A practical multivariate approach to testing volatility spillover[⋆]

Soon Heng Leong[a,*], Giovanni Urga[b]

[a] *Finance Department, ESCP Business School, 527 Finchley Road, London NW3 7BG, UK*
[b] *Centre for Econometric Analysis, Faculty of Finance, Bayes Business School (formerly Cass), City, University of London, UK*

Abstract

We propose an asymptotic $\mathrm{N}(0,1)$ inferential strategy to test for volatility spillover between markets consisting of multiple sectors. First, we use nonparametric kernel method to derive test statistics that assign flexible weight to each lag order and are able to check a growing number of lags as the sample size increases. Second, we propose a practical multivariate volatility modeling approach — which enjoys estimation consistency and simplicity — to facilitate higher dimensional spillover testing. Simulations show the reasonable finite sample performance of the proposed econometric strategy in a relatively large system. An empirical application highlights the merits of the proposed approach.

Keywords: Granger causality in variance; Infinite autoregression; Multivariate analysis; Risk management; Volatility spillover.

JEL Classification: C12; C32; C51; F65.

1. Introduction

Volatility is undoubtedly one of the most informative risk indicators in financial economics as it is fundamentally related to, among others, market liquidity risk (Garbade and Silber, 1979), the interaction between informed and strategic traders (Admati and Pfleiderer, 1988), the rate of information flow to the market (Ross, 1989), revealed private information (Stoll and Whaley, 1990), and the degree of connection between markets as market participants infer information from each other (King and Wadhwani, 1990; Ellington, 2022). Consequently, a statistical tool that detects

[⋆]We are grateful to the Co-Editor Professor Juan Francisco Rubio-Ramirez and an anonymous referee for their comments and reviews. We also thank conference and seminar participants at ESCP Research Workshop (Paris, 8th May 2022) in particular, Paul Karehnke and Howard Zhong, Finance Department, ESCP Business School (Online, 15th June 2021) in particular Santiago Barraza, Panagiotis Dontis-Charitos, Paul Karehnke and Ellapulli Vasudevan, Cass Research Days 2021 (Online, 7th–8th June 2021) in particular Roy Batchelor, Meziane Lasfer and Richard Payne, School of Accounting and Finance, University of Bristol (Online, 10th May 2021) in particular Richard Harris, Guillem Ordonez-Calafi and Manuela Pedio, New York Camp Econometrics XV (Online, 10th–11th April 2021) in particular Kajal Lahiri, Department of Economics, City, University of London (Online, 10th March 2021) in particular Giulia Iori, Keith Pilbeam and Klaus Zauner, University of Nottingham Business School (Online, 5th March 2021) in particular Sanjay Banerji and Evangelos Benos, for useful comments and suggestions. Special thanks to Lynda Khalaf for helpful feedback on previous versions of the paper. Soon Leong acknowledges financial support from the Centre for Econometric Analysis and Bayes Business School including the "2017/2021 PhD Studentship in Memory of Ana Timberlake". The usual disclaimer applies.

[*]Corresponding author

Email addresses: sleong@escp.eu (Soon Heng Leong), G.Urga@city.ac.uk (Giovanni Urga)

volatility spillover between markets provides important decision making information about hedging and managing volatility risk in the commodity market (Sévi, 2014), the energy market (Fianu et al., 2022), as well as the foreign exchange market (Barunik et al., 2016).

Very often, academics and practitioners concern with testing volatility spillover between markets consisting of multiple major sectors. For instance, Bekiros et al. (2017) consider testing volatility spillover between US commodity market (characterized by the energy, metal and agricultural sectors) and US equity market (characterized by 10 major sectors such as financial, health, and technology). In this aspect, analyzing the overall spillover effect between the commodity and stock markets based on a weighted commodity index and a weighted equity index is inadequate because the univariate analysis does not take into account covariances between the individual sectors within each market. As highlighted in our simulation study, covariances between individual sectors can play a nontrivial role in driving overall spillover. In this paper, we propose a new econometric strategy for testing volatility spillover between markets that are characterized by multiple sectors. We begin by generalizing the univariate hypothesis of Hong (2001) to the multivariate setup. We follow the author to define volatility spillover using the notion of Granger (1969, 1980) causality in variance, in that there is volatility spillover from Y_2 to Y_1 if any of the past variances of Y_2 has explanatory power over the current variance of Y_1.[1] Therefore, the terms *variance causality* and *volatility spillover* may be used interchangeably. We derive testable statistics for our hypotheses using normalized cross-spectra and we develop the asymptotic theory. Our test statistics possess several appealing features. First, the computation of our test statistics is relatively straightforward because it requires only the estimation of standardized residuals which are the main event variables. Unlike most existing parameter restriction tests which estimate all series in a global model, our procedures allow the event variables of Y_1 and Y_2 to be estimated separately. Second, our tests check a large number of lag orders M as the sample size T increases. In fact, we allow M to grow with T at a proper rate to ensure power against a broad class of alternatives such as delayed spillover effect. Third, our frequency domain kernel-based procedure allows flexible weighting of the cross-spectrum at each lag order. We show that the conventional Granger-type regression procedure can be viewed as a special case of our approach when the Truncated kernel is used. Both tests assign equal weight to each lag. Instead, we propose to use downward weighting kernels to enhance the power of our tests because empirical stylized facts suggest that market participants discount past information and thus spillover effect is expected to decay over time. Indeed, simulation evidence shows that our downward weighting tests can check a large number of lags without losing significant power when compared with an equally weighted test.

Due to the practical limitations in estimating volatility as its dimension increases, the paper proposes a modeling approach that works coherently with the spillover test. The proposed model takes the form of a modified Bollerslev (1990) variance-covariance structure, where we specify the elements in the diagonal matrix as ARCH(∞) processes to minimize the risk of autocorrelated

[1] For other methodological contributions in Granger causality, see, for instance, Hong (1996); Bouhaddioui and Roy (2006); Hong et al. (2009); Candelon and Tokpavi (2016).

residuals resulting from model inadequacy (see, e.g., Li and Mak, 1994). This will in turn contaminate the resulting test statistics by invalidating its asymptotic property. Indeed, our supplementary empirical study confirms that serial correlation induced by an inadequate GARCH gives misleading inferential result. In this aspect, the proposed long ARCH(∞) process — which is sometimes referred to as a "nonparametric" approach — is more appealing than conditional variance models with a prespecified order. Regarding model estimation, we show that least-squares is feasible and consistent. The proposed nonparametric covariance modeling and least-squares estimation (NC-LS) approach is simple and complements existing literature. For instance, the proposed estimation method enjoys numerical efficiency as we find that it requires only about 5% of the computing time of QMLE in a computational study provided in the supplementary document. The speed advantage effectively makes bootstrap feasible to give a more accurate finite sample performance for the volatility spillover test. Besides, the proposed method allows consistent element-wise estimation of the volatility model that is not limited by certain distributional assumptions (see, e.g., Boudt et al., 2019, p. 217).

Our econometric strategy proceeds in two stages. First, we estimate the event variables by fitting the observed data using our proposed NC-LS approach. Second, we compute our test statistics to draw inference about volatility spillover. Throughout our econometric strategy, we need not perform numerical integration nor optimization. An extensive Monte Carlo study shows that our inferential strategy provides reliable finite sample inference even in the higher dimension up to the case of 10 series while the simulation evidence in most other papers is limited to 2–3 series. We further provide a consistent bootstrap test whose finite sample size is found to converge at a faster rate. We apply our inferential strategy in a multivariate study in which we investigate the distortion in volatility spillover relations between the North America (NA) market and the UK market before and after the Brexit referendum. For a broader study, we also examine the spillover effect on the European Union (EU) market which is represented by eight major economies. First, based on a series of diagnostic examinations we find that the proposed NC-LS approach shows adequacy in modeling and fitting the collected empirical data. Next, using the properly fitted data we apply the proposed multivariate test to examine volatility spillover. Our main findings show that compared with the pre-Brexit sample, the average spillover from UK to NA diminishes in the post-Brexit period. By contrast, we find that the spillover effect from EU to NA is relatively delayed before Brexit but the nexus becomes more immediate after Brexit. Because market indices are driven by the collective actions of the respective market participants, we could infer that on average — in contrast to the pre-Brexit sample — the NA participants appear not to react to the UK market but are following more closely the EU market in the post-Brexit period.

The remainder of this paper is organized as follows. In Section 2, we derive test statistics for the hypotheses of interest and we provide their asymptotic properties. Section 3 presents the nonparametric volatility model and its asymptotic validity. The finite sample performance of our econometric strategy is reported in Section 4 using a Monte Carlo study. In Section 5, we apply the proposed inferential strategy to empirically examine volatility spillover between the North American and European equity markets. Section 6 concludes. Mathematical derivations and proofs are relegated to the Appendices. Additional simulation and empirical results are collected in a supplementary document. Throughout the paper, $\xrightarrow{d}$ and $\xrightarrow{p}$ denote convergences in distribution and probability, respectively. The Euclidean norm is denoted by $||\cdot||$. Unless otherwise indicated, all limits are taken as the sample size $T \to \infty$.

2. Multivariate Granger causality in variance

In this section, we introduce the formal hypotheses for volatility spillover using Granger causality in variance. We then construct kernel-based test statistics for the hypotheses using the quadratic distance between two spectral densities. Finally, we provide the asymptotic properties of the proposed tests.

2.1. From univariate to multivariate Granger causality in variance

For two stationary time series Y_{1t} and Y_{2t}, let I_{1t} and I_{2t} denote the respective information set available at time t. Further let $I_t \equiv (I_{1t}, I_{2t})$ denote the combined information set. Following the definition of Granger (1980), Y_{2t} Granger causes Y_{1t} with respect to I_{t-1} if

$$\mathbb{E}(Y_{1t}|I_{1t-1}) \neq \mathbb{E}(Y_{1t}|I_{t-1}). \tag{1}$$

Granger (1969) introduces a regression-based test for (1), which can be viewed as the *causality in mean* hypothesis. We note that there are other definitions of Granger causality such as those based on projections on Hilbert spaces (see, e.g., Boudjellaba et al., 1992; Comte and Lieberman, 2000). This is not pursued.

Next, Granger et al. (1986) propose the notion of *causality in variance*, which is sometimes referred to as second order causality (see, e.g., Comte and Lieberman, 2000). Let $\mu_{it} \equiv \mathbb{E}(Y_{it}|I_{t-1})$, $i = 1, 2$, the causality in variance hypotheses are given by

$$\mathbb{H}_0 : \mathbb{E}\big[(Y_{1t} - \mu_{1t})^2 | I_{1t-1}\big] = \mathbb{E}\big[(Y_{1t} - \mu_{1t})^2 | I_{t-1}\big], \tag{2}$$

$$\mathbb{H}_A : \mathbb{E}\big[(Y_{1t} - \mu_{1t})^2 | I_{1t-1}\big] \neq \mathbb{E}\big[(Y_{1t} - \mu_{1t})^2 | I_{t-1}\big]. \tag{3}$$

Under the null hypothesis, the variance of Y_{1t} is not affected by I_{2t-1}, we say that Y_{2t} does not Granger cause Y_{1t} in variance. By construction, causality in mean has been filtered out because the hypotheses are not affected by causal relation in the mean equation. Therefore, any remaining causal effect is driven purely by volatility that is unaffected by mean and we follow Hong (2001) to use this information to test for *volatility spillover* from Y_{2t} to Y_{1t} in the higher dimension.

Let us now consider two stationary vectors of time series $(\boldsymbol{Y_{1t}}, \boldsymbol{Y_{2t}})$, where for $i = 1, 2$, $\boldsymbol{Y_{it}} = [Y_{it}(1), ..., Y_{it}(d_i)]'$, $d_i \in \mathbb{Z}^+ < \infty$. Let $\boldsymbol{I_{1t}}$ and $\boldsymbol{I_{2t}}$ denote the information set of $\boldsymbol{Y_{1t}}$ and $\boldsymbol{Y_{2t}}$ available at time t, respectively. The combined information set is denoted by $\boldsymbol{I_t} \equiv (\boldsymbol{I_{1t}}, \boldsymbol{I_{2t}})$. Further let $\boldsymbol{\epsilon_{it}} \equiv \boldsymbol{Y_{it}} - \mathbb{E}(\boldsymbol{Y_{it}}|\boldsymbol{I_{t-1}})$. We suppose that the demeaned series exhibit conditional heteroskedasticity

$$\boldsymbol{\epsilon_{it}} = (\boldsymbol{H^0_{it}})^{1/2}\boldsymbol{\Xi_{it}}, \tag{4}$$

where $\boldsymbol{H^0_{it}}$ is a $(d_i \times d_i)$ positive definite conditional variance-covariance matrix of $\boldsymbol{\epsilon_{it}}$, measurable with respect to $\boldsymbol{I_{it-1}}$. The innovation process $\boldsymbol{\Xi_{it}}$ is such that

$$\mathbb{E}(\boldsymbol{\Xi_{it}}|\boldsymbol{I_{it-1}}) = \boldsymbol{0} \;\text{ a.s.,} \quad \mathbb{E}(\boldsymbol{\Xi_{it}\Xi'_{it}}|\boldsymbol{I_{it-1}}) = \mathrm{I}_{\boldsymbol{di}} \;\text{ a.s.} \tag{5}$$

By generalization of (2)–(3), the multivariate causality in variance hypotheses are given by

$$\mathbb{H}^1_0 : \mathbb{E}\big[\boldsymbol{\Xi_{1t}\Xi'_{1t}}|\boldsymbol{I_{1t-1}}\big] = \mathbb{E}\big[\boldsymbol{\Xi_{1t}\Xi'_{1t}}|\boldsymbol{I_{t-1}}\big], \tag{6}$$

$$\mathbb{H}^1_A : \mathbb{E}\big[\boldsymbol{\Xi_{1t}\Xi'_{1t}}|\boldsymbol{I_{1t-1}}\big] \neq \mathbb{E}\big[\boldsymbol{\Xi_{1t}\Xi'_{1t}}|\boldsymbol{I_{t-1}}\big]. \tag{7}$$

Thus, we can test $\mathbb{H}^1_0$ by checking if $\boldsymbol{\Xi_{2t}\Xi'_{2t}}$ Granger causes $\boldsymbol{\Xi_{1t}\Xi'_{1t}}$ with respect to $\boldsymbol{I_{t-1}}$. In practice, the squared innovations $\boldsymbol{\Xi_{it}\Xi'_{it}}$ can be consistently estimated using the squared standardized residuals based on daily data. This also sets our framework apart from methods requiring intraday data, which may not be readily available to most institutions. Let $\boldsymbol{\theta^0_i}$ denote the true unknown parameters of $\boldsymbol{H^0_{it}}$. Given $\{\boldsymbol{\epsilon_t}\}^T_{t=1}$, where $\boldsymbol{\epsilon_t} = (\boldsymbol{\epsilon_{1t}}, \boldsymbol{\epsilon_{2t}})'$, let $\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}$ denote any $\sqrt{T}$-consistent estimator of $\boldsymbol{\theta^0_i}$, such that $\hat{\boldsymbol{H}}_{\boldsymbol{it}} = \boldsymbol{H_{it}}(\hat{\boldsymbol{\theta}}_{\boldsymbol{i}})$, where $\boldsymbol{H_{it}}$ is the pseudo version of $\boldsymbol{H^0_{it}}$ with initial value that is chosen arbitrarily. For notational simplicity, we further let $\hat{\boldsymbol{Z}}_{\boldsymbol{it}} \equiv \mathrm{vech}\big[(\hat{\boldsymbol{H}}_{\boldsymbol{it}})^{-1/2}\boldsymbol{\epsilon_{it}\epsilon'_{it}}(\hat{\boldsymbol{H}}_{\boldsymbol{it}})^{-1/2}\big]$, a column vector with d^*_i components, where $d^*_i = d_i(d_i + 1)/2$. The vector $\hat{\boldsymbol{Z}}_{\boldsymbol{it}}$ collects the squared standardized residuals and cross products of standardized residuals at time t. The event variables of interest are the centered version of $\hat{\boldsymbol{Z}}_{\boldsymbol{it}}$. We denote by

$$\hat{\boldsymbol{u}}_{\boldsymbol{t}} \equiv \boldsymbol{u_t}(\hat{\boldsymbol{\theta}}_{\boldsymbol{1}}) = \hat{\boldsymbol{Z}}_{\boldsymbol{1t}} - \mathrm{vech}(\mathrm{I}_{\boldsymbol{d1}}), \quad \hat{\boldsymbol{v}}_{\boldsymbol{t}} \equiv \boldsymbol{v_t}(\hat{\boldsymbol{\theta}}_{\boldsymbol{2}}) = \hat{\boldsymbol{Z}}_{\boldsymbol{2t}} - \mathrm{vech}(\mathrm{I}_{\boldsymbol{d2}}), \tag{8}$$

where $\mathrm{I}_{\boldsymbol{di}}$ is an identity matrix of dimension d_i. Similarly, we denote by $\boldsymbol{u^0_t}$ and $\boldsymbol{v^0_t}$ the pseudo version of the event variables based on the true volatility processes $\boldsymbol{H^0_{1t}}$ and $\boldsymbol{H^0_{2t}}$, respectively. Note that we do not require the two event variables to have the same dimension. Essentially, our spillover test allows the number of sectors to vary for the two markets of interest.

2.2. Test statistic

Using cross-spectrum analysis we now derive a test statistic for $\mathbb{H}^1_0$ that reduces to Hong's (2001) statistic when $d_1 = d_2 = 1$. The notion of cross-spectrum is closely related to the concept of Granger (1969) causality. To see the implication of $\mathbb{H}^1_0$ on the cross-spectrum between the event variables $\boldsymbol{u^0_t}$ and $\boldsymbol{v^0_t}$, we first note that the multivariate normalized cross-spectral density of $(\boldsymbol{u^0_t}, \boldsymbol{v^0_t})$ is given by

$$\boldsymbol{f}(\lambda) = \frac{1}{2\pi}\sum_{j=-\infty}^{\infty}\boldsymbol{\rho}(j)e^{-ij\lambda}, \quad \lambda \in [-\pi, \pi], \quad i = \sqrt{-1}, \tag{9}$$

where $\boldsymbol{\rho}(j) \equiv \mathrm{corr}(\boldsymbol{u^0_t}, \boldsymbol{v^0_{t-j}})$. Note that $\boldsymbol{\rho}(j)$ and $\boldsymbol{f}(\lambda)$ contain the same information about the cross-correlation between $\boldsymbol{u^0_t}$ and $\boldsymbol{v^0_{t-j}}$ since they are Fourier transforms of each other. We choose to use the frequency domain $\boldsymbol{f}(\lambda)$ for some desirable properties below. Under $\mathbb{H}_0^1$, we have $\boldsymbol{\rho}(j) = 0$, $\forall j > 0$. As a result, $\boldsymbol{f}(\lambda)$ reduces to

$$\boldsymbol{f}^{\mathbf{0}}(\lambda) = \frac{1}{2\pi} \sum_{j=-\infty}^{0} \boldsymbol{\rho}(j) e^{-ij\lambda}. \tag{10}$$

Therefore, we can test $\mathbb{H}_0^1$ by quantifying the difference between the observed density $\boldsymbol{f}(\lambda)$ and the null density $\boldsymbol{f}^{\mathbf{0}}(\lambda)$ using a proper divergence measure such as the quadratic norm. Any nontrivial deviation between $\boldsymbol{f}(\lambda)$ and $\boldsymbol{f}^{\mathbf{0}}(\lambda)$ is evidence against the null hypothesis.

The true cross-spectra $\boldsymbol{f}(\lambda)$ and $\boldsymbol{f}^{\mathbf{0}}(\lambda)$ are not known but they can be estimated consistently using nonparametric methods. Empirically, return series exhibit the volatility clustering characteristic as a volatile period tends to be followed by another volatile period. This is because markets are generally more influenced by recent information than remote information. Consequently, the magnitude of any economic movement, including volatility spillover, is expected to decay over time. We thus consider the kernel estimator which allows for flexible weighting at each lag order

$$\hat{\boldsymbol{f}}(\lambda) = \frac{1}{2\pi} \sum_{j=-T+1}^{T-1} k(j/M) \hat{\boldsymbol{\rho}}(j) e^{-ij\lambda}, \tag{11}$$

$$\hat{\boldsymbol{f}}^{\mathbf{0}}(\lambda) = \frac{1}{2\pi} \sum_{j=-T+1}^{0} k(j/M) \hat{\boldsymbol{\rho}}(j) e^{-ij\lambda}, \tag{12}$$

where $k(\cdot)$ is a kernel function and M is a truncation point when the kernel is bounded, or a smoothing parameter when the kernel has unbounded supports. The sample cross-correlation matrix $\hat{\boldsymbol{\rho}}(j)$ is given by

$$\hat{\boldsymbol{\rho}}(j) = \mathrm{Diag}\big(\hat{\boldsymbol{C}}_{\boldsymbol{uu}}\big)^{-1/2} \hat{\boldsymbol{C}}_{\boldsymbol{uv}}(j) \mathrm{Diag}\big(\hat{\boldsymbol{C}}_{\boldsymbol{vv}}\big)^{-1/2}, \tag{13}$$

where $\hat{\boldsymbol{C}}_{\boldsymbol{uv}}(j)$ is the sample cross-covariance matrix that is given by

$$\hat{\boldsymbol{C}}_{\boldsymbol{uv}}(j) = \begin{cases} \dfrac{1}{T} \displaystyle\sum_{t=j+1}^{T} \hat{\boldsymbol{u}}_{\boldsymbol{t}} \hat{\boldsymbol{v}}'_{\boldsymbol{t-j}}, & j \geq 0, \\ \dfrac{1}{T} \displaystyle\sum_{t=-j+1}^{T} \hat{\boldsymbol{u}}_{\boldsymbol{t+j}} \hat{\boldsymbol{v}}'_{\boldsymbol{t}}, & j < 0, \end{cases} \tag{14}$$

and $\hat{\boldsymbol{C}}_{\boldsymbol{uu}}$ and $\hat{\boldsymbol{C}}_{\boldsymbol{vv}}$ are the sample covariance matrices of $\hat{\boldsymbol{u}}_{\boldsymbol{t}}$ and $\hat{\boldsymbol{v}}_{\boldsymbol{t}}$, respectively. The function $\mathrm{Diag}(\cdot)$ returns a diagonal matrix consisting of the diagonal elements of the original matrix. Note that $\hat{\boldsymbol{\rho}}(j)$ is a matrix of dimension $(d_1^* \times d_2^*)$.

Recently, Robbins and Fisher (2015) propose a distance measure based on the Toeplitz matrix, though, positive definiteness of the measure cannot be guaranteed and some forms of correction are needed. Instead, we construct our test statistic using the quadratic distance between $\hat{\boldsymbol{f}}(\lambda)$ and $\hat{\boldsymbol{f}}^{\mathbf{0}}(\lambda)$ as with Duchesne and Roy (2004) for improved tractability. The distance measure $\hat{L}^2\big[\hat{\boldsymbol{f}}(\lambda), \hat{\boldsymbol{f}}^{\mathbf{0}}(\lambda)\big]$ is such that $\hat{L}^2\big[\hat{\boldsymbol{f}}(\lambda), \hat{\boldsymbol{f}}^{\mathbf{0}}(\lambda)\big] \geq 0$, and $\hat{L}^2\big[\hat{\boldsymbol{f}}(\lambda), \hat{\boldsymbol{f}}^{\mathbf{0}}(\lambda)\big] = 0$ if and only if $\hat{\boldsymbol{f}}(\lambda) = \hat{\boldsymbol{f}}^{\mathbf{0}}(\lambda)$. We use the quadratic form

$$\hat{L}^2\big[\hat{\boldsymbol{f}}(\lambda), \hat{\boldsymbol{f}}^{\mathbf{0}}(\lambda)\big] \equiv 2\pi \int_{2\pi} \mathrm{vec}\Big[\overline{\hat{\boldsymbol{f}}(\lambda)} - \overline{\hat{\boldsymbol{f}}^{\mathbf{0}}(\lambda)}\Big]' \big(\hat{\mathbf{\Gamma}}_{\boldsymbol{v}}^{-1} \otimes \hat{\mathbf{\Gamma}}_{\boldsymbol{u}}^{-1}\big) \mathrm{vec}\Big[\hat{\boldsymbol{f}}(\lambda) - \hat{\boldsymbol{f}}^{\mathbf{0}}(\lambda)\Big] d\lambda$$
$$= \sum_{j=1}^{T-1} k^2(j/M) \mathrm{vec}\big[\hat{\boldsymbol{\rho}}(j)\big]' \big(\hat{\mathbf{\Gamma}}_{\boldsymbol{v}}^{-1} \otimes \hat{\mathbf{\Gamma}}_{\boldsymbol{u}}^{-1}\big) \mathrm{vec}\big[\hat{\boldsymbol{\rho}}(j)\big], \tag{15}$$

where $\overline{\boldsymbol{f}(\cdot)}$ denotes the complex conjugate of $\boldsymbol{f}(\cdot)$, $\hat{\mathbf{\Gamma}}_{\boldsymbol{u}}$ is the sample correlation matrix of $\hat{\boldsymbol{u}}_{\boldsymbol{t}}$ to the unobserved correlation $\mathbf{\Gamma}_{\boldsymbol{u}}$, and $\hat{\mathbf{\Gamma}}_{\boldsymbol{v}}$ is the sample correlation matrix of $\hat{\boldsymbol{v}}_{\boldsymbol{t}}$ to the unobserved correlation $\mathbf{\Gamma}_{\boldsymbol{v}}$. The equality follows from Paserval's theorem. As a result, numerical integration over λ is not required in terms of the computation of $\hat{L}^2\big[\hat{\boldsymbol{f}}(\lambda), \hat{\boldsymbol{f}}^{\mathbf{0}}(\lambda)\big]$. The derivation of (15) is provided in Appendix A. Compared with Duchesne and Roy (2004), we do not integrate over the angular frequency. We allow for the case where $d_1 \neq d_2$. Besides, the authors work with the unstandardized version of spectral density. As a result, their test is based on covariances rather than correlations. We show in the analysis of Proposition B.2 in Appendix B that there is a cross-covariance representation of (15)

$$\hat{L}^2\big[\hat{\boldsymbol{f}}(\lambda), \hat{\boldsymbol{f}}^{\mathbf{0}}(\lambda)\big] = \sum_{j=1}^{T-1} k^2(j/M) \mathrm{vec}\big[\hat{\boldsymbol{C}}_{\boldsymbol{uv}}(j)\big]' \big(\hat{\boldsymbol{C}}_{\boldsymbol{vv}}^{-1} \otimes \hat{\boldsymbol{C}}_{\boldsymbol{uu}}^{-1}\big) \mathrm{vec}\big[\hat{\boldsymbol{C}}_{\boldsymbol{uv}}(j)\big]. \tag{16}$$

Despite the equivalence, we choose to construct our test statistics using the standardized spectral density so that our measure naturally reduces to that of Hong's (2001) when $d_1 = d_2 = 1$. The proposed test statistic, Q_1, is essentially the centered and scaled version of (15)

$$Q_1 = \frac{T \sum_{j=1}^{T-1} k^2(j/M) \mathrm{vec}\big[\hat{\boldsymbol{\rho}}(j)\big]' \big(\hat{\mathbf{\Gamma}}_{\boldsymbol{v}}^{-1} \otimes \hat{\mathbf{\Gamma}}_{\boldsymbol{u}}^{-1}\big) \mathrm{vec}\big[\hat{\boldsymbol{\rho}}(j)\big] - d_1^* d_2^* C_{1T}(k)}{\big[d_1^* d_2^* D_{1T}(k)\big]^{1/2}}, \tag{17}$$

where $C_{1T}(k)$ and $D_{1T}(k)$ are approximately the centering and scaling factors

$$C_{1T}(k) = \sum_{j=1}^{T-1} (1 - j/T) k^2(j/M), \tag{18}$$
$$D_{1T}(k) = 2 \sum_{j=1}^{T-1} (1 - j/T)[1 - (j+1)/T] k^4(j/M). \tag{19}$$

The constants $C_{1T}(k)$ and $D_{1T}(k)$ are readily computable given $k(\cdot)$ and M. Under some conditions on $k(\cdot)$ and by letting M goes to infinity properly with T, we have $M^{-1}C_{1T}(k) \to \int_0^\infty k^2(z)\mathrm{d}z$ and $M^{-1}D_{1T}(k) \to 2\int_0^\infty k^4(z)\mathrm{d}z$. As a result, $C_{1T}(k)$ and $D_{1T}(k)$ can be replaced, respectively, by $M\int_0^\infty k^2(z)\mathrm{d}z$ and $2M\int_0^\infty k^4(z)\mathrm{d}z$ without affecting the asymptotic properties of Q_1.

2.3. Asymptotic theory

We now establish the asymptotic properties of Q_1. Let $\tilde{\boldsymbol{H}}_{\boldsymbol{it}}$ denote the pseudo version of $\boldsymbol{H_{it}}$ with the true unobserved initial value. Note that $\tilde{\boldsymbol{H}}_{\boldsymbol{it}}(\boldsymbol{\theta_i^0}) = \boldsymbol{H_{it}^0}$, but $\boldsymbol{H_{it}}(\boldsymbol{\theta_i^0}) \neq \boldsymbol{H_{it}^0}$ due to the initial value. As a result, $\tilde{\boldsymbol{u}}_{\boldsymbol{t}}(\boldsymbol{\theta_1^0}) = \boldsymbol{u_t^0}$ and $\tilde{\boldsymbol{v}}_{\boldsymbol{t}}(\boldsymbol{\theta_2^0}) = \boldsymbol{v_t^0}$, but $\boldsymbol{u_t}(\boldsymbol{\theta_1^0}) \neq \boldsymbol{u_t^0}$ and $\boldsymbol{v_t}(\boldsymbol{\theta_2^0}) \neq \boldsymbol{v_t^0}$. This discrepancy is induced by likelihood-based estimation and is properly addressed in the following. To begin with, we present some regularity conditions under the model described by (4)–(5).

