is known to be uniformly zero; see, for example, Zhao (2015). In this case, the effect of using nonparametric estimators to replace the true means is quantified by

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0244

the difference

$$\breve{\gamma}_n(t) - \tilde{\gamma}_n(t) = \frac{1}{nb_n}\sum_{i=1}^{n}\{\hat{\mu}_x(i/n) - \mu_x(i/n)\}\{\hat{\mu}_y(i/n) - \mu_y(i/n)\}K\left(\frac{i/n - t}{b_n}\right) - \frac{1}{nb_n}\sum_{i=1}^{n}\{X_i - \mu_x(i/n)\}\{\hat{\mu}_y(i/n) - \mu_y(i/n)\}K\left(\frac{i/n - t}{b_n}\right) - \frac{1}{nb_n}\sum_{i=1}^{n}\{\hat{\mu}_x(i/n) - \mu_x(i/n)\}\{Y_i - \mu_y(i/n)\}K\left(\frac{i/n - t}{b_n}\right), \tag{2.2}$$

which unfortunately cannot be bounded easily by a negligible stochastic order, as it can in the stationary case. The main reason is that, owing to the time-varying nature, the random weight $\hat{\mu}_y(i/n) - \mu_y(i/n)$ for $X_i - \mu_x(i/n)$ now depends on the index i, and thus cannot be taken outside of the summation, as it can in the stationary case. Because $\hat{\mu}_y(i/n) - \mu_y(i/n)$ and $X_i - \mu_x(i/n)$ can depend on each other in a nontrivial way, it then becomes unclear whether the local averages $(nb_n)^{-1}\sum_{i=1}^{n}\{X_i - \mu_x(i/n)\}\{\hat{\mu}_y(i/n) - \mu_y(i/n)\}K\{(i/n-t)/b_n\}$ will continue to obey the square root rate, let alone a uniform rate, over different time points, which is essential for simultaneous inference. This makes it difficult to obtain a sharp probabilistic bound on the two cross terms in (2.2), and it remains unknown whether they can be treated as negligible in a theoretical analysis. Therefore, although natural, the approach of replacing the unknown mean function with its nonparametric estimator in covariance inference problems is rather ad hoc, and the effects can be difficult to understand theoretically; see also the discussion in Section 3.2 of Zhao (2015). In the following, we propose a solution that enables us to construct simultaneous confidence bands for the time-varying correlation between nonstationary time series, with a solid theoretical guarantee, when their underlying trend functions are unknown.

3. Locally Homogenized Centering: A Fix

3.1 Methodology: The Fundamental Idea

The main problem with the natural centering scheme in (2.1) for covariance inference problems is the local inhomogeneity of $\hat{\mu}_x(i/n)$ and $\hat{\mu}_y(i/n)$, for $i = 1, \ldots, n$. To solve this problem, we propose a locally homogenized centering (LHC) method that instead of using the inhomogeneous centering $\hat{\mu}_x(i/n)$ and $\hat{\mu}_y(i/n)$, which align with the time point at which the data are observed, uses their locally homogenous counterparts $\hat{\mu}_x(t)$ and $\hat{\mu}_y(t)$, which align with the time point at which the local correlation estimation is performed to achieve the local centering. This leads to the locally homogenized centered nonparametric covariance estimator

$$\check{\gamma}_n(t) = \frac{1}{nb_n} \sum_{i=1}^{n} \{X_i - \hat{\mu}_x(t)\}\{Y_i - \hat{\mu}_y(t)\} K\left(\frac{i/n - t}{b_n}\right).$$

In this case, the centering scheme differs for the time points at which the local covariance is calculated. The effect of this new centering scheme can now be quantified by the difference

$$\begin{aligned}\check{\gamma}_n(t) - \tilde{\gamma}_n(t) &= \frac{1}{nb_n}\sum_{i=1}^{n}\{\hat{\mu}_x(t) - \mu_x(i/n)\}\{\hat{\mu}_y(t) - \mu_y(i/n)\}K\left(\frac{i/n - t}{b_n}\right) \\ &\quad -\frac{1}{nb_n}\sum_{i=1}^{n}\{X_i - \mu_x(i/n)\}\{\hat{\mu}_y(t) - \mu_y(i/n)\}K\left(\frac{i/n - t}{b_n}\right) \\ &\quad -\frac{1}{nb_n}\sum_{i=1}^{n}\{\hat{\mu}_x(t) - \mu_x(i/n)\}\{Y_i - \mu_y(i/n)\}K\left(\frac{i/n - t}{b_n}\right).\end{aligned} \tag{3.1}$$

Compared with the decomposition for the natural centering scheme in (2.2), the key difference here is that the random weight $\hat{\mu}_y(t) - \mu_y(i/n)$ for $X_i - \mu_x(i/n)$ can now be decomposed into a locally homogeneous random part $\hat{\mu}_y(t) - \mu_y(t)$, which can be taken outside of the summation, and a deterministic part $\mu_y(t) - \mu_y(i/n)$, where we handle the cumulative interaction between the latter part and $X_i - \mu_x(i/n)$ using a square root stochastic bound.

3.2 Methodology: Derivative Adjustment

The fundamental idea of using locally homogenized mean functions to perform the local centering, as proposed in Section 3.1, enables us to quantify the effect of centering when the underlying trend functions are unknown, potentially leading to new inference protocols for the time-varying covariance or correlation of nonstationary time series. However, this benefit, though crucial and necessary for covariance inference with unknown trend functions, comes at the price of an additional bias,

$$\frac{1}{nb_n}\sum_{i=1}^{n}\{\mu_x(t)-\mu_x(i/n)\}\{\mu_y(t)-\mu_y(i/n)\}K\left(\frac{i/n-t}{b_n}\right),$$

which is of a comparable order to that of the bias of the mean-oracle estimator $E\{\tilde{\gamma}_n(t)\}-\gamma(t)$. We further propose a derivative adjustment method to remove this additional bias, making the resulting covariance estimator asymptotically equivalent to its mean-oracle counterpart, such that the effect of centering becomes theoretically negligible. Let $\hat{\mu}'_x(t)$ and $\hat{\mu}'_y(t)$ be derivative estimators, which can be obtained by, for example, the popular local linear method of Fan and Gijbels (1996). We propose considering the following covariance estimator with a derivative adjustment:

$$\hat{\gamma}_n(t)=\frac{1}{nb_n}\sum_{i=1}^{n}\{X_i-\hat{\mu}_x(t)-\hat{\mu}'_x(t)(i/n-t)\}\{Y_i-\hat{\mu}_y(t)-\hat{\mu}'_y(t)(i/n-t)\}K\left(\frac{i/n-t}{b_n}\right).$$

