Journal of the Korean Statistical Society (2021) 50:333–353
https://doi.org/10.1007/s42952-020-00073-6

Online ISSN 2005-2863
Print ISSN 1226-3192

RESEARCH ARTICLE

Nonparametric estimation of time varying correlation coefficient

Ji-Eun Choi[1] **· Dong Wan Shin**[2]

Received: 6 November 2019 / Accepted: 9 May 2020 / Published online: 8 June 2020

Abstract

We propose a new time varying correlation coefficient, which is a local correlation measure of a pair of time series. The time varying correlation coefficient is locally estimated using a nonparametric kernel method. Asymptotic normality of the estimated time varying correlation is established, which allows us to construct statistical methods of confidence interval and hypothesis tests. Finite sample validity of the proposed methods are demonstrated by a Monte–Carlo study. The proposed time varying correlation coefficient method is well illustrated by an analysis of five sets of world major stock price index returns.

Keywords Confidence interval · Time varying correlation coefficient · Nonparametric estimation · Statistical test

1 Introduction

Correlation coefficient has been widely used as a standard measure to grasp relationships between two variables. Especially, correlation coefficient analysis has been applied in various areas of finance and economics such as portfolio management, forecasting and risk management. The correlation between financial or the economic variables has dominant characteristics of non-linearity and time dependence, which the usual Pearson correlation coefficient cannot capture. Thus, many studies propose time varying correlation coefficients addressing the dominant features.

The time varying correlation coefficient is usually estimated with multivariate generalized autoregressive conditional heteroscedasticity (GARCH) models: the BEKK model introduced by Engle and Kroner (1995), the GO-GARCH model proposed by

✉ Dong Wan Shin
shindw@ewha.ac.kr

Ji-Eun Choi
choije@pknu.ac.kr

1 Department of Statistics, Pukyong National University, Busan, Korea

2 Department of Statistics, Ewha University, Seoul, Korea

van der Weide (2002) and the dynamic conditional correlation (DCC) model of Engle (2002). The DCC model is an extension of the Bollerslev (1990)'s constant correlation coefficient model and is the most widely used as an estimating method of the time varying correlation coefficient. Since the global financial crisis, these time varying correlation coefficients have received a lot of attention to find out influential financial institutions on financial system and to manage system risk, see Girardi and Ergun (2013), Jin and Simone (2014), Brownlees and Engle (2017), Lucas et al. (2017) and many others.

Many interesting correlation dynamics were found by DCC analysis for financial assets in terms of financial crisis: for example, positive correlation between commodity future markets sharply increased during the global financial crisis, as analyzed by Kang et al. (2017); increasing trend in correlation of Romania and Estonia markets and the US market, as analyzed by Syllignakis and Kouretas (2011). However, these findings were stated without statistical justifications. This is because the DCC analysis does not allow statistical methods of confidence interval and statistical hypothesis testing regarding correlation dynamics.

We consider a time varying correlation coefficient method for statistical analysis of correlation dynamics. Related works are the local correlation methods of Tjostheim and Hufthammer (2013) and Lacal and Tjostheim (2018) in which correlations are estimated in neighborhoods of specific quantiles, but not in neighborhoods of time. They develop confidence intervals for correlation functions across quantiles but not across times. Therefore, their methods cannot be applied for investigating dynamics in terms of time. Time-related correlation dynamics were studied by Wied et al. (2012), Galeano and Wied (2014), Wied (2017) and Choi and Shin (2018) in terms of structural break. These break methods tell us only presence of correlation change, if any, but do not provide us confidence interval and tests for specific time dependent correlation pattern. Some interesting recent studies on the varying coefficient methods are the work of Li et al. (2013) for a categorical model with an application to a wage data set, that of Liu et al. (2014) for a model for case–control data sets, that of Cai et al. (2015) for dynamic panel data models and that of Li et al. (2017) for semi-varying coefficient models with nonstationary regressors.

The time varying correlation coefficient is locally estimated using Nadaraya–Watson-type estimator. Asymptotic normality of the time varying correlation coefficient estimator is established. The asymptotic normality allows us to construct a valid standard error of the estimated local correlation coefficient, yielding a valid confidence interval. Also, based on the time varying correlation coefficient, a test is constructed for trend analysis. The test is for the equality of correlation coefficient at specific time points: for example, by selecting two time points of the first time point and the last time point for a specific interval, a t-type test can provide long-term trend analysis during the interval, which can be used for statistical analysis of the trend claimed by Syllignakis and Kouretas (2011) and others.

A Monte–Carlo experiment reveals finite sample validity of the time varying correlation coefficient: confidence interval has stable coverage and the proposed test has reasonable size and power. The time varying correlation coefficient methods by nonparametric method are applied to five sets of stock returns and are compared with the result by DCC method of Engle (2002). The data analysis gives confidence intervals and test results for trends in correlations. The comparison reveals that the proposed time varying correlation and the DCC correlation give different detailed interpretation for some real data sets, indicating a need for both methods.

The remaining of the paper is organized as follows. Section 2 defines the time varying correlation coefficient and discuss its estimation. Section 3 establishes asymptotic distribution. Section 4 develops statistical inference methods. Section 5 provides a finite sample Monte–Carlo simulation. Section 6 applies the time varying correlation coefficient methods to real data sets. Section 7 gives a conclusion.

2 Time varying correlation and its estimation

Let $(X_t, Y_t)^T,\ t = 1, 2 \ldots$ be a sequence of bivariate random vectors with locally varying mean vector $(\mu_{Xt}, \mu_{Yt})^T$ and variance matrix $\begin{pmatrix} \sigma_{Xt}^2 & \sigma_{Xt}\sigma_{Yt}\rho_t \\ \sigma_{Xt}\sigma_{Yt}\rho_t & \sigma_{Yt}^2 \end{pmatrix}$ so that the correlation coefficient ρ_t of X_t and Y_t may be time dependent. Let $U_t = (X_t^2, Y_t^2, X_t, Y_t, X_tY_t)^T,\ t = 1, 2, \ldots$ and $\mu_t = E[U_t] = (\mu_{X^2t}, \mu_{Y^2t}, \mu_{Xt}, \mu_{Yt}, \mu_{XYt})^T$. Given time $t = 1, 2, \ldots$, the correlation coefficient ρ_t between X_t and Y_t is given by

$$\rho_t = \frac{\mu_{XYt} - \mu_{Xt}\mu_{Yt}}{\sigma_{Xt}\sigma_{Yt}}, \quad \sigma_{Xt}^2 = \mu_{X^2t} - \mu_{Xt}^2, \quad \sigma_{Yt}^2 = \mu_{Y^2t} - \mu_{Yt}^2. \tag{1}$$

We assume that, for $t \in R^+$ other than $\{1, 2, \ldots\}$, μ_t is continuously well defined and hence so is ρ_t. The time varying correlation coefficient ρ_t is a measure of local relation between X and Y at a specific time t, which allows us to see the relation of X and Y in a neighborhood of t.

Assume a sample $\{(X_t, Y_t),\ t = 1, \ldots, N\}$ is given. The time varying correlation coefficient ρ_t is locally estimated by

$$\hat{\rho}_t(h) = \frac{\hat{\mu}_{XYt}(h) - \hat{\mu}_{Xt}(h)\hat{\mu}_{Yt}(h)}{\hat{\sigma}_{Xt}(h)\hat{\sigma}_{Yt}(h)}, \quad \hat{\sigma}_{Xt}^2(h) = \hat{\mu}_{X^2t}(h) - \hat{\mu}_{Xt}^2(h), \quad \hat{\sigma}_{Yt}^2(h) = \hat{\mu}_{Y^2t}(h) - \hat{\mu}_{Yt}^2(h), \tag{2}$$

where

$$\hat{\mu}_t(h) = (\hat{\mu}_{X^2t}(h), \hat{\mu}_{Y^2t}(h), \hat{\mu}_{Xt}(h), \hat{\mu}_{Yt}(h), \hat{\mu}_{XYt}(h))^T = \frac{\sum_{i=1}^N K\left(\frac{i-t}{h}\right)U_i}{\sum_{i=1}^N K\left(\frac{i-t}{h}\right)}, \quad t = 1, \ldots, N, \tag{3}$$

$K(\cdot)$ is a kernel function having compact support $[-1, 1]$ and h is a positive bandwidth.

The kernel estimation enables us to construct statistical methods, such as confidence interval and hypothesis testing for local correlation from tractable asymptotic normality of the kernel estimator, which however is not shared by the DCC method and the BEKK method as will be demonstrated in the example section. These points are useful in identifying the local correlation dynamics between financial asset returns whose relation depends on time.

2.1 Bandwidth selection

We need to specify a bandwidth h. We propose a method for bandwidth selection based on a cross-validation criterion widely used in kernel estimation of time-varying parameters, for example, by Wu et al. (1998), Wu and Chiang (2000), Opsomer et al. (2001), Cai et al. (2000a), Tian et al. (2005), Li et al. (2013) and many others.