Assumption 2.1. *For $i = 1, 2$, $\{\boldsymbol{\Xi_{it}}\}$ is multivariate independent and identically distributed sequence with $\mathbb{E}(\boldsymbol{\Xi_{it}}) = \boldsymbol{0}$, $\mathbb{E}(\boldsymbol{\Xi_{it}}\boldsymbol{\Xi_{it}'}) = I_{\boldsymbol{di}}$ and finite eighth order moment. Besides, $\{\boldsymbol{\Xi_{1t}}\}$ and $\{\boldsymbol{\Xi_{2t}}\}$ are mutually independent under the null hypothesis.*

Assumption 2.2. *For $i = 1, 2$, $\sqrt{T}(\hat{\boldsymbol{\theta}}_{\boldsymbol{i}} - \boldsymbol{\theta_i^0}) = O_p(1)$, $\boldsymbol{\theta_i^0} \in \boldsymbol{\Theta_i}$.*

Assumption 2.3. *For each $\boldsymbol{\theta_i} \in \boldsymbol{\Theta_i}$, $i = 1, 2$, $\sup_{\boldsymbol{\theta_1} \in \boldsymbol{\Theta_1}} T \sum_{t=1}^{T} \mathbb{E}||\boldsymbol{u_t}(\boldsymbol{\theta_1}) - \tilde{\boldsymbol{u}}_{\boldsymbol{t}}(\boldsymbol{\theta_1})||^2 = O(1)$, $\sup_{\boldsymbol{\theta_2} \in \boldsymbol{\Theta_2}} T \sum_{t=1}^{T} \mathbb{E}||\boldsymbol{v_t}(\boldsymbol{\theta_2}) - \tilde{\boldsymbol{v}}_{\boldsymbol{t}}(\boldsymbol{\theta_2})||^2 = O(1)$.*

Assumption 2.4. *Let $\nabla_{\boldsymbol{\theta_i}}$ and $\nabla^2_{\boldsymbol{\theta_i}}$ denote, respectively, the gradient and Hessian operators w.r.t. $\boldsymbol{\theta_i}$. Then, $\sup_{\boldsymbol{\theta_1} \in \boldsymbol{\Theta_1}} T^{-1} \sum_{t=1}^{T} \mathbb{E}\left|\left|\nabla_{\boldsymbol{\theta_1}} \tilde{\boldsymbol{u}}_{\boldsymbol{t}}(\boldsymbol{\theta_1})\right|\right|^2 = O(1)$, $\sup_{\boldsymbol{\theta_2} \in \boldsymbol{\Theta_2}} T^{-1} \sum_{t=1}^{T} \mathbb{E}\left|\left|\nabla_{\boldsymbol{\theta_2}} \tilde{\boldsymbol{v}}_{\boldsymbol{t}}(\boldsymbol{\theta_2})\right|\right|^2 = O(1)$, $\sup_{\boldsymbol{\theta_1} \in \boldsymbol{\Theta_1}} T^{-1} \sum_{t=1}^{T} \mathbb{E}\left|\left|\nabla^2_{\boldsymbol{\theta_1}} \tilde{\boldsymbol{u}}_{\boldsymbol{t}}(\boldsymbol{\theta_1})\right|\right|^2 = O(1)$, $\sup_{\boldsymbol{\theta_2} \in \boldsymbol{\Theta_2}} T^{-1} \sum_{t=1}^{T} \mathbb{E}\left|\left|\nabla^2_{\boldsymbol{\theta_2}} \tilde{\boldsymbol{v}}_{\boldsymbol{t}}(\boldsymbol{\theta_2})\right|\right|^2 = O(1)$.*

Assumption 2.5. *The kernel $k : \mathbb{R} \to [-1, 1]$ is symmetric about 0, and is continuous at 0 and at all points except for a finite number of points, with $k(0) = 1$ and $\int_0^\infty k^2(z)dz < \infty$.*

Assumption 2.6. *$M \equiv M(T)$, and $M/T \to 0$ as $M \to \infty$ and $T \to \infty$.*

Assumption 2.1 is the multivariate generalization of the conditions required in Hong (2001). The required moment condition is in line with literature concerning the analysis of variance. For instance, McCloud and Hong (2011) work with the same moment condition in the context of testing structural changes in multivariate GARCH models. The i.i.d. condition on $\boldsymbol{\Xi_{it}}$ corresponds to the "strong GARCH" process defined in Hafner (2008) which is frequently used for estimation and inference in practice. Under this condition, Chan et al. (2007) derive the limiting distribution of the value-at-risk estimate in a GARCH process while Gao and Song (2008) extend the relevant works to cover the expected shortfall estimate. In this paper, the i.i.d. assumption ensures condition (5) that $\mathbb{E}(\boldsymbol{\Xi_{it}}|\boldsymbol{I_{it-1}}) = \boldsymbol{0}$ a.s. and $\mathbb{E}(\boldsymbol{\Xi_{it}}\boldsymbol{\Xi_{it}'}|\boldsymbol{I_{it-1}}) = \mathrm{I}_{\boldsymbol{di}}$ a.s., and it also reduces the complexity of the asymptotic analysis.[2] In Assumption 2.2, we do not impose any estimation restriction. Specifically, we allow for any $\sqrt{T}$-consistent estimator $\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}$. Assumption 2.3 requires that the initial condition of the variance-covariance process is asymptotically negligible. In particular, we require that the difference between $\boldsymbol{u_t}(\boldsymbol{\theta_1})$ and $\tilde{\boldsymbol{u}}_{\boldsymbol{t}}(\boldsymbol{\theta_1})$ goes to zero in probability at proper speed. Note that Assumption 2.3 becomes redundant under our nonparametric volatility specification discussed in Section 3. Assumption 2.4 requires that the event variables are twice continuously differentiable,

[2] We have explored the martingale condition under $\mathbb{H}_0^1$, but we find that it is not feasible without imposing more restrictive moment conditions.

with bounded derivatives. Assumption 2.5 is a standard regularity condition on the kernel function $k(\cdot)$. Most kernels used in spectral analysis satisfy this condition (see, e.g., Priestley, 1981; Andrews, 1991). Assumption 2.6 requires that M goes to infinity as T increases, at a speed slower than T. Finally, we have thus far assume that the demeaned series $\boldsymbol{\epsilon_{it}}$ is observable for ease of exposition, but it can be replaced by any $\sqrt{T}$-consistent estimate without affecting the asymptotic properties of Q_1 given Assumptions 2.1–2.6.

We now state the asymptotic normality of Q_1 under $\mathbb{H}_0^1$.

Theorem 2.1. *Suppose Assumptions 2.1–2.6 hold under the model described by (4)–(5). Then, $Q_1 \xrightarrow{d} \mathrm{N}(0,1)$ under $\mathbb{H}_0^1$.*

The proof of Theorem 2.1 is reported in Appendix B. When the Truncated kernel is used to compute Q_1, our test can be viewed as the Granger (1969)-type procedure. To see the intuition, we first note that the Truncated kernel is given by $k(z) = \mathbb{1}(|z| \leq 1)$, where $\mathbb{1}(\cdot)$ is the indicator function. For the purpose of illustration, suppose $d_1 = 1$ and $d_2 = 2$, we have the following test statistic[3]

$$Q_{1\mathrm{TR}} = \left\{ T \sum_{j=1}^{M} \mathrm{vec}\big[\hat{\boldsymbol{\rho}}(j)\big]' \big(\hat{\boldsymbol{\Gamma}}_{\boldsymbol{u}}^{-1}\big) \mathrm{vec}\big[\hat{\boldsymbol{\rho}}(j)\big] - 3M \right\} \Big/ (6M)^{1/2}, \tag{20}$$

where $\hat{\boldsymbol{\rho}}(j)$ is a (1×3) vector and $\hat{\boldsymbol{\Gamma}}_{\boldsymbol{u}}^{-1}$ is a (3×3) matrix. On the other hand, the Granger-type procedure is based on the following regression

$$\hat{u}_t = \phi_0 + \sum_{j=1}^{M} \boldsymbol{\phi_j} \hat{\boldsymbol{v}}_{\boldsymbol{t-j}} + w_t, \tag{21}$$

which checks whether the (1×3) parameter vector $\{\boldsymbol{\phi_j}\}_{j=1}^M$ are jointly zero. We do not have to include in the auxiliary regression (21) the lagged variables of $\hat{u}_t$ given Assumption 2.1. There is evidence that $\hat{\boldsymbol{v}}_{\boldsymbol{t}}$ Granger causes $\hat{u}_t$ with respect to $\boldsymbol{I_{t-1}}$ if at least one coefficient in $\{\boldsymbol{\phi_j}\}_{j=1}^M$ is significantly different from zero. A typical test statistic GR for this hypothesis obtained from, for instance, the Wald's procedure is asymptotically $\chi^2(3M)$ under $\mathbb{H}_0^1$ (see, e.g., Bauer and Maynard, 2012, Theorem 1). Now, for $Q_{1\mathrm{TR}}$ in (20), we know that under $\mathbb{H}_0^1$, $\sqrt{T}\mathrm{vec}\big[\hat{\boldsymbol{\rho}}(j)\big]$ generally converges to a three dimensional zero mean normal distribution at each lag j. Then, $\sum_{j=1}^M T\mathrm{vec}\big[\hat{\boldsymbol{\rho}}(j)\big]'\big(\hat{\boldsymbol{\Gamma}}_{\boldsymbol{u}}^{-1}\big)\mathrm{vec}\big[\hat{\boldsymbol{\rho}}(j)\big]$ being the M sum of the properly standardized independent $\chi^2(3)$ quantity is also asymptotically $\chi^2(3M)$ under $\mathbb{H}_0^1$. To ensure power of the Granger regression-based test against a large class of alternatives such as delayed spillover, we allow M to grow with the sample size T properly. Using the well-known approximation of $\chi^2(3M)$ when M is large, we obtain the asymptotic normality of GR and $\sum_{j=1}^M T\mathrm{vec}\big[\hat{\boldsymbol{\rho}}(j)\big]'\big(\hat{\boldsymbol{\Gamma}}_{\boldsymbol{u}}^{-1}\big)\mathrm{vec}\big[\hat{\boldsymbol{\rho}}(j)\big]$. With proper transformations, we have under $\mathbb{H}_0^1$, $Q_{1\mathrm{REG}} \equiv (GR - 3M)/(6M)^{1/2} \xrightarrow{d} \mathrm{N}(0,1)$ as well as

[3] Given the Truncated kernel function, we have $C_{1T}(k) = M[1-(1+M)/(2T)]$ and $D_{1T}(k) = 2M[1-(2+M)/T + (M+1)(M+2)/(3T^2)]$. Using a more stringent condition on M such that $M^{3/2}/T = o(1)$, we can conveniently approximate $C_{1T}(k)$ and $D_{1T}(k)$ by M and $2M$, respectively.

$Q_{1\mathrm{TR}} = \{T\sum_{j=1}^{M}\mathrm{vec}[\hat{\boldsymbol{\rho}}(j)'(\hat{\boldsymbol{\Gamma}}_{\boldsymbol{u}}^{-1})\mathrm{vec}[\hat{\boldsymbol{\rho}}(j)] - 3M\}/(6M)^{1/2} \xrightarrow{d} \mathrm{N}(0,1)$.

When M is large, both $Q_{1\mathrm{REG}}$ and $Q_{1\mathrm{TR}}$ may not yield a good power against the alternatives of practical importance. Given that market participants tend to discount past information, the effect of volatility spillover will fade as lag order j increases. Therefore, we propose to use downward weighting kernels such as the Bartlett, Daniell and Quadratic-Spectral kernels to increase the power performance of our Q_1 test. See Section 4 for more discussion and the Monte Carlo study.

To investigate the asymptotic behavior of Q_1 under the alternative hypothesis, we impose a condition on the cross-correlation $\boldsymbol{\rho}(j)$ and a fourth order cumulant condition.

Assumption 2.7. *The two event variables $\boldsymbol{u_t^0}$ and $\boldsymbol{v_t^0}$ are jointly fourth order stationary and their cross-correlation structure is such that $\boldsymbol{\rho}(j) \neq 0$ for at least one $j > 0$ and*

$$\sum_{j=1}^{\infty} ||\boldsymbol{\rho}(j)||^2 < \infty, \quad \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\sum_{l=1}^{\infty} |\kappa_{rsrs}(i,j,l)| < \infty,$$

where $\kappa_{rsrs}(i,j,l)$ is the fourth order cumulant of the distribution of $u_{r,t}^0$, $v_{s,t-i}^0$, $u_{r,t-j}^0$, $v_{s,t-l}^0$, with $r \in \{1,...,d_1^*\}$ and $s \in \{1,...,d_2^*\}$.

The condition $\sum_{j=1}^{\infty} ||\boldsymbol{\rho}(j)||^2 < \infty$ implies that the dependence of $\boldsymbol{u_t^0}$ on $\boldsymbol{v_{t-j}^0}$ decays to zero at a proper speed. However, it still permits a pair of highly cross-dependent processes whose cross-correlation decays to zero at a gradual hyperbolic rate. The cumulant condition is trivially satisfied if the joint process $\{\boldsymbol{u_t^0}, \boldsymbol{v_t^0}\}$ is Gaussian which implies zero fourth order cumulants. Fourth order stationary linear processes with absolutely summable coefficients and with innovations whose fourth order moment exists, also satisfy the cumulant condition (Hannan, 1970, p. 211).

The following theorem states the consistency of Q_1.

Theorem 2.2. *Suppose Assumptions 2.1–2.7 hold under the model described by (4)–(5). Then*

$$\frac{M^{1/2}}{T} Q_1 \xrightarrow{p} \left[2\int_0^{\infty} k^4(z)\mathrm{d}z\right]^{-1/2} \sum_{j=1}^{\infty} \mathrm{vec}\big[\boldsymbol{\rho}(j)\big]'\big(\boldsymbol{\Gamma}_{\boldsymbol{v}}^{-1} \otimes \boldsymbol{\Gamma}_{\boldsymbol{u}}^{-1}\big)\mathrm{vec}\big[\boldsymbol{\rho}(j)\big].$$

The proof of Theorem 2.2 is reported in Appendix C. Theorem 2.2 implies that Q_1 goes to infinity at rate $T/M^{1/2}$ provided $||\boldsymbol{\rho}(j)|| \neq 0$ for any $j > 0$. In the limit, negative values of Q_1 can only occur under $\mathbb{H}_0^1$. Therefore, Q_1 is a one-sided test; upper-tailed N(0,1) critical values should be used. Besides, the faster T grows, the quicker Q_1 will approach infinity and the test will become more powerful. In other words, Q_1 has asymptotic unit power against any linear pairwise volatility spillover. However, it should be noted that Q_1 has no power against the alternatives with zero cross-correlation for all values of $j > 0$, that is, $||\boldsymbol{\rho}(j)|| = 0, \forall j > 0$, though we expect such highly nonlinear alternatives to be empirically rare in economics and finance.

2.4. Bidirectional Granger causality in variance

The proposed Q_1 test is readily extendable to testing bilateral variance causality. We consider the multivariate version of Hong's (2001) bidirectional hypothesis that neither $\boldsymbol{Y_{2t}}$ causes $\boldsymbol{Y_{1t}}$ in variance with respect to $(\boldsymbol{I_{1t}}, \boldsymbol{I_{2t-1}})$ nor $\boldsymbol{Y_{1t}}$ causes $\boldsymbol{Y_{2t}}$ in variance with respect to $(\boldsymbol{I_{1t-1}}, \boldsymbol{I_{2t}})$. This extension is convenient when the existence of spillover relation between $\boldsymbol{Y_{1t}}$ and $\boldsymbol{Y_{2t}}$ is not known nor supported empirically. Essentially, we examine the following bidirectional hypotheses

$$\mathbb{H}_0^2 : \mathbb{E}\big[\boldsymbol{\Xi_{it}}\boldsymbol{\Xi'_{it}}|\boldsymbol{I_{it-1}}\big] = \mathbb{E}\big[\boldsymbol{\Xi_{it}}\boldsymbol{\Xi'_{it}}|\boldsymbol{I_{it}}, \boldsymbol{I_{jt-1}}\big], \quad i = j = 1, 2,\ i \neq j, \tag{22}$$

$$\mathbb{H}_A^2 : \mathbb{E}\big[\boldsymbol{\Xi_{it}}\boldsymbol{\Xi'_{it}}|\boldsymbol{I_{it-1}}\big] \neq \mathbb{E}\big[\boldsymbol{\Xi_{it}}\boldsymbol{\Xi'_{it}}|\boldsymbol{I_{it}}, \boldsymbol{I_{jt-1}}\big], \quad \text{for at least one } i,\ i = j = 1, 2,\ i \neq j. \tag{23}$$

This set of hypotheses can be viewed as complete Granger causality because it checks for non Granger causality between $\boldsymbol{Y_{1t}}$ and $\boldsymbol{Y_{2t}}$ as well as no instantaneous causality between $\boldsymbol{Y_{1t}}$ and $\boldsymbol{Y_{2t}}$. Under $\mathbb{H}_0^2$, we have $\boldsymbol{\rho}(j) = 0$, $\forall j$. As a result, the cross-spectrum $\boldsymbol{f}(\lambda)$ reduces to zero. The normalized quadratic distance between the kernel-based spectral density estimator and the null spectral density is given by

$$\hat{L}_2^2\big[\hat{\boldsymbol{f}}(\lambda), \hat{\boldsymbol{f}}^{\boldsymbol{0}}(\lambda)\big] = \sum_{j=-T+1}^{T-1} k^2(j/M)\text{vec}\big[\hat{\boldsymbol{\rho}}(j)\big]'\big(\hat{\boldsymbol{\Gamma}}_{\boldsymbol{v}}^{-1} \otimes \hat{\boldsymbol{\Gamma}}_{\boldsymbol{u}}^{-1}\big)\text{vec}\big[\hat{\boldsymbol{\rho}}(j)\big]. \tag{24}$$

The proposed bidirectional test statistic Q_2 is a centered and scaled version of $\hat{L}_2^2\big[\hat{\boldsymbol{f}}(\lambda), \hat{\boldsymbol{f}}^{\boldsymbol{0}}(\lambda)\big]$

$$Q_2 = \frac{T\sum_{j=-T+1}^{T-1} k^2(j/M)\text{vec}\big[\hat{\boldsymbol{\rho}}(j)\big]'\big(\hat{\boldsymbol{\Gamma}}_{\boldsymbol{v}}^{-1} \otimes \hat{\boldsymbol{\Gamma}}_{\boldsymbol{u}}^{-1}\big)\text{vec}\big[\hat{\boldsymbol{\rho}}(j)\big] - d_1^* d_2^* C_{2T}(k)}{\big[d_1^* d_2^* D_{2T}(k)\big]^{1/2}}, \tag{25}$$

where $C_{2T}(k)$ and $D_{2T}(k)$ are the bidirectional centering and scaling factors with

$$C_{2T}(k) = \sum_{j=-T+1}^{T-1} (1 - |j|/T)k^2(j/M), \tag{26}$$

$$D_{2T}(k) = 2\sum_{j=-T+1}^{T-1} (1 - |j|/T)[1 - (|j| + 1)/T]k^4(j/M). \tag{27}$$

Similar to Q_1, the bidirectional test statistic Q_2 converges in distribution to N(0,1) under the bilateral null hypothesis and it has asymptotic unit power whenever $||\boldsymbol{\rho}(j)|| \neq 0$ for at least one j. Likewise, upper-tailed N(0,1) critical values should be used for Q_2. The mathematical proof involved is similar to that of Theorems 2.1 and 2.2 by considering both positive and negative lag order j's, and we refrain from repeating the details here. In summary, when prior knowledge about the spillover relation is not available, one may first test the bidirectional hypothesis that neither $\boldsymbol{Y_{2t}}$ Granger causes $\boldsymbol{Y_{1t}}$ nor $\boldsymbol{Y_{1t}}$ Granger causes $\boldsymbol{Y_{2t}}$ in variance completely.

3. Estimation

The proposed Q_1 and Q_2 multivariate tests depend on $\boldsymbol{H_{it}^0}$ which has to be specified. In practice, modeling multivariate volatility is challenging as we face complications such as convergence, dimensionality and negative definiteness issues. Besides, to be coherent with the proposed spillover test, the volatility structure should also show adequacy in modeling empirical data. We now put forward a modeling approach which intersects these features in a practical manner. We suppose that $\boldsymbol{H_{it}^0}$ can be decomposed as follows

$$\boldsymbol{H_{it}^0} = (\boldsymbol{D_{it}^0})^{1/2} \boldsymbol{R_i^0} (\boldsymbol{D_{it}^0})^{1/2}, \text{ for } i = 1, 2, \tag{28}$$

where $\boldsymbol{D_{it}^0} = \text{diag}(h_{i,1t}^0, ..., h_{i,d_it}^0)$ is a diagonal matrix with univariate conditional variances and $\boldsymbol{R_i^0} = \mathbb{E}[(\boldsymbol{D_{it}^0})^{-1/2} \boldsymbol{\epsilon_{it}} \boldsymbol{\epsilon_{it}'} (\boldsymbol{D_{it}^0})^{-1/2}]$ is the covariance matrix of the vector of element-wise standardized residuals by construction. Under this structure, $\boldsymbol{H_{it}^0}$ is positive semidefinite provided that the elements in $\boldsymbol{D_{it}^0}$ are nonnegative and that $\boldsymbol{R_i^0}$ is positive semidefinite.

We specify $\boldsymbol{D_{it}^0}$ as an infinite order ARCH process. For $n = 1, .., d_i$, we denote the n-th elements in $\boldsymbol{D_{it}^0}$ and $\boldsymbol{\epsilon_{it}}$ by $h_{i,n,t}^0$ and $\epsilon_{i,n,t}$, respectively. The ARCH(∞) representation takes the form

$$h_{i,n,t}^0 = \omega_{i,n}^0 + \sum_{j=1}^{\infty} a_{i,n,j}^0 \epsilon_{i,n,t-j}^2. \tag{29}$$

This general process includes Engle's (1982) ARCH(q), Bollerslev's (1986) GARCH(p, q), and Engle and Bollerslev's (1986) integrated GARCH and fractionally differenced GARCH models. We can rewrite (29) as an AR(∞) process

$$\epsilon_{i,n,t}^2 = \omega_{i,n}^0 + \sum_{j=1}^{\infty} a_{i,n,j}^0 \epsilon_{i,n,t-j}^2 + e_{i,n,t}, \tag{30}$$

where $e_{i,n,t} \equiv \epsilon_{i,n,t}^2 - h_{i,n,t}^0$ is such that $\mathbb{E}(e_{i,n,t} | I_{i,n,t-1}) = 0$. Because the assumption of an infinite autoregressive process is rather mild, this is sometimes referred to as a "nonparametric" approach (see, e.g., Lewis and Reinsel, 1985). More recently, Dufour and Pelletier (2021) develop the asymptotic properties for estimating this process under conditions that are weaker than those in Lewis and Reinsel (1985). Given realization $\{\epsilon_{i,n,t}^2\}_{t=1}^T$, we can approximate (30) by a finite order AR(p) process, where p is a function of T

$$\epsilon_{i,n,t}^2 = \omega_{i,n}^{(p)} + \sum_{j=1}^{p} a_{i,n,j}^{(p)} \epsilon_{i,n,t-j}^2 + e_{i,n,t}^{(p)}. \tag{31}$$

We propose to estimate (31) using least-squares following Dufour and Pelletier (2021). Here, we face the added complexity in the form of nonnegative definiteness condition and we shall show how this issue is tackled as we proceed. Although least-squares estimation may give larger standard errors than likelihood-based estimation, it is free from the complications of numerical optimization and likelihood misspecification (see, e.g., Newey and Steigerwald, 1997). Besides, it makes possible element-wise estimation of $\boldsymbol{H_{it}^0}$ without requiring certain distributional assumptions (see, e.g., Boudt et al., 2019, p. 217). It is also computationally less demanding. As we show in a computational study provided in the supplementary document, our method requires only about 5% of the computing time of the QMLE.

For $i = 1, 2$, we let $\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{\boldsymbol{(p)}} \equiv [(\hat{\omega}_{i,1}^{(p)}, ..., \hat{\omega}_{i,d_i}^{(p)}), (\hat{a}_{i,1,1}^{(p)}, ..., \hat{a}_{i,d_i,1}^{(p)}), ...(\hat{a}_{i,1,p}^{(p)}, ..., \hat{a}_{i,d_i,p}^{(p)})]'$ collects the least-squares estimator of the vector of parameters $\boldsymbol{\theta_i} \equiv [(\omega_{i,1}, ..., \omega_{i,d_i}), (a_{i,1,1}, ..., a_{i,d_i,1}), ...(a_{i,1,p}, ..., a_{i,d_i,p})]'$ with true value $\boldsymbol{\theta_i^0} \equiv [(\omega_{i,1}^0, ..., \omega_{i,d_i}^0), (a_{i,1,1}^0, ..., a_{i,d_i,1}^0), ...(a_{i,1,p}^0, ..., a_{i,d_i,p}^0)]'$. We now provide regularity conditions under which $\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{\boldsymbol{(p)}}$ is a consistent estimator of $\boldsymbol{\theta_i^0}$.

Assumption 3.1. *For $i = 1, 2$, $n = 1, ...d_i$, (a) $\{\epsilon_{i,n,t}^2\}$ and $\{e_{i,n,t}\}$ are strictly stationary and ergodic; (b) $\{e_{i,n,t}\}$ is strong mixing with $\mathbb{E}(e_{i,n,t}^2) = \Sigma_{e,i,n}$ and has finite fourth order moment.*

Assumption 3.2. *The lag order p is chosen such that $p = o(T^{1/2}/M^{1/4})$ and $p/log(T) \to \infty$.*

Although the strict stationarity condition is maintained throughout the paper as in Assumption 3.1(a), it should be noted that integrated process is allowed in (29)–(30). Such flexibility is similar to the surplus lag approach in Bauer and Maynard (2012). Assumption 3.1(b) requires the process $\{e_{i,n,t}\}$ is strong mixing and has finite fourth order moment. The former requirement is less restrictive than the martingale difference property of $\{e_{i,n,t}\}$ whereas the latter is of equal order to the moment condition in Assumption 2.1. Assumption 3.2 is a standard condition on p in the long VAR literature. The condition $p = o(T^{1/2}/M^{1/4})$ requires that p not to grow too fast, whereas the condition $p/\log(T) \to \infty$ imposes a lower bound on the growth rate of p. The following proposition states the consistency of $\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{\boldsymbol{(p)}}$.

Proposition 3.1. *Let the conditional variance process of model (4)–(5) be defined by (28)–(31). Suppose Assumptions 3.1 and 3.2 hold, then*

$$||\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{\boldsymbol{(p)}} - \boldsymbol{\theta_i^0}|| = O_p(p^{1/2}T^{-1/2}), \text{ for } i = 1, 2.$$

The proof of Proposition 3.1 is provided in Appendix D. We have shown the consistency of $\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{\boldsymbol{(p)}}$ but it does not converge at the required rate of $\sqrt{T}$. With the current speed, we can provide further conditions such that the spillover test is consistent but the asymptotic normality may not hold under the null hypothesis. We therefore invoke Theorem 5.52 in van der Vaart (1998) to provide conditions for the least-squares criterion function $\boldsymbol{m}(\boldsymbol{\theta_i}, \epsilon_{i,n,t}^2) \equiv (\epsilon_{i,n,t}^2 - \omega_{i,n} - \sum_{j=1}^{\infty} a_{i,n,j}\epsilon_{i,n,t-j}^2)^2$ such that $\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{\boldsymbol{(p)}}$ achieves the required rate of convergence.

Assumption 3.3. *For $i = 1, 2$, $n = 1, ...d_i$, let $\boldsymbol{m}(\boldsymbol{\theta_i}, \epsilon_{i,n,t}^2)$ be any measurable function parameterized by $\boldsymbol{\theta_i}$ such that for fixed constants Δ and $\alpha > \beta$, and for every sufficiently small $\zeta > 0$,*
(a) $\sup_{||\boldsymbol{\theta_i} - \boldsymbol{\theta_i^0}|| < \zeta} \mathbb{E}[\boldsymbol{m}(\boldsymbol{\theta_i}, \epsilon_{i,n,t}^2) - \boldsymbol{m}(\boldsymbol{\theta_i^0}, \epsilon_{i,n,t}^2)] \leq -\Delta\zeta^\alpha$;
(b) $\mathbb{E}\big\{\sup_{||\boldsymbol{\theta_i} - \boldsymbol{\theta_i^0}|| < \zeta} |\mathbb{G}_T[\boldsymbol{m}(\boldsymbol{\theta_i}, \epsilon_{i,n,t}^2) - \boldsymbol{m}(\boldsymbol{\theta_i^0}, \epsilon_{i,n,t}^2)]|\big\} \leq \Delta\zeta^\beta$;
(c) $T^{-1}\sum_{t=1}^T \boldsymbol{m}(\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{\boldsymbol{(p)}}, \epsilon_{i,n,t}^2) \geq T^{-1}\sum_{t=1}^T \boldsymbol{m}(\boldsymbol{\theta_i^0}, \epsilon_{i,n,t}^2) - O_p(T^{\alpha/(2\beta - 2\alpha)})$,
where $\mathbb{G}_T[\boldsymbol{m}(\boldsymbol{\theta_i}, \epsilon_{i,n,t}^2)] = \sqrt{T}\{T^{-1}\sum_{t=1}^T \boldsymbol{m}(\boldsymbol{\theta_i}, \epsilon_{i,n,t}^2) - \mathbb{E}[\boldsymbol{m}(\boldsymbol{\theta_i}, \epsilon_{i,n,t}^2)]\}$.