Compared with how we handle the terms in (3.1), here we decompose the random weight $\hat{\mu}_y(t)+\hat{\mu}'_y(t)(i/n-t)-\mu_y(i/n)$ into three terms: $\hat{\mu}_y(t)-\mu_y(t)$, $\hat{\mu}'_y(t)(i/n-t)-\mu'_y(t)(i/n-t)$, and $\mu_y(t)+\mu'_y(t)(i/n-t)-\mu_y(i/n)$. The first term is random, but does not depend on the index i, and thus can be taken out of the summation for a better bound. The last term depends on the index i, but is deterministic, and thus can be handled by a bound on linear combinations of nonstationary processes. The key difference here is the second term, $\hat{\mu}'_y(t)(i/n - t) - \mu'_y(t)(i/n - t)$, which is random and involves the summation index i. However, it is still different to the natural centering scheme in (2.2) in the sense that we can write it as the product $\{\hat{\mu}'_y(t) - \mu'_y(t)\} \times (i/n - t)$, where the first part can be taken out of the summation, and the second part can be combined with $X_i - \mu_x(i/n)$ into a linear combination of nonstationary processes. This enables us to derive an explicit bound on the difference $\hat{\gamma}_n(t) - \tilde{\gamma}_n(t)$, and makes it asymptotically equivalent to the mean-oracle covariance estimator.

We call this the locally homogenized centering with derivative adjustment (LHC-DA) method, and we can apply it to the time-varying correlation

$$\hat{\rho}_n(t) = \frac{\hat{\gamma}_n(t)}{\hat{\sigma}_{x,n}(t)\hat{\sigma}_{y,n}(t)},$$

where

$$\begin{aligned}
\hat{\sigma}^2_{x,n}(t) &= \frac{1}{nb_n}\sum_{i=1}^{n}\{X_i - \hat{\mu}_x(t) - \hat{\mu}'_x(t)(i/n - t)\}^2 K\left(\frac{i/n - t}{b_n}\right), \\
\hat{\sigma}^2_{y,n}(t) &= \frac{1}{nb_n}\sum_{i=1}^{n}\{Y_i - \hat{\mu}_y(t) - \hat{\mu}'_y(t)(i/n - t)\}^2 K\left(\frac{i/n - t}{b_n}\right).
\end{aligned}$$

Note that a time-varying correlation analysis may be more suitable than a

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0244

covariance analysis to understand the time-varying relationship between two nonstationary time series, because a change in the covariance may simply be because of changes in the variance, whereas the correlation would remain constant. In the following section, we provide the asymptotic theory for the proposed LHC-DA covariance and correlation estimators, based on which, we can construct simultaneous confidence bands as a visualization tool when analyzing the time-varying covariance or correlation for a general class of nonstationary processes.

3.3 Asymptotic Theory

Suppose we observe the time series X_i and Y_i, for $i = 1, \ldots, n$, according to

$$X_i = G(i/n, \boldsymbol{\mathcal{F}}_i), \quad Y_i = H(i/n, \boldsymbol{\mathcal{F}}_i), \quad \boldsymbol{\mathcal{F}}_i = (\ldots, \boldsymbol{\epsilon}_{i-1}, \boldsymbol{\epsilon}_i), \tag{3.2}$$

where $(\boldsymbol{\epsilon}_i)$ is a sequence of independent and identically distributed innovations, and G and H are measurable functions that depend on the time points $t_{i,n} = i/n$, for $i = 1, \ldots, n$. The framework (3.2) covers a wide range of nonstationary processes, and naturally extends many existing stationary time series models to their nonstationary counterparts; see Draghicescu et al. (2009), Zhou and Wu (2010), Zhang and Wu (2011), Degras et al. (2012), and Zhang (2015) for additional discussions. Other contributions on nonstationary time series can be found in Dahlhaus (1997), Cheng and

Tong (1998), Nason et al. (2000), Giurcanu and Spokoiny (2004), Ombao et al. (2005), Zhang (2016a), and the references therein. Let $(\boldsymbol{\epsilon}_i^\star)$ be a sequence of random vectors that share the same distribution as, but are independent of the sequence $(\boldsymbol{\epsilon}_i)$. Then, we can define the coupled shift process $\boldsymbol{\mathcal{F}}_{i,\{0\}} = (\ldots, \boldsymbol{\epsilon}_{-1}, \boldsymbol{\epsilon}_0^\star, \boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_i)$. For a random vector $\boldsymbol{Z}$, we write $\|\boldsymbol{Z}\|_q = \{E(|\boldsymbol{Z}|^q)\}^{1/q}$, for $q > 0$, where $|\boldsymbol{Z}|$ is the Euclidean norm, and denote $\|\boldsymbol{Z}\| = \|\boldsymbol{Z}\|_2$. For a generic process $L(t, \boldsymbol{\mathcal{F}}_i)$, with $t \in [0,1]$ and $i \in \mathbb{Z}$, assuming that $\sup_{t\in[0,1]} \|L(t, \boldsymbol{\mathcal{F}}_0)\|_q$ for some $q > 0$, we define the dependence measure

$$\theta_{i,q}(L) = \sup_{t\in[0,1]} \|L(t, \boldsymbol{\mathcal{F}}_i) - L(t, \boldsymbol{\mathcal{F}}_{i,\{0\}})\|_q,$$

which measures the dependence of $L(t, \boldsymbol{\mathcal{F}}_i)$ on the single innovation $\boldsymbol{\epsilon}_0$ over $t \in [0,1]$. Then, the quantity

$$\Theta_{m,q}(L) = \sum_{i=m}^{\infty} \theta_{i,q}(L)$$

measures the cumulative influence of $\boldsymbol{\epsilon}_0$ on future observations with a gap at least m, and we can interpret $\Theta_{0,q}(L) < \infty$ as a short-range dependence condition (Zhang, 2015). The process $L(t, \boldsymbol{\mathcal{F}}_i)$, with $t \in [0,1]$ and $i \in \mathbb{Z}$, is said to be stochastic Lipschitz continuous, or $L \in \mathrm{SLC}_q$, if there exists a constant $c_q < \infty$ such that

$$\|L(t_1, \boldsymbol{\mathcal{F}}_i) - L(t_2, \boldsymbol{\mathcal{F}}_i)\|_q \leq c_q |t_1 - t_2|$$