Let h_M be a given upper bound of h. We try to choose the bandwidth h which minimize the mean squares of cross validation errors

$$CV(h) = \frac{1}{N}\sum_{t=1}^{N}\left(\hat{\rho}_{t(-t)}(h) - \frac{(X_t - \hat{\mu}_{X,t(-t)}(h))(Y_t - \hat{\mu}_{Y,t(-t)}(h))}{\hat{\sigma}_{X,t(-t)}(h)\hat{\sigma}_{Y,t(-t)}(h)}\right)^2, \tag{4}$$

over a range $[1, h_M]$ of h, where $\hat{\rho}_{t(-t)}(h), \hat{\mu}_{X,t(-t)}(h), \hat{\mu}_{Y,t(-t)}(h), \hat{\sigma}^2_{X,t(-t)}(h), \hat{\sigma}^2_{Y,t(-t)}(h)$ are the "leave-one-out" estimators of ρ_t, μ_{Xt}, μ_{Yt}, σ^2_{Xt}, σ^2_{Yt}, respectively, which are constructed from the kernel method (3) with t-th observation (X_t, Y_t) removed from the estimation data set. We consider the bandwidth obtained by the "elbow criterion" method of Salvador and Chan (2004). As h moves from 1 to h_M, either $CV(h)$ decreases for all h or not. In the latter case, $h_{opt} = argmin_{h\in(0,h_M]}CV(h)$ is taken to minimize $CV(h)$. In the former case of decreasing $CV(h)$, instead of considering $h_M = argmin_{h\in(0,h_M]}CV(h)$, optimal bandwidth $h_{opt} = argmin_{h_0\in[0,h_M]}RMSE_{h_0}$ is taken to minimize

$$RMSE_{h_0} = \frac{h_0 - 1}{h_M - 1}RMSE(L_{h_0}) + \frac{h_M - h_0}{h_M - 1}RMSE(R_{h_0}), \tag{5}$$

where $RMSE(L_{h_0})$ is the root mean squared errors (RMSE) obtained from the linear regression of $CV(h)$ on h for $h \in L_{h_0} = \{1, \dots, h_0\}$, and $RMSE(R_{h_0})$ is similarly defined with $R_{h_0} = \{h_0 + 1, \dots, h_M\}$, $h_0 = 3, \dots, h_M - 1$. We consider $h_M = 8N^{1/2}$. The elbow method provides h_{opt} at which the decrement of $CV(h)$ begins to get small. The proposed bandwidth method will give a good performance for $\hat{\rho}_t(h)$, as will be demonstrated in the Monte–Carlo section below.

3 Asymptotic normality

We prove asymptotic normality of $\hat{\rho}_t(h)$, which enables us to develop statistical inference for ρ_t such as confidence interval and hypothesis testing. We assume the following conditions (A1)–(A6):

(A1) μ_t has first and second order derivatives μ_t' and μ_t'', say, in a neighborhood of t for any $t \in R^+$.

(A2) Let $e_t = (\epsilon_{Xt}, \epsilon_{Yt}, \epsilon^2_{Xt} - 1, \epsilon^2_{Yt} - 1, \epsilon_{Xt}\epsilon_{Yt})^T$,

$$(\epsilon_{Xt}, \epsilon_{Yt}) = \left(\frac{X_t - \mu_{Xt}}{\sigma_{Xt}}, \frac{1}{\sqrt{1-\rho_t^2}} \left(\frac{Y_t - \mu_{Yt}}{\sigma_{Yt}} - \rho_t \frac{X_t - \mu_{Xt}}{\sigma_{Xt}} \right) \right).$$

e_t is a strict α-mixing stationary process with mixing coefficient α_j satisfying $\sum_{j=1}^{\infty} j^{\kappa} \alpha_j^{(\delta-2)/\delta} < \infty$ for some real numbers $\delta > 2$ and $\kappa > (\delta - 2)/\delta$.

(A3) $\sum_{l=-\infty}^{\infty} l||\Gamma_l|| < \infty$ and Γ^{∞} is positive definite, $||\cdot||$ denotes a matrix norm, where $\Gamma^{\infty} = \sum_{l=-\infty}^{\infty} \Gamma_l$ and $\Gamma_l = Cov(e_t, e_{t+l})$.

(A4) The kernel function $K(u)$ has a compact support on $[-1, 1]$ and is differentiable in a neighborhood of u for any $u \in (-1, 1)$. The first derivative $K'(u)$ is bounded.

(A5) The kernel function $K(\cdot)$ satisfies $\int K(u)du = 1, \int K^2(u)du < \infty, \int uK(u)du = 0$ and $\int u^2 K(u) < \infty$.

(A6) The bandwidth satisfies $h/N \to 0$ and $h \to \infty$.

Condition (A1) makes ρ_t be differentiable up to the second order and is generally assumed in the studies for nonparametric kernel estimator together with kernel conditions (A4) and (A5), see Wu et al. (1998) and Cai and Xu (2008) and many others. The α-mixing condition (A2) allows us to construct asymptotic normality of $\hat{\rho}_t(h)$ under weak dependence such as stationary ARMA-type serial correlation and the GARCH-type conditional heteroscedasticity (X_t, Y_t).

The kernel conditions (A4) and (A5) are standard ones for nonparametric kernel estimators and are satisfied for most commonly used kernels, such as uniform, triangle, Epanechnikov, quadratic kernel, etc. Condition (A6) makes asymptotic bias b_t be vanished and is satisfied for optimal bandwidth h_{opt} in (7) below.

In Theorems 3.2 and 3.4, we establish limiting normality of $\hat{\rho}_t(h)$ for $t = cN$, $0 < c < 1$ and for ($t \in [0, ch)$ or $t \in (N - ch, N]$, $0 < c < 1$), which are away from boundaries and near the boundaries, respectively.

Lemma 3.1 *Assume conditions* (A1)–(A6). *Then, for* $t = cN$, $c \in (0, 1)$ *fixed, as* $N \to \infty$, *we have*

$$\sqrt{h} D_{Mt}^{-1/2} \left[\hat{\mu}_t(h) - \mu_t - b_{Mt} + o\left(\frac{h^2}{N^2} \right) \right] \xrightarrow{\text{d}} N(0, I_5),$$

where I_p is $p \times p$ identity matrix, $p = 5$,

$$D_{Mt} = H_t^T \Gamma^{\infty} H_t \int_{-1}^{1} K^2(u)du, \quad b_{Mt} = \frac{h^2}{2N^2} \mu_t'' \int_{-1}^{1} u^2 K(u)du,$$

and

$$H_t = \begin{pmatrix} 2\mu_{Xt}\sigma_{Xt} & 2\mu_{Yt}\sigma_{Yt}\rho_t & \sigma_{Xt} & \sigma_{Yt}\rho_t & \mu_{Xt}\sigma_{Yt}\rho_t + \mu_{Yt}\sigma_{Xt} \\ 0 & 2\mu_{Yt}\sigma_{Yt}\sqrt{1-\rho_t^2} & 0 & \sigma_{Yt}\sqrt{1-\rho_t^2} & \mu_{Xt}\sigma_{Yt}\sqrt{1-\rho_t^2} \\ \sigma_{Xt}^2 & \sigma_{Yt}^2\rho_t^2 & 0 & 0 & \sigma_{Xt}\sigma_{Yt}\rho_t \\ 0 & \sigma_{Yt}^2(\rho_t^2-1) & 0 & 0 & 0 \\ 0 & 2\sigma_{Yt}^2\rho_t\sqrt{1-\rho_t^2} & 0 & 0 & \sqrt{1-\rho_t^2}\sigma_{Xt}\sigma_{Yt} \end{pmatrix}. \tag{6}$$

From Lemma 3.1, applying the multivariate δ-method, we get the asymptotic normality of $\hat{\rho}_t(h)$.

Theorem 3.2 *Assume conditions (A1)–(A6). Then, for $t = cN$, $c \in (0,1)$ fixed, as $N \to \infty$, we have*

$$\sqrt{\frac{h}{D_{Mt}^*}}\left[\hat{\rho}_t(h) - \rho_t - \frac{\partial \rho_t}{\partial \mu_t^T} b_{Mt} + o\left(\frac{h^2}{N^2}\right)\right] \xrightarrow{\mathrm{d}} N(0,1),$$

where $D_{Mt}^* = D_t^T D_{Mt} D_t$ and

$$D_t = \left(-\frac{1}{2}\frac{\mu_{XYt}-\mu_{Xt}\mu_{Yt}}{\sigma_{Xt}^3\sigma_{Yt}},\ -\frac{1}{2}\frac{\mu_{XYt}-\mu_{Xt}\mu_{Yt}}{\sigma_{Xt}\sigma_{Yt}^3},\ \frac{-\mu_{Yt}\sigma_{Xt}^2+\mu_{Xt}(\mu_{XYt}-\mu_{Xt}\mu_{Yt})}{\sigma_{Xt}^3\sigma_{Yt}},\ \frac{-\mu_{Xt}\sigma_{Yt}^2+\mu_{Yt}(\mu_{XYt}-\mu_{Xt}\mu_{Yt})}{\sigma_{Xt}\sigma_{Yt}^3},\ \frac{1}{\sigma_{Xt}\sigma_{Yt}}\right)^T.$$

From Theorem 3.2, for $t = cN$, $0 < c < 1$, the optimal bandwidth h_{opt} which minimizes the asymptotic mean squared error (AMSE) of $\hat{\rho}_t(h)$, $AMSE = b_{Mt}^2 + D_{Mt}^*/h$, is

$$h_{opt} = (ND_{Mt}^*/b_{Mt})^{1/5}. \tag{7}$$

However, this optimal bandwidth h_{opt} is difficult to implement owing to the "hard-to-estimate" term μ_t'' for b_{Mt}. Moreover, it depends on time t. One may apply the method of Ruppert et al. (1995) for estimation of μ_t'' and the resulting "plug-in" bandwidth. Instead, one may use the cross validation bandwidth h in Sect. 2 as a practical alternative to the optimal bandwidth h_{opt}.

We next derive the asymptotic distribution of $\hat{\rho}_t(h)$ near the boundaries. We consider lower boundary point $t_L = ch$ and upper boundary point $t_U = N - ch$, $0 < c < 1$ fixed. We first construct limiting normality of $\hat{\mu}_t(h)$ near the boundaries, which in turn yields the asymptotic normality for $\hat{\rho}_t(h)$ near the boundaries.