In general, $\boldsymbol{m}(\boldsymbol{\theta_i}, \epsilon_{i,n,t}^2)$ can be the criterion function of any other M-estimators. See, for instance, Antoine and Renault (2012) for a comprehensive analysis in the context of GMM estimation. According to Assumption 3.3: provided that (a) the deterministic map $\mathbb{E}[\boldsymbol{m}(\boldsymbol{\theta_i}, \epsilon_{i,n,t}^2)]$ reacts rapid enough as $\boldsymbol{\theta_i}$ moves away from $\boldsymbol{\theta_i^0}$; (b) the random fluctuation between $T^{-1}\sum_{t=1}^{T}\boldsymbol{m}(\boldsymbol{\theta_i}, \epsilon_{i,n,t}^2)$ and $\mathbb{E}[\boldsymbol{m}(\boldsymbol{\theta_i}, \epsilon_{i,n,t}^2)]$ is sufficiently small, then $\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{(\boldsymbol{p})}$ has a high rate of convergence if its distance with $\boldsymbol{\theta_i^0}$ is properly bounded according to (c). For instance, with $\alpha = 1.5$ and $\beta = 0.5$, condition (c) is satisfied using the fact that the squared residuals are bounded by $O_p(p/T) = O_p(T^{-3/4})$, where the equality follows from Assumption 3.2. The desired convergence rate of $\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{(\boldsymbol{p})}$ then follows. The following proposition states the formal result.

Proposition 3.2. *Let the conditional variance process of model (4)–(5) be defined by (28)–(31). Suppose Assumptions 3.1–3.3 hold with $\alpha = 1.5$ and $\beta = 0.5$, then*

$$||\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{(\boldsymbol{p})} - \boldsymbol{\theta_i^0}|| = O_p(T^{-1/2}), \ \textit{for } i = 1, 2.$$

The proof of Proposition 3.2 is provided in Appendix E. With this speed, our test remains valid in the limit, although negative volatilities are not precluded by $\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{(\boldsymbol{p})}$. To adjust for this, in the following we provide the adjusted least-squares estimator $\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{(\boldsymbol{p})\boldsymbol{a}}$ that ensures positive semidefiniteness of $\boldsymbol{D_{it}^0}$. We show that the adjusted estimator can be computed based on an ex post estimate of $\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{(\boldsymbol{p})}$. We require the following additional conditions to hold.

Assumption 3.4. *For $i = 1, 2$, the true parameter vector $\boldsymbol{\theta_i^0}$ lies in the parameter space $[\boldsymbol{\mathcal{R}}_{\boldsymbol{i}}^{min}, \boldsymbol{\mathcal{R}}_{\boldsymbol{i}}^{max})$ such that $\boldsymbol{D_{it}^0}$ is positive semidefinite.*

Assumption 3.5. *For $i = 1, 2$, there exists a vector $\boldsymbol{\delta_i}$ with nonnegative entries such that $\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{(\boldsymbol{p})} + \boldsymbol{\delta_i} \in [\boldsymbol{\mathcal{R}}_{\boldsymbol{i}}^{min}, \boldsymbol{\mathcal{R}}_{\boldsymbol{i}}^{max})$ and $(\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{(\boldsymbol{p})} + \boldsymbol{\delta_i})\mathbb{1}(\boldsymbol{\delta_i} > \boldsymbol{0}) = \boldsymbol{\mathcal{R}}_{\boldsymbol{i}}^{min}\mathbb{1}(\boldsymbol{\delta_i} > \boldsymbol{0})$.*

Assumption 3.4 is a standard condition that restricts the true parameter such that $\boldsymbol{D_{it}^0}$ is positive semidefinite. The sufficient condition is that each element in $\boldsymbol{\theta_i^0}$ is nonnegative (i.e. $\boldsymbol{\mathcal{R}}_{\boldsymbol{i}}^{\min} = \boldsymbol{0}$, $\boldsymbol{\mathcal{R}}_{\boldsymbol{i}}^{\max} = \boldsymbol{\infty}$). Assumption 3.5 requires the existence of a lower bound vector $\boldsymbol{\delta_i}$ with nonnegative entries such that $\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{(\boldsymbol{p})} + \boldsymbol{\delta_i}$ yields a positive semidefinite $\boldsymbol{D_{it}^0}$. In practice, we can replace the negative entries in $\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{(\boldsymbol{p})}$ by zeros such that they corresponds to $\boldsymbol{\mathcal{R}}_{\boldsymbol{i}}^{\min}$, that is, $\boldsymbol{\delta_i} = -\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{(\boldsymbol{p})}\mathbb{1}(\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{(\boldsymbol{p})} < 0)$. Note that $\boldsymbol{\delta_i}$ is simply a vector of zeros if the unadjusted estimator lies in the desired parameter space.

Given $\boldsymbol{\delta_i}$, the adjusted least-squares estimators can be computed, on an ex post basis, by $\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{(\boldsymbol{p})\boldsymbol{a}} = \hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{(\boldsymbol{p})} + \boldsymbol{\delta_i}$. The following proposition states the consistency of $\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{(\boldsymbol{p})\boldsymbol{a}}$.

Proposition 3.3. *Let the conditional variance process of model (4)–(5) be defined by (28)–(31). Suppose Assumptions 3.1–3.5 hold, then*

$$||\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{(\boldsymbol{p})\boldsymbol{a}} - \boldsymbol{\theta_i^0}|| = O_p(T^{-1/2}), \ \textit{for } i = 1, 2.$$

The proof of Proposition 3.3 is provided in Appendix F. The proof uses the fact that the ex post adjustment does not affect the asymptotic properties of $\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{(\boldsymbol{p})}$ because it is only applied to the entries that are outside of the true neighborhood of $\boldsymbol{\theta_i^0}$.

To establish the asymptotic validity of (28)–(31) for our Q_1 and Q_2 tests, a final condition is required for the proper convergence of $\boldsymbol{R_i^0}$.

Assumption 3.6. *For $i = 1, 2$, $(\boldsymbol{D}_{it}^{\mathbf{0}})^{-1/2}\boldsymbol{\epsilon}_{it}$ maintains the same stochastic properties as $\boldsymbol{\Xi}_{it}$ with covariance $\mathbb{E}[(\boldsymbol{D}_{it}^{\mathbf{0}})^{-1/2}\boldsymbol{\epsilon}_{it}\boldsymbol{\epsilon}_{it}'(\boldsymbol{D}_{it}^{\mathbf{0}})^{-1/2}] = \boldsymbol{R}_i^{\mathbf{0}}$.*

It is evident that $(\boldsymbol{D}_{it}^{\mathbf{0}})^{-1/2}\boldsymbol{\epsilon}_{it}$ belongs to a special case of $\boldsymbol{\Xi}_{it} = (\boldsymbol{H}_{it}^{\mathbf{0}})^{-1/2}\boldsymbol{\epsilon}_{it}$ with diagonal $\boldsymbol{H}_{it}^{\mathbf{0}}$; it is therefore natural for the former to inherit the stochastic properties of the latter but with covariance $\boldsymbol{R}_i^{\mathbf{0}}$ instead of identity covariance. Then, the estimation of (28) proceeds in two steps. First, we estimate (31) for each of the diagonal elements in $\boldsymbol{D}_{it}^{\mathbf{0}}$ using the adjusted least-squares estimator $\hat{\boldsymbol{\theta}}_i^{(\boldsymbol{p})\boldsymbol{a}}$. The estimated positive semidefinite process is denoted by $\hat{\boldsymbol{D}}_{it}$. In the second step, $\boldsymbol{R}_i^{\mathbf{0}}$ is estimated using the sample covariance of $\hat{\boldsymbol{D}}_{it}^{-1/2}\boldsymbol{\epsilon}_{it}$, which we denote by $\hat{\boldsymbol{R}}_i$. Because $\hat{\boldsymbol{R}}_i$ is always positive semidefinite, this ensures that the estimated time-varying covariance matrix is always positive semidefinite.

The following proposition states the validity of Theorems 2.1 and 2.2 under the proposed volatility model.

Proposition 3.4. *Let the conditional variance process of model (4)–(5) be defined by (28)–(31). Suppose Assumptions 2.1–2.2, 2.4–2.6, 3.1–3.6 hold, then the results of Theorem 2.1 remain valid. Additionally, suppose Assumption 2.7 holds, then the results of Theorem 2.2 remain valid.*

The proof of Proposition 3.4 is provided in Appendix G. The key is to show that the second step estimator $\hat{\boldsymbol{R}}_i$ is $\sqrt{T}$-consistent for $\boldsymbol{R}_i^{\mathbf{0}}$. Given this result and by collecting $\hat{\boldsymbol{R}}_i$ in the estimator vector, the results of Theorems 2.1 and 2.2 continue to hold under their respective conditions. Note that Assumption 2.3 is not needed here since we do not have to specify an initial value for our model. Besides, when the true data generating process has finite autoregressive order, we have $\sqrt{T}$-consistent estimators regardless of Assumption 3.3. We provide Assumption 3.3 as a formal condition to maintain the generality of our approach where we allow p to grow with T.

In summary, a multivariate volatility model is proposed to facilitate the estimation of Q_1 and Q_2. The proposed specification enjoys estimation simplicity and computational efficiency. The approach is somewhat "nonparametric" in that it imposes minimal assumption on the structure of $\boldsymbol{D}_{it}^{\mathbf{0}}$. We also do not impose any parametric assumption on $\boldsymbol{R}_i^{\mathbf{0}}$. The structure of (28) is a modified version of the variance-covariance matrix in Bollerslev (1990), where we specify the elements in $\boldsymbol{D}_{it}^{\mathbf{0}}$ using a more general volatility process and we propose least-squares estimation. We also demonstrate the consistency of our two-steps estimators. For notational simplicity, we denote our approach in short as the NC-LS approach, where the acronym highlights the nonparametric modeling of covariance and its least-squares estimation. The NC-LS approach is readily extendable to the case where $\boldsymbol{R}_i^{\mathbf{0}}$ has a dynamic structure using (say) the Exponentially Weighted Moving Average (EWMA) method, while keeping its estimation properties. However, the simulation and empirical results in the next sections suggest that the proposed NC-LS approach works coherently with the spillover test to deliver reliable inference and shows adequacy in fitting real world data. Thus, such extension is best left for future work.

4. Monte Carlo simulations

In this section, we investigate the finite sample performance of the proposed econometric strategy using Monte Carlo simulations. We first consider a bivariate setup (i.e., $d_1 = d_2 = 2$), where we conduct experiments to study the effect of covariance intensity on the finite sample size and power of our testing strategy. Then, we study the behavior of our method with increasing dimension. To save space, we focus here on the unidirectional test statistic Q_1, and we report and discuss in the supplementary document the full results based on the bidirectional statistic Q_2.

4.1. The bivariate case

We work with the following bivariate data generating process with persistent conditional means and variances to reflect the stylized features of empirical data

$$
\begin{aligned}
\boldsymbol{Y_{it}} = \begin{pmatrix} Y_{i,1t} \\ Y_{i,2t} \end{pmatrix} &= \begin{pmatrix} 1 + m_{i,1t} \\ 1 + m_{i,2t} \end{pmatrix} + \begin{pmatrix} \epsilon_{i,1t} \\ \epsilon_{i,2t} \end{pmatrix}, \quad i = 1, 2, \quad t = 1, ..., T, \\
\begin{pmatrix} \epsilon_{i,1t} \\ \epsilon_{i,2t} \end{pmatrix} &\overset{\text{iid}}{\sim} \mathrm{N}\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} h^0_{i,1t} & r_i (h^0_{i,1t})^{1/2} (h^0_{i,2t})^{1/2} \\ r_i (h^0_{i,2t})^{1/2} (h^0_{i,1t})^{1/2} & h^0_{i,2t} \end{pmatrix} \right], \\
m_{i,1t} &= 0.8 m_{i,1t-1} + e_{i,1t}, \quad m_{i,2t} = 0.8 m_{i,2t-1} + e_{i,2t}, \quad e_{i,1t}, e_{i,2t} \overset{\text{iid}}{\sim} \mathrm{N}(0, 4), \\
h^0_{i,1t} &= 0.1 + 0.8 h^0_{i,1t-1} + 0.05 \epsilon^2_{i,1t-1}, \quad h^0_{i,2t} = 0.1 + 0.8 h^0_{i,2t-1} + 0.05 \epsilon^2_{i,2t-1}.
\end{aligned} \tag{32}
$$

We consider the following correlation structures

$$
\begin{aligned}
&\text{NullA: } r_1 = r_2 = 0.2, \quad \text{NullB: } r_1 = r_2 = 0.5, \\
&\text{NullC: } r_1 = r_2 = r_t = 0.2 + 0.1 \times 0.2 \cos[2\pi t/(T/4)].
\end{aligned}
$$

Under NullA, we have a relatively moderate correlation between the conditional variances in both $\boldsymbol{Y_{1t}}$ and $\boldsymbol{Y_{2t}}$. Combination NullB increases the correlation magnitude. Under NullC, we have a stable time-varying correlation structure that is generated by a cosine function with four periods over sample size T. To study the power of our testing strategy, we simulate the effect of volatility spillover by generating correlated squared innovation $\tilde{\epsilon}^2_{1,jt}$ and $\tilde{\epsilon}^2_{2,jt}$ using Cholesky transformation, where for $j = 1, 2$, $\tilde{\epsilon}^2_{1,jt} = s_2 \epsilon^2_{2,jt-1} + (1 - s_2^2)^{1/2} \epsilon^2_{1,jt-1}$, $\tilde{\epsilon}^2_{2,jt} = \epsilon^2_{2,jt}$. The parameter $s_2 \in [0, 1]$ controls the intensity of volatility spillover from $\boldsymbol{Y_{2t}}$ to $\boldsymbol{Y_{1t}}$ with respect to $\boldsymbol{I_{t-1}}$. We consider the following parameter combinations

$$
\text{AlterA: } r_1 = r_2 = 0.2, s_2 = 0.35, \quad \text{AlterB: } r_1 = 0.2, r_2 = 0.5, s_2 = 0.35.
$$

Both AlterA and AlterB generate equal spillover intensity ($s_2 = 0.35$) from $\boldsymbol{Y_{2t}}$ to $\boldsymbol{Y_{1t}}$ with respect to $\boldsymbol{I_{t-1}}$. This allows examining the power of our test. We increase the covariance of the risk transmitter $\boldsymbol{Y_{2t}}$ under AlterB to study the role it plays in driving volatility spillover.

For each data generating process, we conduct 10000 Monte Carlo simulations with sample size T = 1000 and 1500, which correspond to approximately four and six years of daily data, respectively.

For each T, we generate $T + 1000$ observations and then we discard the first 1000 observations to reduce possible effects from the chosen starting values $(h^0_{i,10}, h^0_{i,20}, m_{i,10}, m_{i,20}) = [0.1/(1 - 0.05 - 0.8), 0.1/(1 - 0.05 - 0.8), 0, 0]$. We consider the following three downward weighting kernel functions $k(\cdot)$.

The Bartlett (BAR) kernel,

$$k(z) = \begin{cases} 1 - |z|, & \text{if } |z| \leq 1, \\ 0, & \text{otherwise.} \end{cases}$$

The Daniell (DAN) kernel,

$$k(z) = \frac{\sin(\pi z)}{\pi z}, \quad z \in \mathbb{R}.$$

The Quadratic-Spectral (QS) kernel,

$$k(z) = \frac{25}{12\pi^2 z^2}\left[\frac{\sin(6\pi z/5)}{6\pi z/5} - \cos(6\pi z/5)\right], \quad z \in \mathbb{R}.$$

For comparison with an equally weighted test, we also include the Truncated (TR) kernel. Note that the selected kernels satisfy the requirements in Assumption 2.5. Among the downward weighting kernels, the BAR kernel gives the most intuitive empirical interpretation due to its linear weighting scheme and its compact support. For instance, when the BAR kernel is used, we are checking for spillover up to M lags with linearly decreasing weight to each lag order. Thus, in practice M can be selected based on the scope of the empirical examination. For instance, a relative small (large) M can be set to examine short-run (long-run) volatility spillover. Considering the empirical implication, we assess the sensitivity of our inferential strategy by considering a wide $M = 10$, 20 and 30. All spillover tests are carried out at the 5% significance level.

For each simulation, the estimation and testing procedure proceeds in steps. First, we filter the conditional mean of $\boldsymbol{Y_{it}}$ using least-squares which yields $\sqrt{T}$-consistent residuals (see, e.g., White, 2001, Theorem 5.11). Then, based on the procedures described in Section 3 we fit the NC-LS model. We select for the order of every diagonal ARCH process in the volatility structure using the Bayesian information criteria up to the 25$^{\text{th}}$ order. Finally, we test for spillover by computing Q_1 and Q_2.

In addition to asymptotic critical values, we also consider a nonparametric bootstrap in which we randomly re-sample the estimated residuals with replacement. As is well known, the bootstrap procedure can often yield a more accurate finite sample size (see, e.g., Chen and Hong, 2012a,b, 2016). We denote the bootstrap statistic using asterisk by Q_1^*. Step (i), retain fitted series and residuals $\boldsymbol{\hat{Y}_{1t}}$, $\boldsymbol{\hat{Y}_{2t}}$, $\boldsymbol{\hat{\epsilon}_{1t}}$ and $\boldsymbol{\hat{\epsilon}_{2t}}$. Step (ii), compute Q_1. Step (iii), obtain bootstrap residuals $\boldsymbol{\hat{\epsilon}^*_{1t}}$ and $\boldsymbol{\hat{\epsilon}^*_{2t}}$ and construct bootstrap sample $\boldsymbol{Y^*_{1t}} = \boldsymbol{\hat{Y}_{1t}} + \boldsymbol{\hat{\epsilon}^*_{1t}}$ and $\boldsymbol{Y^*_{2t}} = \boldsymbol{\hat{Y}_{2t}} + \boldsymbol{\hat{\epsilon}^*_{2t}}$. Step (iv), compute the

b-th statistic Q_1^{*b} in the same way as we compute Q_1 but with $\{\boldsymbol{Y_{1t}^*}, \boldsymbol{Y_{2t}^*}\}_{t=1}^T$ replacing the original sample $(\boldsymbol{\mathcal{Y}_1}, \boldsymbol{\mathcal{Y}_2}) \equiv \{\boldsymbol{Y_{1t}}, \boldsymbol{Y_{2t}}\}_{t=1}^T$. Step (v), repeat steps (iii) to (iv) B times to obtain B bootstrap test statistics $\{Q_1^{*b}\}_{b=1}^B$. Step (vi), compute bootstrap p-value by $B^{-1}\sum_{b=1}^B \mathbb{1}(Q_1^{*b} > Q_1)$. We set $B = 499$ and we maintain 10000 simulations.

By design, the bootstrap approach ensures that the null hypothesis always holds in the bootstrap world since the two series $(\boldsymbol{\mathcal{Y}_1}, \boldsymbol{\mathcal{Y}_2})$ are re-sampled independently. This ensures the asymptotic normality of Q_1^*. Under the alternatives, our bootstrap approach has asymptotic unit power. This follows from the fact that while Q_1^* remains converging in distribution to $\mathrm{N}(0, 1)$, Q_1 converges to positive infinity in probability following Theorem 2.2. Therefore, we obtain consistent bootstrap p-value.

Table 1 reports the empirical sizes of our volatility spillover tests under NullA, NullB and NullC based on the NC-LS modeling. In general, we find that Q_1 tends to over reject the null a little but not excessively. The size improves gradually as T increases. We find the rejection rates of Q_1 to be stable across the three parameter combinations. This implies that the size of our inferential strategy is not affected by increasing portfolio correlation and the time-varying cosine case. As expected, our bootstrap test Q_1^* yields a more accurate finite sample size than Q_1, and it too is robust to changing correlations. Overall, we find the proposed econometric strategy to be reasonably sized. This result appears to hold across the kernel functions and the value of their smoothing parameter M.

We report the empirical powers of our testing approach in Table 2. For Q_1, we use empirical critical values that are computed from the 10000 simulations under NullA. This gives size-adjusted powers. In general, we find that our inferential strategy becomes more powerful as T increases. We also find that both Q_1 and Q_1^* give rather similar power. The rejection rates of Q_1 and Q_1^* decrease in M. This is because under AlterA and AlterB, we have one-period lagged volatility spillover. Therefore, we expect a test that focuses on recent events to give better power. Besides, we find that downward weighting kernels often yield better power than the TR kernel, and they are less affected by a large M. These results confirm our expectation that, compared with an equally weighted test, downward weighting tests alleviate the impact of choosing a relatively large M because they discount higher order lags. Interestingly, we find that the rejection rates of Q_1 and Q_1^* are higher under parameter combination AlterB. This implies that, other things being equal, an increase in the correlation within the risk transmitter $\boldsymbol{Y_{2t}}$ can drive the overall effect of volatility spillover. This result highlights the nontrivial role covariance can play in driving spillover.

Finally, we carry out likelihood-based sensitivity analysis, that is, we investigate the finite sample performance of the volatility spillover tests based on model described in (28)–(29) where the ARCH(∞) process (29) is estimated using QMLE.[4] Simulation results (deferred to the supplementary document) are very comparable to those reported in Table 1 and Table 2. This is consistent with Theorem 2.1 and Theorem 2.2 that, given the regularity conditions, the asymptotic properties

[4]We thank an anonymous referee for this suggestion.

of the spillover test are robust to estimation procedures.

Table 1: Empirical sizes of Q_1 and Q_1^*

T	M	NullA 10	NullA 20	NullA 30	NullB 10	NullB 20	NullB 30	NullC 10	NullC 20	NullC 30
Rejection rates based on asymptotic critical values										
1000	$Q_{1\text{BAR}}$	7.1	6.8	6.9	7.0	6.9	6.8	7.2	6.7	6.7
	$Q_{1\text{DAN}}$	7.1	6.8	6.9	7.0	6.7	7.1	7.0	6.7	6.8
	$Q_{1\text{QS}}$	7.1	6.9	6.8	7.1	6.8	7.1	6.8	6.7	6.9
	$Q_{1\text{TR}}$	7.1	6.9	7.0	7.0	7.3	7.1	6.6	6.6	6.7
1500	$Q_{1\text{BAR}}$	6.7	6.5	6.4	6.7	6.4	6.3	6.8	6.7	6.5
	$Q_{1\text{DAN}}$	6.8	6.6	6.4	6.7	6.3	6.5	6.7	6.6	6.2
	$Q_{1\text{QS}}$	6.8	6.5	6.3	6.6	6.4	6.5	6.8	6.5	6.3
	$Q_{1\text{TR}}$	6.5	6.4	6.4	6.1	6.3	6.3	6.3	6.4	6.4
Rejection rates based on bootstrap critical values										
1000	$Q^*_{1\text{BAR}}$	5.3	5.6	5.4	5.3	5.5	5.1	5.3	5.5	5.5
	$Q^*_{1\text{DAN}}$	5.4	5.3	5.2	5.3	5.3	5.1	5.4	5.3	5.2
	$Q^*_{1\text{QS}}$	5.5	5.4	5.0	5.5	5.2	5.2	5.5	5.4	5.1
	$Q^*_{1\text{TR}}$	5.0	5.0	5.2	5.0	4.9	5.1	5.1	5.1	5.1
1500	$Q^*_{1\text{BAR}}$	5.2	5.3	5.4	5.3	5.3	5.3	5.3	5.2	5.4
	$Q^*_{1\text{DAN}}$	5.3	5.1	5.3	5.2	5.2	5.3	5.3	5.2	5.3
	$Q^*_{1\text{QS}}$	5.2	5.2	5.3	5.3	5.3	5.2	5.3	5.2	5.3
	$Q^*_{1\text{TR}}$	4.9	5.3	5.3	4.9	5.5	5.2	4.8	5.3	5.3

NOTES: The table reports empirical sizes (in %) of Q_1 under NullA, NullB and NullC at the 5% significance level based on NC-LS modeling. Number of simulations = 10000. $Q_{1\text{BAR}}$, $Q_{1\text{DAN}}$, $Q_{1\text{QS}}$, $Q_{1\text{TR}}$ and $Q^*_{1\text{BAR}}$, $Q^*_{1\text{DAN}}$, $Q^*_{1\text{QS}}$, $Q^*_{1\text{TR}}$ denote the rejection rates of Q_1 using asymptotic and bootstrap critical values, respectively; the subscripts BAR, DAN, QS and TR denote, respectively, the Bartlett kernel, the Daniell kernel, the Quadratic-Spectral kernel and the Truncated kernel. Number of bootstraps = 499. T and M denote the sample size and kernel smoothing parameter, respectively.

4.2. Higher dimensions

The finite sample performance of existing multivariate dependence tests is often demonstrated up to the case of three series. For instance, the bivariate case is examined in Duchesne and Roy (2004), whereas Robbins and Fisher (2015) study the relations between bivariate and trivariate processes, that is $d_1 = 3$, $d_2 = 2$. By contrast, we now demonstrate the finite sample performance of our multivariate approach in higher dimensions, which is made feasible thanks to the proposed NC-LS modeling. We focus on the case where we analyze spillover effects on a relatively large market covering multiple countries such as the European Union. In particular, we study $d_1 = 3, 4, ..., 10$ and $d_2 = 2$. We expect our approach to perform similarly given the opposite relation or any combinations of d_i so long as they are of equal complexity.

We study the size of our inferential strategy under combination NullD, where with increasing d_1, we retain the correlation intensity r_i of NullA because our bivariate simulations show stability across combinations NullA–NullC. In spite of that, we perform a sensitivity check to find that

Table 2: Empirical powers of Q_1 and Q_1^*

T	M	AlterA			AlterB		
		10	20	30	10	20	30
Rejection rates based on empirical critical values							
1000	$Q_{1\text{BAR}}$	78.2	69.0	61.9	95.0	90.1	85.1
	$Q_{1\text{DAN}}$	74.5	62.1	53.4	93.3	85.1	78.2
	$Q_{1\text{QS}}$	73.1	59.8	51.7	92.7	83.8	76.6
	$Q_{1\text{TR}}$	51.5	38.4	32.0	76.8	61.1	51.7
1500	$Q_{1\text{BAR}}$	92.1	85.4	80.0	99.4	98.0	96.7
	$Q_{1\text{DAN}}$	89.8	79.6	72.3	99.0	96.6	93.7
	$Q_{1\text{QS}}$	88.9	78.2	70.9	98.8	96.2	92.8
	$Q_{1\text{TR}}$	71.1	55.1	46.4	92.9	83.3	73.2
Rejection rates based on bootstrap critical values							
1000	$Q^*_{1\text{BAR}}$	76.1	66.0	59.1	94.2	88.4	83.1
	$Q^*_{1\text{DAN}}$	71.7	58.8	49.7	92.4	82.9	75.1
	$Q^*_{1\text{QS}}$	70.3	57.0	48.2	91.6	81.4	73.7
	$Q^*_{1\text{TR}}$	47.9	34.8	29.0	73.6	58.3	49.3
1500	$Q^*_{1\text{BAR}}$	91.2	83.9	78.0	99.4	98.0	96.2
	$Q^*_{1\text{DAN}}$	88.4	77.9	69.1	99.0	96.2	92.8
	$Q^*_{1\text{QS}}$	87.2	76.3	67.6	98.7	95.6	91.9
	$Q^*_{1\text{TR}}$	67.6	52.0	42.8	92.0	80.9	72.0

NOTES: The table reports empirical powers (in %) of Q_1 under AlterA and AlterB at the 5% significance level based on NC-LS modeling. Number of simulations = 10000. $Q_{1\text{BAR}}$, $Q_{1\text{DAN}}$, $Q_{1\text{QS}}$, $Q_{1\text{TR}}$ and $Q^*_{1\text{BAR}}$, $Q^*_{1\text{DAN}}$, $Q^*_{1\text{QS}}$, $Q^*_{1\text{TR}}$ denote the rejection rates of Q_1 using empirical and bootstrap critical values, respectively; the subscripts BAR, DAN, QS and TR denote, respectively, the Bartlett kernel, the Daniell kernel, the Quadratic-Spectral kernel and the Truncated kernel. Number of bootstraps = 499. T and M denote the sample size and kernel smoothing parameter, respectively.

the performance of our approach in the higher dimensions is robust to the time-varying cosine correlation of NullC. For power study, we maintain the covariance structure in AlterA but we reduce the spillover intensity s_2 to 0.15 to highlight the power effects as d_1 increases. Given $d_1 > d_2$, we generate spillover to each series in $\boldsymbol{Y_{1t}}$ by repeating the influence of $\boldsymbol{Y_{2t}}$. This ensures that every risk recipients in $\boldsymbol{Y_{1t}}$ is equally affected by the spillover effect. For instance, when $d_1 = 4$, $d_2 = 2$, $Y_{1,3t}$ and $Y_{1,4t}$ will be affected by $Y_{2,1t}$ and $Y_{2,2t}$, respectively. We denote this parameter combination by AlterC. Because the overall performance of our test is stable across M, we only report here the case where $M = 20$ to save space. The full set of results are reported in the supplementary document.