holds for all $t_1, t_2 \in [0,1]$. Let

$$\varpi_L(t) = \sum_{k \in \mathbb{Z}} \mathrm{cov}\{L(t, \boldsymbol{\mathcal{F}}_0), L(t, \boldsymbol{\mathcal{F}}_k)\},$$

which is a well-defined and finite quantity when $\Theta_{0,q}(L) < \infty$, for some $q \geq 2$. Write

$$\mu_x(t) = E\{G(t, \boldsymbol{\mathcal{F}}_i)\}, \quad \mu_y(t) = E\{H(t, \boldsymbol{\mathcal{F}}_i)\},$$

$$\sigma_x^2(t) = \mathrm{var}\{G(t, \boldsymbol{\mathcal{F}}_i)\}, \quad \sigma_y^2(t) = \mathrm{var}\{H(t, \boldsymbol{\mathcal{F}}_i)\},$$

$$\gamma(t) = \mathrm{cov}\{G(t, \boldsymbol{\mathcal{F}}_i), H(t, \boldsymbol{\mathcal{F}}_i)\}, \quad \rho(t) = \mathrm{cor}\{G(t, \boldsymbol{\mathcal{F}}_i), H(t, \boldsymbol{\mathcal{F}}_i)\},$$

and denote

$$U(t, \boldsymbol{\mathcal{F}}_i) = [G(t, \boldsymbol{\mathcal{F}}_i) - E\{G(t, \boldsymbol{\mathcal{F}}_i)\}][H(t, \boldsymbol{\mathcal{F}}_i) - E\{H(t, \boldsymbol{\mathcal{F}}_i)\}]$$

and

$$V(t, \boldsymbol{\mathcal{F}}_i) = \frac{U(t, \boldsymbol{\mathcal{F}}_i)}{\sigma_x(t)\sigma_y(t)} - \gamma(t) \left\{ \frac{[G(t, \boldsymbol{\mathcal{F}}_i) - E\{G(t, \boldsymbol{\mathcal{F}}_i)\}]^2}{2\sigma_x^3(t)\sigma_y(t)} + \frac{[H(t, \boldsymbol{\mathcal{F}}_i) - E\{H(t, \boldsymbol{\mathcal{F}}_i)\}]^2}{2\sigma_x(t)\sigma_y^3(t)} \right\}.$$

Throughout this section, we assume that the kernel function $K \in \mathcal{K}$, the collection of symmetric functions in $\mathcal{C}^1[-1,1]$ with $\int_{-1}^1 K(v)dv = 1$, where $\mathcal{C}^k$ denotes the collection of functions with k continuous derivatives. Let $\mathcal{T}_n = [b_n, 1-b_n]$, $\kappa_2 = \int_{-1}^1 v^2 K(v)dv$, and $\phi_2 = \int_{-1}^1 K(v)^2 dv$. The following theorem provides the central limit theorem for the LHC-DA covariance estimator and the asymptotic distribution of the associated maximal deviation, which is useful for constructing simultaneous confidence bands for the underlying covariance function.

Theorem 1. *Assume that* $\mu_x, \mu_y, \gamma \in \mathcal{C}^3$, $\theta_{k,4}(G) + \theta_{k,4}(H) + \theta_{k,4}(U) = O(k^{-2})$, $G, H, U \in \mathrm{SLC}_2$, *and that* $\varpi_U(t)$ *is Lipschitz continuous and bounded away from zero on* $[0,1]$. *If* $n^{-2/5}b_n^{-1}(\log n)^3 + nb_n^7 \log n \to 0$, *then*

$$(nb_n)^{1/2}\{\hat{\gamma}_n(t) - \gamma(t) - 2^{-1}\kappa_2 b_n^2 \gamma''(t)\} \to_d N\{0, \varpi_U(t)\phi_2\},$$

and

$$\mathrm{pr}\left\{\frac{(nb_n)^{1/2}}{\phi_2^{1/2}} \sup_{t\in\mathcal{T}_n}\left|\frac{\hat{\gamma}_n(t) - \gamma(t) - 2^{-1}\kappa_2 b_n^2\gamma''(t)}{\varpi_U(t)^{1/2}}\right| - (-2\log b_n)^{1/2} - \frac{C_K}{(-2\log b_n)^{1/2}} \le \frac{z}{(-2\log b_n)^{1/2}}\right\} \to \exp\{-2\exp(-z)\},$$

where $C_K = 2^{-1}\log\{(4\pi^2\phi_2)^{-1}\int_{-1}^1 |K'(v)|^2 dv\}$.

Theorem 1 concerns the covariance case. Theorem 2 provides results on the LHC-DA correlation estimator. Compared with the covariance case, the proof in the correlation case is more technically involved. The major difference stems from the fact that, for correlation estimators, the asymptotic behavior is affected by the covariance part and the variance part, and neither is negligible relative to the other; see also Zhao (2015), who considers an autocorrelation inference for processes with a known zero mean and geometrically decaying dependence. Here, we deal with the more general time-varying correlation for processes with unknown trend functions and only algebraically decaying dependence.

Theorem 2. *Assume that* $\mu_x, \mu_y, \gamma, \rho \in \mathcal{C}^3$, $\theta_{k,8}(G) + \theta_{k,8}(H) = O(k^{-2})$, $G, H \in \mathrm{SLC}_4$, *and that* $\varpi_V(t)$, $\sigma_x(t)$, *and* $\sigma_y(t)$ *are Lipschitz continuous and bounded away from zero on* $[0,1]$. *If* $n^{-2/5}b_n^{-1}(\log n)^3 + nb_n^7 \log n \to 0$, *then*

$$(nb_n)^{1/2}[\hat{\rho}_n(t) - \rho(t) - 2^{-1}\kappa_2 b_n^2\{\sigma_x(t)\sigma_y(t)\}^{-1}\gamma''(t)] \to_d N\{0, \varpi_V(t)\phi_2\},$$

and

$$\mathrm{pr}\left[\frac{(nb_n)^{1/2}}{\phi_2^{1/2}} \sup_{t\in\mathcal{T}_n}\left|\frac{\hat{\rho}_n(t) - \rho(t) - 2^{-1}\kappa_2 b_n^2\{\sigma_x(t)\sigma_y(t)\}^{-1}\gamma''(t)}{\varpi_V(t)^{1/2}}\right| - (-2\log b_n)^{1/2} \right.$$
$$\left. - \frac{C_K}{(-2\log b_n)^{1/2}} \leq \frac{z}{(-2\log b_n)^{1/2}}\right] \to \exp\{-2\exp(-z)\}.$$

4. Numerical Experiments

4.1 Implementation: Algorithm and Visualization

In this section, we provide a detailed algorithm that implements the results in Section 3 to construct simultaneous confidence bands for the time-varying correlation between X_i and Y_i, for $i = 1, \ldots, n$, when their trend functions are unknown. If one of them is taken as the lagged version of the other, then the algorithm provides simultaneous confidence bands for the corresponding autocorrelation. To alleviate the problem of slow convergence to the extreme value distribution, we also use a simulation-assisted procedure to improve the finite-sample performance. The detailed implementation is as follows:

(i) Select the bandwidth b_n using the dependence-adjusted generalized cross-validation method of Zhang and Wu (2012) by viewing (2.1) as a kernel regression on time.