Lemma 3.3 *Assume conditions (A1)–(A6). Then, for $t = ch$ or $N - ch$, $0 < c < 1$ fixed, as $N \to \infty$, we have*

$$\sqrt{h}D_{Jt}^{-1/2}\left(\hat{\mu}_t(h) - \mu_t - b_{Jt} + o\left(\frac{h^2}{N^2}\right)\right) \xrightarrow{\mathrm{d}} N(0, I_5),$$

where $J = L$ if $t = ch$, $J = U$ if $t = N - ch$,

$$D_{Lt} = \left[\int_{-c}^{1} K(u)du\right]^{-2} \left[H_t^T \Gamma^\infty H_t \int_{-c}^{1} K^2(u)du\right], b_{Lt} = \left[\int_{-c}^{1} K(u)du\right]^{-1}$$
$$\left[\frac{h}{N}\mu_t' \int_{-c}^{1} uK(u)du + \frac{h^2}{2N^2}\mu_t'' \int_{-c}^{1} u^2K(u)du\right],$$
$$D_{Ut} = \left[\int_{-1}^{c} K(u)du\right]^{-2} \left[H_t^T \Gamma^\infty H_t \int_{-1}^{c} K^2(u)du\right], b_{Ut} = \left[\int_{-1}^{c} K(u)du\right]^{-1}$$
$$\left[\frac{h}{N}\mu_t' \int_{-1}^{c} uK(u)du + \frac{h^2}{2N^2}\mu_t'' \int_{-1}^{c} u^2K(u)du\right].$$

Theorem 3.4 *Assume conditions (A1)–(A6). Then, for $t = ch$ or $N - ch$, $0 < c < 1$ fixed, as $N \to \infty$, we have*

$$\sqrt{\frac{h}{D_{Jt}^*}}\left[\hat{\rho}_t(h) - \rho_t - \frac{\partial \rho_t}{\partial \mu_t^T} b_{Jt} + o\left(\frac{h^2}{N^2}\right)\right] \xrightarrow{\mathrm{d}} N(0, 1),$$

where $D_{Jt}^* = D_t^T D_{Jt} D_t$, $J = L$ if $t = ch$; $J = U$ if $t = N - ch$.

The asymptotic bias b_{Mt}, b_{Lt}, or b_{Ut} disappear under mild conditions such as the bandwidth condition $h = o(N)$ in (A6). Note that the asymptotic variance D_t^*, say, of $\sqrt{h}\hat{\rho}_t(h)$ depend on whether t is at the boundaries $[0, h)$, $(N - h, N]$ or not:

$$D_t^* = D_t^T D_{Lt} D_t \text{ for } t \in [0, h),\ D_t^* = D_t^T D_{Ut} D_t \text{ for } t \in (N - h, N],\ D_t^* = D_t^T D_{Mt} D_t \text{ for } t = cN,\ 0 < c < 1.$$

From Theorems 3.2 and 3.4, we can construct a consistent estimator $\hat{D}_t^*$ of asymptotic variance D_t^* of $\sqrt{h}\hat{\rho}_t(h)$ with the consistent estimator $\hat{\mu}_t(h)$ in place of μ_t. The standard error is

$$se(\hat{\rho}_t(h)) = \sqrt{\hat{D}_t^T \hat{D}_{Lt} \hat{D}_t / h} \text{ if } t < h;\ \sqrt{\hat{D}_t^T \hat{D}_{Ut} \hat{D}_t / h} \text{ if } t > N - h;\ \sqrt{\hat{D}_t^T \hat{D}_{Mt} \hat{D}_t / h} \text{ otherwise,} \tag{8}$$

where $\hat{D}_t$, $\hat{D}_{Lt}$, $\hat{D}_{Ut}$, $\hat{D}_{Mt}$ are obtained from D_t, D_{Lt}, D_{Ut}, D_{Mt} by replacing μ_t with the kernel estimator $\hat{\mu}_t(h)$ and Γ^∞ with

$$\hat{\Gamma}^\infty = \hat{\Gamma}_0 + \sum_{l=1}^{L}\left(1 - \frac{l}{L}\right)(\hat{\Gamma}_l + \hat{\Gamma}_l^T), \quad \hat{\Gamma}_l = \frac{1}{N}\sum_{t=1}^{N-l} \hat{e}_t \hat{e}_{t+l}^T, \quad \hat{e}_t = (\hat{H}_t^T)^{-1}(U_t - \hat{\mu}_t(h)), \tag{9}$$

L is a bandwidth. Note that $\hat{\Gamma}^\infty$ is a Bartlett-type consistent estimator of the long run variance $\Gamma^\infty = \lim_{N\to\infty} Var(\frac{1}{\sqrt{N}}\sum_{t=1}^{N} e_t)$. The estimator $\hat{D}_t^*$ is consistent for serially correlated and/or conditionally heteroscedastic samples under mild conditions because of consistency of $\hat{\mu}_t$.

We extend asymptotic normality of $\hat{\rho}_t(h)$ to joint asymptotic normality of $\hat{R}_{\{t_1,\dots,t_k\}}(h) = (\hat{\rho}_{t_1}(h), \dots, \hat{\rho}_{t_k}(h))^T$ at times $t_1 < \cdots < t_k$, $1 < k < N$ apart from each other.

Theorem 3.5 *Assume conditions (A1)–(A6). Then, for* $(t_1 = c_1N$ *or* $c_1h)$, $t_2 = c_2N, \dots, t_{k-1} = c_{k-1}N$, $(t_k = c_kN$ *or* $N - c_kh)$, $c_1, \dots, c_k \in (0,1)$ *fixed and distinct, as* $N \to \infty$, *we have*

$$\sqrt{h}D^{*-1/2}_{\{t_1,\dots,t_k\}}\left(\hat{R}_{\{t_1,\dots,t_k\}}(h) - R_{\{t_1,\dots,t_k\}} - B_{\{t_1,\dots,t_k\}} + o\left(\frac{h^2}{N^2}\right)\right) \xrightarrow{\mathrm{d}} N(0, I_k),$$

where $R_{\{t_1,\dots,t_k\}} = (\rho_{t_1}, \dots, \rho_{t_k})^T$, $D^*_{\{t_1,\dots,t_k\}} = diag(D^*_{t_1}, \dots, D^*_{t_k})$ *and* $B_{\{t_1,\dots,t_k\}} = \left(\frac{\partial \rho_{t_1}}{\partial \mu^T_{t_1}} b_{t_1}, \dots, \frac{\partial \rho_{t_k}}{\partial \mu^T_{t_k}} b_{t_k}\right)^T$.

Similarly to the time varying correlation estimator $\hat{\rho}_t(h)$, $\hat{R}_{\{t_1,\dots,t_k\}}(h)$ has asymptotic distribution with bias $B_{\{t_1,\dots,t_k\}}$ and variance $D^*_{\{t_1,\dots,t_k\}}/h$ depending on whether $\{t_1, t_k\}$ is near boundaries $[0, h)$, $(N-h, N]$ or not. Theorem 3.5 allows us to develop hypothesis test for varying correlation at several specific time points, as will be provided in Sect. 4. The choice $(t_1 = c_1N$ or $c_1h)$, $(t_k = c_kN$ or $N - c_kh)$ allow t_1 and t_k to be chosen near the boundaries and also to be chosen away from boundaries. Other time points $t_2, \dots, t_{k-1}$ are chosen away from boundaries because comparing more than two time points near the same boundary would rarely be of interest in practice.

4 Statistical inference

We develop confidence interval and hypothesis test for the time varying correlation coefficient.

4.1 Confidence interval

A pointwise $(1-\alpha)\%$ confidence interval of ρ_t for a given t is developed from Theorems 3.2 and 3.4 as

$$CI_t = (\hat{\rho}_t(h) + \Phi^{-1}(\alpha/2)se(\hat{\rho}_t(h)),\ \hat{\rho}_t(h) + \Phi^{-1}(1-\alpha/2)se(\hat{\rho}_t(h))), \quad \alpha \in (0,1). \tag{10}$$

The p value of $\hat{\rho}_t(h)$ is computed as $p = 2\Phi(-|\hat{\rho}_t(h)|/se(\hat{\rho}_t(h)))$, where $\Phi(\cdot)$ is the distribution function of standard normal distribution.

Equality test at specific time points

Let $t_1, \dots, t_k$ be k-time points of special interests. The null hypothesis of equal local correlation coefficients across the specific time points is

$$H_0 : \rho_{t_1} = \cdots = \rho_{t_k} \tag{11}$$

for fixed $t_1/N, \dots, t_k/N \in [0,1]$. We consider the Wald test statistic

$$W_{\{t_1,\ldots,t_k\}} = [C\hat{R}_{\{t_1,\ldots,t_k\}}(h)]^T[C(\hat{D}^*_{\{t_1,\ldots,t_k\}}/h)C^T]^{-1}[C\hat{R}_{\{t_1,\ldots,t_k\}}(h)],$$
$$C = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \\ 0 & 0 & 0 & \cdots & 1 & -1 \end{pmatrix}, \tag{12}$$

where C is $(k-1)\times k$, $\hat{D}^*_{\{t_1,\ldots,t_k\}} = diag(\hat{D}^*_{t_1},\ldots,\hat{D}^*_{t_k})$ and $t_1,\ldots,t_k$ are sufficiently apart from each other. Of special interest is the case of two time comparison in which $k=2$. Then $W_{\{t_1,t_2\}}$ is the square of a t-type test

$$T_{\{t_1,t_2\}} = \frac{\hat{\rho}_{t_1}(h) - \hat{\rho}_{t_2}(h)}{se(\hat{\rho}_{t_1}(h) - \hat{\rho}_{t_2}(h))},$$

where $se(\hat{\rho}_{t_1}(h) - \hat{\rho}_{t_2}(h)) = \sqrt{se^2(\hat{\rho}_{t_1}(h)) + se^2(\hat{\rho}_{t_2}(h))}$ denotes the standard error of $\hat{\rho}_{t_1}(h) - \hat{\rho}_{t_2}(h)$. By taking $k=2$, $t_1 = 1$ and $t_2 = N$, we can test long-term trend by $T_{\{1,N\}}$ whether local correlation is significantly increased (or decreased) over the entire period or not. The test $T_{\{1,N\}}$ can give statistical justification for the trend claims by Syllignakis and Kouretas (2011).

