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
Notice: Accepted version subject to English editing.	

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 impact from precentering using nonparametric trend function estimators. This is mainly due to the trend estimators being time-varying across different time points, which makes it difficult to quantify their cumulative interaction with the error process in the time series setting. We propose to fix this unpleasant issue by using a centering scheme that, instead of aligning with the time point at which the data is observed, aligns with the time point at which the local correlation estimation is performed. We show that this new centering scheme can lead 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. Numerical examples including a real data analysis are provided to illustrate the proposed method.

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

1. Introduction

The correlation coefficient has been a prominent metric to quantify the dependence between two variables. In the time series setting, one may be interested in the correlation between two observed time series to understand their relationship or co-movement over time, or the correlation between a time series and its lagged version to study the underlying dependence structure. The latter is often referred to as the autocorrelation, and a nice survey can be found in Wu and Xiao (2012). In addition, one can be interested in the correlation between a time series and the lagged version of another time series to understand the lagged effect of one on the other, which is often referred to as the Granger causality in time series analysis. The problem of estimating the correlation and autocorrelation has been extensively studied 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 references therein. In the aforementioned results, the underlying process is mostly assumed to be stationary and as a result the correlation coefficient is a constant that does not change over time, which largely facilitated its estimation and statistical inference.

In nonstationary time series applications, however, it is generally expected that certain aspects of the observed data can evolve over time, which makes it more desirable to consider time-varying correlations as a function of time. For this, Mallat et al. (1998) considered covariance estimation through a local cosine basis approximation for locally stationary processes. Dahlhaus (2012) considered a data tapering method for covariance estimation of locally stationary processes using kernel functions. Fu et al. (2014) considered estimating the time-varying covariance between two locally stationary biological processes, and provided an asymptotic analysis on the resulting estimation bias and variance. Choi and Shin (2021) considered nonparametric estimation of the time-varying correlation coefficient and established its asymptotic normality when the joint error process is strong mixing and stationary except for a scale difference. The aforementioned results mostly only concern the estimation or pointwise inference of the time-varying covariance at a given time point, while the difficult task of developing a simultaneous inference procedure for the time-varying correlation as a function of time has been much less explored.

In an important work, Zhao (2015) provided a solution to this problem by constructing simultaneous confidence bands for local autocorrelations of locally stationary time series. Their theory and methods, however, rely on the assumption that the mean trend function of the underlying process is known to be uniformly zero, and it was argued in Zhao (2015) that the daily or weekly data on financial returns are generally perceived to satisfy this assumption. For data with potentially nonzero trend functions, Zhao (2015) proposed to first precenter the data using parametric or nonparametric trend estimators and then apply their methods to the residual process. The impact from such a precentering procedure on the subsequent correlation inference procedure, however, is nontrivial to quantify and was left as an open problem; see the discussion in Section 3.2 of Zhao (2015). In Section 2, we present a detailed investigation about why the effect of precentering was trivially negligible in the stationary case but suddenly becomes difficult to understand in the nonstationary case. It is suggested that this is largely due to the time-varying nature of the trend function that makes it difficult to quantify its cumulative interaction with the error process. We then in Section 3 propose a fix, called the locally homogenized centering method, that is able to alleviate the issue from traditional centering schemes and lead to simultaneous inference of time-varying correlations with a solid theoretical guarantee when the underlying trend functions are unknown. In addition to the allowance of unknown trend functions, we also consider the more general setting when one is interested in the time-varying correlation between two time series that can depend on each other in a nontrivial way. In particular, when one time series is taken as the lagged version of the other, then it reduces to the autocorrelation setting as considered in Zhao (2015). Additionally, Zhao (2015) requires a geometric moment contraction condition under which the dependence decays geometrically quickly, while the current results 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 Issue

We shall 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 $\bar{Y}_n = n^{-1}\sum_{i=1}^{n} Y_i$, which then leads to the covariance estimator

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

The effect of using the sample mean to replace the true mean can then be quantified 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 $\gamma(t)$, $t \in [0, 1]$. Note that the scaling device itself does not impose any additional assumption on the underlying dynamics, but it can work well with certain smoothness conditions to provide 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) for more discussions. Assuming that the true mean functions $\mu_x(\cdot)$ and $\mu_y(\cdot)$ are known, then 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 $\hat{\mu}_y(\cdot)$, such as the Nadaraya (1964) and Watson (1964) estimator, the Priestley and Chao (1972) estimator, or the local linear estimator of Fan and Gijbels (1996), which then 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 precentered data $X_i - \hat{\mu}_x(i/n)$ and $Y_i - \hat{\mu}_y(i/n)$ to compute the covariance as 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 then quantified by the difference