(ii) Compute the trend estimators $\hat{\mu}_x(t)$ and $\hat{\mu}_y(t)$ and their derivative estimators $\hat{\mu}'_x(t)$ and $\hat{\mu}'_y(t)$, respectively, using the local linear method of Fan and Gijbels (1996), with $K(\cdot)$ being the Epanechnikov kernel.

(iii) Use the LHC-DA method proposed in Section 3 to compute the time-varying covariance and correlation $\hat{\gamma}_n(t)$ and $\hat{\rho}_n(t)$, respectively, for each time point, and use a higher-order kernel $K^\star(v) = 2^{3/2}K(2^{1/2}v) - K(v)$ for bias correction.

(iv) Obtain an estimate $\hat{\varpi}_V(t)$ of the asymptotic variance using the banding estimator of Zhang and Wu (2012); see also Zhang (2016b) for a uniform consistency result on such a variance estimator.

(v) Generate independent standard normal random variables $X_i^\diamond$ and $Y_i^\diamond$,

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0244

for $i = 1, \ldots, n$, and compute the associated $\hat{\rho}_n^{\diamond}(t)$ and $\hat{\varpi}_V^{\diamond}(t)$ to calculate

$$T_n^{\diamond} = \frac{(nb_n)^{1/2}(-2\log b_n)^{1/2}}{\phi_2^{1/2}} \sup_{t \in \mathcal{T}_n} \left| \frac{\hat{\rho}_n^{\diamond}(t)}{\varpi_V^{\diamond}(t)^{1/2}} \right|.$$

(vi) Repeat (v) many times to obtain the $(1-\alpha)$th quantile of $T_n^{\diamond}$, denoted by $\hat{q}_{1-\alpha}^{\diamond}$.

(vii) Construct the $(1-\alpha)$th simultaneous confidence band of $\rho(t)$ by

$$\hat{\rho}_n(t) \pm \hat{q}_{1-\alpha}^{\diamond} \frac{\phi_2^{1/2} \hat{\varpi}_V(t)^{1/2}}{(nb_n)^{1/2}(-2\log b_n)^{1/2}},$$

which can be visualized by plotting it against time, using a solid curve for $\hat{\rho}_n(t)$ and dashed curves for the upper and lower simultaneous confidence bands.

The above algorithm can be implemented for the LHC-DA covariance as well, if needed, enabling us to examine the time-varying covariance or correlation when the observed data contain an unknown trend in the mean. Similarly to the bootstrap method, the simulation-assisted procedure approximates the distribution of the test statistic using that of generated data. The difference is that bootstrapped data are often generated by resampling from the original data, whereas the simulation-assisted procedure generates data as independent normal random variables. As a result, the correlation $\rho^{\diamond}(t)$ between the generated data $(X_i^{\diamond})$ and $(Y_i^{\diamond})$ holds conveniently at zero under this simulation-assisted mechanism, which is used in step (v) of the above algorithm. By Theorems 1 and 2, the simulation-assisted procedure may also work when the simulated data are generated independently using marginal distributions other than the normal; see also the additional simulation results provided in the Supplementary Material, which suggest robustness to the distributional choice, as long as the conditions in Theorems 1 and 2 are satisfied. We use the normal distribution to generate the simulated data because of the connection with the Gaussian approximation (Wu, 2007; Berkes et al., 2014), which states that the partial sum distribution can be well approximated by that of normal random variables. Such a Gaussian approximation can improve the finite-sample performance; see, for example, the discussions in Zhang and Wu (2011), Zhang and Wu (2012) and Zhang (2016b).

4.2 A Monte Carlo Simulation Study

However, we present a simulation study to examine the finite-sample performance of the proposed simulation-assisted LHC-DA method for the simultaneous inference of time-varying correlations. Let $(\epsilon_{i,1})$ be a sequence of independent standard normal random variables, and let $(\epsilon_{i,2})$ be a sequence of independent Rademacher random variables that is also independent of

$(\epsilon_{i,1})$. Let

$$X_i = \mu_x(i/n) + 3\sin(1.5\pi i/n)\{|\epsilon_{i,1}| - (2/\pi)^{1/2}\} + 2\cos(1.5\pi i/n)\epsilon_{i,2} + \sum_{j=1}^{\infty} j^{-2}\epsilon_{i-j,2};$$

$$Y_i = \mu_y(i/n) + \{1.5 - (i/n)^2\}\epsilon_{i,1} + (i/n)\epsilon_{i,2} + \sum_{j=1}^{\infty} 2^{-j}\epsilon_{i-j,1},$$