By Theorem 3.5, the Wald test $W_{\{t_1,\ldots,t_k\}}$ converges to a chi-square distribution with $k-1$ degrees of freedom under the null hypothesis, as stated in the following theorem.

Theorem 4.1 *Assume conditions (A1)–(A6). Then, for* $(t_1 = c_1N$ *or* $c_1h)$, $t_2 = c_2N$, $\ldots, t_{k-1} = c_{k-1}N$, $(t_k = c_kN$ *or* $N - c_kh)$, $c_1, c_2, \ldots, c_k \in (0,1)$ *fixed and distinct, as* $N \to \infty$, *under the null hypothesis* H_0,

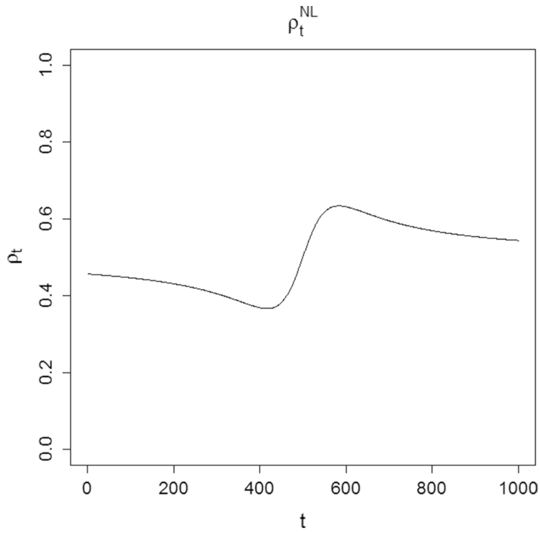
$$W_{\{t_1,\ldots,t_k\}} \xrightarrow{\text{d}} \chi^2_{k-1}.$$

5 Monte–Carlo experiment

We investigate finite sample coverage of $(1-\alpha) = 95\%$ confidence interval CI_t of ρ_t and finite sample sizes and powers of the level-α test $T_{\{t_1,t_2\}}$ proposed in Sect. 4.

The bivariate sample $(X_t, Y_t)^T = (\sigma_{Xt}\epsilon_{Xt}, \sigma_{Yt}(\rho_t\epsilon_{Xt} + \sqrt{1-\rho_t^2}\epsilon_{Yt}))^T$, $t = 1,\ldots,N$ is generated with iid errors $\epsilon_t = (\epsilon_{Xt}, \epsilon_{Yt})^T$ having zero mean and variance matrix $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, and with the variances σ^2_{Xt} and σ^2_{Yt} from GARCH(1,1) model, $\sigma^2_{Xt} = \alpha_0 + \alpha_1\epsilon^2_{X,t-1} + \beta_1\sigma^2_{X,t-1}$ and $\sigma^2_{Yt} = \alpha_0 + \alpha_1\epsilon^2_{Y,t-1} + \beta_1\sigma^2_{Y,t-1}$. The error ϵ_t is constructed from two independent errors generated from the standard normal distribution and the standardized t distribution, t_8, with 8 degrees of freedom. The correlation function ρ_t is specified to be either $\rho_t = \rho = 0.5, 0.8$ or $\rho_t^{NL} = 3.2(t/N - 0.5)/(1 + 144(t/N - 0.5^2)) + 0.5$, $0 \le t \le N$, which has a down-up-down shape as in Fig. 1. The normal and t_8 random variables are generated from

Fig. 1 Nonlinear function ρ_t^{NL}, $N = 1000$



R-package "rmvnorm" and "rstd". For the other parameters, we consider $\alpha_0 = 0.05$, $\alpha_1 = 0.1$, $\beta_1 = 0.85$.

The long run variance estimator $\hat{\Gamma}^\infty$ of $e_t = (H_t^T)^{-1}(U_t - \mu_t)$ in (9) is constructed with the Bartlett kernel and the optimal bandwidth parameter L_{And} of Andrews (1991),

$$L_{And} = 1.1447(4N\hat{\gamma}_1^2/(1 - \hat{\gamma}_1^2)^2)^{1/3},$$

where $\hat{\gamma}_1$ is the estimated AR(1) coefficient in the AR(1) fitting to $\{\sum_{j=1}^5 \hat{e}_{jt}, t = 1, \dots, N\}$, $\hat{e}_t = (\hat{e}_{1t}, \dots, \hat{e}_{5t})^T = (\hat{H}_t^T)^{-1}(U_t - \hat{\mu}_t)$, and $\hat{H}_t$ is obtained by replacing μ_t with $\hat{\mu}_t$ in (6).

Empirical coverage of the 95% confidence interval CI_t and rejection fraction of level 5% tests $T_{\{t_1,t_2\}}$ are computed from $M = 1000$ random samples of $\{(X_t, Y_t), t = 1, \dots, N\}$. We consider the Epanechnikov kernel. For each sample, the optimal bandwidth h_{opt} in kernel estimation is chosen by the method discussed at the later part of Sect. 2.

Table 1 provides the coverage of the CI_t of ρ_t, $t = 1, 0.1N, \dots, 0.9N, N$. Generally, the confidence interval has reasonable coverage even though somewhat under coverage. In the case of constant correlation $\rho = 0.5, 0.8$, CI_t has coverage reasonably close to the nominal coverage $1 - \alpha = 95\%$ for the normal distribution and has slight under-coverage for $N = 500, 1000$, but not for $N = 2000$ for the t_8 distribution. In the case of the nonlinear correlation ρ_t^{NL}, CI_t has reasonable coverages. Even for $t = 0.4N$, $0.6N$, at which ρ_t has sudden changes as shown in Fig. 1(a), coverage of CI_t for ρ_t^{NL} is not bad.

Therefore, we can say that, for all DGPs considered, the confidence interval CI_t of ρ_t has reasonable finite sample coverage. The confidence interval with coverage close to 95% demonstrates finite sample validity of the asymptotic normality of $\hat{\rho}_t(h)$ in Theorems 3.2 and 3.4.

Table 1 Empirical coverages (%) of the 95% confidence interval CI_t of time varying correlation coefficient

N	DGP		t										
	Error dist	ρ_t	1	0.1N	0.2N	0.3N	0.4N	0.5N	0.6N	0.7N	0.8N	0.9N	N
500	Normal	0.5	91.3	92.8	91.1	91.4	91.9	92.8	91.4	91.4	91.8	91.3	89.7
		0.8	89.6	92.6	91.3	90.7	91.7	92.8	91.7	92.9	93.5	92.9	90.4
		ρ_t^{NL}	91.3	93.0	90.4	90.6	90.2	90.2	88.3	91.5	91.7	91.4	89.8
	t_8	0.5	90.7	91.8	90.1	91.8	91.3	92.1	90.8	92.5	91.8	91.5	89.9
		0.8	89.0	91.5	91.7	92.1	91.4	90.3	90.7	90.7	91.8	91.1	89.1
		ρ_t^{NL}	91.5	92.1	91.2	91.6	91.2	90.0	86.6	92.6	91.5	91.0	90.1
1000	Normal	0.5	92.7	93.7	92.5	91.5	93.1	94.4	92.0	92.4	91.8	92.7	91.8
		0.8	92.3	93.6	93.0	93.9	94.7	95.5	94.0	92.7	93.1	93.9	93.1
		ρ_t^{NL}	92.4	94.1	92.7	92.2	92.9	92.9	92.0	92.4	92.3	92.8	91.4
	t_8	0.5	89.6	92.4	92.5	92.4	91.9	92.9	91.7	91.9	92.8	91.1	91.8
		0.8	91.3	91.4	92.0	91.1	90.9	92.3	92.0	90.8	91.7	90.4	91.1
		ρ_t^{NL}	90.1	92.4	92.5	92.6	91.9	91.0	90.8	91.4	92.4	90.5	91.5
2000	Normal	0.5	92.3	93.0	94.8	94.3	93.2	93.8	93.6	93.3	93.6	93.5	92.9
		0.8	92.5	93.5	95.1	94.7	94.2	94.3	94.3	94.0	94.8	94.7	93.2
		ρ_t^{NL}	91.9	93.6	94.9	94.7	93.2	93.9	93.4	93.8	94.1	93.0	93.3
	t_8	0.5	91.8	92.7	92.9	93.4	92.8	91.8	92.7	92.2	92.1	91.3	91.2
		0.8	90.1	92.8	92.6	93.2	92.5	92.3	93.8	92.8	93.1	92.7	90.3
		ρ_t^{NL}	92.1	92.2	92.9	93.7	93.1	91.6	93.2	93.5	92.1	91.9	90.9

Finite sample size and power of the test $T_{\{t_1,t_2\}}$ are next investigated. For the test $T_{\{t_1,t_2\}}$, we consider the two time points t_1 and t_2: both of the time points at the boundaries $(t_1, t_2) = (1, N)$; both of the time points away from boundaries $(t_1, t_2) = (3/8N, 5/8N)$; one time point at the boundary and the other one away from boundary $(t_1, t_2) = (1, 5/8N)$.