Table 3 reports the empirical sizes of our inferential approach. We find that the size of Q_1 increases in dimension, but not overly excessive nor rapid. The size of Q_1 generally improves and stabilizes as T increases. The rejection rates of our bootstrap approach Q_1^* also tend to increase in dimension when $T = 1000$, but they become very stable as T approaches 1500. Table 4 reports the power study. As with the bivariate study, we use empirical critical values for Q_1. In general, our approach has power despite a rather low spillover intensity $s_2 = 0.15$. Both Q_1 and Q_1^* give similar rejection rates, and they become more powerful as T increases. We find that the power of our tests grows with d_1. Because the number of risk recipients increases as d_1 increases, this yields a stronger evidence of spillover and thus increase the rejection rates Q_1 and Q_1^*.

Table 3: Empirical sizes of Q_1 and Q_1^*

		NullD							
T	d_1	3	4	5	6	7	8	9	10
Rejection rates based on asymptotic critical values									
1000	$Q_{1\text{BAR}}$	7.1	7.1	7.1	7.2	7.4	7.9	7.9	7.6
	$Q_{1\text{DAN}}$	7.3	6.9	7.1	7.3	7.3	7.9	8.2	7.9
	$Q_{1\text{QS}}$	7.1	7.0	7.3	7.2	7.4	7.8	8.3	7.9
	$Q_{1\text{TR}}$	7.2	7.3	7.4	7.4	7.5	8.1	8.8	8.9
1500	$Q_{1\text{BAR}}$	6.7	7.0	6.6	6.8	6.3	7.2	7.1	6.8
	$Q_{1\text{DAN}}$	6.6	6.9	6.6	6.7	6.4	7.2	7.2	7.0
	$Q_{1\text{QS}}$	6.7	6.9	6.6	6.7	6.4	7.2	7.1	7.1
	$Q_{1\text{TR}}$	6.7	6.9	7.1	7.0	7.0	6.8	7.0	7.3
Rejection rates based on bootstrap critical values									
1000	$Q^*_{1\text{BAR}}$	4.8	5.5	5.5	5.2	5.5	5.2	5.3	5.8
	$Q^*_{1\text{DAN}}$	4.9	5.6	5.3	5.5	5.4	5.3	5.4	5.7
	$Q^*_{1\text{QS}}$	5.1	5.5	5.5	5.5	5.5	5.5	5.3	5.8
	$Q^*_{1\text{TR}}$	5.1	5.8	5.7	5.3	5.7	5.7	5.7	5.6
1500	$Q^*_{1\text{BAR}}$	5.1	5.2	4.8	5.0	5.4	5.5	5.4	5.1
	$Q^*_{1\text{DAN}}$	5.1	5.3	4.8	5.2	5.3	5.5	5.4	5.1
	$Q^*_{1\text{QS}}$	5.2	5.2	4.8	5.2	5.4	5.5	5.4	5.1
	$Q^*_{1\text{TR}}$	5.3	5.4	5.2	5.3	5.5	5.6	5.3	5.5

NOTES: The table reports empirical sizes (in %) of Q_1 under NullD at the 5% significance level based on NC-LS modeling. Number of simulations = 10000. $Q_{1\text{BAR}}$, $Q_{1\text{DAN}}$, $Q_{1\text{QS}}$, $Q_{1\text{TR}}$ and $Q^*_{1\text{BAR}}$, $Q^*_{1\text{DAN}}$, $Q^*_{1\text{QS}}$, $Q^*_{1\text{TR}}$ denote the rejection rates of Q_1 using asymptotic and bootstrap critical values, respectively; the subscripts BAR, DAN, QS and TR denote, respectively, the Bartlett kernel, the Daniell kernel, the Quadratic-Spectral kernel and the Truncated kernel. Number of bootstraps = 499. T and d_1 denote the sample size and dimension of portfolio 1, respectively.

5. Empirical application

The North America (NA) has historically maintained a strong economic partnership with the UK but Cumming and Zahra (2016) suggest that this relation is to be challenged after the UK voted to leave the European Union on 23rd June 2016. In this section, we use the new inferential strategy to study, before and after the Brexit referendum, the spillover relations between the North American and the UK equity markets. We use the American S&P-500 and the Canadian S&P-TSX stock indices for the NA market, and we use the FTSE-All index for the UK market. To examine possible Brexit effect on the broader European market, we also study its spillover relations with the NA market. Regarding the former, we use the European Union (EU) portfolio previously constructed by Baele (2005): Austria, Belgium, France, Germany, Ireland, Italy, the Netherlands and Spain, where the market indices are taken as ATX, Bel-20, FrCAC-40, DAX-30, ISEQ-All, FTSE-MIB, AEX and IBEX-35, respectively. It is worth highlighting that this is the first study

Table 4: Empirical powers of Q_1 and Q_1^*

T	d_1	AlterC 3	4	5	6	7	8	9	10
Rejection rates based on empirical critical values									
1000	$Q_{1\text{BAR}}$	27.6	40.8	55.2	65.5	73.9	77.1	80.9	83.0
	$Q_{1\text{DAN}}$	23.0	33.4	44.8	54.9	62.5	65.9	70.3	71.4
	$Q_{1\text{QS}}$	22.1	31.6	43.2	52.9	60.1	64.1	67.4	69.0
	$Q_{1\text{TR}}$	14.3	18.8	23.7	30.0	34.2	35.5	37.1	37.8
1500	$Q_{1\text{BAR}}$	42.6	63.0	84.2	91.4	96.0	97.7	98.6	99.1
	$Q_{1\text{DAN}}$	34.1	51.9	73.2	82.4	90.7	93.6	95.0	96.8
	$Q_{1\text{QS}}$	33.1	49.5	70.7	80.3	88.9	92.2	93.8	95.8
	$Q_{1\text{TR}}$	18.8	28.1	39.2	47.8	57.2	61.9	66.3	67.8
Rejection rates based on bootstrap critical values									
1000	$Q^*_{1\text{BAR}}$	23.8	36.6	52.0	62.2	70.1	74.9	78.6	80.9
	$Q^*_{1\text{DAN}}$	19.6	29.6	42.1	50.9	58.1	63.3	67.2	69.8
	$Q^*_{1\text{QS}}$	19.0	28.2	40.3	48.4	55.6	60.4	64.6	67.1
	$Q^*_{1\text{TR}}$	12.4	15.6	21.7	26.4	30.4	32.9	35.9	37.0
1500	$Q^*_{1\text{BAR}}$	38.7	62.3	81.3	89.9	94.8	97.3	98.4	98.9
	$Q^*_{1\text{DAN}}$	31.2	50.9	70.2	80.3	87.9	92.2	94.4	95.9
	$Q^*_{1\text{QS}}$	29.6	48.4	67.4	77.8	86.1	90.5	93.0	94.8
	$Q^*_{1\text{TR}}$	16.9	26.0	36.4	44.5	52.0	58.3	61.8	65.9

NOTES: The table reports empirical powers (in %) of Q_1 under AlterC at the 5% significance level based on NC-LS modeling. Number of simulations = 10000. $Q_{1\text{BAR}}$, $Q_{1\text{DAN}}$, $Q_{1\text{QS}}$, $Q_{1\text{TR}}$ and $Q^*_{1\text{BAR}}$, $Q^*_{1\text{DAN}}$, $Q^*_{1\text{QS}}$, $Q^*_{1\text{TR}}$ denote the rejection rates of Q_1 using empirical and bootstrap critical values, respectively; the subscripts BAR, DAN, QS and TR denote, respectively, the Bartlett kernel, the Daniell kernel, the Quadratic-Spectral kernel and the Truncated kernel. Number of bootstraps = 499. T and d_1 denote the sample size and dimension of portfolio 1, respectively.

to provide insights into the distortion of spillover relation between the North American and the European equity markets in the aftermath of the Brexit referendum. Given the historical economic relations between the regions, the one-way volatility spillover test is appropriate for this study.

We sample our data centering the referendum event from 2nd January 2012 to 31st December 2019 at the daily frequency from *Datastream*. This gives 2087 observations. Then, we divide the sample into two subperiods: the pre–Brexit sample (2nd January 2012 – 23rd June 2016) and the post–Brexit sample (24th June 2016 – 31st December 2019). Using the referendum date as a reference point is appropriate because the voting outcome had immediately affected economic outlooks and expectations (see, e.g., Breinlich et al., 2022, p. 66).[5] We collect all data in US dollar to minimize potential bias due to currency risk. Return series are calculated by taking the first difference of the price indices in natural logarithm. Although the markets being studied are partially synchronous, our analyses are not driven by spurious spillover effects (type 1 error) on the same calendar day

[5]We also perform a sensitivity analysis (results reported in the supplementary document) in which we curtailed 8 observations before and after the Brexit referendum date to find that our main conclusions remain robust. We choose 8 observations as the Vote Leave Campaign presented its major plan 8 days before the referendum date.

because the Q_1 test checks for the lagged spillover effect of at least one day. Besides, both the EU and UK markets share the similarity that they are mostly synchronous and partially overlaps with the NA market. Thus, our setup is impartial to UK and EU to study and compare individually their connections with the NA market.

We begin with the NA–UK study. First, we estimate our nonparametric conditional variance model for both subsamples. To filter out possible mean causality, each return variable is regressed on the remaining lagged series. With the residuals that are free from mean causality, we estimate the volatility model using the proposed NC-LS approach. The best fitting model lag orders are selected on the basis of Bayesian information criteria and diagnostic examinations. Regarding the pre–Brexit period, we obtain orders 9, 3 and 11 for the conditional variances of the UK, US and Canada markets, respectively. As for the post–Brexit sample, we obtain orders 4, 16 and 12 for the conditional variances of the UK, US and Canada series, respectively. We carry out for the NA portfolio the Engle and Sheppard's (2001) diagnosis of stable correlation — which shows better power to regime switching in economic condition such as Brexit (McCloud and Hong, 2011, Tables 6–7) — to find that we cannot reject the null of stable correlation structure at the usual significance level, with p-values of 0.8969 and 0.1325 for the pre–Brexit and post–Brexit samples, respectively. The optimal orders of the correlation stability test are automatically selected based on Bayesian information criteria. This, together with a series of conventional Ljung-Box examinations reported in Table 5, suggests the adequacy of our NC-LS modeling.

Next, we compute our bootstrap Q_1^* tests using the Bartlett kernel since simulations suggest similar performance across the downward weighting kernels. We report the p-values in Table 6. In the pre–Brexit sample, we find that the spillover effect from the NA market to the UK market is statistically significant at the 5% level for all M's. This finding implies that the NA market has a significant influence on the UK market in both the short term and the long run. In the other direction, we find evidence of spillover effect from the UK market to the NA market at the 10% level for all M's. Our findings imply feedback spillovers in the NA–UK nexus. This interdependent relation, however, diminishes in the post–Brexit period.

Before Brexit, the feedback spillover in the NA–UK nexus can be explained by the closely interconnected economic activities in the two regions. Since the two markets rely on each other, the market participants in the two regions tend to follow each other closely. Therefore, an increase in uncertainty or volatility of one market would inevitably affect the other. Interestingly, the spillover effects between NA and UK disappear after Brexit. In other words, the NA (UK) market is no longer significantly affected by the volatility in the UK (NA) market. After the Brexit referendum, market participants in the UK may be discouraged to infer information from the NA market because they are less confident about the UK's bargaining position in the international market especially among major players such as NA. Consequently, uncertainly in NA does not spill to UK. In the other direction, the evidence that market participants in the NA region do not respond to the UK could be due to them weighing the merits of further expanding into the UK market. This is consistent with the reported survey in Cumming and Zahra (2016) that 20% of UK entrepreneurs are considering a move to other countries probably elsewhere in the EU.

Table 5: Diagnostic tests (UK–NA)

	LB(10)	LB(20)	LB(30)	$LB^2(10)$	$LB^2(20)$	$LB^2(30)$
Pre–Brexit (2^{nd} January 2012 – 23^{rd} June 2016)						
UK	7.155	22.093	36.494	6.177	18.319	25.026
	[0.711]	[0.335]	[0.192]	[0.800]	[0.566]	[0.724]
US	13.169	20.312	32.243	12.124	20.539	30.504
	[0.214]	[0.439]	[0.356]	[0.277]	[0.425]	[0.440]
Canada	4.224	20.656	34.183	2.911	23.436	38.206
	[0.937]	[0.418]	[0.274]	[0.983]	[0.268]	[0.145]
Post–Brexit (24^{th} June 2016 – 31^{st} December 2019)						
UK	11.071	19.740	27.036	10.468	14.336	16.783
	[0.352]	[0.474]	[0.621]	[0.400]	[0.813]	[0.975]
US	15.436	18.869	23.736	10.546	13.020	17.188
	[0.117]	[0.530]	[0.784]	[0.394]	[0.877]	[0.970]
Canada	6.779	18.437	22.004	4.432	18.823	21.513
	[0.746]	[0.559]	[0.854]	[0.926]	[0.533]	[0.871]

NOTES: The table reports diagnostic analyses for all fitted series. LB(M) and $LB^2(M)$ are the Ljung-Box tests for the null of no serial correlation (up to lag order M) on the standardized and squared standardized residuals, respectively. The values in the squared parentheses are the p-values of the tests.

Table 6: Spillover results (UK–NA)

	Pre–Brexit			*Post–Brexit*		
M	10	20	30	10	20	30
Q^*_{1BAR}	0.028	0.036	0.038	0.168	0.154	0.186
Q^*_{-1BAR}	0.088	0.058	0.054	0.729	0.677	0.774

NOTES: The table reports bootstrap p-values of the proposed spillover tests. Number of bootstraps = 499. Q^*_{1BAR} denotes the one-way test for the null hypothesis of no volatility spillover from the NA market to the UK market. Q^*_{-1BAR} denotes the one-way test for the null hypothesis of no volatility spillover from the UK market to the NA market. The subscript BAR denotes the Bartlett kernel. M denotes the kernel smoothing parameter.

We now examine the NA–EU spillover relation. In the pre–Brexit sample, we obtain orders 16, 14, 12, 11, 7, 18, 9, 15, 3 and 15 for the conditional variances of the Austria, Belgium, France, Germany, Ireland, Italy, the Netherlands, Spain, US and Canada markets, respectively. As for the post–Brexit period, we obtain orders 11, 16, 10, 12, 9, 10, 11, 8, 9 and 10 for the conditional variances of the Austria, Belgium, France, Germany, Ireland, Italy, the Netherlands, Spain, US and Canada series, respectively. As with the NA–UK study, we perform the Engle and Sheppard's (2001) diagnosis to find that we cannot reject the null of stable correlation structure at the usual significance level for both subsamples and for both portfolios. Regarding the EU portfolio, we obtain p-values of 0.4118 and 0.1183 for the pre–Brexit and post–Brexit samples, respectively. As for the NA portfolio, we obtain p-values of 0.6592 and 0.5472 for the pre–Brexit and post–Brexit samples, respectively. These examinations, along with a series of Ljung-Box diagnoses reported in Tables 7 and 8, confirm the adequacy of our NC-LS model parameterizations.

Table 9 reports the volatility spillover test results. In the pre–Brexit period, we find that the spillover effect from the NA market to the EU market is statistically significant at the 10% level for $M = 10, 30$. This finding suggests that the NA market has nontrivial influences on the broader EU market in the short and long terms. In the opposite direction, the spillover effect from the EU market is significant at the 10% level for $M = 30$. In the post–Brexit sample, the spillover effect between the NA market and the EU market persists, with the impact from the latter occurs at a lower M.

Before Brexit, the feedback spillover in the NA–EU nexus can be largely attributed to the interconnected economic activities in the two regions. Therefore, uncertainty in one market would naturally spill to the other. Interestingly, the spillover effect from the EU is somehow delayed as the effect is not felt immediately by the NA market. One possible explanation for this finding is that, before Brexit, market participants in the NA can access the European Single Market seamlessly through the UK. Thus, they may tend to focus primarily on the UK, which results in their delayed response to the volatility in the EU market. However, we find that the NA–EU spillover nexus becomes more immediate after Brexit. This follows as NA participants switch their attention to EU directly for the Single Market, and naturally they show a more immediate response to EU.

In summary, our findings suggest that, before the Brexit referendum, participants in the NA market pay a relatively closer attention to the UK than the EU market. Consequently, the NA is driven more immediately by the volatility in the UK. After Brexit, the NA tends not to focus on the UK, and it prefers to follow the EU more closely. As a result, uncertainty in the UK does not significantly affect the NA while that in the EU has a more immediate effect on the NA. Our findings are consistent with the stylized fact that to a certain degree in the short term, Brexit has reduced UK's economic and geographic attracting power while the neighboring EU receives more attention (see, e.g., Hamre and Wright, 2021). Besides, our findings also agree with the conclusions in Breinlich et al. (2022) that Brexit has adversely affected the living costs in the UK, which could well be one of the main reasons that renders the region less attractive to businesses.

Table 7: Diagnostic tests (EU–NA)

	LB(10)	LB(20)	LB(30)	$LB^2(10)$	$LB^2(20)$	$LB^2(30)$
Pre–Brexit (2nd January 2012 – 23rd June 2016)						
Austria	8.839	11.259	29.522	1.821	2.540	15.942
	[0.547]	[0.939]	[0.490]	[0.998]	[1.000]	[0.983]
Belgium	7.327	17.561	39.203	2.063	14.773	31.616
	[0.694]	[0.616]	[0.121]	[0.996]	[0.789]	[0.386]
France	7.405	18.745	38.255	5.597	16.625	27.864
	[0.687]	[0.538]	[0.143]	[0.848]	[0.677]	[0.578]
Germany	4.424	20.098	37.650	1.810	16.234	26.026
	[0.926]	[0.452]	[0.159]	[0.998]	[0.702]	[0.674]
Ireland	7.429	21.941	37.667	4.151	21.944	40.053
	[0.684]	[0.344]	[0.158]	[0.940]	[0.344]	[0.104]
Italy	5.208	24.126	39.677	4.861	18.723	29.314
	[0.877]	[0.237]	[0.111]	[0.900]	[0.540]	[0.501]
Netherlands	7.255	19.837	37.514	2.467	14.265	21.344
	[0.701]	[0.468]	[0.163]	[0.991]	[0.817]	[0.877]
Spain	11.198	22.854	39.655	8.825	18.703	37.979
	[0.342]	[0.296]	[0.112]	[0.549]	[0.541]	[0.150]
US	13.084	21.296	36.398	15.271	22.915	34.113
	[0.219]	[0.380]	[0.195]	[0.122]	[0.293]	[0.276]
Canada	3.926	17.647	30.166	3.113	20.511	33.009
	[0.951]	[0.611]	[0.457]	[0.979]	[0.426]	[0.322]

NOTES: The table reports diagnostic analyses for all fitted series. LB(M) and $LB^2(M)$ are the Ljung-Box tests for the null of no serial correlation (up to lag order M) on the standardized and squared standardized residuals, respectively. The values in the squared parentheses are the p-values of the tests.

Table 8: Diagnostic tests (EU–NA)

	LB(10)	LB(20)	LB(30)	$LB^2(10)$	$LB^2(20)$	$LB^2(30)$
Post–Brexit (24th June 2016 – 31st December 2019)						
Austria	3.371	15.052	19.516	2.601	13.909	17.353
	[0.971]	[0.773]	[0.929]	[0.989]	[0.835]	[0.968]
Belgium	7.672	18.799	24.333	1.892	5.901	7.232
	[0.661]	[0.535]	[0.757]	[0.997]	[0.999]	[1.000]
France	10.970	16.590	27.751	5.436	9.278	14.264
	[0.360]	[0.679]	[0.584]	[0.860]	[0.979]	[0.993]
Germany	2.087	15.826	24.130	1.233	10.628	14.429
	[0.996]	[0.727]	[0.766]	[1.000]	[0.955]	[0.993]
Ireland	4.952	15.662	29.053	0.465	15.670	22.324
	[0.894]	[0.737]	[0.515]	[1.000]	[0.737]	[0.842]
Italy	5.015	21.839	26.462	2.595	11.920	18.647
	[0.890]	[0.349]	[0.651]	[0.989]	[0.919]	[0.947]
Netherlands	4.078	7.539	17.054	1.409	8.056	12.786
	[0.944]	[0.995]	[0.972]	[0.999]	[0.991]	[0.997]
Spain	6.255	21.479	23.440	2.728	15.164	25.255
	[0.793]	[0.369]	[0.797]	[0.987]	[0.767]	[0.713]
US	7.289	21.190	30.560	15.012	23.508	25.919
	[0.698]	[0.386]	[0.437]	[0.132]	[0.265]	[0.679]
Canada	5.511	27.150	33.990	3.446	26.552	30.657
	[0.855]	[0.131]	[0.281]	[0.969]	[0.148]	[0.432]

NOTES: The table reports diagnostic analyses for all fitted series. LB(M) and $LB^2(M)$ are the Ljung-Box tests for the null of no serial correlation (up to lag order M) on the standardized and squared standardized residuals, respectively. The values in the squared parentheses are the p-values of the tests.

Table 9: Spillover results (EU–NA)

	Pre–Brexit			*Post–Brexit*		
M	10	20	30	10	20	30
$Q^*_{1\mathrm{BAR}}$	0.046	0.124	0.070	0.078	0.128	0.182
$Q^*_{-1\mathrm{BAR}}$	0.210	0.144	0.072	0.132	0.080	0.078

NOTES: The table reports bootstrap p-values of the proposed spillover tests. Number of bootstraps = 499. $Q^*_{1\mathrm{BAR}}$ denotes the one-way test for the null hypothesis of no volatility spillover from the NA market to the EU market. $Q^*_{-1\mathrm{BAR}}$ denotes the one-way test for the null hypothesis of no volatility spillover from the EU market to the NA market. The subscript BAR denotes the Bartlett kernel. M denotes the kernel smoothing parameter.

6. Conclusions

We proposed a class of asymptotic $N(0,1)$ multivariate econometric strategy for testing volatility spillover. The test statistics were constructed based on the quadratic distance between a kernel-based spectral density estimator and the null spectral density. The proposed test statistics are straightforward to compute and they check a growing number of lags as the sample size increases. The Granger regression-type method can be viewed as a special case of the proposed procedure under the uniformly weighted Truncated kernel, but downward weighting kernels were proposed to be in line with stylized facts and thus to improve the power performance of the tests. Considering practical limitations in estimating multivariate volatility such as numerical convergence and dimensionality, we proposed a modeling approach that worked coherently with the spillover test. Consistent least-squares estimators that are computationally efficient were provided. Throughout the proposed econometric strategy, numerical optimization and integration are not required. The capacity of the multivariate testing strategy was highlighted using Monte Carlo experiments. First, it can check a large number of lags without losing significant power thanks to the use of downward weighting kernel functions. Second, it was found that the testing strategy performed reasonably well up to the tenth dimension. Furthermore, the paper provided a bootstrap version of the spillover tests whose size was found to converge at a faster speed. Finally, the paper included a empirical study in which the volatility spillover relations between the North America (NA) market and the greater European market (both UK and EU) before and after the Brexit referendum were examined. First, the proposed volatility modeling strategy adequately fitted the collected empirical data. Next, we found that the NA was driven more immediately by UK volatility than EU volatility in the pre-Brexit sample. After Brexit, it was found that volatility in the UK did not spill to NA while that in the EU had a more immediate spillover effect on NA. As market indices reflect the collective actions of the market participants, we could infer that on average — in contrast to the pre-Brexit sample — the NA participants appeared not to respond to the UK market but were following more closely the EU market in the post-Brexit period.

References

Admati, A.R., Pfleiderer, P., 1988. A theory of intraday patterns: Volume and price variability. Review of Financial Studies 1, 3–40.

Andrews, D.W., 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. Econometrica , 817–858.

Antoine, B., Renault, E., 2012. Efficient minimum distance estimation with multiple rates of convergence. Journal of Econometrics 170, 350–367.

Baele, L., 2005. Volatility spillover effects in European equity markets. Journal of Financial and Quantitative Analysis 40, 373–401.

Barunik, J., Krehlik, T., Vacha, L., 2016. Modeling and forecasting exchange rate volatility in time-frequency domain. European Journal of Operational Research 251, 329–340.

Bauer, D., Maynard, A., 2012. Persistence-robust surplus-lag Granger causality testing. Journal of Econometrics 169, 293–300.

Bekiros, S., Nguyen, D.K., Junior, L.S., Uddin, G.S., 2017. Information diffusion, cluster formation and entropy-based network dynamics in equity and commodity markets. European Journal of Operational Research 256, 945–961.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307–327.

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

Boudjellaba, H., Dufour, J.M., Roy, R., 1992. Testing causality between two vectors in multivariate autoregressive moving average models. Journal of the American Statistical Association 87, 1082–1090.

Boudt, K., Ghalanos, A., Payseur, S., Zivot, E., 2019. Multivariate GARCH models for large-scale applications: A survey, in: Vinod, H.D., Rao, C. (Eds.), Handbook of Statistics. Elsevier Science BV, Amsterdam. chapter 7, pp. 193–242.

Bouhaddioui, C., Roy, R., 2006. A generalized portmanteau test for independence of two infinite-order vector autoregressive series. Journal of Time Series Analysis 27, 505–544.

Breinlich, H., Leromain, E., Novy, D., Sampson, T., 2022. The Brexit vote, inflation and UK living standards. International Economic Review 63, 63–93.

Candelon, B., Tokpavi, S., 2016. A nonparametric test for Granger causality in distribution with application to financial contagion. Journal of Business & Economic Statistics 34, 240–253.

Chan, N.H., Deng, S.J., Peng, L., Xia, Z., 2007. Interval estimation of value-at-risk based on GARCH models with heavy-tailed innovations. Journal of Econometrics 137, 556–576.

Chen, B., Hong, Y., 2012a. Testing for smooth structural changes in time series models via nonparametric regression. Econometrica 80, 1157–1183.

Chen, B., Hong, Y., 2012b. Testing for the Markov property in time series. Econometric Theory 28, 130–178.

Chen, B., Hong, Y., 2016. Detecting for smooth structural changes in GARCH models. Econometric Theory 32, 740–791.

Comte, F., Lieberman, O., 2000. Second-order noncausality in multivariate GARCH processes. Journal of Time Series Analysis 21, 535–557.

Cumming, D.J., Zahra, S.A., 2016. International business and entrepreneurship implications of Brexit. British Journal of Management 27, 687–692.

Duchesne, P., Roy, R., 2004. On consistent testing for serial correlation of unknown form in vector time series models. Journal of Multivariate Analysis 89, 148–180.

Dufour, J.M., Pelletier, D., 2021. Practical methods for modeling weak VARMA processes: Identification, estimation and specification with a macroeconomic application. Journal of Business & Economic Statistics , 1–13.

Ellington, M., 2022. Fat tails, serial dependence, and implied volatility index connections. European Journal of Operational Research 299, 768–779.

Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica 50, 987–1007.

Engle, R.F., Bollerslev, T., 1986. Modelling the persistence of conditional variances. Econometric Reviews 5, 1–50.

Engle, R.F., Sheppard, K., 2001. Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH. Technical Report. National Bureau of Economic Research.

Fianu, E.S., Ahelegbey, D.F., Grossi, L., 2022. Modeling risk contagion in the Italian zonal electricity market. European Journal of Operational Research 298, 656–679.

Gao, F., Song, F., 2008. Estimation risk in GARCH VaR and ES estimates. Econometric Theory , 1404–1424.

Garbade, K.D., Silber, W.L., 1979. Structural organization of secondary markets: Clearing frequency, dealer activity and liquidity risk. Journal of Finance 34, 577–593.

Granger, C.W.J., 1969. Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37, 424–438.

Granger, C.W.J., 1980. Testing for causality: A personal viewpoint. Journal of Economic Dynamics and Control 2, 329–352.

Granger, C.W.J., Robins, R., Engle, R.F., 1986. Wholesale and retail prices: Bivariate time-series modeling with forecastable error variances, in: Belsley, D.A., Kuh, E. (Eds.), Model Reliability. MIT Press, Cambridge, MA. chapter 1, pp. 1–17.

Hafner, C.M., 2008. Temporal aggregation of multivariate GARCH processes. Journal of Econometrics 142, 467–483.

Hamre, E.F., Wright, W., 2021. Brexit & The City: The Impact So Far. New Financial URL: https://newfinancial.org/brexit-the-city-the-impact-so-far/.

Hannan, E.J., 1970. *Multiple Time Series.* John Wiley and Sons. New York.

Harville, D.A., 1997. *Matrix Algebra From a Statistician's Perspective.* Springer. New York.

Hong, Y., 1996. Testing for independence between two covariance stationary time series. Biometrika 83, 615–625.

Hong, Y., 2001. A test for volatility spillover with application to exchange rates. Journal of Econometrics 103, 183–224.

Hong, Y., Liu, Y., Wang, S., 2009. Granger causality in risk and detection of extreme risk spillover between financial markets. Journal of Econometrics 150, 271–287.

King, M.A., Wadhwani, S., 1990. Transmission of volatility between stock markets. Review of Financial Studies 3, 5–33.

Lewis, R., Reinsel, G.C., 1985. Prediction of multivariate time series by autoregressive model fitting. Journal of Multivariate Analysis 16, 393–411.

Li, W.K., Mak, T., 1994. On the squared residual autocorrelations in non-linear time series with conditional heteroskedasticity. Journal of Time Series Analysis 15, 627–636.

McCloud, N., Hong, Y., 2011. Testing the structure of conditional correlations in multivariate GARCH models: A generalized cross-spectrum approach. International Economic Review 52, 991–1037.