where $\mu_x(t) = 2t^2 + 2t$ and $\mu_y(t) = 2\{\sin(1.5\pi t) + t\}$. We perform a simultaneous inference on the time-varying covariance and on the correlation between the two time series; see the Supplementary Material for expressions of these quantities. Zhao (2015) considers autocorrelations when the underlying trend function is known to be zero. Here, we consider an inference of the first-order autocorrelation of (X_i). Note that the method of Zhao (2015) requires that the underlying process be precentered using the true mean function. Thus, we precenter the data using the local linear trend estimate; see the discussion in Section 3.2 of Zhao (2015) about the theoretical gap of such a heuristic approach. For the proposed method, precentering is not necessary, because the mean trend is automatically nullified by the LHC-DA method, with a solid theoretic guarantee. Let $n \in \{500, 1000\}$ and $b_n \in \{0.1, 0.15, 0.2, 0.25, 0.3\}$; the results are summarized in Tables 1 and 2 for the correlation case and the autocorrelation case, respectively. The proposed method can be applied to a general covariance or correlation between two time series, including when one is the lagged value of the other. In contrast, the method of Zhao (2015), denoted by Z15, was developed specifically for the autocorrelation, and is therefore only reported in the second portion of Table 2, when applicable. We also report results for the LHC method without the derivative adjustment, as a comparison. Tables 1 and 2 show that the proposed LHC-DA method performs reasonably well, because the empirical coverage probabilities are mostly close to their nominal levels when a suitable bandwidth is used. In general, it outperforms the LHC method, indicating that the derivative adjustment scheme described in Section 3.2 solves the bias problem from a theoretical point of view, and improves the finite-sample performance. As discussed in Section 3.2, the LHC-DA method makes the procedure asymptotically equivalent to that using the mean-oracle covariance estimators. This is not true of the LHC method, owing to the existence of an additional bias from the trend estimation. The method of Zhao (2015) applies only to the autocorrelation part in Table 2, and does not seem to be very robust with respect to the bandwidth choice when compared with the proposed LHC-DA method. Therefore, in addition to being applicable to a broader setting and successfully handling noncentered data, the proposed LHC-DA method leads to empirical tools that provide more robust finite-sample performance. Additional simulation results can be found in the Supplementary Material; these consider different data generations (e.g., time-varying autoregressions with potentially heavier tails) and produce qualitatively similar findings, as long as the conditions of Theorems 1 and 2 are satisfied.

4.3 Application to Financial Data

Correlation analysis between international stock markets is an important topic in economics and finance, and has been studied by Lin et al. (1994), Longin and Solnik (1995), Karolyi and Stulz (1996), Chesnay and Jondeau (2001), Engle (2002), Forbes and Rigobon (2002), Evans and McMillan (2009), and Madaleno and Pinho (2012), among many others. The assumption of a constant correlation has been challenged, and proven to be unsuitable in many studies; see, for example, Longin and Solnik (1995), Chesnay and Jondeau (2001), Engle (2002), Choi and Shin (2021), and the references therein. Here, we focus on the U.S. and Germany stock markets, and consider weekly return data of the U.S. S&P 500 index and the Germany DAX index from 01/01/1995 to 12/28/2020, with a total of $n = 1357$ data points. The data are available from Yahoo! Finance, and a time series plot is given in Figure 1. We allow the underlying correlation to change over time, and obtain a nonparametric estimate and its associated simultaneous confidence band for uncertainty quantification. Because two time

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0244

Table 1: Empirical coverage probabilities of the simultaneous confidence bands for the time-varying covariance and time-varying correlation between (X_i) and (Y_i) as functions of time.

		Z15			LHC			LHC-DA		
n	b_n	90%	95%	99%	90%	95%	99%	90%	95%	99%
					covariance					
500	0.1	-	-	-	0.879	0.938	0.979	0.888	0.940	0.978
	0.15	-	-	-	0.879	0.941	0.991	0.899	0.952	0.993
	0.2	-	-	-	0.898	0.952	0.990	0.911	0.957	0.993
	0.25	-	-	-	0.906	0.953	0.990	0.914	0.955	0.992
	0.3	-	-	-	0.909	0.957	0.994	0.916	0.961	0.993
1000	0.1	-	-	-	0.884	0.946	0.988	0.901	0.945	0.990
	0.15	-	-	-	0.892	0.943	0.988	0.904	0.951	0.990
	0.2	-	-	-	0.899	0.941	0.989	0.905	0.951	0.991
	0.25	-	-	-	0.909	0.953	0.988	0.926	0.962	0.992
	0.3	-	-	-	0.914	0.960	0.992	0.921	0.966	0.993
					correlation					
500	0.1	-	-	-	0.874	0.940	0.987	0.875	0.943	0.985
	0.15	-	-	-	0.839	0.917	0.986	0.852	0.918	0.983
	0.2	-	-	-	0.856	0.916	0.982	0.875	0.925	0.986
	0.25	-	-	-	0.866	0.924	0.985	0.887	0.939	0.989
	0.3	-	-	-	0.881	0.934	0.987	0.900	0.937	0.989
1000	0.1	-	-	-	0.851	0.911	0.982	0.852	0.913	0.984
	0.15	-	-	-	0.866	0.922	0.979	0.871	0.930	0.978
	0.2	-	-	-	0.867	0.923	0.974	0.881	0.933	0.976
	0.25	-	-	-	0.878	0.933	0.979	0.897	0.942	0.980
	0.3	-	-	-	0.883	0.942	0.991	0.895	0.950	0.990

Table 2: Empirical coverage probabilities of the simultaneous confidence bands for the first-order autocovariance and autocorrelation functions of (X_i).

		Z15			LHC			LHC-DA		
n	b_n	90%	95%	99%	90%	95%	99%	90%	95%	99%
					autocovariance					
500	0.1	-	-	-	0.867	0.925	0.960	0.861	0.924	0.957
	0.15	-	-	-	0.868	0.923	0.979	0.876	0.926	0.978
	0.2	-	-	-	0.885	0.933	0.984	0.891	0.937	0.981
	0.25	-	-	-	0.892	0.935	0.991	0.898	0.948	0.991
	0.3	-	-	-	0.897	0.941	0.994	0.903	0.948	0.996
1000	0.1	-	-	-	0.912	0.959	0.993	0.919	0.964	0.993
	0.15	-	-	-	0.901	0.955	0.992	0.907	0.957	0.993
	0.2	-	-	-	0.903	0.942	0.994	0.904	0.952	0.996
	0.25	-	-	-	0.905	0.953	0.991	0.921	0.957	0.996
	0.3	-	-	-	0.901	0.953	0.988	0.930	0.962	0.994
					autocorrelation					
500	0.1	0.891	0.937	0.978	0.911	0.965	0.998	0.908	0.965	0.997
	0.15	0.971	0.988	0.999	0.878	0.936	0.993	0.881	0.939	0.991
	0.2	1.000	1.000	1.000	0.874	0.939	0.985	0.876	0.946	0.987
	0.25	1.000	1.000	1.000	0.860	0.922	0.987	0.879	0.936	0.987
	0.3	1.000	1.000	1.000	0.839	0.901	0.979	0.872	0.929	0.988
1000	0.1	0.905	0.944	0.983	0.871	0.931	0.991	0.873	0.930	0.989
	0.15	0.996	1.000	1.000	0.895	0.938	0.989	0.900	0.948	0.992
	0.2	1.000	1.000	1.000	0.878	0.948	0.990	0.889	0.952	0.994
	0.25	1.000	1.000	1.000	0.861	0.930	0.981	0.887	0.948	0.985
	0.3	1.000	1.000	1.000	0.853	0.922	0.977	0.879	0.941	0.981

series are involved in this application, the method of Zhao (2015) is not directly applicable. Using the simulation-assisted algorithm in Section 4.1, the time-varying correlation and its 95% simultaneous confidence band are

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0244

visualized in Figure 2, from which we can see that the correlation between the U.S. and Germany stock markets indeed changes over time. In particular, there is a long-term increasing trend in the correlation between the two markets, indicating that the economies of the two countries are, in general, becoming increasingly interdependent over time; see also the discussion in Longin and Solnik (1995). The increasing trend peaked around 2008–2009, at the time of the financial crisis, with both countries subsequently relying on their own monetary policies to recover, which may explain the decrease in the correlation during that period. Once they had recovered, the correlation between the two markets experiences another increasing trend, similar to that before the financial crisis.