Rejection rates of the tests are computed from 1000 repetition and are displayed in Table 2. Consider first the result for the test $T_{\{t_1,t_2\}}$ for equality of correlation at the two time points. The test $T_{\{t_1,t_2\}}$ shows reasonable size for the time points t_1 and t_2 away from boundaries in all considered DGPs. In the cases of t_1 or/and t_2 at the boundary, the test $T_{\{t_1,t_2\}}$ tends to have minor oversize, which however improves as N increases from 500 and 2000. The power of the test $T_{\{t_1,t_2\}}$ increases under the alternative hypothesis of ρ_t^{NL} as N increases, indicating consistency of the test. The power difference among the three cases of (t_1, t_2) are due to differences of $|\rho_{t_1}^{NL} - \rho_{t_2}^{NL}|$ for the two time points (t_1,t_2), as shown in Fig. 1: $|\rho_{t_1}^{NL} - \rho_{t_2}^{NL}| \cong 0.1$ for $(t_1, t_2) = (1, N)$; $|\rho_{t_1}^{NL} - \rho_{t_2}^{NL}| \cong 0.25$ for $(t_1, t_2) = (3/8N, 5/8N)$; $|\rho_{t_1}^{NL} - \rho_{t_2}^{NL}| \cong 0.15$ for $(t_1, t_2) = (1, 5/8N)$.

Now, the results of Monte–Carlo simulation are summarized. The confidence interval CI_t of ρ_t and the proposed test $T_{\{t_1,t_2\}}$ have acceptable coverage, size and

Table 2 Sizes (%) and powers (%) of level 5% test $T_{\{t_1,t_2\}}$

N	Error dist	ρ_t		$T_{\{t_1,t_2\}}$ $(t_1,t_2)=(1,N)$	$(t_1,t_2)=(3/8N,5/8N)$	$(t_1,t_2)=(1,5/8N)$
500	Normal	0.5	Size	7.6	6.9	7.9
		0.8	Size	5.0	5.9	7.1
		ρ_t^{NL}	Power	12.4	50.4	23.6
	t_8	0.5	Size	8.2	7.8	9.3
		0.8	Size	5.4	7.6	7.2
		ρ_t^{NL}	Power	12.4	48.4	17.5
1000	Normal	0.5	Size	6.5	5.5	7.2
		0.8	Size	4.5	5.3	5.3
		ρ_t^{NL}	Power	12.1	66.5	27.3
	t_8	0.5	Size	7.3	7.2	8.6
		0.8	Size	4.8	6.3	7.0
		ρ_t^{NL}	Power	10.9	61.2	25.9
2000	Normal	0.5	Size	5.9	6.7	5.8
		0.8	Size	4.3	3.8	5.5
		ρ_t^{NL}	Power	13.4	71.7	32.2
	t_8	0.5	Size	6.6	6.8	4.1
		0.8	Size	6.5	6.5	7.2
		ρ_t^{NL}	Power	10.6	68.3	28.4

power performance, verifying finite sample validity of the asymptotic theory of the time varying correlation and of the proposed statistical methods.

6 Example

The proposed time varying correlation coefficient method is applied to the stock price return data sets of the US S&P 500 index, the US NASDAQ index, the United Kingdom FTSE 100 index, the Germany DAX index and the Japan Nikkei 225 index for the period of 01/04/2000–12/30/2017 with $n = 4080$, the number of days of the data sets. The stock price return data sets are obtained from the Oxford-Man realized library (https://realized.oxford-man.ox.ac.uk/). We examine the correlation dynamics between stock price returns over time by the nonparametric kernel method (2) and by the dynamic conditional correlation (DCC) method of Engle (2002). We also give statistical justification for some interesting findings for correlation dynamics for the periods of the major financial events considered by Lee and Chang (2013), Antonakakis and Vergos (2013) and Kang et al. (2017) and displayed in beige shades in Fig. 2: dot com bubble (01/04/2001–12/27/2001), global financial crisis (11/21/2007–05/22/2009) and European debt crisis (05/06/2011–05/30/2012).

Figure 2 shows the time varying correlation coefficients $\hat{\rho}_t(h)$ estimated from nonparametric method (2) and the conditional correlations $\hat{r}_t$ estimated from DCC model of Engle (2002). The dynamic correlation r_t is estimated by fitting the DCC model,

$$S_t = \begin{pmatrix} 1 & r_t \\ r_t & 1 \end{pmatrix}, S_t = (diag(Q_t))^{-1/2} Q_t (diag(Q_t))^{-1/2}, \tag{13}$$

$$Q_t = (1 - \delta_1 - \delta_2)\bar{Q} + \delta_1 (u_{t-1} u_{t-1}^T) + \delta_2 Q_{t-1}, \tag{14}$$

$$u_t = ((X_t - \hat{\mu}_X))/\hat{\sigma}_{Xt},\ (Y_t - \hat{\mu}_Y)/\hat{\sigma}_{Yt})^T = (u_{1t}, u_{2t})^T, \tag{15}$$

$$z_t = S_t^{-1/2} u_t \overset{\text{iid}}{\sim} N_2(0, I_2), \tag{16}$$

where $\bar{Q}$ is unconditional sample correlation matrix of $(u_{1t}, u_{2t})^T$, $N_2(0, I_2)$ is bivariate normal distribution with zero mean and 2×2 identity variance matrix, $diag(A)$ is the matrix with diagonal of A on the diagonal, zeros off-diagonal, $\hat{\mu}_X$ and $\hat{\mu}_Y$ are the sample mean of X_t and Y_t, $t = 1, \ldots, n$, and $\hat{\sigma}^2_{Xt}$ and $\hat{\sigma}^2_{Yt}$ are the variances estimated from the GARCH model,

$$\sigma^2_{Xt} = \alpha_{X0} + \alpha_{X1}(X_t - \hat{\mu}_X)^2 + \beta_{X1}\sigma^2_{X,t-1}, \quad \sigma^2_{Yt} = \alpha_{Y0} + \alpha_{Y1}(Y_t - \hat{\mu}_Y)^2 + \beta_{Y1}\sigma^2_{Y,t-1}.$$

Note that $r_t = E[u_{1,t-1}u_{2,t-1}|(X_s, Y_s),\ s = 1, \ldots, t-1]$ is a conditional correlation of X_t and Y_t under the assumption of constant mean of X_t and Y_t. For nonparametric estimation, we consider the Epanechnikov kernel is used and the optimal bandwidth h_{opt} is selected as h by the method discussed in Sect. 2.

The proposed time varying correlation $\hat{\rho}_t(h)$ shows diverse interesting points for the relationship between the stock price returns. For all the five plots, we see that the time varying correlation coefficient $\hat{\rho}_t(h)$ is more strongly fluctuating after the global financial crisis than before. The time varying correlation coefficient $\hat{\rho}_t(h)$ of the S&P 500 and the NASDAQ shows that the correlation $\hat{\rho}_t(h)$ between the US stock price indexes is stronger in the crisis times than in the boom times for the US financial market. The correlation $\hat{\rho}_t(h)$ between the S&P 500 and the FTSE 100 is stronger in the European debt crisis than in the global financial crisis, implying that the correlation $\hat{\rho}_t(h)$ between the US return index and the European return index is more influenced by the European financial shock than by the US financial shock. The correlation $\hat{\rho}_t(h)$ between the FTSE 100 and the DAX is stronger in the European debt crisis than in the global financial crisis. The Japan index return has a relatively weak relationship with both the S&P 500 and the FTSE 100 return.

The DCC correlation $\hat{r}_t$ shows movements similar to those of the proposed time varying correlation for some pairs but different for the other pairs. For the three pairs (S&P 500, FTSE 100), (FTSE 100, DAX), (FTSE 100, Nikkei), the DCC correlation $\hat{r}_t$ has the same long-term shapes as those of $\hat{\rho}_t(h)$ while having larger high frequency noise than $\hat{\rho}_t(h)$. For the other two pairs (S&P 500, Nikkei), (FTSE 100, Nikkei), $\hat{r}_t$ has very small changes over the time indicating no conditional correlation changes, while $\hat{\rho}_t(h)$ indicates some unconditional

Table 3 Estimated $\Delta\hat{\rho}$ and t test statistics $T_{\{1,N\}}$

Data		$\hat{\rho}_1(h)$	$\hat{\rho}_N(h)$	$\Delta\hat{\rho}$	$T_{\{1,N\}}$	p value
S&P 500	NASDAQ	0.69	0.83	0.14	1.70	0.09
S&P 500	FTSE 100	0.33	0.34	0.01	0.09	0.93
S&P 500	Nikkei 225	0.00	0.17	0.17	1.47	0.14
FTSE 100	DAX	0.69	0.47	− 0.22	− 1.72	0.09
FTSE 100	Nikkei 225	0.15	0.05	− 0.10	− 0.93	0.35