Newey, W.K., Steigerwald, D.G., 1997. Asymptotic bias for quasi-maximum-likelihood estimators in conditional heteroskedasticity models. Econometrica , 587–599.

Priestley, M.B., 1981. *Spectral Analysis and Time Series: Probability and Mathematical Statistics.* Academic press.

Robbins, M.W., Fisher, T.J., 2015. Cross-correlation matrices for tests of independence and causality between two multivariate time series. Journal of Business & Economic Statistics 33, 459–473.

Ross, S.A., 1989. Information and volatility: The no-arbitrage martingale approach to timing and resolution irrelevancy. Journal of Finance 44, 1–17.

Sévi, B., 2014. Forecasting the volatility of crude oil futures using intraday data. European Journal of Operational Research 235, 643–659.

Stoll, H.R., Whaley, R.E., 1990. Stock market structure and volatility. Review of Financial Studies 3, 37–71.

van der Vaart, A.W., 1998. *Asymptotic Statistics.* Cambridge University Press.

White, H., 2001. *Asymptotic Theory for Econometricians.* Academic press.

Wiener, N., Masani, P., 1957. The prediction theory of multivariate stochastic processes. Acta Mathematica 98, 111–150.

Appendix A Derivation of the normalized quadratic distance

Derivation of (15). Recall that the normalized quadratic distance is given as

$$\begin{aligned}
\hat{L}^2\big[\hat{\boldsymbol{f}}(\lambda), \hat{\boldsymbol{f}}^{\boldsymbol{0}}(\lambda)\big] &= 2\pi \int_{2\pi} \mathrm{vec}\Big[\overline{\hat{\boldsymbol{f}}(\lambda)} - \overline{\hat{\boldsymbol{f}}^{\boldsymbol{0}}(\lambda)}\Big]'\Big(\hat{\boldsymbol{\Gamma}}_{\boldsymbol{v}}^{-1} \otimes \hat{\boldsymbol{\Gamma}}_{\boldsymbol{u}}^{-1}\Big)\mathrm{vec}\Big[\hat{\boldsymbol{f}}(\lambda) - \hat{\boldsymbol{f}}^{\boldsymbol{0}}(\lambda)\Big]d\lambda \\
&= 2\pi \int_{2\pi} \mathrm{vec}\Big[\overline{\hat{\boldsymbol{f}}(\lambda) - \hat{\boldsymbol{f}}^{\boldsymbol{0}}(\lambda)}\Big]'\Big(\hat{\boldsymbol{\Gamma}}_{\boldsymbol{v}}^{-1} \otimes \hat{\boldsymbol{\Gamma}}_{\boldsymbol{u}}^{-1}\Big)\mathrm{vec}\Big[\hat{\boldsymbol{f}}(\lambda) - \hat{\boldsymbol{f}}^{\boldsymbol{0}}(\lambda)\Big]d\lambda \\
&= 2\pi \int_{2\pi} \mathrm{tr}\bigg\{\Big[\overline{\hat{\boldsymbol{f}}(\lambda) - \hat{\boldsymbol{f}}^{\boldsymbol{0}}(\lambda)}\Big]'\hat{\boldsymbol{\Gamma}}_{\boldsymbol{u}}^{-1}\Big[\hat{\boldsymbol{f}}(\lambda) - \hat{\boldsymbol{f}}^{\boldsymbol{0}}(\lambda)\Big]\hat{\boldsymbol{\Gamma}}_{\boldsymbol{v}}^{-1}\bigg\}d\lambda \\
&= 2\pi \int_{2\pi} \mathrm{tr}\bigg\{\overline{\Big[\hat{\boldsymbol{f}}(\lambda) - \hat{\boldsymbol{f}}^{\boldsymbol{0}}(\lambda)\Big]'}\hat{\boldsymbol{\Gamma}}_{\boldsymbol{u}}^{-1}\Big[\hat{\boldsymbol{f}}(\lambda) - \hat{\boldsymbol{f}}^{\boldsymbol{0}}(\lambda)\Big]\hat{\boldsymbol{\Gamma}}_{\boldsymbol{v}}^{-1}\bigg\}d\lambda \\
&= 2\pi \int_{2\pi} \mathrm{tr}\bigg\{\overline{\Big[\hat{\boldsymbol{f}}(\lambda) - \hat{\boldsymbol{f}}^{\boldsymbol{0}}(\lambda)\Big]'\hat{\boldsymbol{\Gamma}}_{\boldsymbol{u}}^{-1}}\Big[\hat{\boldsymbol{f}}(\lambda) - \hat{\boldsymbol{f}}^{\boldsymbol{0}}(\lambda)\Big]\hat{\boldsymbol{\Gamma}}_{\boldsymbol{v}}^{-1}\bigg\}d\lambda, \qquad \text{(A.1)}
\end{aligned}$$

where $\overline{\boldsymbol{f}}$ denotes the complex conjugate of $\boldsymbol{f}$. The second equality follows from the fact that complex conjugate of a sum of individuals is the sum of the complex conjugate of the individuals. The third equality follows from the matrix relation $\mathrm{tr}(\boldsymbol{A}'\boldsymbol{B}\boldsymbol{C}\boldsymbol{D}') = \big[\mathrm{vec}(\boldsymbol{A})\big]'(\boldsymbol{D} \otimes \boldsymbol{B})\big[\mathrm{vec}(\boldsymbol{C})\big]$, (see, e.g., Harville, 1997, Theorem 16.2.2). The fourth equality follows from the interchangeability of complex conjugation and transposition. The fifth equality follows from the fact that the complex conjugate of real matrix is the real matrix itself. Let $\boldsymbol{C}(\lambda) \equiv \big[\hat{\boldsymbol{f}}(\lambda) - \hat{\boldsymbol{f}}^{\boldsymbol{0}}(\lambda)\big]'\hat{\boldsymbol{\Gamma}}_{\boldsymbol{u}}^{-1}$, $\boldsymbol{D}(\lambda) \equiv \big[\hat{\boldsymbol{f}}(\lambda) - \hat{\boldsymbol{f}}^{\boldsymbol{0}}(\lambda)\big]\hat{\boldsymbol{\Gamma}}_{\boldsymbol{v}}^{-1}$, and put $\boldsymbol{A_j} \equiv (2\pi)^{-1}k(j/M)\hat{\boldsymbol{\rho}}(j)'\hat{\boldsymbol{\Gamma}}_{\boldsymbol{u}}^{-1}$, $\boldsymbol{B_j} \equiv (2\pi)^{-1}k(j/M)\hat{\boldsymbol{\rho}}(j)\hat{\boldsymbol{\Gamma}}_{\boldsymbol{v}}^{-1}$. We have $\boldsymbol{C}(\lambda) = \sum_{j=1}^{T-1}\boldsymbol{A_j}e^{-ij\lambda}$ and $\boldsymbol{D}(\lambda) = \sum_{j=1}^{T-1}\boldsymbol{B_j}e^{-ij\lambda}$, for $\lambda \in [-\pi, \pi]$ and $i = \sqrt{-1}$. We can rewrite (A.1) as

$$\begin{aligned}
\hat{L}^2\big[\hat{\boldsymbol{f}}(\lambda), \hat{\boldsymbol{f}}^{\boldsymbol{0}}(\lambda)\big] &= 2\pi \int_{2\pi} \mathrm{tr}\Big[\overline{\boldsymbol{C}(\lambda)}\boldsymbol{D}(\lambda)\Big]d\lambda \\
&= 2\pi \mathrm{tr}\bigg[\int_{2\pi} \overline{\boldsymbol{C}(\lambda)}\boldsymbol{D}(\lambda)d\lambda\bigg] \\
&= 2\pi \mathrm{tr}\bigg(2\pi \sum_{j=1}^{T-1} \overline{\boldsymbol{A_j}}\boldsymbol{B_j}\bigg) \\
&= (2\pi)^2 \mathrm{tr}\bigg(\sum_{j=1}^{T-1} \overline{\boldsymbol{A_j}}\boldsymbol{B_j}\bigg),
\end{aligned}$$

where the second equality follows from the interchangeability of trace and integral and the third equality follows from Parseval's identity (see, e.g., Wiener and Masani, 1957, Theorem 3.9). Substituting the relevant terms back into the normalized quadratic equation, we have

$$\begin{aligned}\hat{L}^2\big[\hat{\boldsymbol{f}}(\lambda), \hat{\boldsymbol{f}}^{\mathbf{0}}(\lambda)\big] &= (2\pi)^2 \mathrm{tr}\left[\sum_{j=1}^{T-1} \overline{\frac{1}{2\pi} k(j/M)\hat{\boldsymbol{\rho}}(j)'\hat{\boldsymbol{\Gamma}}_{\boldsymbol{u}}^{-1}} \frac{1}{2\pi} k(j/M)\hat{\boldsymbol{\rho}}(j)\hat{\boldsymbol{\Gamma}}_{\boldsymbol{v}}^{-1}\right] \\ &= \mathrm{tr}\left[\sum_{j=1}^{T-1} k^2(j/M)\hat{\boldsymbol{\rho}}(j)'\hat{\boldsymbol{\Gamma}}_{\boldsymbol{u}}^{-1}\hat{\boldsymbol{\rho}}(j)\hat{\boldsymbol{\Gamma}}_{\boldsymbol{v}}^{-1}\right] \\ &= \sum_{j=1}^{T-1} k^2(j/M)\mathrm{tr}\Big\{\hat{\boldsymbol{\rho}}(j)'\hat{\boldsymbol{\Gamma}}_{\boldsymbol{u}}^{-1}\hat{\boldsymbol{\rho}}(j)\hat{\boldsymbol{\Gamma}}_{\boldsymbol{v}}^{-1}\Big\} \\ &= \sum_{j=1}^{T-1} k^2(j/M)\mathrm{vec}\big[\hat{\boldsymbol{\rho}}(j)\big]'\big(\hat{\boldsymbol{\Gamma}}_{\boldsymbol{v}}^{-1} \otimes \hat{\boldsymbol{\Gamma}}_{\boldsymbol{u}}^{-1}\big)\mathrm{vec}\big[\hat{\boldsymbol{\rho}}(j)\big],\end{aligned}$$

where the fourth equality follows from the interchangeability of trace and summation. This completes the derivation. ■

Appendix B Proof of Theorem 2.1

Throughout Appendices B–G, the following notations are adopted. The inner product between vector $\boldsymbol{x_1}$ and vector $\boldsymbol{x_2}$ is denoted by $\langle \boldsymbol{x_1}, \boldsymbol{x_2}\rangle$. The Euclidean norm is denoted using $||\cdot||$. The notations O_p and o_p are the usual order in probability notations. The scalar Δ represents a positive finite generic constant that may differ at every occurrence.

Proof of Theorem 2.1. We let $\hat{S} \equiv T\hat{L}^2\big[\hat{\boldsymbol{f}}(\lambda), \hat{\boldsymbol{f}}^{\mathbf{0}}(\lambda)\big]$, $\boldsymbol{C}^{\mathbf{0}}_{\boldsymbol{uu}} \equiv \mathbb{E}[\boldsymbol{u}^{\mathbf{0}}_{\boldsymbol{t}}(\boldsymbol{u}^{\mathbf{0}}_{\boldsymbol{t}})']$ and $\boldsymbol{C}^{\mathbf{0}}_{\boldsymbol{vv}} \equiv \mathbb{E}[\boldsymbol{v}^{\mathbf{0}}_{\boldsymbol{t}}(\boldsymbol{v}^{\mathbf{0}}_{\boldsymbol{t}})']$, the proof begins by defining the following pseudo statistic

$$S^* = T\sum_{j=1}^{T-1} k^2(j/M)\mathrm{vec}\big[\hat{\boldsymbol{\rho}}^*(j)\big]'\big(\boldsymbol{\Gamma}_{\boldsymbol{v}}^{-1} \otimes \boldsymbol{\Gamma}_{\boldsymbol{u}}^{-1}\big)\mathrm{vec}\big[\hat{\boldsymbol{\rho}}^*(j)\big], \tag{B.1}$$

where $\hat{\boldsymbol{\rho}}^*(j) = \mathrm{Diag}\big(\boldsymbol{C}^{\mathbf{0}}_{\boldsymbol{uu}}\big)^{-1/2}\hat{\boldsymbol{C}}_{\boldsymbol{uv}}(j)\mathrm{Diag}\big(\boldsymbol{C}^{\mathbf{0}}_{\boldsymbol{vv}}\big)^{-1/2}$; and $\boldsymbol{\Gamma}_{\boldsymbol{u}} = \mathrm{Diag}\big(\boldsymbol{C}^{\mathbf{0}}_{\boldsymbol{uu}}\big)^{-1/2}\boldsymbol{C}^{\mathbf{0}}_{\boldsymbol{uu}}\mathrm{Diag}\big(\boldsymbol{C}^{\mathbf{0}}_{\boldsymbol{uu}}\big)^{-1/2}$ and $\boldsymbol{\Gamma}_{\boldsymbol{v}} = \mathrm{Diag}\big(\boldsymbol{C}^{\mathbf{0}}_{\boldsymbol{vv}}\big)^{-1/2}\boldsymbol{C}^{\mathbf{0}}_{\boldsymbol{vv}}\mathrm{Diag}\big(\boldsymbol{C}^{\mathbf{0}}_{\boldsymbol{vv}}\big)^{-1/2}$ are the true correlation matrices of true $\boldsymbol{u}^{\mathbf{0}}_{\boldsymbol{t}}$ and $\boldsymbol{v}^{\mathbf{0}}_{\boldsymbol{t}}$, respectively. We can decompose Q_1 as

$$Q_1 = \frac{S^* - d_1^* d_2^* C_{1T}(k)}{\big[d_1^* d_2^* D_{1T}(k)\big]^{1/2}} + \frac{\hat{S} - S^*}{\big[d_1^* d_2^* D_{1T}(k)\big]^{1/2}}. \tag{B.2}$$

Then, the result of Theorem 2.1 follows from Propositions B.1–B.2. ■

Proposition B.1. *Suppose the conditions of Theorem 2.1 hold, we have that*

$$\frac{S^* - d_1^* d_2^* C_{1T}(k)}{\big[d_1^* d_2^* D_{1T}(k)\big]^{1/2}} \xrightarrow{d} \mathrm{N}(0,1).$$

Proposition B.2. *Suppose the conditions of Theorem 2.1 hold, we have that*

$$\frac{\hat{S}-S^*}{\big[d_1^*d_2^*D_{1T}(k)\big]^{1/2}} \xrightarrow{p} 0.$$

Proof of Proposition B.1. Let $\hat{\boldsymbol{C}}^{\boldsymbol{0}}_{\boldsymbol{uv}}$ denotes the sample cross-covariance matrix in (14) with true $\boldsymbol{u}_{\boldsymbol{t}}^{\boldsymbol{0}}$ and $\boldsymbol{v}_{\boldsymbol{t-j}}^{\boldsymbol{0}}$. The proof of Proposition B.1 begins by defining the another pseudo statistic

$$S = T\sum_{j=1}^{T-1} k^2(j/M)\mathrm{vec}\big[\hat{\boldsymbol{\rho}}^{\boldsymbol{0}}(j)\big]'\big(\boldsymbol{\Gamma}_{\boldsymbol{v}}^{-1}\otimes\boldsymbol{\Gamma}_{\boldsymbol{u}}^{-1}\big)\mathrm{vec}\big[\hat{\boldsymbol{\rho}}^{\boldsymbol{0}}(j)\big], \tag{B.3}$$

where $\hat{\boldsymbol{\rho}}^{\boldsymbol{0}}(j) = \mathrm{Diag}\big(\boldsymbol{C}^{\boldsymbol{0}}_{\boldsymbol{uu}}\big)^{-1/2}\hat{\boldsymbol{C}}^{\boldsymbol{0}}_{\boldsymbol{uv}}(j)\mathrm{Diag}\big(\boldsymbol{C}^{\boldsymbol{0}}_{\boldsymbol{vv}}\big)^{-1/2}$. We consider a similar decomposition

$$\frac{S^*-d_1^*d_2^*C_{1T}(k)}{\big[d_1^*d_2^*D_{1T}(k)\big]^{1/2}} = \frac{S-d_1^*d_2^*C_{1T}(k)}{\big[d_1^*d_2^*D_{1T}(k)\big]^{1/2}} + \frac{S^*-S}{\big[d_1^*d_2^*D_{1T}(k)\big]^{1/2}}. \tag{B.4}$$

The result of Proposition B.1 is given by Lemmas B.1 and B.2. ■

Lemma B.1. *Suppose the conditions of Theorem 2.1 hold, we have that*

$$\frac{S-d_1^*d_2^*C_{1T}(k)}{\big[d_1^*d_2^*D_{1T}(k)\big]^{1/2}} \xrightarrow{d} \mathrm{N}(0,1).$$

Lemma B.2. *Suppose the conditions of Theorem 2.1 hold, we have that*

$$S^*-S = o_p(M^{1/2}).$$

Proof of Lemma B.1. We begin by showing the covariance representation of S. Using the properties $\mathrm{vec}(\boldsymbol{AXB}) = (\boldsymbol{B}'\otimes\boldsymbol{A})\mathrm{vec}(\boldsymbol{X})$; $(\boldsymbol{A}\otimes\boldsymbol{B})' = \boldsymbol{A}'\otimes\boldsymbol{B}'$; $(\boldsymbol{A}')^{-1} = (\boldsymbol{A}^{-1})'$; $(\boldsymbol{A}\otimes\boldsymbol{B})(\boldsymbol{C}\otimes\boldsymbol{D}) = (\boldsymbol{AC})\otimes(\boldsymbol{BD})$; $(\boldsymbol{ABC})^{-1} = \boldsymbol{C}^{-1}\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$; $(\boldsymbol{A}\otimes\boldsymbol{B})^{-1} = \boldsymbol{A}^{-1}\otimes\boldsymbol{B}^{-1}$, we write for the correlation components of S

$$\begin{aligned}
&\mathrm{vec}\big[\hat{\boldsymbol{\rho}}^{\boldsymbol{0}}(j)\big]'\big(\boldsymbol{\Gamma}_{\boldsymbol{v}}^{-1}\otimes\boldsymbol{\Gamma}_{\boldsymbol{u}}^{-1}\big)\mathrm{vec}\big[\hat{\boldsymbol{\rho}}^{\boldsymbol{0}}(j)\big]\\
&\quad= \mathrm{vec}\big[\hat{\boldsymbol{C}}^{\boldsymbol{0}}_{\boldsymbol{uv}}(j)\big]'\big[\mathrm{Diag}(\boldsymbol{C}^{\boldsymbol{0}}_{\boldsymbol{vv}})^{-1/2}\otimes\mathrm{Diag}(\boldsymbol{C}^{\boldsymbol{0}}_{\boldsymbol{uu}})^{-1/2}\big]\big(\boldsymbol{\Gamma}_{\boldsymbol{v}}^{-1}\otimes\boldsymbol{\Gamma}_{\boldsymbol{u}}^{-1}\big)\mathrm{vec}\big[\hat{\boldsymbol{\rho}}^{\boldsymbol{0}}(j)\big]\\
&\quad= \mathrm{vec}\big[\hat{\boldsymbol{C}}^{\boldsymbol{0}}_{\boldsymbol{uv}}(j)\big]'\big[\mathrm{Diag}(\boldsymbol{C}^{\boldsymbol{0}}_{\boldsymbol{vv}})^{-1/2}\boldsymbol{\Gamma}_{\boldsymbol{v}}^{-1}\big]\otimes\big[\mathrm{Diag}(\boldsymbol{C}^{\boldsymbol{0}}_{\boldsymbol{uu}})^{-1/2}\boldsymbol{\Gamma}_{\boldsymbol{u}}^{-1}\big]\mathrm{vec}\big[\hat{\boldsymbol{\rho}}^{\boldsymbol{0}}(j)\big]\\
&\quad= \mathrm{vec}\big[\hat{\boldsymbol{C}}^{\boldsymbol{0}}_{\boldsymbol{uv}}(j)\big]'\big\{(\boldsymbol{C}^{\boldsymbol{0}}_{\boldsymbol{vv}})^{-1}\big[\mathrm{Diag}(\boldsymbol{C}^{\boldsymbol{0}}_{\boldsymbol{vv}})^{-1/2}\big]^{-1}\big\}\otimes\big\{(\boldsymbol{C}^{\boldsymbol{0}}_{\boldsymbol{uu}})^{-1}\big[\mathrm{Diag}(\boldsymbol{C}^{\boldsymbol{0}}_{\boldsymbol{uu}})^{-1/2}\big]^{-1}\big\}\mathrm{vec}\big[\hat{\boldsymbol{\rho}}^{\boldsymbol{0}}(j)\big]\\
&\quad= \mathrm{vec}\big[\hat{\boldsymbol{C}}^{\boldsymbol{0}}_{\boldsymbol{uv}}(j)\big]'\big[(\boldsymbol{C}^{\boldsymbol{0}}_{\boldsymbol{vv}})^{-1}\otimes(\boldsymbol{C}^{\boldsymbol{0}}_{\boldsymbol{uu}})^{-1}\big]\big\{\big[\mathrm{Diag}(\boldsymbol{C}^{\boldsymbol{0}}_{\boldsymbol{vv}})^{-1/2}\big]^{-1}\otimes\big[\mathrm{Diag}(\boldsymbol{C}^{\boldsymbol{0}}_{\boldsymbol{uu}})^{-1/2}\big]^{-1}\big\}\mathrm{vec}\big[\hat{\boldsymbol{\rho}}^{\boldsymbol{0}}(j)\big]\\
&\quad= \mathrm{vec}\big[\hat{\boldsymbol{C}}^{\boldsymbol{0}}_{\boldsymbol{uv}}(j)\big]'\big[(\boldsymbol{C}^{\boldsymbol{0}}_{\boldsymbol{vv}})^{-1}\otimes(\boldsymbol{C}^{\boldsymbol{0}}_{\boldsymbol{uu}})^{-1}\big]\big\{\big[\mathrm{Diag}(\boldsymbol{C}^{\boldsymbol{0}}_{\boldsymbol{vv}})^{-1/2}\big]\otimes\big[\mathrm{Diag}(\boldsymbol{C}^{\boldsymbol{0}}_{\boldsymbol{uu}})^{-1/2}\big]\big\}^{-1}\mathrm{vec}\big[\hat{\boldsymbol{\rho}}^{\boldsymbol{0}}(j)\big]\\
&\quad= \mathrm{vec}\big[\hat{\boldsymbol{C}}^{\boldsymbol{0}}_{\boldsymbol{uv}}(j)\big]'\big[(\boldsymbol{C}^{\boldsymbol{0}}_{\boldsymbol{vv}})^{-1}\otimes(\boldsymbol{C}^{\boldsymbol{0}}_{\boldsymbol{uu}})^{-1}\big]\big\{\big[\mathrm{Diag}(\boldsymbol{C}^{\boldsymbol{0}}_{\boldsymbol{vv}})^{-1/2}\big]\otimes\big[\mathrm{Diag}(\boldsymbol{C}^{\boldsymbol{0}}_{\boldsymbol{uu}})^{-1/2}\big]\big\}^{-1}\\
&\qquad\times\big[\mathrm{Diag}(\boldsymbol{C}^{\boldsymbol{0}}_{\boldsymbol{vv}})^{-1/2}\otimes\mathrm{Diag}(\boldsymbol{C}^{\boldsymbol{0}}_{\boldsymbol{uu}})^{-1/2}\big]\mathrm{vec}\big[\hat{\boldsymbol{C}}^{\boldsymbol{0}}_{\boldsymbol{uv}}(j)\big]\\
&\quad= \mathrm{vec}\big[\hat{\boldsymbol{C}}^{\boldsymbol{0}}_{\boldsymbol{uv}}(j)\big]'\big[(\boldsymbol{C}^{\boldsymbol{0}}_{\boldsymbol{vv}})^{-1}\otimes(\boldsymbol{C}^{\boldsymbol{0}}_{\boldsymbol{uu}})^{-1}\big]\mathrm{vec}\big[\hat{\boldsymbol{C}}^{\boldsymbol{0}}_{\boldsymbol{uv}}(j)\big].
\end{aligned} \tag{B.5}$$

We have

$$S = T\sum_{j=1}^{T-1} k^2(j/M)\text{vec}\big[\hat{\boldsymbol{\rho}}^{\mathbf{0}}(j)\big]'\big(\boldsymbol{\Gamma}_{\boldsymbol{v}}^{-1}\otimes\boldsymbol{\Gamma}_{\boldsymbol{u}}^{-1}\big)\text{vec}\big[\hat{\boldsymbol{\rho}}^{\mathbf{0}}(j)\big]$$
$$= T\sum_{j=1}^{T-1} k^2(j/M)\text{vec}\big[\hat{\boldsymbol{C}}^{\mathbf{0}}_{\boldsymbol{uv}}(j)\big]'\big[(\boldsymbol{C}^{\mathbf{0}}_{\boldsymbol{vv}})^{-1}\otimes(\boldsymbol{C}^{\mathbf{0}}_{\boldsymbol{uu}})^{-1}\big]\text{vec}\big[\hat{\boldsymbol{C}}^{\mathbf{0}}_{\boldsymbol{uv}}(j)\big]. \quad \text{(B.6)}$$

With this representation, the result of Lemma B.1 follows from Lemma 2.2 in Candelon and Tokpavi (2016), which is based on Lemma 1 in Bouhaddioui and Roy (2006). In both papers, the asymptotic normality result is obtained under the following conditions: (i) The event variables $\boldsymbol{u}_t^{\mathbf{0}}$ and $\boldsymbol{v}_t^{\mathbf{0}}$ are multivariate i.i.d. sequences with finite fourth order moment. (ii) Mutual independence between $\boldsymbol{u}_t^{\mathbf{0}}$ and $\boldsymbol{v}_{t-j}^{\mathbf{0}}$ for $j > 0$. In our framework, condition (i) is satisfied given Assumption 2.1, and condition (ii) is satisfied under the null hypothesis, this completes the proof. ■

Proof of Lemma B.2. We begin by defining the following notations. Let $\boldsymbol{b_{1t}} \equiv (\boldsymbol{C}^{\mathbf{0}}_{\boldsymbol{uu}})^{-1/2}\boldsymbol{u}_t^{\mathbf{0}}$ and $\boldsymbol{b_{2t}} \equiv (\boldsymbol{C}^{\mathbf{0}}_{\boldsymbol{vv}})^{-1/2}\boldsymbol{v}_t^{\mathbf{0}}$. Similarly, we let $\hat{\boldsymbol{b}}_{\boldsymbol{1t}} \equiv (\boldsymbol{C}^{\mathbf{0}}_{\boldsymbol{uu}})^{-1/2}\hat{\boldsymbol{u}}_t$ and $\hat{\boldsymbol{b}}_{\boldsymbol{2t}} \equiv (\boldsymbol{C}^{\mathbf{0}}_{\boldsymbol{vv}})^{-1/2}\hat{\boldsymbol{v}}_t$, denote the analogues of $\boldsymbol{b_{1t}}$ and $\boldsymbol{b_{2t}}$ based on estimated event variables $\hat{\boldsymbol{u}}_t$ and $\hat{\boldsymbol{v}}_t$. Then, we obtain $\boldsymbol{C}_{\hat{\boldsymbol{b}}}(j) \equiv (\boldsymbol{C}^{\mathbf{0}}_{\boldsymbol{uu}})^{-1/2}\hat{\boldsymbol{C}}_{\boldsymbol{uv}}(j)(\boldsymbol{C}^{\mathbf{0}}_{\boldsymbol{vv}})^{-1/2}$, the sample cross-covariance matrix between $\hat{\boldsymbol{b}}_{\boldsymbol{1t}}$ and $\hat{\boldsymbol{b}}_{\boldsymbol{2t}}$ at lag order j. Similarly, we have $\boldsymbol{C}_{\boldsymbol{b}}(j) \equiv (\boldsymbol{C}^{\mathbf{0}}_{\boldsymbol{uu}})^{-1/2}\hat{\boldsymbol{C}}^{\mathbf{0}}_{\boldsymbol{uv}}(j)(\boldsymbol{C}^{\mathbf{0}}_{\boldsymbol{vv}})^{-1/2}$, the sample cross-covariance matrix between $\boldsymbol{b_{1t}}$ and $\boldsymbol{b_{2t}}$ at lag order j. By reasonings similar to the derivation of (B.5) in Lemma B.1, we write S^* in terms of covariances

$$S^* = T\sum_{j=1}^{T-1} k^2(j/M)\text{vec}\big[\hat{\boldsymbol{\rho}}^{*}(j)\big]'\big(\boldsymbol{\Gamma}_{\boldsymbol{v}}^{-1}\otimes\boldsymbol{\Gamma}_{\boldsymbol{u}}^{-1}\big)\text{vec}\big[\hat{\boldsymbol{\rho}}^{*}(j)\big]$$
$$= T\sum_{j=1}^{T-1} k^2(j/M)\text{vec}\big[\hat{\boldsymbol{C}}_{\boldsymbol{uv}}(j)\big]'\big[(\boldsymbol{C}^{\mathbf{0}}_{\boldsymbol{vv}})^{-1}\otimes(\boldsymbol{C}^{\mathbf{0}}_{\boldsymbol{uu}})^{-1}\big]\text{vec}\big[\hat{\boldsymbol{C}}_{\boldsymbol{uv}}(j)\big]. \quad \text{(B.7)}$$