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

which unfortunately cannot be easily bounded by a negligible stochastic order as in the stationary case. The main reason here is that, due 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 in the stationary case. Since $\hat{\mu}_y(i/n) - \mu_y(i/n)$ and $X_i - \mu_x(i/n)$ can depend on each other in a nontrivial way, it then becomes unclear if 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 along their uniform rate over different time points which is however 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 if they can be treated as negligible in the theoretical analysis. Therefore, although being natural, the approach of replacing the unknown mean function by its nonparametric estimator in covariance inference problems is rather ad hoc whose impact can be theoretically difficult to understand; see also the discussion in Section 3.2 of Zhao (2015). We shall in the following propose a fix to this unpleasant issue, which then 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

Our investigation above indicates that the major issue of the natural centering scheme as in (2.1) for covariance inference problems is the local inhomogeneity of $\hat{\mu}_x(i/n)$ and $\hat{\mu}_y(i/n)$ that are different for different indices $i = 1, \ldots, n$. To alleviate the issue, we propose a locally homogenized centering (LHC) method which, instead of using the inhomogeneous centering $\hat{\mu}_x(i/n)$ and $\hat{\mu}_y(i/n)$ that align with the time point at which the data is observed, uses their locally homogenous counterparts $\hat{\mu}_x(t)$ and $\hat{\mu}_y(t)$ that 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 will be different for different time points at which the local covariance is calculated. The effect of this new centering scheme can now be quantified by the difference

$$\begin{aligned}\check{\gamma}_n(t) - \tilde{\gamma}_n(t) \quad = \quad & \frac{1}{nb_n}\sum_{i=1}^{n}\{\hat{\mu}_x(t) - \mu_x(i/n)\}\{\hat{\mu}_y(t) - \mu_y(i/n)\}K\left(\frac{i/n - t}{b_n}\right) \\ & -\frac{1}{nb_n}\sum_{i=1}^{n}\{X_i - \mu_x(i/n)\}\{\hat{\mu}_y(t) - \mu_y(i/n)\}K\left(\frac{i/n - t}{b_n}\right) \\ & -\frac{1}{nb_n}\sum_{i=1}^{n}\{\hat{\mu}_x(t) - \mu_x(i/n)\}\{Y_i - \mu_y(i/n)\}K\left(\frac{i/n - t}{b_n}\right).\end{aligned} \tag{3.3}$$

Compared with the decomposition for the natural centering scheme as 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 $\hat{\mu}_y(t) - \mu_y(t)$ that can be taken outside of the summation and a deterministic part $\mu_y(t) - \mu_y(i/n)$ whose cumulative interaction with $X_i - \mu_x(i/n)$ can be handled by 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, which can then lead to new inference protocols for the time-varying covariance or correlation of nonstationary time series. The merit, though being 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 the bias of the mean-oracle estimator $E\{\tilde{\gamma}_n(t)\}-\gamma(t)$. We shall here further propose a derivative adjustment method to get rid of this additional bias and make the resulting covariance estimator asymptotically equivalent to its mean-oracle counterpart so 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 to consider the covariance estimator with 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 to how we handle the terms in (3.3), we in this case 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 at the same time involves the summation index i. However, it is still different from 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 process. 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 name it the locally homogenized centering with derivative adjustment (LHC-DA), 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

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

Note that a time-varying correlation analysis may be more suitable than the covariance to understand the time-varying relationship between two nonstationary time series, as the change in covariance can be simply due to changes in the variance while the correlation can remain as a constant. We shall in the following provide an asymptotic theory for the proposed LHC-DA covariance and correlation estimators, based on which simultaneous confidence bands can be constructed as a visualization tool to analyze 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, $i = 1, \ldots, n$, according to

$$X_i = G(i/n, \boldsymbol{\mathcal{F}}_i), \quad Y_i = H(i/n, \boldsymbol{\mathcal{F}}_i), \quad \boldsymbol{\mathcal{F}}_i = (\ldots, \boldsymbol{\epsilon}_{i-1}, \boldsymbol{\epsilon}_i), \tag{3.4}$$