4.4 Application to COVID Data

The recent COVID-19 pandemic has become a major concern for policymakers and researchers. Cross-country studies have shown that the virus spread rate and pattern are affected by local cultures, government responses, and economic developments, among other factors; see, for example, Balmford et al. (2020), Middelburg and Rosendaal (2020), Rypdal and Rypdal (2020), Zarikas et al. (2020), Vampa (2021), and the references therein. Mahmoudi et al. (2021) and Nobi et al. (2021) examined the correlation between case

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0244

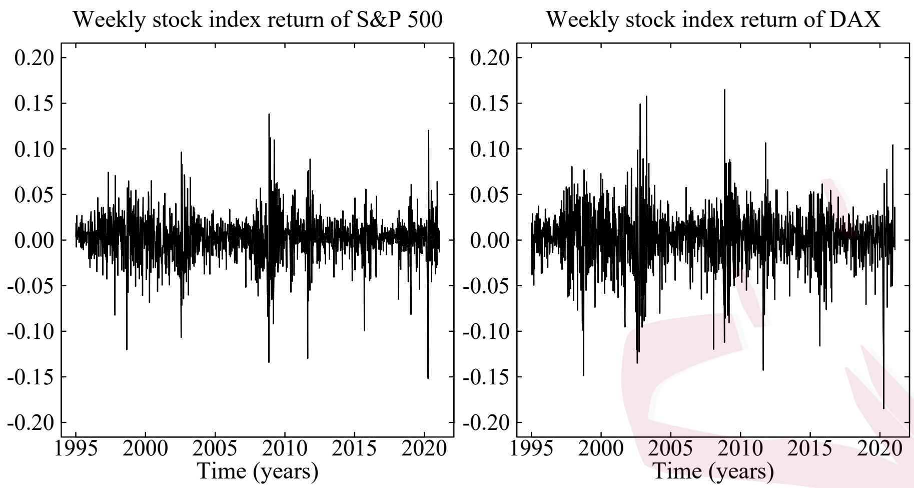

Figure 1: Time series plots for weekly returns of the U.S. S&P 500 index and the Germany DAX index from 01/01/1995 to 12/28/2020.

numbers from different countries, and Sulyok et al. (2021) showed that the correlation can vary at different times during the pandemic. We model the underlying correlation as a nonparametric function of time, to be more flexible and less vulnerable to parametric models, and apply our results to obtain a simultaneous confidence band to help examine the pattern. For this, we consider the log daily new cases per million people in Germany and the United Kingdom from 06/01/2020 to 12/31/2021, with a total of $n = 572$ data points. The data are available from Ritchie et al. (2020), and a time series plot is provided in Figure 3. The time-varying correlation and its 95% simultaneous confidence band are visualized in Figure 4, from

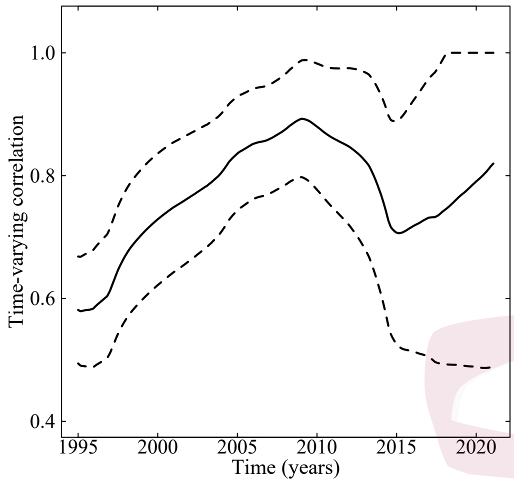

Figure 2: The time-varying correlation (solid curve) and its associated 95% simultaneous confidence band (dashed curve) between weekly returns of the U.S. S&P 500 index and the Germany DAX index from 01/01/1995 to 12/28/2020.

which we can see that Germany and the United Kingdom began with a relatively stable correlation, which then decreased to a negative value in the early part of 2021. During this period, the number of daily new COVID-19 cases seems to exhibit an increase in Germany, but continued to decrease in the United Kingdom, which may be related to the different degrees of vaccine intervention in the two countries around that time. In particular, during the first few months of 2021, the United Kingdom experienced a much more rapid increase in its vaccination rates compared with Germany, which potentially helped the United Kingdom and differentiated it from

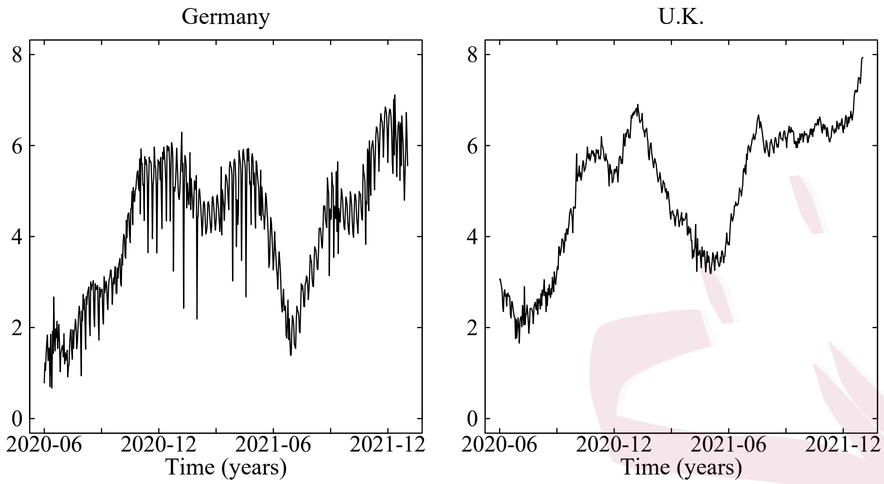

Figure 3: Time series plots of log daily new COVID-19 cases per million people in Germany and the United Kingdom from 06/01/2020 to 12/31/2021.