We can express $S^* - S$ as

$$S^* - S = T\sum_{j=1}^{T-1} k^2(j/M)\big|\big|\text{vec}[\boldsymbol{C}_{\hat{\boldsymbol{b}}}(j)] - \text{vec}[\boldsymbol{C}_{\boldsymbol{b}}(j)]\big|\big|^2$$
$$+ 2T\sum_{j=1}^{T-1} k^2(j/M)\big\langle\text{vec}[\boldsymbol{C}_{\boldsymbol{b}}(j)], \text{vec}[\boldsymbol{C}_{\hat{\boldsymbol{b}}}(j)] - \text{vec}[\boldsymbol{C}_{\boldsymbol{b}}(j)]\big\rangle$$
$$= \mathcal{A}_{1T} + 2\mathcal{A}_{2T}, \text{ say.} \quad \text{(B.8)}$$

We shall show that both $\mathcal{A}_{1T}$ and $\mathcal{A}_{2T}$ are $o_p(M^{1/2})$. The first term $\mathcal{A}_{1T}$ can be written as

$$\mathcal{A}_{1T} = T\sum_{j=1}^{T-1} k^2(j/M)$$

$$\times \left|\left|(\boldsymbol{C^0_{vv}})^{-1/2} \otimes (\boldsymbol{C^0_{uu}})^{-1/2} \,\mathrm{vec}[\hat{\boldsymbol{C}}_{\boldsymbol{uv}}(j)] - (\boldsymbol{C^0_{vv}})^{-1/2} \otimes (\boldsymbol{C^0_{uu}})^{-1/2} \,\mathrm{vec}[\hat{\boldsymbol{C}}^{\boldsymbol{0}}_{\boldsymbol{uv}}(j)]\right|\right|^2$$

$$= T \sum_{j=1}^{T-1} k^2(j/M) \left|\left| \big[(\boldsymbol{C^0_{vv}})^{-1/2} \otimes (\boldsymbol{C^0_{uu}})^{-1/2}\big] \big\{\mathrm{vec}[\hat{\boldsymbol{C}}_{\boldsymbol{uv}}(j)] - \mathrm{vec}[\hat{\boldsymbol{C}}^{\boldsymbol{0}}_{\boldsymbol{uv}}(j)]\big\} \right|\right|^2$$

$$\leq T \sum_{j=1}^{T-1} k^2(j/M) \left|\left|(\boldsymbol{C^0_{vv}})^{-1/2} \otimes (\boldsymbol{C^0_{uu}})^{-1/2}\right|\right|^2 \left|\left|\mathrm{vec}[\hat{\boldsymbol{C}}_{\boldsymbol{uv}}(j)] - \mathrm{vec}[\hat{\boldsymbol{C}}^{\boldsymbol{0}}_{\boldsymbol{uv}}(j)]\right|\right|^2, \quad \text{(B.9)}$$

which we make use the property $\mathrm{vec}(\boldsymbol{AXB}) = (\boldsymbol{B'} \otimes \boldsymbol{A})\mathrm{vec}(\boldsymbol{X})$ and Cauchy-Schwarz inequality. Because $\big|\big|(\boldsymbol{C^0_{vv}})^{-1/2} \otimes (\boldsymbol{C^0_{uu}})^{-1/2}\big|\big|^2 = O_p(1)$ by Assumption 2.1, it suffices to show that $\mathcal{A}_{11T} = o_p(M^{1/2})$, with

$$\mathcal{A}_{11T} = T \sum_{j=1}^{T-1} k^2(j/M) \left|\left|\mathrm{vec}[\hat{\boldsymbol{C}}_{\boldsymbol{uv}}(j)] - \mathrm{vec}[\hat{\boldsymbol{C}}^{\boldsymbol{0}}_{\boldsymbol{uv}}(j)]\right|\right|^2$$

$$= T \sum_{m=1}^{d_1^*} \sum_{n=1}^{d_2^*} \sum_{j=1}^{T-1} k^2(j/M) \big[\hat{C}^{m,n}_{uv}(j) - \hat{C}^{0,m,n}_{uv}(j)\big]^2, \quad \text{(B.10)}$$

where $\hat{C}^{m,n}_{uv}(j)$ and $\hat{C}^{0,m,n}_{uv}(j)$ are the (m,n)-th elements of matrices $\hat{\boldsymbol{C}}_{\boldsymbol{uv}}(j)$ and $\hat{\boldsymbol{C}}^{\boldsymbol{0}}_{\boldsymbol{uv}}(j)$, respectively. It suffices to show that $\sum_{j=1}^{T-1} k^2(j/M)\big[\hat{C}^{m,n}_{uv}(j) - \hat{C}^{0,m,n}_{uv}(j)\big]^2 = o_p(M^{1/2}/T)$. Let $u^0_{m,t}$ and $\hat{u}_{m,t}$ denote the m-th element of $\boldsymbol{u^0_t}$ and $\hat{\boldsymbol{u}}_{\boldsymbol{t}}$, respectively. Similarly, let $v^0_{n,t}$ and $\hat{v}_{n,t}$ denote the n-th element of $\boldsymbol{v^0_t}$ and $\hat{\boldsymbol{v}}_{\boldsymbol{t}}$, respectively. We have

$$\hat{C}^{m,n}_{uv}(j) - \hat{C}^{0,m,n}_{uv}(j) = \frac{1}{T} \sum_{t=j+1}^{T} \hat{u}_{m,t}\hat{v}_{n,t-j} - u^0_{m,t}v^0_{n,t-j}$$

$$= \frac{1}{T} \sum_{t=j+1}^{T} \big(\hat{u}_{m,t} - u^0_{m,t}\big)v^0_{n,t-j} + \frac{1}{T} \sum_{t=j+1}^{T} u^0_{m,t}\big(\hat{v}_{n,t-j} - v^0_{n,t-j}\big)$$

$$+ \frac{1}{T} \sum_{t=j+1}^{T} \big(\hat{u}_{m,t} - u^0_{m,t}\big)\big(\hat{v}_{n,t-j} - v^0_{n,t-j}\big)$$

$$= \mathcal{B}_{1T}(j) + \mathcal{B}_{2T}(j) + \mathcal{B}_{3T}(j), \text{ say.} \quad \text{(B.11)}$$

It follows that

$$\sum_{j=1}^{T-1} k^2(j/M)\big[\hat{C}^{m,n}_{uv}(j) - \hat{C}^{0,m,n}_{uv}(j)\big]^2 \leq \Delta \sum_{j=1}^{T-1} k^2(j/M)\big[\mathcal{B}^2_{1T}(j) + \mathcal{B}^2_{2T}(j) + \mathcal{B}^2_{3T}(j)\big]. \quad \text{(B.12)}$$

Applying Cauchy-Schwarz inequality to the last term $\mathcal{B}^2_{3T}(j)$, we have

$$\sup_{1\leq j\leq T-1} \mathcal{B}^2_{3T}(j) \leq \left[\frac{1}{T} \sum_{t=1}^{T} \big(\hat{u}_{m,t} - u^0_{m,t}\big)^2\right]\left[\frac{1}{T} \sum_{t=1}^{T} \big(\hat{v}_{n,t} - v^0_{n,t}\big)^2\right].$$

We shall show that $T^{-1}\sum_{t=1}^{T}\left(\hat{u}_{m,t}-u^0_{m,t}\right)^2 = O_p(T^{-1})$. The proof for $T^{-1}\sum_{t=1}^{T}\left(\hat{v}_{n,t}-v^0_{n,t}\right)^2$ is the same. Using Cauchy-Schwarz inequality and noting that $\hat{u}_{m,t}-u^0_{m,t} = u_{m,t}(\hat{\boldsymbol{\theta}}_{\mathbf{1}})-u^0_{m,t} = \left[u_{m,t}(\hat{\boldsymbol{\theta}}_{\mathbf{1}})-\tilde{u}_{m,t}(\hat{\boldsymbol{\theta}}_{\mathbf{1}})\right]+\left[\tilde{u}_{m,t}(\hat{\boldsymbol{\theta}}_{\mathbf{1}})-u^0_{m,t}\right]$, we have

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T}\left[u_{m,t}(\hat{\boldsymbol{\theta}}_{\mathbf{1}})-u^0_{m,t}\right]^2 &\leq 2\frac{1}{T}\sum_{t=1}^{T}\left[u_{m,t}(\hat{\boldsymbol{\theta}}_{\mathbf{1}})-\tilde{u}_{m,t}(\hat{\boldsymbol{\theta}}_{\mathbf{1}})\right]^2\\ &\quad+2\frac{1}{T}\sum_{t=1}^{T}\left[\tilde{u}_{m,t}(\hat{\boldsymbol{\theta}}_{\mathbf{1}})-u^0_{m,t}\right]^2\\ &= 2\mathcal{B}_{31T}(j)+2\mathcal{B}_{32T}(j), \text{ say.}\end{aligned} \tag{B.13}$$

We have $\mathcal{B}_{31T}(j) = O_p(T^{-2})$ by Assumption 2.3, it remains to show that $\mathcal{B}_{32T}(j) = O_p(T^{-1})$. By the mean value theorem and Cauchy-Schwarz inequality, we have

$$\mathcal{B}_{32T}(j) \leq \left|\left|\hat{\boldsymbol{\theta}}_{\mathbf{1}}-\boldsymbol{\theta}^{\mathbf{0}}_{\mathbf{1}}\right|\right|^2\left(\frac{1}{T}\sum_{t=1}^{T}\left|\left|\nabla_{\boldsymbol{\theta}_{\mathbf{1}}}\tilde{u}_{m,t}(\bar{\boldsymbol{\theta}}_{\mathbf{1}})\right|\right|^2\right), \tag{B.14}$$

where $\nabla_{\boldsymbol{\theta}_{\mathbf{1}}}$ is the gradient operator with respect to $\boldsymbol{\theta}_{\mathbf{1}}$ and $\bar{\boldsymbol{\theta}}_{\mathbf{1}}$ lies in the segment between $\hat{\boldsymbol{\theta}}_{\mathbf{1}}$ and $\boldsymbol{\theta}^{\mathbf{0}}_{\mathbf{1}}$. Given Assumption 2.2, we have $\left|\left|\hat{\boldsymbol{\theta}}_{\mathbf{1}}-\boldsymbol{\theta}^{\mathbf{0}}_{\mathbf{1}}\right|\right|^2 = O_p(T^{-1})$. Given Assumption 2.4, we have $T^{-1}\sum_{t=1}^{T}\left|\left|\nabla_{\boldsymbol{\theta}_{\mathbf{1}}}\tilde{u}_{m,t}(\bar{\boldsymbol{\theta}}_{\mathbf{1}})\right|\right|^2 = O_p(1)$ by Markov's inequality. Therefore, we have $\mathcal{B}_{32T}(j) = O_p(T^{-1})$. Subsequently,

$$\begin{aligned}\sum_{j=1}^{T-1}k^2(j/M)\mathcal{B}^2_{3T}(j) &\leq M\sup_{1\leq j\leq T-1}\mathcal{B}^2_{3T}(j)\left[\frac{1}{M}\sum_{j=1}^{T-1}k^2(j/M)\right]\\ &= O_p(M/T^2),\end{aligned} \tag{B.15}$$

where $M^{-1}\sum_{j=1}^{T-1}k^2(j/M)\to\int_0^\infty k^2(z)dz<\infty$ follows by Assumptions 2.5-2.6.

Next, we rewrite $\mathcal{B}_{1T}(j)$ as

$$\begin{aligned}\mathcal{B}_{1T}(j) &= \frac{1}{T}\sum_{t=j+1}^{T}\left[u_{m,t}(\hat{\boldsymbol{\theta}}_{\mathbf{1}})-\tilde{u}_{m,t}(\hat{\boldsymbol{\theta}}_{\mathbf{1}})\right]v^0_{n,t-j}+\frac{1}{T}\sum_{t=j+1}^{T}\left[\tilde{u}_{m,t}(\hat{\boldsymbol{\theta}}_{\mathbf{1}})-u^0_{m,t}\right]v^0_{n,t-j}\\ &= \mathcal{B}_{11T}(j)+\mathcal{B}_{12T}(j), \text{ say.}\end{aligned} \tag{B.16}$$

Applying Cauchy-Schwarz inequality to the first term $\mathcal{B}_{11T}(j)$, we have

$$
\begin{aligned}
&\sum_{j=1}^{T-1} k^2(j/M)\mathcal{B}_{11T}^2(j) \\
&\quad \le \frac{1}{T^2}\sum_{j=1}^{T-1} k^2(j/M)\left\{\sum_{t=j+1}^{T}\left[u_{m,t}(\hat{\boldsymbol{\theta}}_{\mathbf{1}}) - \tilde{u}_{m,t}(\hat{\boldsymbol{\theta}}_{\mathbf{1}})\right]^2\right\}\left[\sum_{t=j+1}^{T}\left(v_{n,t-j}^0\right)^2\right] \\
&\quad \le \frac{1}{T}\sum_{j=1}^{T-1} k^2(j/M)\left\{\sum_{t=1}^{T}\left[u_{m,t}(\hat{\boldsymbol{\theta}}_{\mathbf{1}}) - \tilde{u}_{m,t}(\hat{\boldsymbol{\theta}}_{\mathbf{1}})\right]^2\right\}\left[\frac{1}{T}\sum_{t=1}^{T}\left(v_{n,t}^0\right)^2\right] \\
&\quad = O_p(M/T^2), \qquad \text{(B.17)}
\end{aligned}
$$

given Assumption 2.3, and $T^{-1}\sum_{t=1}^{T}\left(v_{n,t}^0\right)^2 = O_p(1)$ by Markov's inequality. Applying two-term Taylor expansion to the second term $\mathcal{B}_{12T}(j)$, we have

$$
\begin{aligned}
\mathcal{B}_{12T}(j) &= \left(\hat{\boldsymbol{\theta}}_{\mathbf{1}} - \boldsymbol{\theta}_{\mathbf{1}}^{\mathbf{0}}\right)'\frac{1}{T}\sum_{t=j+1}^{T}\nabla_{\boldsymbol{\theta}_{\mathbf{1}}}\tilde{u}_{m,t}(\boldsymbol{\theta}_{\mathbf{1}}^{\mathbf{0}})v_{n,t-j}^0 \\
&\quad + \frac{1}{2}\left(\hat{\boldsymbol{\theta}}_{\mathbf{1}} - \boldsymbol{\theta}_{\mathbf{1}}^{\mathbf{0}}\right)'\left[\frac{1}{T}\sum_{t=j+1}^{T}\nabla_{\boldsymbol{\theta}_{\mathbf{1}}}^2\tilde{u}_{m,t}(\bar{\boldsymbol{\theta}}_{\mathbf{1}})v_{n,t-j}^0\right]\left(\hat{\boldsymbol{\theta}}_{\mathbf{1}} - \boldsymbol{\theta}_{\mathbf{1}}^{\mathbf{0}}\right) \\
&= \mathcal{B}_{121T}(j) + \mathcal{B}_{122T}(j), \text{ say,} \qquad \text{(B.18)}
\end{aligned}
$$

where $\nabla_{\boldsymbol{\theta}_{\mathbf{1}}}^2$ is the Hessian operator with respect to $\boldsymbol{\theta}_{\mathbf{1}}$ and $\bar{\boldsymbol{\theta}}_{\mathbf{1}}$ lies in the segment between $\hat{\boldsymbol{\theta}}_{\mathbf{1}}$ and $\boldsymbol{\theta}_{\mathbf{1}}^{\mathbf{0}}$. By Cauchy-Schwarz inequality, we obtain for the first term

$$
\begin{aligned}
&\sum_{j=1}^{T-1} k^2(j/M)\mathcal{B}_{121T}^2(j) \\
&\quad \le \left|\left|\hat{\boldsymbol{\theta}}_{\mathbf{1}} - \boldsymbol{\theta}_{\mathbf{1}}^{\mathbf{0}}\right|\right|^2\left\{\sum_{j=1}^{T-1} k^2(j/M)\frac{1}{T^2}\left[\sum_{t=j+1}^{T}\left|\left|\nabla_{\boldsymbol{\theta}_{\mathbf{1}}}\tilde{u}_{m,t}(\boldsymbol{\theta}_{\mathbf{1}}^{\mathbf{0}})\right|\right|v_{n,t-j}^0\right]^2\right\} \\
&\quad = O_p(M/T^2),
\end{aligned}
$$

given Assumption 2.2 and $\sum_{j=1}^{T-1} k^2(j/M)T^{-2}\left[\sum_{t=j+1}^{T}\left|\left|\nabla_{\boldsymbol{\theta}_{\mathbf{1}}}\tilde{u}_{m,t}(\boldsymbol{\theta}_{\mathbf{1}}^{\mathbf{0}})\right|\right|v_{n,t-j}^0\right]^2 = O_p(M/T)$, which follows from Markov's inequality, Assumption 2.4 and

$$
\begin{aligned}
\mathbb{E}\left\{\frac{1}{T^2}\left[\sum_{t=j+1}^{T}\left|\left|\nabla_{\boldsymbol{\theta}_{\mathbf{1}}}\tilde{u}_{m,t}(\boldsymbol{\theta}_{\mathbf{1}}^{\mathbf{0}})\right|\right|v_{n,t-j}^0\right]^2\right\} &= \frac{1}{T^2}\sum_{t=j+1}^{T}\mathbb{E}\left\{\left[\left|\left|\nabla_{\boldsymbol{\theta}_{\mathbf{1}}}\tilde{u}_{m,t}(\boldsymbol{\theta}_{\mathbf{1}}^{\mathbf{0}})\right|\right|v_{n,t-j}^0\right]^2\right\} \\
&= \frac{1}{T^2}\sum_{t=j+1}^{T}\mathbb{E}\left[\left|\left|\nabla_{\boldsymbol{\theta}_{\mathbf{1}}}\tilde{u}_{m,t}(\boldsymbol{\theta}_{\mathbf{1}}^{\mathbf{0}})\right|\right|^2\right]\mathbb{E}\left[\left(v_{n,t-j}^0\right)^2\right] \\
&= O(T^{-1}),
\end{aligned}
$$

where the first equality follows from Assumption 2.1 and the second equality follows from the independence between $\boldsymbol{\Xi_{1t}\Xi'_{1t}}$ and $\boldsymbol{\Xi_{2t-j}\Xi'_{2t-j}}$ under the null hypothesis. By Cauchy-Schwarz inequality, we can write the second term $\mathcal{B}_{122T}(j)$ as

$$
\begin{aligned}
&\sum_{j=1}^{T-1} k^2(j/M)\mathcal{B}_{122T}^2(j) \\
&\quad \le \frac{1}{T^2}\left|\left|\hat{\boldsymbol{\theta}}_{\mathbf{1}} - \boldsymbol{\theta}_{\mathbf{1}}^{\mathbf{0}}\right|\right|^4 \sum_{j=1}^{T-1} k^2(j/M)\left[\sum_{t=j+1}^{T} \left|\left|\nabla^2_{\boldsymbol{\theta_1}} \tilde{u}_{m,t}(\bar{\boldsymbol{\theta}}_{\mathbf{1}})\right|\right| v^0_{n,t-j}\right]^2 \\
&\quad \le \frac{1}{T^2}\left|\left|\hat{\boldsymbol{\theta}}_{\mathbf{1}} - \boldsymbol{\theta}_{\mathbf{1}}^{\mathbf{0}}\right|\right|^4 \sum_{j=1}^{T-1} k^2(j/M)\left[\sum_{t=j+1}^{T} \left|\left|\nabla^2_{\boldsymbol{\theta_1}} \tilde{u}_{m,t}(\bar{\boldsymbol{\theta}}_{\mathbf{1}})\right|\right|^2\right]\left[\sum_{t=j+1}^{T} (v^0_{n,t-j})^2\right] \\
&\quad = \left|\left|\hat{\boldsymbol{\theta}}_{\mathbf{1}} - \boldsymbol{\theta}_{\mathbf{1}}^{\mathbf{0}}\right|\right|^4 \sum_{j=1}^{T-1} k^2(j/M)\left[\frac{1}{T}\sum_{t=j+1}^{T} \left|\left|\nabla^2_{\boldsymbol{\theta_1}} \tilde{u}_{m,t}(\bar{\boldsymbol{\theta}}_{\mathbf{1}})\right|\right|^2\right]\left[\frac{1}{T}\sum_{t=j+1}^{T} (v^0_{n,t-j})^2\right] \\
&\quad = O_p(M/T^2),
\end{aligned}
$$

having used Assumptions 2.2 and 2.4 with Markov's inequality. Therefore,

$$
\sum_{j=1}^{T-1} k^2(j/M)\mathcal{B}_{1T}^2(j) = O_p(M/T^2) = o_p(M^{1/2}/T). \tag{B.19}
$$

By the same reasonings, we also have

$$
\sum_{j=1}^{T-1} k^2(j/M)\mathcal{B}_{2T}^2(j) = O_p(M/T^2) = o_p(M^{1/2}/T). \tag{B.20}
$$

Collecting (B.10), (B.11), (B.15), (B.19) and (B.20), we have $\mathcal{A}_{1T} = o_p(M^{1/2})$.

For the second term $\mathcal{A}_{2T}$ in (B.8), we have

$$
\begin{aligned}
\mathcal{A}_{2T} &= T\sum_{j=1}^{T-1} k^2(j/M)\mathrm{vec}[\boldsymbol{C_b}(j)]'\big(\mathrm{vec}[\boldsymbol{C_{\hat{b}}}(j)] - \mathrm{vec}[\boldsymbol{C_b}(j)]\big) \\
&= T\sum_{j=1}^{T-1} k^2(j/M)\Big\{(\boldsymbol{C^0_{vv}})^{-1/2} \otimes (\boldsymbol{C^0_{uu}})^{-1/2}\mathrm{vec}[\boldsymbol{\hat{C}^0_{uv}}(j)]\Big\}'\Big[(\boldsymbol{C^0_{vv}})^{-1/2} \otimes (\boldsymbol{C^0_{uu}})^{-1/2}\Big] \\
&\quad \times \big\{\mathrm{vec}[\boldsymbol{\hat{C}_{uv}}(j)] - \mathrm{vec}[\boldsymbol{\hat{C}^0_{uv}}(j)]\big\} \\
&= T\sum_{j=1}^{T-1} k^2(j/M)\mathrm{vec}[\boldsymbol{\hat{C}^0_{uv}}(j)]'\Big[(\boldsymbol{C^0_{vv}})^{-1} \otimes (\boldsymbol{C^0_{uu}})^{-1}\Big]\big\{\mathrm{vec}[\boldsymbol{\hat{C}_{uv}}(j)] - \mathrm{vec}[\boldsymbol{\hat{C}^0_{uv}}(j)]\big\},
\end{aligned}
$$

having used again the properties $\mathrm{vec}(\boldsymbol{ABC}) = (\boldsymbol{C}' \otimes \boldsymbol{A})\mathrm{vec}(\boldsymbol{B})$ and $(\boldsymbol{A} \otimes \boldsymbol{B})(\boldsymbol{C} \otimes \boldsymbol{D}) = (\boldsymbol{AC}) \otimes (\boldsymbol{BD})$. For $p = 1, ..., d_1^* d_2^*$, let $\hat{C}^p_{uv}(j)$ and $\hat{C}^{0,p}_{uv}(j)$ denote the p-th element of $\mathrm{vec}[\boldsymbol{\hat{C}_{uv}}(j)]$ and $\mathrm{vec}[\boldsymbol{\hat{C}^0_{uv}}(j)]$, respectively. Similarly, we denote by $\hat{C}^q_{uv}(j)$ and $\hat{C}^{0,q}_{uv}(j)$ the q-th element of $\mathrm{vec}[\boldsymbol{\hat{C}_{uv}}(j)]$

and $\text{vec}[\hat{\boldsymbol{C}}^{\mathbf{0}}_{\boldsymbol{uv}}(j)]$, respectively. Further let $G^{p,q}$ denotes the (p,q)-th element of $\boldsymbol{G}$, where $\boldsymbol{G} \equiv (\boldsymbol{C}^{\mathbf{0}}_{\boldsymbol{vv}})^{-1} \otimes (\boldsymbol{C}^{\mathbf{0}}_{\boldsymbol{uu}})^{-1}$. Then,

$$\begin{aligned}\mathcal{A}_{2T} &= T\sum_{j=1}^{T-1} k^2(j/M)\left\{\sum_{p=1}^{d_1^*d_2^*}\sum_{q=1}^{d_1^*d_2^*} \hat{C}^{0,p}_{uv}(j)\Big[\hat{C}^{q}_{uv}(j) - \hat{C}^{0,q}_{uv}(j)\Big]G^{p,q}\right\} \\ &= T\sum_{p=1}^{d_1^*d_2^*}\sum_{q=1}^{d_1^*d_2^*} G^{p,q}\sum_{j=1}^{T-1} k^2(j/M)\hat{C}^{0,p}_{uv}(j)\Big[\hat{C}^{q}_{uv}(j) - \hat{C}^{0,q}_{uv}(j)\Big]. \end{aligned} \tag{B.21}$$

Because $G^{p,q} = O_p(1)$ by Assumption 2.1, it suffices to show that $\sum_{j=1}^{T-1} k^2(j/M)\hat{C}^{0,p}_{uv}(j)[\hat{C}^{q}_{uv}(j) - \hat{C}^{0,q}_{uv}(j)] = O_p(M/T^{3/2})$. By Cauchy-Schwarz inequality, we can write

$$\begin{aligned}&\left|\sum_{j=1}^{T-1} k^2(j/M)\hat{C}^{0,p}_{uv}(j)\Big[\hat{C}^{q}_{uv}(j) - \hat{C}^{0,q}_{uv}(j)\Big]\right| \\ &\qquad \le \left[\sum_{j=1}^{T-1} k^2(j/M)\hat{C}^{0,p}_{uv}(j)^2\right]^{1/2}\left\{\sum_{j=1}^{T-1} k^2(j/M)\Big[\hat{C}^{q}_{uv}(j) - \hat{C}^{0,q}_{uv}(j)\Big]^2\right\}^{1/2}.\end{aligned}$$

We have $\sum_{j=1}^{T-1} k^2(j/M)[\hat{C}^{q}_{uv}(j) - \hat{C}^{0,q}_{uv}(j)]^2 = O_p(M/T^2)$ from the proof of (B.12) and we have $\sum_{j=1}^{T-1} k^2(j/M)\hat{C}^{0,p}_{uv}(j)^2 = O_p(M/T)$ by Markov's inequality and

$$\begin{aligned}\sum_{j=1}^{T-1} k^2(j/M)\mathbb{E}\big[\hat{C}^{0,p}_{uv}(j)^2\big] &= \frac{M}{T}C^{0,p}_{uu}C^{0,p}_{vv}\left(\frac{1}{M}\sum_{j=1}^{T-1}(1 - j/T)k^2(j/M)\right) \\ &= O(M/T),\end{aligned} \tag{B.22}$$

where $C^{0,p}_{uu}$ and $C^{0,p}_{vv}$ denote the p-th entries of $\text{vec}(\boldsymbol{C}^{\mathbf{0}}_{\boldsymbol{uu}})$ and $\text{vec}(\boldsymbol{C}^{\mathbf{0}}_{\boldsymbol{vv}})$, respectively. This gives $\mathcal{A}_{2T} = O_p(M/T^{1/2}) = o_p(M^{1/2})$ and completes the proof. ■

Proof of Proposition B.2. Recall Proposition B.2 is stated as

Proposition B.2. *Suppose the conditions of Theorem 2.1 hold, we have that*

$$\frac{\hat{S} - S^*}{\big[d_1^*d_2^*D_{1T}(k)\big]^{1/2}} \xrightarrow{p} 0.$$

Given Assumption 2.5 and since $M \to \infty$ as $T \to \infty$, it follows that

$$D_{1T}(k) = M\int_0^\infty k^4(z)\mathrm{d}z[1 + o(1)].$$

Therefore, the result of Proposition B.2 can be obtained by showing that $\hat{S} - S^* = O_p(M/T^{1/2})$.