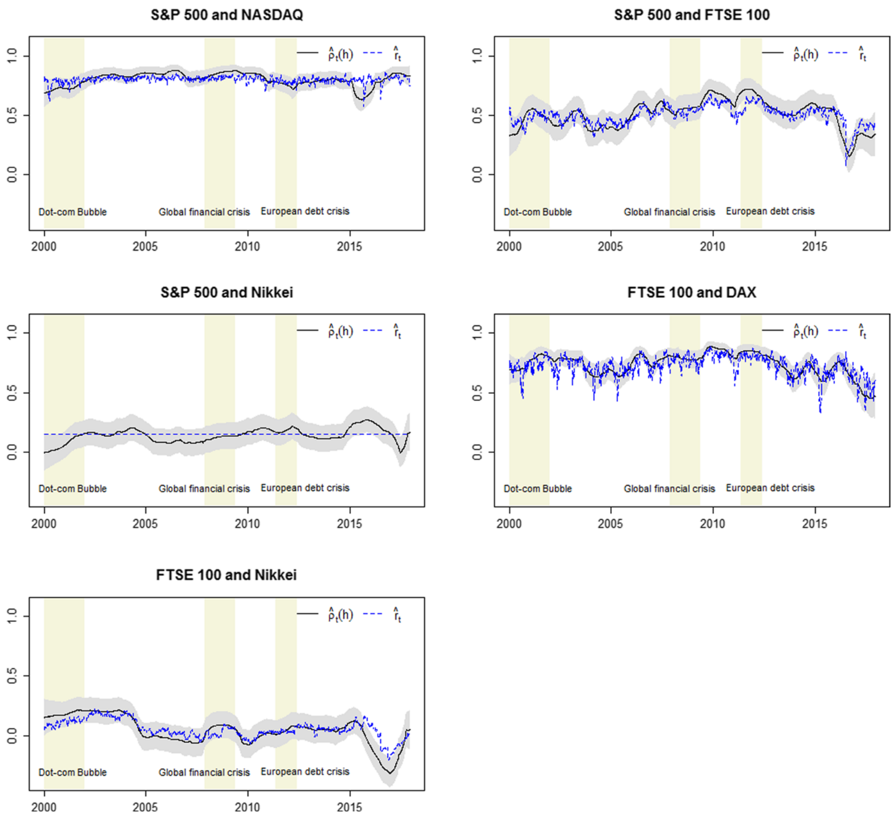


Fig. 2 Sample time varying correlations $\hat{\rho}_t(h)$, DCC correlation $\hat{r}_t$ and 95% confidence interval (CI) for $\hat{\rho}_t(h)$ of pairs of stock price returns

correlation changes. Especially, for the pair (S&P 500, Nikkei), $\hat{r}_t$ tells us no change in conditional correlation over time but $\hat{\rho}_t(h)$ indicates change in unconditional correlation. The reason for the no correlation change result for (S&P 500, Nikkei) is that DCC parameter estimators $(\hat{\delta}_1, \hat{\delta}_2)$ are (0.000, 0.994) indicating $\hat{r}_t \cong \hat{r}_{t-1} = 0.151$. This close to zero estimated value 0.000 of δ_1 and near-unit estimated value 0.994 of δ_2 is partly a consequence of sharply increased pairs $(u_{1t}, u_{2t}) = ((X_t - \hat{\mu}_X)/\hat{\sigma}_{Xt}, (Y_t - \hat{\mu}_Y)/\hat{\sigma}_{Yt}) = (-5.84, -7.33),\ (2.12, -9.65)$ on

06/24/2016, 11/09/2016. In a time-series plot of $u_{1t}u_{2t}$, the two cases look like additive outliers. We observe the study of Grane et al. (2014) that DCC estimators are sensitive to outlier. The two outlier cases affect the DCC r_t globally, making it be nearly constant all the time, while they affect $\hat{\rho}_t(h)$ locally making it have local peak and valley near the two times as revealed in Fig. 2. The findings indicate that the conditional correlation $\hat{r}_t$ and unconditional correlation $\hat{\rho}_t(h)$ may give different detailed interpretation for the real data sets and the proposed kernel method is a more robust method. Therefore both two methods need to be considered to grasp correlation movements over time.

We next make some statistical inference for correlation dynamics. Since the DCC correlation r_t is conditional correlation, unlike ρ_t, the DCC method does not allow us confidence interval and hypothesis testing. We therefore can continue on confidence interval analysis and hypothesis testing for correlation dynamics ρ_t only by the proposed kernel method. The confidence intervals of ρ_t in Fig. 2 tell us that the time varying correlation coefficients are statistical significant at 5% level for all period and for all pairs of data set, except for some periods for (FTSE 100, Nikkei) and (S&P 500, Nikkei): (01/04/2000–01/07/2000, 06/28/2004–12/22/2014, 05/15/2015–02/02/2016, 06/28/2017–12/04/2017), (01/04/2000–05/14/2001, 04/06/2005–03/18/2008, 06/28/2013–10/09/2013, 03/10/2017–10/24/2017), respectively. The findings indicates some insignificant correlation between Asia stock return and (European, US) stock returns during some periods.

We investigate whether correlations have trend during the whole period between 01/04/2000–12/30/2017. The trend is measured by $\Delta\hat{\rho}$, the difference between the first time varying correlation $\hat{\rho}_1(h)$ and the last time varying correlation $\hat{\rho}_N(h)$. Syllignakis and Kouretas (2011) reported some interesting findings regarding the trend $\Delta\hat{\rho}$ computed from a DCC analysis without a statistical significance analysis. For our data sets, $\Delta\hat{\rho}$ provides approximate long-term correlation trend over the 17 years of the sample period. Significance of $\Delta\hat{\rho}$ can be tested by $T_{\{1,N\}}$. Table 3 reports estimated $\Delta\hat{\rho}$ and its t test statistics. From the table, we observe that the correlation of the pair (S&P 500, NASDAQ) seems to have increased tendency (p value = 0.09) in the long term, and the pair (FTSE100, DAX) seems to have decreased tendency (p value = 0.09).

7 Conclusion

We have proposed a local estimator for time varying correlation coefficient based on a nonparametric kernel method. The local correlation coefficient estimation allows us to construct a confidence interval and a hypothesis test. We have established asymptotic normality for the estimated time varying correlation coefficient and the asymptotic null distribution of the proposed test. A Monte–Carlo simulation has demonstrated good performance of the proposed confidence interval and hypothesis test in terms of stable empirical coverage, size and reasonable power. The local time varying correlation coefficient method is applied to various pairs of stock price log returns: the US S&P500 index, the US NASDAQ index, the United Kingdom

FTSE100 index, the Germany DAX index and the Japan Nikkei 225 index. Some extensions may be good topics of future works. Extension to the time varying correlation at irregular or unsynchronized time points might be an interesting issue for further research. The proposed time varying correlation can also be extended to time varying correlation matrix for high-dimensional data.

Acknowledgements The authors are very thankful of two anonymous referees whose comments improved the paper considerably. This study was supported by a Grant from the National Research Foundation of Korea (2019R1A2C1004679) and by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2019R1A6A1A11051177).

Appendix: Proofs

For any function f_t with subscript $t \in [0, N]$, define function $f(\tau)$ of argument $\tau \in [0, 1]$ by $f(\tau) = f_t$, $\tau = \frac{t}{N}$. More specifically, for $\tau = \frac{1}{N}, \frac{2}{N}, \dots, \frac{N}{N}$, with $t = N\tau$, let $X(\tau) = X_t$, $Y(\tau) = Y_t$, $U(\tau) = U_t$. Then sample points $t = 1, 2, \dots, N$ are mapped to $\tau = \frac{1}{N}, \frac{2}{N}, \dots, \frac{N}{N} \in [0, 1]$ by $\tau = \frac{t}{N}$. Also for $\tau \in [0, 1]$, let $\mu(\tau) = \mu_t$, $\rho(\tau) = \rho_t$, $e(\tau) = e_t = U(\tau) - \mu(\tau)$, $b_M(\tau) = b_{Mt}$, $b_L(\tau) = b_{Lt}$, $D_M(\tau) = D_{Mt}$, $D_L(\tau) = D_{Lt}$, $D^*_M(\tau) = D^*_{Mt}$, $D(\tau) = D_t$, $\mu''(\tau) = \mu''_t$. With $\tau = \frac{t}{N}$ and $\hbar = \frac{h}{N}$, define $\hat{\rho}(\tau, \hbar) = \hat{\rho}_t(h)$, $\hat{\mu}(\tau, \hbar) = \hat{\mu}_t(h)$.

Proof of Lemma 3.1 Let $S_{KN}(\tau) = \sum_{j=1}^{N} K\left(\frac{j/N-\tau}{\hbar}\right)\frac{1}{N\hbar}$. Then, for any $\tau \in (0, 1)$, $S_{KN}(\tau) \to \int_{-1}^{1} K(u)du = 1$. Under assumptions (A2), (A5), from (2) we have, for $t \in [0, N]$, with $\tau = t/N \in [0, 1]$,

$$
\begin{aligned}
\hat{\mu}_t(h) - \mu_t &= \frac{1}{N\hbar}\left[\sum_{j=1}^{N} K\left(\frac{j/N-\tau}{\hbar}\right)\left(U\left(\frac{j}{N}\right) - \mu(\tau)\right)\right] / S_{KN}(\tau) \\
&= \frac{1}{N\hbar}\left[\sum_{j=1}^{N} K\left(\frac{j/N-\tau}{\hbar}\right)\left(\mu\left(\frac{j}{N}\right) - \mu(\tau)\right)\right] / S_{KN}(\tau) \\
&\quad + \frac{1}{N\hbar}\left[\sum_{j=1}^{N} K\left(\frac{j/N-\tau}{\hbar}\right) H^T\left(\frac{j}{N}\right) e\left(\frac{j}{N}\right)\right] / S_{KN}(\tau) \\
&= B_{1N}(\tau, \hbar) + B_{2N}(\tau, \hbar).
\end{aligned} \tag{17}
$$