where $(\boldsymbol{\epsilon}_i)$ is a sequence of independent and identically distributed innovations, and G and H are measurable functions that can depend on the time points $t_{i,n} = i/n$, $i = 1, \ldots, n$. The framework (3.4) 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 references therein. Let $(\boldsymbol{\epsilon}_i^\star)$ be a sequence of random vectors that share the same distribution as, but 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 $\|\boldsymbol{Z}\|_q = \{E(|\boldsymbol{Z}|^q)\}^{1/q}$, $q > 0$, where $|\boldsymbol{Z}|$ is the Euclidean norm, and denote $\|\boldsymbol{Z}\| = \|\boldsymbol{Z}\|_2$. For a generic process $L(t, \boldsymbol{\mathcal{F}}_i)$, $t \in [0, 1]$, $i \in \mathbb{Z}$, 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 $L(t, \boldsymbol{\mathcal{F}}_i)$, $t \in [0,1]$, $i \in \mathbb{Z}$, 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 we 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 can be 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* $n^{-2/5} b_n^{-1} (\log n)^3 + n b_n^7 \log n \to 0$, *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

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

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, and we shall in Theorem 2 provide 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 not only the covariance part but also the variance part, and neither is negligible comparing to the other; see also Zhao (2015) which considers the autocorrelation inference for processes with a known zero mean and geometrically decaying dependence. We shall here 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

We shall here provide a detailed algorithm that implements the developed results in Section 3 to construct simultaneous confidence bands for the time-varying correlation between X_i and Y_i, $i = 1, \ldots, n$, when their trend functions are unknown. If one of them is taken as the lagged version of the other, then the algorithm will provide simultaneous confidence bands for the corresponding autocorrelation. To alleviate the issue of slow convergence to the extreme value distribution, we also consider the use of a simulation-assisted procedure to help 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 $\hat{\mu}_x(t)$, $\hat{\mu}_y(t)$ and their derivative estimators $\hat{\mu}'_x(t)$, $\hat{\mu}'_y(t)$ by using the local linear method of Fan and Gijbels (1996) with $K(\cdot)$ being the Epanechnikov kernel.

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

(iv) Obtain an estimate $\hat{\varpi}_V(t)$ of the asymptotic variance using the band-ing 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 $Y_i^\diamond$, $i = 1, \ldots, n$, 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) for a large number of 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 against time while 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, and the resulting tool can be useful for practitioners to examine the time-varying covariance or correlation when the observed data contains an unknown trend in the mean. Similar to the bootstrap, the simulation-assisted procedure aims at approximating the distribution of the test statistic by that of a generated data. The difference, however, is that bootstrapped data are often generated by resampling from the original data, while 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 continue to work when the simulated data are independently generated using marginal distributions other than the normal; see also additional simulation results provided in the supplementary materials, which suggest robustness to the distributional choice as long as conditions in Theorems 1 and 2 are satisfied. The use of the normal distribution to generate the simulated data, however, is due to 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 often help lead to improvements in 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

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

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

where $\mu_x(t) = 2t^2 + 2t$ and $\mu_y(t) = 2\{\sin(1.5\pi t) + t\}$, and we consider making simultaneous inference on the time-varying covariance and correlation between the two time series; see the supplementary materials for expressions of these quantities. Zhao (2015) considered the situation of autocorrelations when the underlying trend function is known to be zero, and we shall here also make a comparison by considering inference of the first-order autocorrelation of (X_i). Note that the method of Zhao (2015) requires the underlying process to be precentered by the true mean function, and we shall here follow their heuristic suggestion and precenter the data by 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, as the mean trend will be 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 autocorrelation case respectively. Unlike the proposed method that can be applied to a general covariance or correlation between two time series including when one is the lagged value of the other, the method of Zhao (2015), denoted by Z15, was specifically developed for the autocorrelation and is therefore only reported in the second portion of Table 2 when it is applicable. We also report results for the LHC method without the derivative adjustment as a comparison. It can be seen from Tables 1 and 2 that the proposed LHC-DA method performs reasonably well as the empirical coverage probabilities are mostly close to their nominal levels when a suitable bandwidth is used. It generally outperforms the LHC method, indicating that the derivative adjustment scheme as described in Section 3.2 not only addresses the unpleasant bias issue from a theoretical point of view but also leads to improvements in 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, while this benefit is generally not shared by the LHC method due to the existence of an additional bias from trend estimation. For the method of Zhao (2015), it only applies to the autocorrelation part in Table 2, and it 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 addressing the gap about handling noncentered data that was left as an open problem in Zhao (2015), the proposed LHC-DA method also seems to be able to lead to empirical tools that have improved or more robust finite-sample performance. Additional simulation results that supplement the current simulation study can be found in the supplementary materials; these consider different data generations (e.g., time-varying autoregressions with potentially heavier tails) and produce qualitatively similar findings as long as conditions of Theorems 1 and 2 are satisfied.