By reasonings similar to the derivations of (B.5) and (B.7), we write for $\hat{S}$

$$\hat{S} = T\sum_{j=1}^{T-1} k^2(j/M)\text{vec}\big[\hat{\boldsymbol{\rho}}(j)\big]'\big(\hat{\boldsymbol{\Gamma}}_{\boldsymbol{v}}^{-1}\otimes\hat{\boldsymbol{\Gamma}}_{\boldsymbol{u}}^{-1}\big)\text{vec}\big[\hat{\boldsymbol{\rho}}(j)\big]$$
$$= T\sum_{j=1}^{T-1} k^2(j/M)\text{vec}\big[\hat{\boldsymbol{C}}_{\boldsymbol{uv}}(j)\big]'\big(\hat{\boldsymbol{C}}_{\boldsymbol{vv}}^{-1}\otimes\hat{\boldsymbol{C}}_{\boldsymbol{uu}}^{-1}\big)\text{vec}\big[\hat{\boldsymbol{C}}_{\boldsymbol{uv}}(j)\big]. \tag{B.23}$$

Then, $\hat{S}-S^*$ is equal to

$$\hat{S}-S^* = T\sum_{j=1}^{T-1} k^2(j/M)\text{vec}\big[\hat{\boldsymbol{C}}_{\boldsymbol{uv}}(j)\big]'\big[\hat{\boldsymbol{C}}_{\boldsymbol{vv}}^{-1}\otimes\hat{\boldsymbol{C}}_{\boldsymbol{uu}}^{-1}-(\boldsymbol{C}_{\boldsymbol{vv}}^{\boldsymbol{0}})^{-1}\otimes(\boldsymbol{C}_{\boldsymbol{uu}}^{\boldsymbol{0}})^{-1}\big]\text{vec}\big[\hat{\boldsymbol{C}}_{\boldsymbol{uv}}(j)\big]$$
$$= T\sum_{j=1}^{T-1} k^2(j/M)\text{vec}\big[\hat{\boldsymbol{C}}_{\boldsymbol{uv}}(j)\big]'\big[(\hat{\boldsymbol{C}}_{\boldsymbol{vv}}\otimes\hat{\boldsymbol{C}}_{\boldsymbol{uu}})^{-1}-(\boldsymbol{C}_{\boldsymbol{vv}}^{\boldsymbol{0}}\otimes\boldsymbol{C}_{\boldsymbol{uu}}^{\boldsymbol{0}})^{-1}\big]\text{vec}\big[\hat{\boldsymbol{C}}_{\boldsymbol{uv}}(j)\big].$$

We proceed by showing that $\hat{\boldsymbol{C}}_{\boldsymbol{vv}}\otimes\hat{\boldsymbol{C}}_{\boldsymbol{uu}}-\boldsymbol{C}_{\boldsymbol{vv}}^{\boldsymbol{0}}\otimes\boldsymbol{C}_{\boldsymbol{uu}}^{\boldsymbol{0}} = O_p(T^{-1/2})$. Its inverse counterpart has the same stochastic order due to a continuous transformation. For $m,n=1,...,d_1^*$, we let $\hat{C}_{vv}^{m,n}$ and $C_{vv}^{0,m,n}$ denote the (m,n)-th elements of matrices $\hat{\boldsymbol{C}}_{\boldsymbol{vv}}$ and $\boldsymbol{C}_{\boldsymbol{vv}}^{\boldsymbol{0}}$, respectively. Similarly, for $r,s=1,...,d_2^*$, we let $\hat{C}_{uu}^{r,s}$ and $C_{uu}^{0,r,s}$ denote the (r,s)-th entries of matrices $\hat{\boldsymbol{C}}_{\boldsymbol{uu}}$ and $\boldsymbol{C}_{\boldsymbol{uu}}^{\boldsymbol{0}}$, respectively. Then, we have the $[d_1^*(m-1)+r, d_1^*(n-1)+s]$-th entry in $\hat{\boldsymbol{C}}_{\boldsymbol{vv}}\otimes\hat{\boldsymbol{C}}_{\boldsymbol{uu}}$ and $\boldsymbol{C}_{\boldsymbol{vv}}^{\boldsymbol{0}}\otimes\boldsymbol{C}_{\boldsymbol{uu}}^{\boldsymbol{0}}$ be given, respectively, by $\hat{C}_{vv}^{m,n}\hat{C}_{uu}^{r,s}$ and $C_{vv}^{0,m,n}C_{uu}^{0,r,s}$. It suffices to show that $\hat{C}_{vv}^{m,n}\hat{C}_{uu}^{r,s}-C_{vv}^{0,m,n}C_{uu}^{0,r,s} = O_p(T^{-1/2})$. Since $\hat{C}_{vv}^{m,n}\hat{C}_{uu}^{r,s}-C_{vv}^{0,m,n}C_{uu}^{0,r,s} = (\hat{C}_{vv}^{m,n}-C_{vv}^{0,m,n})C_{uu}^{0,r,s}+(\hat{C}_{uu}^{r,s}-C_{uu}^{0,r,s})C_{vv}^{0,m,n}+(\hat{C}_{vv}^{m,n}-C_{vv}^{0,m,n})(\hat{C}_{uu}^{r,s}-C_{uu}^{0,r,s})$ and given $C_{uu}^{0,r,s}=O_p(1)$ and $C_{vv}^{0,m,n}=O_p(1)$ by Assumption 2.1, we know that the controlling terms in $\hat{C}_{vv}^{m,n}\hat{C}_{uu}^{r,s}-C_{vv}^{0,m,n}C_{uu}^{0,r,s}$ are $\hat{C}_{vv}^{m,n}-C_{vv}^{0,m,n}$ and $\hat{C}_{uu}^{r,s}-C_{uu}^{0,r,s}$. It therefore suffices to show that $\hat{C}_{uu}^{r,s}-C_{uu}^{0,r,s}=O_p(T^{-1/2})$, the proof for $\hat{C}_{vv}^{m,n}-C_{vv}^{0,m,n}$ is similar. It follows from the triangle inequality that

$$|\hat{C}_{uu}^{r,s}-C_{uu}^{0,r,s}| \le |\hat{C}_{uu}^{r,s}-\hat{C}_{uu}^{0,r,s}|+|\hat{C}_{uu}^{0,r,s}-C_{uu}^{0,r,s}|,$$

where we have $\hat{C}_{uu}^{0,r,s}-C_{uu}^{0,r,s}=O_p(T^{-1/2})$ by Chebyshev's inequality and Assumption 2.1. Following the same reasonings to the proof of (B.11), we have for the first term

$$\begin{aligned}\hat{C}_{uu}^{r,s}-\hat{C}_{uu}^{0,r,s} &= \frac{1}{T}\sum_{t=1}^{T}\hat{u}_{m,t}\hat{u}_{m,t}-u_{m,t}^0u_{m,t}^0\\ &= \frac{1}{T}\sum_{t=1}^{T}(\hat{u}_{m,t}-u_{m,t}^0)u_{m,t}^0+\frac{1}{T}\sum_{t=1}^{T}(\hat{u}_{m,t}-u_{m,t}^0)u_{m,t}^0\\ &\quad+\frac{1}{T}\sum_{t=1}^{T}(\hat{u}_{m,t}-u_{m,t}^0)(\hat{u}_{m,t}-u_{m,t}^0)\\ &= O_p(T^{-1/2}).\end{aligned}$$

Therefore we have

$$
\begin{aligned}
\hat{S} - S^* &= T \sum_{j=1}^{T-1} k^2(j/M) \text{vec}\big[\hat{\boldsymbol{C}}_{\boldsymbol{uv}}(j)\big]' \big[O_p(T^{-1/2})\big] \text{vec}\big[\hat{\boldsymbol{C}}_{\boldsymbol{uv}}(j)\big] \\
&= O_p(T^{1/2}) \sum_{j=1}^{T-1} k^2(j/M) \text{vec}\big[\hat{\boldsymbol{C}}_{\boldsymbol{uv}}(j)\big]' \text{vec}\big[\hat{\boldsymbol{C}}_{\boldsymbol{uv}}(j)\big]. \qquad \text{(B.24)}
\end{aligned}
$$

For the rest of the proof, it suffices to show that $\mathcal{F}_T = O_p(M/T)$, where

$$
\begin{aligned}
\mathcal{F}_T &= \sum_{j=1}^{T-1} k^2(j/M) \text{vec}\big[\hat{\boldsymbol{C}}_{\boldsymbol{uv}}(j)\big]' \text{vec}\big[\hat{\boldsymbol{C}}_{\boldsymbol{uv}}(j)\big] \\
&= \sum_{j=1}^{T-1} k^2(j/M) \text{vec}\big[\hat{\boldsymbol{C}}_{\boldsymbol{uv}}(j)\big]' \text{vec}\big[\hat{\boldsymbol{C}}_{\boldsymbol{uv}}(j)\big] - \text{vec}\big[\hat{\boldsymbol{C}}^{\boldsymbol{0}}_{\boldsymbol{uv}}(j)\big]' \text{vec}\big[\hat{\boldsymbol{C}}^{\boldsymbol{0}}_{\boldsymbol{uv}}(j)\big] \\
&\quad + \sum_{j=1}^{T-1} k^2(j/M) \text{vec}\big[\hat{\boldsymbol{C}}^{\boldsymbol{0}}_{\boldsymbol{uv}}(j)\big]' \text{vec}\big[\hat{\boldsymbol{C}}^{\boldsymbol{0}}_{\boldsymbol{uv}}(j)\big] \\
&= \mathcal{F}_{1T} + \mathcal{F}_{2T}, \text{ say.} \qquad \text{(B.25)}
\end{aligned}
$$

We write for the first term $\mathcal{F}_{1T}$

$$
\begin{aligned}
\mathcal{F}_{1T} &= \sum_{j=1}^{T-1} k^2(j/M) \sum_{m=1}^{d_1^*} \sum_{n=1}^{d_2^*} \Big[\hat{C}_{uv}^{m,n}(j)^2 - \hat{C}_{uv}^{0,m,n}(j)^2\Big] \\
&= \sum_{m=1}^{d_1^*} \sum_{n=1}^{d_2^*} \sum_{j=1}^{T-1} k^2(j/M) \Big[\hat{C}_{uv}^{m,n}(j)^2 - \hat{C}_{uv}^{0,m,n}(j)^2\Big], \qquad \text{(B.26)}
\end{aligned}
$$

where $\hat{C}_{uv}^{m,n}(j)$ and $\hat{C}_{uv}^{0,m,n}(j)$ are the (m,n)-th elements of matrices $\hat{\boldsymbol{C}}_{\boldsymbol{uv}}(j)$ and $\hat{\boldsymbol{C}}^{\boldsymbol{0}}_{\boldsymbol{uv}}(j)$, respectively. It suffices to show that $\sum_{j=1}^{T-1} k^2(j/M)\big[\hat{C}_{uv}^{m,n}(j)^2 - \hat{C}_{uv}^{0,m,n}(j)^2\big] = O_p(M/T)$. We write

$$
\begin{aligned}
\sum_{j=1}^{T-1} k^2(j/M) \Big[\hat{C}_{uv}^{m,n}(j)^2 - \hat{C}_{uv}^{0,m,n}(j)^2\Big] &= \sum_{j=1}^{T-1} k^2(j/M) \Big[\hat{C}_{uv}^{m,n}(j) - \hat{C}_{uv}^{0,m,n}(j)\Big]^2 \\
&\quad + 2 \sum_{j=1}^{T-1} k^2(j/M) \hat{C}_{uv}^{0,m,n}(j) \Big[\hat{C}_{uv}^{m,n}(j) - \hat{C}_{uv}^{0,m,n}(j)\Big] \\
&= \mathcal{F}_{11T} + 2\mathcal{F}_{12T}, \text{ say.}
\end{aligned}
$$

We have $\mathcal{F}_{11T} = O_p(M/T^2)$ from the proof of (B.12) and we have $\mathcal{F}_{12T} = O_p(M/T^{3/2})$ from the proof of (B.21). Thus, $\mathcal{F}_{1T} = O_p(M/T)$.

Next, we rewrite the second term $\mathcal{F}_{2T}$ as

$$\mathcal{F}_{2T} = \sum_{j=1}^{T-1} k^2(j/M)\left[\sum_{p=1}^{d_1^* d_2^*} \hat{C}_{uv}^{0,p}(j)^2\right] = \sum_{p=1}^{d_1^* d_2^*}\sum_{j=1}^{T-1} k^2(j/M)\hat{C}_{uv}^{0,p}(j)^2, \tag{B.27}$$

where $\hat{C}_{uv}^{0,p}(j)$ is the p-th element in $\text{vec}[\hat{\boldsymbol{C}}_{\boldsymbol{uv}}^{\mathbf{0}}(j)]$, $p = 1, ..., d_1^* d_2^*$. By Markov's inequality and (B.22), we have $\mathcal{F}_{2T} = O_p(M/T)$. This completes the proof. ■

Appendix C Proof of Theorem 2.2

Proof of Theorem 2.2. Recall that $C_{1T}(k) = O(M)$ and $D_{1T}(k) = 2M\int_0^\infty k^4(z)\mathrm{d}z[1 + o(1)]$ as $M \to \infty$ and $M/T \to 0$, we have

$$\begin{aligned}
\frac{M^{1/2}}{T}Q_1 &= \frac{\sum_{j=1}^{T-1} k^2(j/M)\text{vec}\big[\hat{\boldsymbol{\rho}}(j)\big]'\big(\hat{\boldsymbol{\Gamma}}_{\boldsymbol{v}}^{-1} \otimes \hat{\boldsymbol{\Gamma}}_{\boldsymbol{u}}^{-1}\big)\text{vec}\big[\hat{\boldsymbol{\rho}}(j)\big]}{\Big[2\int_0^\infty k^4(z)\mathrm{d}z\Big]^{1/2}}\big[1 + o(1)\big] + o(1) \\
&= T^{-1}\hat{S}\left[2\int_0^\infty k^4(z)\mathrm{d}z\right]^{-1/2}\big[1 + o(1)\big] + o(1).
\end{aligned}$$

Therefore, the proof of Theorem 2.2 is given by Lemmas C.1–C.3. ■

Lemma C.1. *Suppose the conditions of Theorem 2.2 hold, then* $T^{-1}(S^* - S) = o_p(1)$.

Lemma C.2. *Suppose the conditions of Theorem 2.2 hold, then* $T^{-1}(\hat{S} - S^*) = o_p(1)$.

Lemma C.3. *Suppose the conditions of Theorem 2.2 hold, then*

$$\begin{aligned}
\frac{1}{T}S &= \sum_{j=1}^{T-1} k^2(j/M)\text{vec}\big[\hat{\boldsymbol{\rho}}^{\mathbf{0}}(j)\big]'\big(\boldsymbol{\Gamma}_{\boldsymbol{v}}^{-1} \otimes \boldsymbol{\Gamma}_{\boldsymbol{u}}^{-1}\big)\text{vec}\big[\hat{\boldsymbol{\rho}}^{\mathbf{0}}(j)\big] \\
&\xrightarrow{p} \sum_{j=1}^{\infty}\text{vec}\big[\boldsymbol{\rho}(j)\big]'\big(\boldsymbol{\Gamma}_{\boldsymbol{v}}^{-1} \otimes \boldsymbol{\Gamma}_{\boldsymbol{u}}^{-1}\big)\text{vec}\big[\boldsymbol{\rho}(j)\big].
\end{aligned}$$

Proof of Lemma C.1. The proof of Lemma C.1 follows largely from the proof of Lemma B.2 with some modifications as we are now under the alternative hypothesis. We have $T^{-1}(S^* - S) = T^{-1}(\mathcal{A}_{1T} + 2\mathcal{A}_{2T})$ as in (B.8). We shall show that $T^{-1}\mathcal{A}_{1T} = o_p(1)$ and $T^{-1}\mathcal{A}_{2T} = o_p(1)$. We begin with $\mathcal{A}_{1T}$, we have from (B.9) that

$$T^{-1}\mathcal{A}_{1T} \leq \sum_{j=1}^{T-1} k^2(j/M)\Big|\Big|(\boldsymbol{C}_{\boldsymbol{vv}}^{\mathbf{0}})^{-1/2} \otimes (\boldsymbol{C}_{\boldsymbol{uu}}^{\mathbf{0}})^{-1/2}\Big|\Big|^2\ \Big|\Big|\text{vec}[\hat{\boldsymbol{C}}_{\boldsymbol{uv}}(j)] - \text{vec}[\hat{\boldsymbol{C}}_{\boldsymbol{uv}}^{\mathbf{0}}(j)]\Big|\Big|^2. \tag{C.1}$$

It suffices to show that $T^{-1}\mathcal{A}_{11T} = o_p(1)$, where

$$T^{-1}\mathcal{A}_{11T} = \sum_{j=1}^{T-1} k^2(j/M)\Big|\Big|\text{vec}[\hat{\boldsymbol{C}}_{\boldsymbol{uv}}(j)] - \text{vec}[\hat{\boldsymbol{C}}^{\boldsymbol{0}}_{\boldsymbol{uv}}(j)]\Big|\Big|^2$$
$$= \sum_{m=1}^{d_1^*}\sum_{n=1}^{d_2^*}\sum_{j=1}^{T-1} k^2(j/M)\big[\hat{C}_{uv}^{m,n}(j) - \hat{C}_{uv}^{0,m,n}(j)\big]^2. \tag{C.2}$$

As in (B.12), we have

$$\sum_{j=1}^{T-1} k^2(j/M)\big[\hat{C}_{uv}^{m,n}(j) - \hat{C}_{uv}^{0,m,n}(j)\big]^2 \leq \Delta\sum_{j=1}^{T-1} k^2(j/M)\big[\mathcal{B}_{1T}^2(j) + \mathcal{B}_{2T}^2(j) + \mathcal{B}_{3T}^2(j)\big]. \tag{C.3}$$

It suffices to show that $\sum_{j=1}^{T-1} k^2(j/M)\big[\mathcal{B}_{iT}^2(j)\big] = o_p(1)$, for $i = 1,2,3$. As shown in (B.15), we have $\sum_{j=1}^{T-1} k^2(j/M)\mathcal{B}_{3T}^2(j) = O_p(M/T^2) = o_p(1)$ under Assumptions 2.1-2.6. Next, from (B.16), we have $\mathcal{B}_{1T}(j) = \mathcal{B}_{11T}(j)+\mathcal{B}_{12T}(j)$. We know from (B.17) that $\sum_{j=1}^{T-1} k^2(j/M)\mathcal{B}_{11T}^2(j) = O_p(M/T^2) = o_p(1)$ under Assumptions 2.1-2.6. Applying Cauchy-Schwarz inequality to the second term $\mathcal{B}_{12T}(j)$, we have

$$\begin{aligned}
&\sum_{j=1}^{T-1} k^2(j/M)\mathcal{B}_{12T}^2(j)\\
&\leq \frac{1}{T^2}\sum_{j=1}^{T-1} k^2(j/M)\left\{\sum_{t=j+1}^{T}\big[\tilde{u}_{m,t}(\hat{\boldsymbol{\theta}}_{\boldsymbol{1}}) - u_{m,t}^0\big]^2\right\}\left[\sum_{t=j+1}^{T}\big(v_{n,t-j}^0\big)^2\right]\\
&\leq \sum_{j=1}^{T-1} k^2(j/M)\left\{\frac{1}{T}\sum_{t=j+1}^{T}\big[\tilde{u}_{m,t}(\hat{\boldsymbol{\theta}}_{\boldsymbol{1}}) - u_{m,t}^0\big]^2\right\}\left[\frac{1}{T}\sum_{t=1}^{T}\big(v_{n,t}^0\big)^2\right]\\
&= O_p(M/T),
\end{aligned} \tag{C.4}$$

given $\sum_{j=1}^{T-1} k^2(j/M) = O(M)$, $T^{-1}\sum_{t=j+1}^{T}\big[\tilde{u}_{m,t}(\hat{\boldsymbol{\theta}}_{\boldsymbol{1}}) - u_{m,t}^0\big]^2 = O_p(T^{-1})$ from (B.14) and $T^{-1}\sum_{t=1}^{T}\big(v_{n,t}^0\big)^2 = O_p(1)$ by Markov's inequality. It follows that $\sum_{j=1}^{T-1} k^2(j/M)\mathcal{B}_{1T}^2(j) = O_p(M/T) = o_p(1)$. By the same reasonings, we also have $\sum_{j=1}^{T-1} k^2(j/M)\mathcal{B}_{2T}^2(j) = o_p(1)$. Therefore, $T^{-1}\mathcal{A}_{1T} = o_p(1)$.

Next, we shall show that $T^{-1}\mathcal{A}_{2T} = o_p(1)$. From (B.21), we have

$$T^{-1}\mathcal{A}_{2T} = \sum_{p=1}^{d_1^*d_2^*}\sum_{q=1}^{d_1^*d_2^*} G^{p,q}\sum_{j=1}^{T-1} k^2(j/M)\hat{C}_{uv}^{0,p}(j)\Big[\hat{C}_{uv}^{q}(j) - \hat{C}_{uv}^{0,q}(j)\Big]. \tag{C.5}$$

It suffices to show that $\sum_{j=1}^{T-1} k^2(j/M)\hat{C}_{uv}^{0,p}(j)\big[\hat{C}_{uv}^{q}(j) - \hat{C}_{uv}^{0,q}(j)\big] = o_p(1)$. By Cauchy-Schwarz inequality, we have

$$\left|\sum_{j=1}^{T-1} k^2(j/M)\hat{C}_{uv}^{0,p}(j)\left[\hat{C}_{uv}^{q}(j)-\hat{C}_{uv}^{0,q}(j)\right]\right| \\ \leq \left[\sum_{j=1}^{T-1} k^2(j/M)\hat{C}_{uv}^{0,p}(j)^2\right]^{1/2}\left\{\sum_{j=1}^{T-1} k^2(j/M)\left[\hat{C}_{uv}^{q}(j)-\hat{C}_{uv}^{0,q}(j)\right]^2\right\}^{1/2}.$$

We have $\sum_{j=1}^{T-1} k^2(j/M)\left[\hat{C}_{uv}^{q}(j)-\hat{C}_{uv}^{0,q}(j)\right]^2 = o_p(1)$ from (C.2) and $\sum_{j=1}^{T-1} k^2(j/M)\hat{C}_{uv}^{0,p}(j)^2 = O_p(1)$ by Lemma C.3 and $\sum_{j=1}^{\infty} ||\boldsymbol{\rho}(j)||^2 < \infty$. This completes the proof. ■

Proof of Lemma C.2. The proof of Lemma C.2 can be readily deduced from the proofs of Proposition B.2 and Lemma C.1. Based on Assumptions 2.1-2.6, we have shown that $T^{-1}(\hat{S}-S^*) = T^{-1}O_p(T^{1/2})\mathcal{F}_T = O_p(T^{-1/2})\mathcal{F}_T$ in (B.24) and (B.25). It suffices to show that $\mathcal{F}_T = \mathcal{F}_{1T}+\mathcal{F}_{2T} = O_p(1)$. Using the results in (C.2) and (C.5), we have $\mathcal{F}_{1T} = O_p(1)$ under the alternative hypothesis. Finally, $\mathcal{F}_{2T} = O_p(1)$ by Lemma C.3 and $\sum_{j=1}^{\infty} ||\boldsymbol{\rho}(j)||^2 < \infty$. This completes the proof. ■

Proof of Lemma C.3. First, we write

$$\begin{aligned}\frac{1}{T}S &= \sum_{j=1}^{T-1} k^2(j/M)\mathrm{vec}\big[\boldsymbol{\rho}(j)\big]'\big(\boldsymbol{\Gamma}_{\boldsymbol{v}}^{-1}\otimes\boldsymbol{\Gamma}_{\boldsymbol{u}}^{-1}\big)\mathrm{vec}\big[\boldsymbol{\rho}(j)\big] \\ &\quad+\Bigg\{\sum_{j=1}^{T-1} k^2(j/M)\mathrm{vec}\big[\hat{\boldsymbol{\rho}}^{\mathbf{0}}(j)\big]'\big(\boldsymbol{\Gamma}_{\boldsymbol{v}}^{-1}\otimes\boldsymbol{\Gamma}_{\boldsymbol{u}}^{-1}\big)\mathrm{vec}\big[\hat{\boldsymbol{\rho}}^{\mathbf{0}}(j)\big] \\ &\qquad-\sum_{j=1}^{T-1} k^2(j/M)\mathrm{vec}\big[\boldsymbol{\rho}(j)\big]'\big(\boldsymbol{\Gamma}_{\boldsymbol{v}}^{-1}\otimes\boldsymbol{\Gamma}_{\boldsymbol{u}}^{-1}\big)\mathrm{vec}\big[\boldsymbol{\rho}(j)\big]\Bigg\} \\ &= \mathcal{G}_{1T}+\mathcal{G}_{2T}, \text{ say.}\end{aligned} \tag{C.6}$$

For the first term $\mathcal{G}_{1T}$ in (C.6), we have

$$\begin{aligned}\mathcal{G}_{1T} &= \sum_{j=1}^{\infty} \mathrm{vec}\big[\boldsymbol{\rho}(j)\big]'\big(\boldsymbol{\Gamma}_{\boldsymbol{v}}^{-1}\otimes\boldsymbol{\Gamma}_{\boldsymbol{u}}^{-1}\big)\mathrm{vec}\big[\boldsymbol{\rho}(j)\big] \\ &\quad+\sum_{j=1}^{T-1}\big(k^2(j/M)-1\big)\mathrm{vec}\big[\boldsymbol{\rho}(j)\big]'\big(\boldsymbol{\Gamma}_{\boldsymbol{v}}^{-1}\otimes\boldsymbol{\Gamma}_{\boldsymbol{u}}^{-1}\big)\mathrm{vec}\big[\boldsymbol{\rho}(j)\big] \\ &\quad-\sum_{j=T}^{\infty} \mathrm{vec}\big[\boldsymbol{\rho}(j)\big]'\big(\boldsymbol{\Gamma}_{\boldsymbol{v}}^{-1}\otimes\boldsymbol{\Gamma}_{\boldsymbol{u}}^{-1}\big)\mathrm{vec}\big[\boldsymbol{\rho}(j)\big] \\ &\xrightarrow{p} \sum_{j=1}^{\infty} \mathrm{vec}\big[\boldsymbol{\rho}(j)\big]'\big(\boldsymbol{\Gamma}_{\boldsymbol{v}}^{-1}\otimes\boldsymbol{\Gamma}_{\boldsymbol{u}}^{-1}\big)\mathrm{vec}\big[\boldsymbol{\rho}(j)\big].\end{aligned} \tag{C.7}$$

The convergence follows because the third term $\sum_{j=T}^{\infty}\mathrm{vec}\big[\boldsymbol{\rho}(j)\big]'\big(\boldsymbol{\Gamma}_{\boldsymbol{v}}^{-1}\otimes\boldsymbol{\Gamma}_{\boldsymbol{u}}^{-1}\big)\mathrm{vec}\big[\boldsymbol{\rho}(j)\big]$ goes to zero as a consequence of the absolute summable condition $\sum_{j=1}^{\infty} ||\boldsymbol{\rho}(j)||^2 < \infty$; and the second term $\sum_{j=1}^{T-1}\big[k^2(j/M)-1\big]\mathrm{vec}\big[\boldsymbol{\rho}(j)\big]'\big(\boldsymbol{\Gamma}_{\boldsymbol{v}}^{-1}\otimes\boldsymbol{\Gamma}_{\boldsymbol{u}}^{-1}\big)\mathrm{vec}\big[\boldsymbol{\rho}(j)\big]\to 0$, which follows from the dominated convergence theorem, $\lim_{M\to\infty}\big[k^2(j/M)-1\big]\to 0$ and $\sum_{j=1}^{\infty}||\boldsymbol{\rho}(j)||^2<\infty$.