Germany when the delta variant hit both countries around that time. On the other hand, there seems to be a decrease in the correlation around the end of 2021, as shown in Figure 4. This time, the number of daily new COVID-19 cases seems to decrease in Germany, but continued to rise in the United Kingdom, which is the opposite of what happened in the early part of 2021. This may be related to the different lockdown policies of the two governments. In particular, Germany cancelled their Christmas markets and imposed local lockdowns when the highly contagious omicron variant hit both countries.

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0244

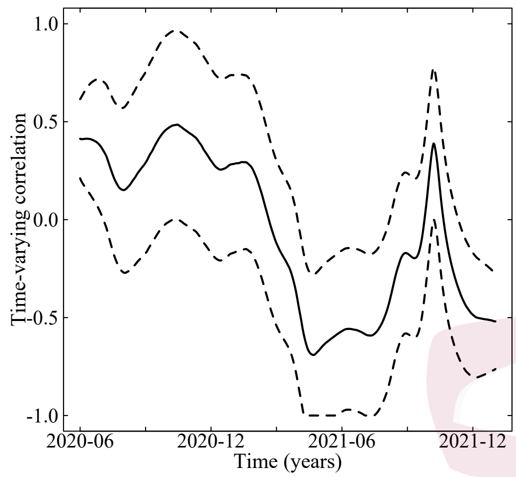

Figure 4: The time-varying correlation (solid curve) and its associated 95% simultaneous confidence band (dashed curve) between log daily new cases per million people in Germany and the United Kingdom from 06/01/2020 to 12/31/2021.

5. Conclusion

We have considered a simultaneous inference of the nonparametric correlation curve between two nonstationary time series. Compared with the result of Zhao (2015), which was developed specifically for autocorrelations of a univariate time series, our results can be applied to the broader setting in which one time series is not necessarily a lagged version of the other. In addition, we address the problem discussed in Zhao (2015) about how to handle the nuisance unknown trend function when making an inference about the correlation curve. Unlike the stationary setting, the straightfor-

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0244

ward precentering approach in the current time-varying setting can result in estimators with theoretical properties that are difficult to understand. To address this, we propose an LHC scheme that, instead of aligning with the time point at which the data are observed, aligns with the time point at which the local correlation estimation is performed. Although this newly proposed centering scheme makes it possible to quantify the effect of trend estimation in correlation inference, it comes at the cost of an additional bias term, making the effect of trend estimation not asymptotically negligible. We then propose a further derivative adjustment scheme, which is able to make the bias term asymptotically negligible, so that the resulting correlation estimators can be asymptotically equivalent to the mean-oracle ones, obtained as if we know the true mean functions. Our simulation results in Section 4.2 show that, in addition to being applicable to a broader setting and successfully handling noncentered data, the proposed LHC-DA method delivers an improved and more robust finite-sample performance. We expect that the proposed method will become a useful tool for examining correlations that are not constant, but change over time.

Acknowledgements

We are grateful to the editor, associate editor, and anonymous referee for their helpful comments and suggestions. This research was supported by the NSF CAREER Award DMS-2131821.

Supplementary Material

The online Supplementary Material provides technical proofs of our main results in Section 3 and additional simulation results for the simulation study in Section 4.2.

References

Anderson, T. W. (1971). *The Statistical Analysis of Time Series.* John Wiley & Sons.

Balmford, B., Annan, J. D., Hargreaves, J. C., Altoè, M. and Bateman, I. J. (2020). Cross-country comparisons of COVID-19: policy, politics and the price of life. *Environmental and Resource Economics*, **76**, 525–551.

Berkes, I., Liu, W. and Wu, W. B. (2014). Komlós-Major-Tusnády approximation under dependence. *The Annals of Probability*, **42**, 794–817.

Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods.* Springer.

Cai, Z. (2007). Trending time-varying coefficient time series models with serially correlated errors. *Journal of Econometrics*, **136**, 163–188.

Cheng, B. and Tong, H. (1998). K-stationarity and wavelets. *Journal of Statistical Planning and Inference*, **68**, 129–144.

Chesnay, F. and Jondeau, E. (2001). Does correlation between stock returns really increase during turbulent periods? *Economic Notes*, **30**, 53–80.

Choi, J.-E. and Shin, D. W. (2021). Nonparametric estimation of time varying correlation coefficient. *Journal of the Korean Statistical Society*, **50**, 333–353.

Dahlhaus, R. (1996). On the Kullback-Leibler information divergence of locally stationary processes. *Stochastic Processes and their Applications*, **62**, 139–168.

Dahlhaus, R. (1997). Fitting time series models to nonstationary processes. *The Annals of Statistics*, **25**, 1–37.

Dahlhaus, R. (2012). Locally stationary processes. In T. S. Rao, S. S. Rao and C. R. Rao, editors, *Handbook of Statistics*, vol. 30. Elsevier, pp. 351–413.

Degras, D., Xu, Z., Zhang, T. and Wu, W. B. (2012). Testing for parallelism among trends in multiple time series. *IEEE Transactions on Signal Processing*, **60**, 1087–1097.

Draghicescu, D., Guillas, S. and Wu, W. B. (2009). Quantile curve estimation and visualization for non-stationary time series. *Journal of Computational and Graphical Statistics*, **18**, 1–20.

Engle, R. F. (2002). Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, **20**, 339–350.

Evans, T. and McMillan, D. G. (2009). Financial co-movement and correlation: evidence from 33 international stock market indices. *International Journal of Banking, Accounting and Finance*, **1**, 215–241.

Fan, J. and Gijbels, I. (1996). *Local Polynomial Modeling and its Applications.* Chapman & Hall, London.

Forbes, K. J. and Rigobon, R. (2002). No contagion, only interdependence: measuring stock market comovements. *The Journal of Finance*, **57**, 2223–2261.

Fu, Z., Chan, S.-C., Di, X., Biswal, B. and Zhang, Z. (2014). Adaptive covariance estimation of non-stationary processes and its application to infer dynamic connectivity from fMRI. *IEEE Transactions on Biomedical Circuits and Systems*, **8**, 228–239.

Giurcanu, M. and Spokoiny, V. (2004). Confidence estimation of the covariance function of stationary and locally stationary processes. *Statistics & Decisions*, **22**, 283–300.

Hall, P. and Heyde, C. C. (1980). *Martingale Limit Theory and its Application.* Academic Press.