4.3 Application to Financial Data

Correlation analysis between international stock markets has been an important topic in economics and finance, which 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

Table 1: Empirical coverage probabilities of 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 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

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 references therein. Here we focus on the U.S. and Germany stock markets, and consider the 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 is available from Yahoo! Finance, and a time series plot is given in Figure 1. We shall here allow the underlying correlation to change over time and apply the developed results to obtain a nonparametric estimate and its associated simultaneous confidence band for uncertainty quantification. Since two time series are involved in this application, the method of Zhao (2015) is not directly applicable. By the simulation-assisted algorithm in Section 4.1, the time-varying correlation and its 95% simultaneous confidence band are visualized in Figure 2, from which we can see that the correlation between the U.S. and Germany stock markets is indeed changing over time. In particular, a long-term increasing trend can be observed in the correlation between the two markets indicating that the economy of the two countries tend to depend more and more on each other under globalization in general; see also the discussion in Longin and Solnik (1995). Such an increasing trend peaked around 2008–2009, at which time both countries start to suffer from the financial crisis and may rely more on their own monetary policies to recover, which can help explain the decrease

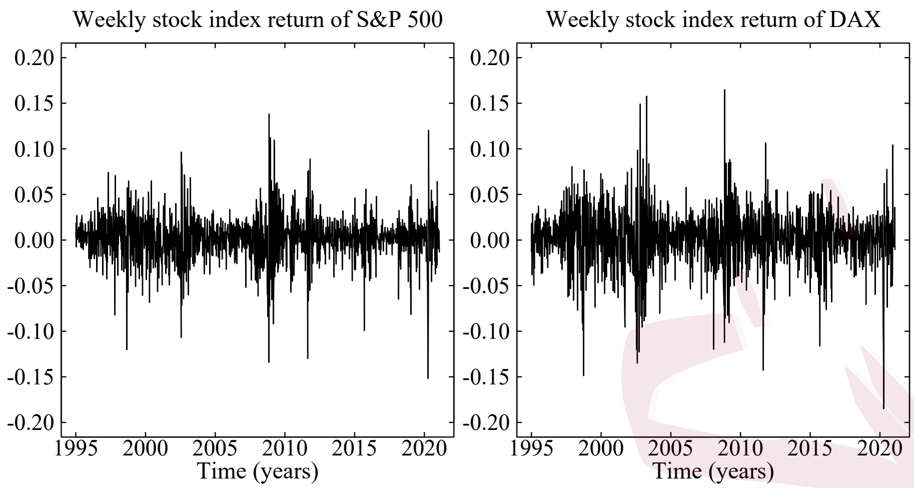

Figure 1: Time series plots for weekly returns of the U.S. S&P 500 index and the Germany DAX index from 01/01/1995 to 12/28/2020.

in correlation during that period. Once recovered, the correlation between the two markets seems to experience another increasing trend similar to what happened before the financial crisis.

4.4 Application to COVID Data

The recent pandemic of COVID-19 has become a major concern for policy makers all over the globe and has been involving researchers from various disciplines. Cross-country studies have shown that the virus spread rate and pattern can be affected by local cultures, government responses, and economic developments, among other factors; see for example Balmford

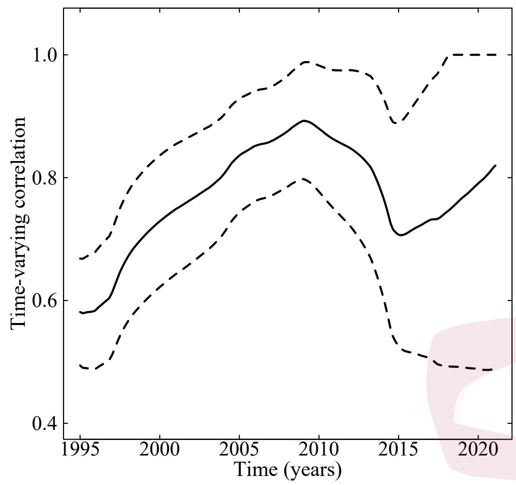

Figure 2: The time-varying correlation (solid curve) and its associated 95% simultaneous confidence band (dashed curve) between weekly returns of the U.S. S&P 500 index and the Germany DAX index from 01/01/1995 to 12/28/2020.