We show

$$
E\left[\hat{\mu}_t(h) - \mu_t - b_{Mt} + o\left(\frac{h^2}{N^2}\right)\right] \to 0 \tag{18}
$$

and

$$
Var[\sqrt{h} D_{Mt}^{-1/2}(\hat{\mu}_t(h) - \mu_t)] \to I_5. \tag{19}
$$

Then, we observe the proof of Theorem 2 of Cai et al. (2000b) for asymptotic normality of their Q_n in their (A.7) under the strong α-mixing condition of (A2). If we replace $(n, X_i, c_0, c_1, Y_i - m(U_i, X_i))$ in Q_n by $(N, H_j, 1, 0, e_j)$, we have $Q_n = B_{2N}(\tau, \hbar)$ and the proof gives asymptotic normality of $B_{2N}(\tau, \hbar)$. Now, we get the desired result, as $N \to \infty$,

$$\sqrt{h} D_{Mt}^{-1/2}\left[\hat{\mu}_t(h) - \mu_t - b_{Mt} + o\left(\frac{h^2}{N^2}\right)\right] \xrightarrow{\mathrm{d}} N(0, I_5). \tag{20}$$

It remains to prove (18) and (19). We first show (18). Applying Taylor approximation, under assumptions (A1), (A2), (A6), we have

$$\begin{aligned} E[\hat{\mu}_t(h) - \mu_t] = B_{1N}(\tau, \hbar) &= \frac{1}{N\hbar}\left[\sum_{i=-N\hbar}^{N\hbar} K\left(\frac{i}{N\hbar}\right)\left(\mu\left(\tau + \frac{i}{N}\right) - \mu(\tau)\right)\right] / S_{KN}(\tau) \\ &= \frac{1}{N\hbar}\left[\sum_{i=-N\hbar}^{N\hbar} K\left(\frac{i}{N\hbar}\right)\left(\frac{i}{N\hbar}\mu'(\tau) + \frac{\hbar^2}{2}\left(\frac{i}{N\hbar}\right)^2 \mu''(\tau)\right) + o(\hbar^2)\right] / S_{KN}(\tau), \\ &= \left[\frac{\hbar^2}{2}\mu''(\tau)\int_{-1}^{1} u^2 K(u) du + o(\hbar^2)\right](1 + o(1)) = b_M(\tau) + o(\hbar^2), \end{aligned} \tag{21}$$

arriving at (18).

We next show (19). From (17), we have

$$\begin{aligned} & hVar[\hat{\mu}_t(h) - \mu_t] \\ &= N\hbar Var[\hat{\mu}(\tau, \hbar) - \mu(\tau)] = N\hbar Var[B_{2N}(\tau, h)] \\ &= \frac{1}{N\hbar} \sum_{i_1, i_2 \in \{-N\hbar, \ldots, N\hbar\}} K\left(\frac{i_1}{N\hbar}\right) H^T\left(\tau + \frac{i_1}{N}\right) \\ &\quad Cov\left[e\left(\tau + \frac{i_1}{N}\right), e\left(\tau + \frac{i_2}{N}\right)\right] H\left(\tau + \frac{i_2}{N}\right) K\left(\frac{i_2}{N\hbar}\right) / S_{KN}^2(\tau) \\ &= \frac{1}{N\hbar} \sum_{i=-N\hbar}^{N\hbar} \sum_{l=i-N\hbar}^{i+N\hbar} K\left(\frac{i}{N\hbar}\right) H^T\left(\tau + \frac{i}{N}\right) \Gamma_l H\left(\tau + \frac{i-l}{N}\right) K\left(\frac{i-l}{N\hbar}\right) / S_{KN}^2(\tau). \end{aligned} \tag{22}$$

Since $S_{KN}(\tau) \to 1$ and

$$\begin{aligned} & \frac{1}{N\hbar} \sum_{i=-N\hbar}^{N\hbar} \sum_{l=i-N\hbar}^{i+N\hbar} K\left(\frac{i}{N\hbar}\right) H^T\left(\tau + \frac{i}{N}\right) \Gamma_l H\left(\tau + \frac{i}{N}\right) K\left(\frac{i}{N\hbar}\right) \\ & \to H^T(\tau) \sum_{l=-\infty}^{\infty} \Gamma_l H(\tau) \int_{-1}^{1} K^2(u) du = D_M(\tau), \end{aligned}$$

we obtain the result (19) if we show

$$R_N(\tau,\hbar) = \frac{1}{N\hbar}\sum_{i=-N\hbar}^{N\hbar}\sum_{l=i-N\hbar}^{i+N\hbar} ||K\Big(\frac{i}{N\hbar}\Big)H^T\Big(\tau+\frac{i}{N}\Big)\Gamma_l \\ \Big\{H\Big(\tau+\frac{i-l}{N}\Big)K\Big(\frac{i-l}{N\hbar}\Big) - H\Big(\tau+\frac{i}{N}\Big)K\Big(\frac{i}{N\hbar}\Big)\Big\}|| \to 0. \tag{23}$$

By the mean value theorem and boundednesses of $H(\tau)$, $H'(\tau)$, $K(u)$, and $K'(u)$ from assumption (A1) and (A4), we have $||H(\tau+\frac{i-l}{N})K(\frac{i-l}{N\hbar}) - H(\tau+\frac{i}{N})K(\frac{i}{N\hbar})|| \le \frac{Ml}{N\hbar}$ for some $M<\infty$. Therefore,

$$\begin{aligned} &R_N(\tau,\hbar) \\ &\le (N\hbar)^{-1}\sum_{i=-N\hbar}^{N\hbar}\sum_{l=i-N\hbar}^{i+N\hbar} ||K\Big(\frac{i}{N\hbar}\Big)H^T\Big(\tau+\frac{i}{N}\Big)\Gamma_l|| ||H\Big(\tau+\frac{i-l}{N}\Big) \\ &K\Big(\frac{i-l}{N\hbar}\Big) - H\Big(\tau+\frac{i}{N}\Big)K\Big(\frac{i}{N\hbar}\Big)|| \\ &\le M(N\hbar)^{-1}\sum_{i=-N\hbar}^{N\hbar}\sum_{l=i-N\hbar}^{i+N\hbar} ||K\Big(\frac{i}{N\hbar}\Big)H^T\Big(\tau+\frac{i}{N}\Big)l\Gamma_l||\frac{1}{N\hbar} = O\Big(\frac{1}{N\hbar}\Big) = o(1), \text{ by (A3)}, \end{aligned}$$

where $||\cdot||$ is the matrix norm, and we obtain the desired result. □

Proof of Theorem 3.2 The result is derived easily by applying the multivariate δ -method to Lemma 3.1. □

Proof of Lemma 3.3 We first consider the left boundary point $\tau = c\hbar$, $0<c<1$. Note that $S_{KN}(c\hbar) \to \int_{-c}^{1} K(u)du$. From (17), for $0<c<1$, we have

$$\hat{\mu}(c\hbar,\hbar) - \mu(c\hbar) = B_{1N}(c\hbar,\hbar) + B_{2N}(c\hbar,\hbar).$$

We obtain the result from

$$\begin{aligned} &E[\hat{\mu}(c\hbar,\hbar) - \mu(c\hbar)] = B_{1N}(c\hbar,\hbar) \\ &= \left[\frac{1}{N\hbar}\sum_{i=-N\hbar c}^{N\hbar} K\Big(\frac{i}{N\hbar}\Big)\Big(\frac{i}{N\hbar}\mu'(c\hbar)\Big) + \frac{\hbar^2}{2}\Big(\frac{i}{N\hbar}\Big)^2\mu''(c\hbar)) + o(\hbar^2)\right] / S_{KN}(c\hbar) \\ &= \left[\int_{-c}^{1} K(u)du\right]^{-1}\left[\hbar\mu'(c\hbar)\int_{-c}^{1} uK(u)du + \frac{\hbar^2}{2}\mu''(c\hbar)\int_{-c}^{1} u^2K(u)du + o(\hbar^2)\right](1+o(1)) \\ &= b_L(\tau) + o(\hbar^2), \end{aligned}$$

and

$$\begin{aligned} &hVar[\hat{\mu}(c\hbar,\hbar) - \mu(c\hbar)] = N\hbar Var[B_{2N}(c\hbar,\hbar)] \\ &= \left[\frac{1}{N\hbar}\sum_{i=-Nc\hbar}^{N\hbar}\sum_{l=i-N\hbar}^{i+Nc\hbar} K(\frac{i}{N\hbar})H^T(c\hbar+\frac{i}{N})\Gamma_l H\Big(c\hbar+\frac{i}{N}\Big)K\Big(\frac{i}{N\hbar}\Big)\right] / S_{KN}^2(c\hbar) \\ &= \left[\int_{-c}^{1} K(u)du\right]^{-2}\left[\int_{-c}^{1} K^2(u)duH^T(c\hbar)\sum_{l=-\infty}^{\infty}\Gamma_l H(c\hbar)\right](1+o(1)) = D_L(\tau) + o(1). \end{aligned}$$

The result for the right boundary point $\tau = 1 - c\hbar$, $0 < c < 1$ is similarly obtained. □

Proof of Theorem 3.4 The result is derived easily by applying the multivariate δ -method to Lemma 3.3. □

Proof of Theorem 3.5 From Theorem 3.2, the theorem can be obtained by extending the asymptotic normality of one mean vector $\hat{\mu}(\tau, \hbar)$ to k mean vector $(\hat{\mu}(\tau_1, \hbar), \ldots, \hat{\mu}(\tau_k, \hbar))^T$ and by applying the multivariate δ-method, where $\tau = t/N$, $\tau_j = t_j/N,\ j = 1, \ldots, k$. Therefore, by Lemma 3.1 and Theorem 3.2, it suffices to show

$$Cov[\sqrt{h}(\hat{\mu}(\tau_1, \hbar) - \mu(\tau_1)), \sqrt{h}(\hat{\mu}(\tau_2, \hbar) - \mu(\tau_2))] = o(1).$$