Hannan, E. J. (1976). The asymptotic distribution of serial covariances. *The Annals of Statistics*, **4**, 396–399.

Hosking, J. R. (1996). Asymptotic distributions of the sample mean, autocovariances, and autocorrelations of long-memory time series. *Journal of Econometrics*, **73**, 261–284.

Karolyi, G. A. and Stulz, R. M. (1996). Why do markets move together? An investigation of US-Japan stock return comovements. *The Journal of Finance*, **51**, 951–986.

LIN, W.-L., ENGLE, R. F. AND ITO, T. (1994). Do bulls and bears move across borders? International transmission of stock returns and volatility. *Review of Financial Studies*, **7**, 507–538.

LONGIN, F. AND SOLNIK, B. (1995). Is the correlation in international equity returns constant: 1960–1990? *Journal of International Money and Finance*, **14**, 3–26.

MADALENO, M. AND PINHO, C. (2012). International stock market indices comovements: a new look. *International Journal of Finance & Economics*, **17**, 89–102.

MAHMOUDI, M. R., HEYDARI, M. H., QASEM, S. N., MOSAVI, A. AND BAND, S. S. (2021). Principal component analysis to study the relations between the spread rates of COVID-19 in high risks countries. *Alexandria Engineering Journal*, **60**, 457–464.

MALLAT, S., PAPANICOLAOU, G. AND ZHANG, Z. (1998). Adaptive covariance estimation of locally stationary processes. *The Annals of Statistics*, **26**, 1–47.

MIDDELBURG, R. A. AND ROSENDAAL, F. R. (2020). COVID-19: how to make between-country comparisons. *International Journal of Infectious Diseases*, **96**, 477–481.

NADARAYA, E. A. (1964). On estimating regression. *Theory of Probability & Its Applications*, **9**, 141–142.

NASON, G. P., VON SACHS, R. AND KROISANDT, G. (2000). Wavelet processes and adaptive estimation of the evolutionary wavelet spectrum. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **62**, 271–292.

NOBI, A., TUHIN, K. H. AND LEE, J. W. (2021). Application of principal component analysis on temporal evolution of COVID-19. *Plos One*, **16**, e0260899.

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

Phillips, P. C. and Solo, V. (1992). Asymptotics for linear processes. *The Annals of Statistics*, 971–1001.

Priestley, M. B. (1981). *Spectral Analysis and Time Series: Probability and Mathematical Statistics*. Academic Press, New York.

Priestley, M. B. and Chao, M. T. (1972). Non-parametric function fitting. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **34**, 385–392.

Ritchie, H., Mathieu, E., Rodés-Guirao, L., Appel, C., Giattino, C., Ortiz-Ospina, E., Hasell, J., Macdonald, B., Beltekian, D. and Roser, M. (2020). Coronavirus pandemic (COVID-19). *Our World in Data*. Https://ourworldindata.org/coronavirus.

Robinson, P. M. (1989). Nonparametric estimation of time-varying parameters. In P. Hackl, editor, *Statistical Analysis and Forecasting of Economic Structural Change*. Springer, pp. 253–264.

Robinson, P. M. (1991). Time-varying nonlinear regression. In P. Hackl and A. H. Westlund, editors, *Economic Structural Change*. Springer, pp. 179–190.

Rypdal, K. and Rypdal, M. (2020). A parsimonious description and cross-country analysis of COVID-19 epidemic curves. *International Journal of Environmental Research and Public Health*, **17**, 6487.

Sulyok, M., Ferenci, T. and Walker, M. (2021). Google trends data and COVID-19 in Europe: correlations and model enhancement are European wide. *Transboundary and Emerging Diseases*, **68**, 2610–2615.

Vampa, D. (2021). COVID-19 and territorial policy dynamics in Western Europe: comparing France, Spain, Italy, Germany, and the United Kingdom. *Publius: The Journal of Federalism*, **51**, 601–626.

Watson, G. S. (1964). Smooth regression analysis. *Sankhyā: The Indian Journal of Statistics, Series A*, 359–372.

Wu, W. B. (2007). Strong invariance principles for dependent random variables. *The Annals of Probability*, **35**, 2294–2320.

Wu, W. B. (2009). An asymptotic theory for sample covariances of Bernoulli shifts. *Stochastic Processes and their Applications*, **119**, 453–467.

Wu, W. B. and Min, W. (2005). On linear processes with dependent innovations. *Stochastic Processes and their Applications*, **115**, 939–958.

Wu, W. B. and Xiao, H. (2012). Covariance matrix estimation in time series. In T. S. Rao, S. S. Rao and C. R. Rao, editors, *Handbook of Statistics*, vol. 30. Elsevier, pp. 187–209.

Zarikas, V., Poulopoulos, S. G., Gareiou, Z. and Zervas, E. (2020). Clustering analysis of countries using the COVID-19 cases dataset. *Data in Brief*, **31**, 105787.

Zhang, T. (2013). Clustering high-dimensional time series based on parallelism. *Journal of the American Statistical Association*, **108**, 577–588.

Zhang, T. (2015). Semiparametric model building for regression models with time-varying parameters. *Journal of Econometrics*, **187**, 189–200.

Zhang, T. (2016a). Testing additive assumptions on means of regular monitoring data: a multivariate nonstationary time series approach. *Statistica Sinica*, **26**, 1611–1630.

Zhang, T. (2016b). Testing for jumps in the presence of smooth changes in trends of nonstationary time series. *Electronic Journal of Statistics*, **10**, 706–735.

Zhang, T. (2018). A thresholding-based prewhitened long-run variance estimator and its dependence oracle property. *Statistica Sinica*, **28**, 319–338.

Zhang, T. and Wu, W. B. (2011). Testing parametric assumptions of trends of a nonstationary time series. *Biometrika*, **98**, 599–614.

Zhang, T. and Wu, W. B. (2012). Inference of time-varying regression models. *The Annals of Statistics*, **40**, 1376–1402.

Zhao, Z. (2015). Inference for local autocorrelations in locally stationary models. *Journal of Business & Economic Statistics*, **33**, 296–306.

Zhou, Z. and Wu, W. B. (2010). Simultaneous inference of linear models with time varying coefficients. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **72**, 513–531.

Department of Statistics, University of Georgia, 310 Herty Drive, Athens, GA 30602, U.S.A.

E-mail: tingzhang@uga.edu

Department of Mathematics and Statistics, Boston University, 665 Commonwealth Ave, Boston, MA 02215, U.S.A.

E-mail: yshao19@bu.edu