et al. (2020), Middelburg and Rosendaal (2020), Rypdal and Rypdal (2020), Zarikas et al. (2020), Vampa (2021), and references therein. Mahmoudi et al. (2021) and Nobi et al. (2021) examined the correlation between case numbers from different countries, and it was found by Sulyok et al. (2021) that the correlation can be different at different times in the pandemic. We shall here model the underlying correlation as a nonparametric function of time to be more flexible and less vulnerable to parametric models, and apply the developed 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 U.K. from 06/01/2020 to 12/31/2021, with a total of $n = 572$ data points. The data is 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 which we can see that Germany and U.K. began with a relatively stable correlation, which then decreased to a negative value in Spring 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 U.K., 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, U.K. experienced a much more rapid increase in its vaccination rates compared to Germany, which potentially helped U.K. and differentiated it from Germany when the delta variant hit both countries around that time. On the other hand, there seems to exist another decrease in correlation around the end of 2021 from Figure 4. This time, however, the number of daily new COVID-19 cases seems to exhibit a decrease in Germany but continued to rise in the U.K., which is the opposite of what happened in Spring 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, while around the same time the highly contagious omicron

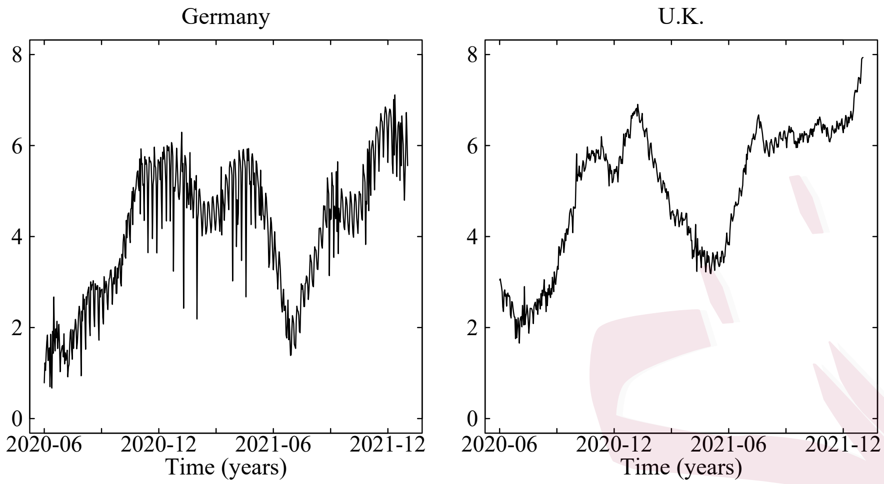

Figure 3: Time series plots of log daily new COVID-19 cases per million people in Germany and U.K. from 06/01/2020 to 12/31/2021.

variant hit both countries.

5. Conclusion

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

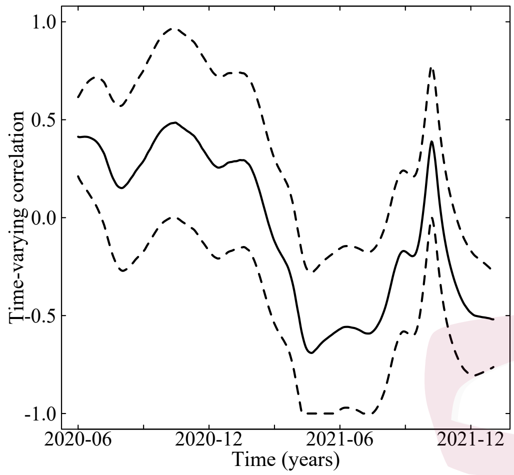

Figure 4: The time-varying correlation (solid curve) and its associated 95% simultaneous confidence band (dashed curve) between log daily new cases per million people in Germany and U.K. from 06/01/2020 to 12/31/2021.

tion curve. It can be seen from the discussion in Section 2 that, unlike the stationary setting, the straightforward precentering approach in the current time-varying setting can result in estimators whose theoretical properties are very difficult to understand. To address this, we propose a locally homogenized centering scheme which, instead of aligning with the time point at which the data is 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. It can be seen from our simulation results in Section 4.2 that, in addition to being applicable to a broader setting and successfully addressing the gap about handling noncentered data that was left as an open problem in Zhao (2015), the proposed LHC-DA method also seems to deliver an improved or more robust finite-sample performance. We expect that it will become a useful tool for practitioners to examine correlations that are not constant but change over time in their applications.

Acknowledgement

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

Supplementary Materials

Technical proofs of our main results in Section 3 and additional simulation results that supplement the simulation study in Section 4.2 are provided in the supplementary materials.

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, 111 Cummington Mall, Boston, MA 02215, U.S.A.

E-mail: yshao19@bu.edu