Recall that $E[e(\tau)] = E[e(\tau + l)] = 0$ and $Var(e(\tau)) < \infty$ by assumption (A2). Assume that for some $\delta > 0$, $E|e_{\tau+l}|^{2+2\delta} < \infty$. Then, for both of the time points τ_1, τ_2 away from boundaries, we have

$$\begin{aligned}
&h||Cov[\hat{\mu}(\tau_1, \hbar) - \mu(\tau_1), \hat{\mu}(\tau_2, \hbar) - \mu(\tau_2)]|| = N\hbar||Cov[B_{2N}(\tau_1, \hbar), B_{2N}(\tau_2, \hbar)]|| \\
&\quad \le \frac{1}{N\hbar} \sum_{i_1, i_2 \in \{-N\hbar, \ldots, N\hbar\}} ||K\left(\frac{i_1}{N\hbar}\right) H^T\left(\tau_1 + \frac{i_1}{N}\right) Cov\left(e\left(\tau_1 + \frac{i_1}{N}\right), e(\tau_2 + \frac{i_2}{N})\right) \\
&\quad\quad \times H\left(\tau_2 + \frac{i_2}{N}\right) K\left(\frac{i_2}{N\hbar}\right)||/(S_{KN}(\tau_1) S_{KN}(\tau_2)) \\
&\quad = \frac{1}{N\hbar} \sum_{i=-2N\hbar}^{2N\hbar} \sum_{l=i-N\hbar}^{i+N\hbar} ||K\left(\frac{i}{N\hbar}\right) H^T\left(\tau + \frac{i}{N}\right) E\left[e\left(\tau + \frac{i}{N}\right) e\left(\tau + \tau_0 + \frac{i-l}{N}\right)\right] \\
&\quad\quad \times H\left(\tau + \tau_0 + \frac{i-l}{N}\right) K\left(\frac{i-l}{N\hbar}\right)||(1 + o(1)), \\
&\quad = \frac{1}{N\hbar} \sum_{i=-2N\hbar}^{2N\hbar} \sum_{l=i-N\hbar}^{i+N\hbar} ||K\left(\frac{i}{N\hbar}\right) H^T\left(\tau + \frac{i}{N}\right) E\left[e\left(\tau + \frac{i}{N}\right) e\left(\tau + \tau_0 + \frac{i-l}{N}\right)\right] \\
&\quad\quad \times H\left(\tau + \tau_0 + \frac{i}{N}\right) K\left(\frac{i}{N\hbar}\right)||(1 + o(1)), \quad \text{by the same argument of (23),} \\
&\le \frac{1}{N\hbar} \sum_{i=-2N\hbar}^{2N\hbar} \sum_{l=i-N\hbar}^{i+N\hbar} K(\frac{i}{N\hbar})||H^T(\tau + \frac{i}{N})||\{2(2^{1/2} + 1)\alpha_{N\tau_0 - l}^{\delta/(2+2\delta)}||Var(e(\tau + \frac{i}{N}))^{1/2}|| \\
&\quad \times E[||e_{t+l}||^{2+2\delta}]^{1/(2+2\delta)}\}||H(\tau + \tau_0 + \frac{i}{N})||K(\frac{i}{N\hbar})(1 + o(1)) \\
&\quad\quad = O(\sum_{l=-3N\hbar}^{3N\hbar} \alpha_{N\tau_0 - l}^{\delta/(2+2\delta)}) = o(1).
\end{aligned} \tag{24}$$

The inequalities in (24) is a consequence of Corollary 6.16 of White (White , p.148). The results for the time points τ_1 or/and τ_2 at the boundary can be derived in a similar way. □

Proof of Theorem 4.1 The result is derived easily by Theorem 3.5 for the joint asymptotic normality of $\hat{R}_{\{t_1,\ldots,t_k\}}(h)$. $\square$

References

Andrews, D. W. K. (1991). Heteroscedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, *59*, 817–854.

Antonakakis, N., & Vergos, K. (2013). Sovereign bond yield spillovers in the Euro zone during the financial and debt crisis. *Journal of International Financial Markets, Institutions and Money*, *26*, 258–272.

Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model. *The Review of Economics and Statistics*, *72*, 498–505.

Brownlees, C., & Engle, R. F. (2017). SRISK: A conditional capital shortfall measure of systemic risk. *The Review of Financial Studies*, *30*, 48–79.

Cai, Z., & Xu, X. (2008). Nonparametric quantile estimations for dynamic smooth coefficient models. *Journal of the American Statistical Association*, *103*, 1595–1608.

Cai, Z., Fan, J., & Li, R. (2000a). Efficient estimation and inferences for varying coefficient models. *Journal of the American Statistical Association*, *95*, 888–902.

Cai, Z., Fan, J., & Yao, Q. (2000b). Functional-coefficient regression models for nonlinear time series. *Journal of the American Statistical Association*, *95*, 941–956.

Cai, Z., Chen, L., & Fang, Y. (2015). Semiparametric estimation of partially varying-coefficient dynamic panel data models. *Econometric Reviews*, *34*, 695–719.

Choi, J. E., & Shin, D. W. (2018). Moving block bootstrapping for a CUSUM test for correlation change. *Computational Statistics and Data Analysis*, *20*, 10.

Engle, R. F. (2002). Dynamic conditional correlation. *Journal of Business and Economic Statistics*, *20*, 339–350.

Engle, R. F., & Kroner, K. F. (1995). Multivariate simultaneous generalized ARCH. *Econometric Theory*, *11*, 122–150.

Filis, G., Degiannakis, S., & Floros, C. (2011). Dynamic correlation between stock market and oil prices: The case of oil-importing and oil-exporting countries. *International Review of Financial Analysis*, *20*, 152–164.

Galeano, P., & Wied, D. (2014). Multiple break detection in the correlation structure of random variables. *Computational Statistics and Data Analysis*, *76*, 262–282.

Girardi, G., & Ergun, A. T. (2013). Systemic risk measurement: Multivariate GARCH estimation of CoVaR. *Journal of Banking and Finance*, *37*, 3169–3180.

Grane, A., Veiga, H., & Martin-Barragan, B. (2014). *Additive level outliers in multivariate GARCH models. Topics in statistical simulation* (Vol. 114, pp. 247–255). New York: Springer.

Jin, X., & Simone, F. A. N. (2014). Banking systemic vulnerabilities: A tail-risk dynamic CIMDO approach. *Journal of Financial Stability*, *14*, 81–101.

Kang, S. H., Mclver, R., & Yoon, S. M. (2017). Dynamic spillover effects among crude oil, precious metal, and agricultural commodity futures markets. *Energy Economics*, *2017*, 19–32.

Lacal, V., & Tjostheim, D. (2018). Estimating and testing nonlinear local dependence between two time series. *Journal of Business and Economic Statistics*,. https://doi.org/10.1080/07350015.2017.1407777.

Lee, H. C., & Chang, S. L. (2013). Spillovers of currency carry trade returns, market risk sentiment, and U.S. market returns. *The North American Journal of Economics and Finance*, *26*, 197–216.

Li, Q., Ouyang, D., & Racine, J. S. (2013). Categorical semiparametric varying-coefficient models. *Journal of Applied Econometrics*, *28*, 551–579.

Li, K., Li, D., Liang, Z., & Hsiao, C. (2017). Estimation of semi-varying coefficient models with nonstationary regressors. *Econometric Reviews*, *36*, 354–369.

Liu, X., Jiang, H., & Zhou, Y. (2014). Local empirical likelihood inference for varying coefficient density-ratio models based on case–control data. *Journal of the American Statistical Association*, *109*, 635–646.

Lucas, A., Schwaab, B., & Zhang, X. (2017). Modeling financial sector joint tail risk in the euro area. *Journal of Applied Econometrics*, *32*, 171–191.

Opsomer, J., Wang, Y., & Yang, Y. (2001). Nonparametric regression with correlated errors. *Statistical Science*, *16*, 134–153.

Ruppert, D., Sheather, S. J., & Wand, M. P. (1995). An effective bandwidth selector for local least squares regression. *Journal of the American Statistical Association*, *90*, 1257–1270.

Salvador, S., & Chan, P. (2004). Determining the number of clusters/segments in hierarchical clustering/segmentation algorithms. In *Proceedings of the 16th IEE international conference on tools with AI*, pp. 576–584.

Syllignakis, M. N., & Kouretas, G. P. (2011). Dynamic correlation analysis of financial contagion: Evidence from the Central and Eastern European markets. *International Review of Economics and Finance*, *20*, 717–732.

Tian, L., Zucker, D., & Wei, L. J. (2005). On the Cox model with time-varying regression coefficients. *Journal of the American Statistical Association*, *469*, 172–183.

Tjostheim, D., & Hufthammer, K. O. (2013). Local Gaussian correlation: A new measure of dependence. *Journal of Econometrics*, *172*, 33–48.

van der Weide, R. (2002). GO-GARCH: A multivariate generalized orthogonal GARCH model. *Journal of Applied Econometrics*, *17*, 549–564.

White, H. (1984). *Asymptotic theory for econometricians*. New York: Academic Press.

Wied, D. (2017). A nonparametric test for a constant correlation matrix. *Econometric Reviews*, *36*, 1157–1172.

Wied, D., Kramer, W., & Dehling, H. (2012). Testing for a chance in correlation at an unknown point in time using an extended functional delta method. *Econometric Theory*, *28*, 570–589.

Wu, C. O., & Chiang, C. T. (2000). Kernal smoothing on varying coefficient models with longitudinal dependent variable. *Statistica Sinica*, *10*, 433–456.

Wu, C. O., Chiang, C. T., & Hoover, D. R. (1998). Asymptotic confidence regions for kernel smoothing of a varying coefficient model with longitudinal data. *Journal of the American Statistical Association*, *93*, 1388–1402.