We now consider the second term $\mathcal{G}_{2T}$ in (C.6). Let $\boldsymbol{C}_{\boldsymbol{uv}}^{\boldsymbol{0}}(j)\equiv\mathbb{E}[\boldsymbol{u}_{\boldsymbol{t}}^{\boldsymbol{0}}(\boldsymbol{v}_{\boldsymbol{t-j}}^{\boldsymbol{0}})']$, we have $\boldsymbol{C}_{\boldsymbol{b}}^{\boldsymbol{0}}(j)\equiv(\boldsymbol{C}_{\boldsymbol{uu}}^{\boldsymbol{0}})^{-1/2}\boldsymbol{C}_{\boldsymbol{uv}}^{\boldsymbol{0}}(j)(\boldsymbol{C}_{\boldsymbol{vv}}^{\boldsymbol{0}})^{-1/2}$ the true cross-covariance between $\boldsymbol{b_{1t}}$ and $\boldsymbol{b_{2t}}$ at lag order j. As in (B.8), we write $\mathcal{G}_{2T}$ as

$$\begin{aligned}\mathcal{G}_{2T}&=\sum_{j=1}^{T-1}k^2(j/M)\big|\big|\mathrm{vec}[\boldsymbol{C}_{\boldsymbol{b}}(j)]-\mathrm{vec}[\boldsymbol{C}_{\boldsymbol{b}}^{\boldsymbol{0}}(j)]\big|\big|^2\\&\quad+2\sum_{j=1}^{T-1}k^2(j/M)\big\langle\mathrm{vec}[\boldsymbol{C}_{\boldsymbol{b}}^{\boldsymbol{0}}(j)],\mathrm{vec}[\boldsymbol{C}_{\boldsymbol{b}}(j)]-\mathrm{vec}[\boldsymbol{C}_{\boldsymbol{b}}^{\boldsymbol{0}}(j)]\big\rangle\\&=\mathcal{G}_{21T}+2\mathcal{G}_{22T},\text{ say.}\end{aligned}\tag{C.8}$$

For the rest of the proof, it suffices to show that the first term $\mathcal{G}_{21T}$ goes to zero in probability because $\mathcal{G}_{22T}$ can be bounded by the product of the first term and a finite constant using Cauchy-Schwarz inequality. Following a similar decomposition as in (B.9), we have

$$\mathcal{G}_{21T}\le\sum_{j=1}^{T-1}k^2(j/M)\Big|\Big|(\boldsymbol{C}_{\boldsymbol{vv}}^{\boldsymbol{0}})^{-1/2}\otimes(\boldsymbol{C}_{\boldsymbol{uu}}^{\boldsymbol{0}})^{-1/2}\Big|\Big|^2\ \Big|\Big|\mathrm{vec}[\hat{\boldsymbol{C}}_{\boldsymbol{uv}}^{\boldsymbol{0}}(j)]-\mathrm{vec}[\boldsymbol{C}_{\boldsymbol{uv}}^{\boldsymbol{0}}(j)]\Big|\Big|^2.\tag{C.9}$$

It suffices to show that $\mathcal{G}_{211T}=o_p(1)$, with

$$\begin{aligned}\mathcal{G}_{211T}&=\sum_{j=1}^{T-1}k^2(j/M)\Big|\Big|\mathrm{vec}[\hat{\boldsymbol{C}}_{\boldsymbol{uv}}^{\boldsymbol{0}}(j)]-\mathrm{vec}[\boldsymbol{C}_{\boldsymbol{uv}}^{\boldsymbol{0}}(j)]\Big|\Big|^2\\&=\sum_{m=1}^{d_1^*}\sum_{n=1}^{d_2^*}\sum_{j=1}^{T-1}k^2(j/M)\big[\hat{C}_{uv}^{0,m,n}(j)-C_{uv}^{0,m,n}(j)\big]^2,\end{aligned}\tag{C.10}$$

where $\hat{C}_{uv}^{0,m,n}(j)$ and $C_{uv}^{0,m,n}(j)$ are the (m,n)-th elements of matrices $\hat{\boldsymbol{C}}_{\boldsymbol{uv}}^{\boldsymbol{0}}(j)$ and $\boldsymbol{C}_{\boldsymbol{uv}}^{\boldsymbol{0}}(j)$, respectively. We have $\sup_{1\le j\le T-1}\mathrm{Var}[\hat{C}_{uv}^{0,m,n}(j)]\le\Delta T^{-1}$ given Assumption 2.7 (see, e.g., Hannan, 1970, p. 209). Therefore, $\mathcal{G}_{21T}=O_p(M/T)=o_p(1)$, where we make use of Markov's inequality and $\sum_{j=1}^{T-1}k^2(j/M)=O(M)$. This completes the proof. ■

Appendix D Proof of Proposition 3.1

Proof of Proposition 3.1. The desired result follows from a simple modification of the proof of Theorem 4.1 in Dufour and Pelletier (2021). The consistency result in Dufour and Pelletier (2021) is obtained under the following conditions: (i) The sequence $\{\epsilon_{i,n,t}^2\}$ and $\{e_{i,n,t}\}$ are strictly stationary and ergodic. (ii) The error term $\{e_{i,n,t}\}$ is strong mixing with finite fourth order moment and regular variance. (iii) The lag order p is such that $p/\log(T)\to\infty$ and $p^2/T\to 0$. (iv) The process of interest has an infinite vector autoregressive representation with zero mean. Given conditions (i)–(iv), it follows from Dufour and Pelletier (2021) that $||\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{\boldsymbol{(p)}}-\boldsymbol{\theta}_{\boldsymbol{i}}^{\boldsymbol{0}}||=O_p(p^{1/2}T^{-1/2})$. In our framework, conditions (i) and (ii) are satisfied under Assumption 3.1, whereas condition (iii) is implied by Assumption 3.2. For condition (iv), we show in (30) that the process of interest has an AR(∞) representation. Besides, it is straightforward to show that the proof in Dufour and Pelletier (2021) holds with the addition of an intercept term by putting $\boldsymbol{Y_{t-1}^{(p)}} \equiv [1, \epsilon^2_{i,n,t-1}, ..., \epsilon^2_{i,n,t-p}]$ and $\hat{\Pi}^{(p)} \equiv [\hat{\omega}^{(p)}_{i,n}, \hat{a}^{(p)}_{i,n,1}, ..., \hat{a}^{(p)}_{i,n,p}]'$. This completes the proof. ■

Appendix E Proof of Proposition 3.2

Proof of Proposition 3.2. Given Assumption 3.3 with $\alpha = 1.5$ and $\beta = 0.5$, we have for $i = 1, 2$, $||\hat{\boldsymbol{\theta}}^{(p)}_{\boldsymbol{i}} - \boldsymbol{\theta}^{\boldsymbol{0}}_{\boldsymbol{i}}|| = O_p(T^{-1/(2\alpha-2\beta)}) = O_p(T^{-1/2})$, which follows from Theorem 5.52 in van der Vaart (1998). This completes the proof. ■

Appendix F Proof of Proposition 3.3

Proof of Proposition 3.3. Given $\hat{\boldsymbol{\theta}}^{(p)a}_{\boldsymbol{i}} = \hat{\boldsymbol{\theta}}^{(p)}_{\boldsymbol{i}} + \boldsymbol{\delta_i}$, we have $||\hat{\boldsymbol{\theta}}^{(p)a}_{\boldsymbol{i}} - \boldsymbol{\theta}^{\boldsymbol{0}}_{\boldsymbol{i}}|| = ||\hat{\boldsymbol{\theta}}^{(p)}_{\boldsymbol{i}} + \boldsymbol{\delta_i} - \boldsymbol{\theta}^{\boldsymbol{0}}_{\boldsymbol{i}}|| \leq ||\hat{\boldsymbol{\theta}}^{(p)}_{\boldsymbol{i}} - \boldsymbol{\theta}^{\boldsymbol{0}}_{\boldsymbol{i}}|| = O_p(T^{-1/2})$. The inequality follows from Assumptions 3.4 and 3.5 that the positive elements of $\boldsymbol{\delta_i}$, when added by $\hat{\boldsymbol{\theta}}^{(p)}_{\boldsymbol{i}}$, is always less than or equal to the corresponding entries of the true $\boldsymbol{\theta}^{\boldsymbol{0}}_{\boldsymbol{i}}$. This completes the proof. ■

Appendix G Proof of Proposition 3.4

Proof of Proposition 3.4. For $i = 1, 2$, recall $\boldsymbol{R}^{\boldsymbol{0}}_{\boldsymbol{i}} = \mathbb{E}[(\boldsymbol{D}^{\boldsymbol{0}}_{\boldsymbol{it}})^{-1/2}\boldsymbol{\epsilon_{it}}\boldsymbol{\epsilon}'_{\boldsymbol{it}}(\boldsymbol{D}^{\boldsymbol{0}}_{\boldsymbol{it}})^{-1/2}]$ and $\hat{\boldsymbol{R}}_{\boldsymbol{i}} = T^{-1}\sum_{t=1}^{T} \hat{\boldsymbol{D}}_{\boldsymbol{it}}^{-1/2}\boldsymbol{\epsilon_{it}}\boldsymbol{\epsilon}'_{\boldsymbol{it}}\hat{\boldsymbol{D}}_{\boldsymbol{it}}^{-}$ We shall show that $\hat{\boldsymbol{R}}_{\boldsymbol{i}} - \boldsymbol{R}^{\boldsymbol{0}}_{\boldsymbol{i}} = O_p(T^{-1/2})$, $i = 1, 2$. Let $\hat{\boldsymbol{R}}^{\boldsymbol{0}}_{\boldsymbol{i}} \equiv T^{-1}\sum_{t=1}^{T}(\boldsymbol{D}^{\boldsymbol{0}}_{\boldsymbol{it}})^{-1/2}\boldsymbol{\epsilon_{it}}\boldsymbol{\epsilon}'_{\boldsymbol{it}}(\boldsymbol{D}^{\boldsymbol{0}}_{\boldsymbol{it}})^{-1/2}$ denote the analogue of $\hat{\boldsymbol{R}}_{\boldsymbol{i}}$ based on true variance. We consider the following decomposition

$$\hat{\boldsymbol{R}}_{\boldsymbol{i}} - \boldsymbol{R}^{\boldsymbol{0}}_{\boldsymbol{i}} = (\hat{\boldsymbol{R}}_{\boldsymbol{i}} - \hat{\boldsymbol{R}}^{\boldsymbol{0}}_{\boldsymbol{i}}) + (\hat{\boldsymbol{R}}^{\boldsymbol{0}}_{\boldsymbol{i}} - \boldsymbol{R}^{\boldsymbol{0}}_{\boldsymbol{i}}).$$

By Chebyshev's inequality and Assumption 3.6, we have for the second term $\hat{\boldsymbol{R}}^{\boldsymbol{0}}_{\boldsymbol{i}} - \boldsymbol{R}^{\boldsymbol{0}}_{\boldsymbol{i}} = O_p(T^{-1/2})$. It remains to show that the first term $\hat{\boldsymbol{R}}_{\boldsymbol{i}} - \hat{\boldsymbol{R}}^{\boldsymbol{0}}_{\boldsymbol{i}} = O_p(T^{-1/2})$. We write

$$\begin{aligned}\hat{\boldsymbol{R}}_{\boldsymbol{i}} - \hat{\boldsymbol{R}}^{\boldsymbol{0}}_{\boldsymbol{i}} &= \frac{1}{T}\sum_{t=1}^{T} \hat{\boldsymbol{D}}_{\boldsymbol{it}}^{-1/2}\boldsymbol{\epsilon_{it}}\boldsymbol{\epsilon}'_{\boldsymbol{it}}\hat{\boldsymbol{D}}_{\boldsymbol{it}}^{-1/2} - (\boldsymbol{D}^{\boldsymbol{0}}_{\boldsymbol{it}})^{-1/2}\boldsymbol{\epsilon_{it}}\boldsymbol{\epsilon}'_{\boldsymbol{it}}(\boldsymbol{D}^{\boldsymbol{0}}_{\boldsymbol{it}})^{-1/2} \\ &= \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{\Psi_{it}}(\hat{\boldsymbol{\theta}}^{(p)a}_{\boldsymbol{i}}) - \boldsymbol{\Psi_{it}}(\boldsymbol{\theta}^{\boldsymbol{0}}_{\boldsymbol{i}}), \text{ say.}\end{aligned}$$

Let $\Psi^{m,n}_{it}(\hat{\boldsymbol{\theta}}^{(p)a}_{\boldsymbol{i}})$ and $\Psi^{m,n}_{it}(\hat{\boldsymbol{\theta}}^{\boldsymbol{0}}_{\boldsymbol{i}})$ denote the (m,n)-th entries of $\boldsymbol{\Psi_{it}}(\hat{\boldsymbol{\theta}}^{(p)a}_{\boldsymbol{i}})$ and $\boldsymbol{\Psi_{it}}(\boldsymbol{\theta}^{\boldsymbol{0}}_{\boldsymbol{i}})$, respectively. It suffices to show that $T^{-1}\sum_{t=1}^{T}\Psi^{m,n}_{it}(\hat{\boldsymbol{\theta}}^{(p)a}_{\boldsymbol{i}}) - \Psi^{m,n}_{it}(\hat{\boldsymbol{\theta}}^{\boldsymbol{0}}_{\boldsymbol{i}}) = O_p(T^{-1/2})$. By the mean value theorem and Cauchy-Schwarz inequality, we have

$$\frac{1}{T}\sum_{t=1}^{T}\Psi_{it}^{m,n}(\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{\boldsymbol{(p)a}})-\Psi_{it}^{m,n}(\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{\boldsymbol{0}})=\frac{1}{T}\sum_{t=1}^{T}\big[\nabla_{\boldsymbol{\theta}_{\boldsymbol{i}}}\Psi_{it}^{m,n}(\bar{\boldsymbol{\theta}}_{\boldsymbol{i}})\big]'\big[\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{\boldsymbol{(p)a}}-\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{\boldsymbol{0}}\big]$$
$$\leq\big|\big|\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{\boldsymbol{(p)a}}-\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{\boldsymbol{0}}\big|\big|\frac{1}{T}\sum_{t=1}^{T}\big|\big|\nabla_{\boldsymbol{\theta}_{\boldsymbol{i}}}\Psi_{it}^{m,n}(\bar{\boldsymbol{\theta}}_{\boldsymbol{i}})\big|\big|,$$

where $\bar{\boldsymbol{\theta}}_{\boldsymbol{i}}$ lies between $\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{\boldsymbol{(p)a}}$ and $\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{\boldsymbol{0}}$. Given Assumption 2.4, we have $T^{-1}\sum_{t=1}^{T}\big|\big|\nabla_{\boldsymbol{\theta}_{\boldsymbol{i}}}\Psi_{it}^{m,n}(\bar{\boldsymbol{\theta}}_{\boldsymbol{i}})\big|\big|=O_p(1)$ by Markov's inequality. Next, we have $\big|\big|\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{\boldsymbol{(p)a}}-\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}^{\boldsymbol{0}}\big|\big|=O_p(T^{-1/2})$ from Proposition 3.3. This gives $\hat{\boldsymbol{R}}_{\boldsymbol{i}}-\hat{\boldsymbol{R}}_{\boldsymbol{i}}^{\boldsymbol{0}}=O_p(T^{-1/2})$ and completes the proof. ■

Supplement to "A practical multivariate approach to testing volatility spillover"

1. QMLE-based simulations

We report in this section the bivariate simulation results based on model described in (28)–(29) where the ARCH(∞) process (29) is estimated using QMLE. That is, we reproduce Table 1 (size study) and Table 2 (power study) of the main paper but with spillover testing based on estimating (29) using QMLE. We denote such modeling procedure by NC-QMLE. Due to the computational cost associated with QMLE, we focus on the asymptotic tests with 1000 number of simulations. Table S1 reports the size results and Table S2 reports the power results.

Table S1: Empirical sizes of Q_1

		NullA			NullB			NullC		
T	M	10	20	30	10	20	30	10	20	30
Rejection rates based on asymptotic critical values										
1000	$Q_{1\mathrm{BAR}}$	7.1	6.9	6.7	7.7	7.4	7.6	6.9	6.9	7.1
	$Q_{1\mathrm{DAN}}$	7.0	6.8	7.4	7.5	7.4	7.0	7.7	6.5	6.4
	$Q_{1\mathrm{QS}}$	7.3	7.0	7.6	7.7	7.6	7.0	7.5	6.8	6.8
	$Q_{1\mathrm{TR}}$	6.8	7.1	6.6	7.5	6.1	7.2	7.1	7.0	6.8
1500	$Q_{1\mathrm{BAR}}$	6.7	5.9	5.3	7.2	6.5	6.0	6.3	6.9	6.8
	$Q_{1\mathrm{DAN}}$	6.6	5.6	5.4	7.1	6.2	5.6	6.9	7.1	6.5
	$Q_{1\mathrm{QS}}$	6.1	5.7	6.2	7.0	6.0	6.1	7.2	6.8	6.2
	$Q_{1\mathrm{TR}}$	6.3	6.4	6.9	6.4	6.6	7.1	6.6	5.7	6.4

NOTES: The table reports empirical sizes (in %) of Q_1 under NullA, NullB and NullC at the 5% significance level based on NC-QMLE modeling. Number of simulations = 1000. $Q_{1\mathrm{BAR}}$, $Q_{1\mathrm{DAN}}$, $Q_{1\mathrm{QS}}$, $Q_{1\mathrm{TR}}$ denote the rejection rates of Q_1 using asymptotic critical values; the subscripts BAR, DAN, QS and TR denote, respectively, the Bartlett kernel, the Daniell kernel, the Quadratic-Spectral kernel and the Truncated kernel. T and M denote the sample size and kernel smoothing parameter, respectively.

From Table S1, we observe that Q_1 over rejects the null a little but not excessively. The size of the test generally improves in T, across different data generating processes. This is consistent with Table 1 of the paper. Similarly, we observe that the power patterns reported in Table S2 is very similar to those reported in Table 2 of the paper. Overall, this study highlights that, under regularity conditions the asymptotic theory of the proposed test statistic is robust to different estimation procedures. Besides, the study also highlights the appealing feature of the proposed NC-LS modeling, because its computational efficiency makes bootstrap approach feasible to yield a more accurate finite sample performance for the test statistic.

Table S2: Empirical powers of Q_1

T	M	AlterA			AlterB		
		10	20	30	10	20	30
Rejection rates based on empirical critical values							
1000	$Q_{1\text{BAR}}$	78.7	67.1	59.9	96.0	88.8	83.5
	$Q_{1\text{DAN}}$	73.8	61.1	52.4	93.8	84.2	77.5
	$Q_{1\text{QS}}$	72.3	60.7	50.6	92.8	83.0	75.3
	$Q_{1\text{TR}}$	52.0	36.1	31.3	77.2	59.2	53.7
1500	$Q_{1\text{BAR}}$	92.4	85.5	79.3	99.7	98.6	96.5
	$Q_{1\text{DAN}}$	89.4	79.9	71.5	99.5	96.7	94.5
	$Q_{1\text{QS}}$	89.0	77.9	69.3	99.5	96.0	93.9
	$Q_{1\text{TR}}$	70.3	51.0	43.1	92.6	80.3	71.3

NOTES: The table reports empirical powers (in %) of Q_1 under AlterA and AlterB at the 5% significance level based on NC-QMLE modeling. Number of simulations = 1000. $Q_{1\text{BAR}}$, $Q_{1\text{DAN}}$, $Q_{1\text{QS}}$, $Q_{1\text{TR}}$ denote the rejection rates of Q_1 using empirical critical values; the subscripts BAR, DAN, QS and TR denote, respectively, the Bartlett kernel, the Daniell kernel, the Quadratic-Spectral kernel and the Truncated kernel. T and M denote the sample size and kernel smoothing parameter, respectively.

2. Simulation results of Tables 3 and 4 for all M

We collect in this section the full set of simulation results for the dimensional study in Section 4.2 of the main paper where only the results for $M = 20$ are reported to save space. It is useful to recall that the size study under NullD and the power study under AlterC are reported, respectively, in Tables 3 and 4 of the paper.

Now, we report for $M = 10, 20, 30$, the size study under NullD in Table S3, and the power study under AlterC in Table S4. Consistent with the paper, we report here both the results based on asymptotic test (Q_1) and bootstrap test (Q_1^*). In general, we find that the size of both Q_1 and Q_1^* to be reasonably stable across M for each of the dimension considered. This result is consistent with the bivariate case reported in Table 1 of the paper. As for the power study, we find that the rejection rates of Q_1 and Q_1^* decrease in M. This is because we have one-period lag in volatility spillover under AlterC. Thus, a test that focuses on recent events is expected give better power. This finding is consistent with the bivariate case reported in Table 2 of the paper. We also observe that the rejection rates of both Q_1 and Q_1^* increase in T and d_1, consistent with the higher dimensional results reported in Table 4 of the paper. This finding appears to hold regardless of the choice M.

Table S3: Empirical sizes

		NullD																							
	d_1	3			4			5			6			7			8			9			10		
T	M	10	20	30	10	20	30	10	20	30	10	20	30	10	20	30	10	20	30	10	20	30	10	20	30
Rejection rates based on asymptotic critical values																									
1000	$Q_{1\mathrm{BAR}}$	7.0	7.1	7.2	7.0	7.1	7.0	7.2	7.1	7.4	7.4	7.2	7.3	7.5	7.4	7.5	7.9	7.9	8.0	7.9	7.9	8.4	7.5	7.6	8.0
	$Q_{1\mathrm{DAN}}$	6.9	7.3	7.2	7.1	6.9	6.9	7.2	7.1	7.5	7.6	7.3	7.5	7.3	7.3	7.7	7.9	7.9	8.2	8.1	8.2	8.7	7.5	7.9	8.6
	$Q_{1\mathrm{QS}}$	6.9	7.1	7.2	7.1	7.0	7.0	7.2	7.3	7.4	7.4	7.2	7.4	7.4	7.4	7.5	7.8	7.8	8.1	8.1	8.3	8.3	7.5	7.9	8.7
	$Q_{1\mathrm{TR}}$	7.1	7.2	7.4	6.5	7.3	7.5	6.9	7.4	8.0	7.5	7.4	8.0	7.3	7.5	8.4	7.6	8.1	8.9	8.0	8.8	9.0	7.8	8.9	9.9
1500	$Q_{1\mathrm{BAR}}$	7.0	6.7	6.8	6.7	7.0	7.1	6.9	6.6	6.7	6.8	6.8	6.7	6.5	6.3	6.6	6.9	7.2	7.1	6.9	7.1	7.2	7.2	6.8	7.1
	$Q_{1\mathrm{DAN}}$	7.0	6.6	6.7	6.8	6.9	7.1	6.7	6.6	6.8	6.8	6.7	6.7	6.5	6.4	6.5	7.0	7.2	7.1	6.9	7.2	7.4	7.0	7.0	7.2
	$Q_{1\mathrm{QS}}$	6.9	6.7	6.7	6.8	6.9	7.0	6.6	6.6	6.9	6.9	6.7	6.7	6.3	6.4	6.6	7.2	7.2	7.2	6.9	7.1	7.4	6.8	7.1	7.1
	$Q_{1\mathrm{TR}}$	6.3	6.7	6.5	7.0	6.9	7.0	7.0	7.1	7.4	6.8	7.0	7.2	6.4	7.0	7.6	7.1	6.8	7.4	7.1	7.0	8.1	7.0	7.3	7.9
Rejection rates based on bootstrap critical values																									
1000	$Q^*_{1\mathrm{BAR}}$	4.8	4.8	5.0	5.3	5.5	5.5	5.0	5.5	5.4	5.1	5.2	5.5	5.5	5.5	5.6	5.0	5.2	5.5	5.3	5.3	5.4	5.8	5.8	5.8
	$Q^*_{1\mathrm{DAN}}$	5.0	4.9	5.1	5.3	5.6	5.7	5.1	5.3	5.6	5.2	5.5	5.3	5.7	5.4	5.7	5.1	5.3	5.5	5.4	5.4	5.6	5.8	5.7	5.7
	$Q^*_{1\mathrm{QS}}$	5.0	5.1	5.2	5.3	5.5	5.7	5.1	5.5	5.5	5.3	5.5	5.4	5.6	5.5	5.7	5.1	5.5	5.6	5.3	5.3	5.6	5.8	5.8	5.7
	$Q^*_{1\mathrm{TR}}$	5.1	5.1	5.5	5.5	5.8	5.9	5.2	5.7	5.3	5.3	5.3	5.6	5.7	5.7	5.9	5.5	5.7	5.6	5.5	5.7	5.7	5.8	5.6	5.7
1500	$Q^*_{1\mathrm{BAR}}$	4.9	5.1	5.3	5.4	5.2	5.3	5.0	4.8	4.9	5.1	5.0	5.2	5.0	5.4	5.5	5.4	5.5	5.4	5.3	5.4	5.4	5.2	5.1	5.3
	$Q^*_{1\mathrm{DAN}}$	5.1	5.1	5.1	5.2	5.3	5.1	5.0	4.8	5.0	5.0	5.2	5.3	5.1	5.3	5.5	5.4	5.5	5.4	5.2	5.4	5.6	5.1	5.1	5.4
	$Q^*_{1\mathrm{QS}}$	5.1	5.2	5.2	5.1	5.2	5.2	4.9	4.8	4.9	5.1	5.2	5.3	5.3	5.4	5.5	5.3	5.5	5.5	5.1	5.4	5.5	5.1	5.1	5.5
	$Q^*_{1\mathrm{TR}}$	5.3	5.3	5.4	5.5	5.4	5.2	4.6	5.2	5.1	5.0	5.3	5.3	5.2	5.5	5.2	5.5	5.6	5.2	5.2	5.3	5.6	5.5	5.5	5.3

NOTES: The table reports empirical sizes (in %) of Q_1 under NullD at the 5% significance level based on NC-LS modeling. Number of simulations = 10000. $Q_{1\mathrm{BAR}}$, $Q_{1\mathrm{DAN}}$, $Q_{1\mathrm{QS}}$, $Q_{1\mathrm{TR}}$ and $Q^*_{1\mathrm{BAR}}$, $Q^*_{1\mathrm{DAN}}$, $Q^*_{1\mathrm{QS}}$, $Q^*_{1\mathrm{TR}}$ denote the rejection rates of Q_1 using asymptotic and bootstrap critical values, respectively; the subscripts BAR, DAN, QS and TR denote, respectively, the Bartlett kernel, the Daniell kernel, the Quadratic-Spectral kernel and the Truncated kernel. Number of bootstraps = 499. M denotes the kernel smoothing parameter. T and d_1 denote the sample size and dimension of portfolio 1, respectively.

Table S4: Empirical powers

		AlterC																							
	d_1	3			4			5			6			7			8			9			10		
T	M	10	20	30	10	20	30	10	20	30	10	20	30	10	20	30	10	20	30	10	20	30	10	20	30
Rejection rates based on empirical critical values																									
1000	$Q_{1\text{BAR}}$	38.4	27.6	23.1	55.3	40.8	33.2	72.9	55.2	45.6	82.7	65.5	55.0	89.3	73.9	62.6	91.5	77.1	66.8	93.7	80.9	70.4	95.3	83.0	71.4
	$Q_{1\text{DAN}}$	33.0	23.0	19.2	48.2	33.4	27.1	66.3	44.8	35.6	75.9	54.9	43.9	83.2	62.5	50.7	86.3	65.9	53.9	89.3	70.3	57.5	91.3	71.4	58.2
	$Q_{1\text{QS}}$	31.9	22.1	18.7	46.0	31.6	25.9	64.0	43.2	33.5	73.4	52.9	41.9	80.9	60.1	48.2	84.3	64.1	51.9	87.4	67.4	54.9	89.4	69.0	55.5
	$Q_{1\text{TR}}$	18.2	14.3	12.7	26.4	18.8	15.7	34.4	23.7	19.5	41.7	30.0	23.3	50.5	34.2	26.9	52.9	35.5	28.7	55.1	37.1	29.5	57.7	37.8	28.8
1500	$Q_{1\text{BAR}}$	57.2	42.6	34.7	82.2	63.0	52.0	95.5	84.2	72.8	98.6	91.4	82.6	99.6	96.0	90.6	99.8	97.7	93.6	100.0	98.6	94.9	99.9	99.1	96.6
	$Q_{1\text{DAN}}$	50.3	34.1	27.1	74.8	51.9	41.3	91.8	73.2	59.5	96.5	82.4	70.6	98.8	90.7	80.3	99.5	93.6	84.2	99.8	95.0	87.4	99.8	96.8	89.6
	$Q_{1\text{QS}}$	48.1	33.1	26.0	72.1	49.5	39.6	90.4	70.7	56.5	95.7	80.3	68.5	98.3	88.9	77.8	99.2	92.2	81.8	99.6	93.8	85.1	99.7	95.8	88.1
	$Q_{1\text{TR}}$	27.9	18.8	16.7	39.7	28.1	23.3	57.2	39.2	30.5	68.5	47.8	37.2	79.2	57.2	45.6	84.0	61.9	49.5	85.7	66.3	51.2	89.0	67.8	54.0
Rejection rates based on bootstrap critical values																									
1000	$Q^*_{1\text{BAR}}$	33.1	23.8	20.0	50.8	36.6	29.4	70.4	52.0	42.2	79.8	62.2	51.1	86.7	70.1	58.2	90.7	74.9	63.0	93.2	78.6	67.2	94.0	80.9	69.4
	$Q^*_{1\text{DAN}}$	28.4	19.6	16.7	43.5	29.6	23.4	62.2	42.1	33.4	72.6	50.9	40.1	80.3	58.1	46.3	84.6	63.3	50.4	88.0	67.2	53.7	89.2	69.8	56.6
	$Q^*_{1\text{QS}}$	27.2	19.0	16.0	41.8	28.2	22.3	59.5	40.3	31.8	70.1	48.4	38.2	77.9	55.6	44.4	82.4	60.4	48.1	85.8	64.6	51.8	87.5	67.1	54.1
	$Q^*_{1\text{TR}}$	15.7	12.4	10.9	22.1	15.6	13.7	31.8	21.7	18.0	38.2	26.4	21.4	43.7	30.4	24.4	48.3	32.9	26.6	52.2	35.9	28.4	54.9	37.0	29.4
1500	$Q^*_{1\text{BAR}}$	54.3	38.7	31.1	80.7	62.3	50.9	94.7	81.3	70.3	98.1	89.9	80.3	99.4	94.8	88.3	99.8	97.3	92.1	99.9	98.4	94.4	100.0	98.9	95.9
	$Q^*_{1\text{DAN}}$	47.0	31.2	24.4	73.3	50.9	39.9	90.0	70.2	56.5	95.7	80.3	67.1	98.3	87.9	76.2	99.3	92.2	81.9	99.6	94.4	85.8	99.9	95.9	88.1
	$Q^*_{1\text{QS}}$	44.9	29.6	23.7	70.7	48.4	38.0	88.3	67.4	54.1	94.7	77.8	64.4	97.7	86.1	73.5	99.0	90.5	79.5	99.5	93.0	83.7	99.8	94.8	86.0
	$Q^*_{1\text{TR}}$	23.5	16.9	14.1	37.6	26.0	20.9	54.1	36.4	28.3	64.8	44.5	35.0	75.0	52.0	41.0	80.8	58.3	46.4	83.8	61.8	49.2	86.8	65.9	52.2

NOTES: The table reports empirical powers (in %) of Q_1 under AlterC at the 5% significance level based on NC-LS modeling. Number of simulations = 10000. $Q_{1\text{BAR}}$, $Q_{1\text{DAN}}$, $Q_{1\text{QS}}$, $Q_{1\text{TR}}$ and $Q^*_{1\text{BAR}}$, $Q^*_{1\text{DAN}}$, $Q^*_{1\text{QS}}$, $Q^*_{1\text{TR}}$ denote the rejection rates of Q_1 using empirical and bootstrap critical values, respectively; the subscripts BAR, DAN, QS and TR denote, respectively, the Bartlett kernel, the Daniell kernel, the Quadratic-Spectral kernel and the Truncated kernel. Number of bootstraps = 499. M denotes the kernel smoothing parameter. T and d_1 denote the sample size and dimension of portfolio 1, respectively.

3. Monte Carlo study of the bidirectional test

We report and discuss in this section the full set of simulation results for our bidirectional spillover test based on the proposed NC-LS modeling. To ensure consistency and comparability, we maintain the same experimental design and parameter combinations as those described in Section 4 of the paper. As per the unidirectional study, we conduct simulations for both the asymptotic test (Q_2) and bootstrap test (Q_2^*). Overall, the finite sample performance of Q_2 and Q_2^* is very similar to that of Q_1 and Q_1^*, as will be discussed in the following.

To keep the presentation consistent, we tabulate and present the simulation results of the bidirectional tests in the same ordering as their unidirectional counterparts. First, we report the bivariate simulation results in Tables S5 and S6, which represents the bidirectional counterparts of Tables 1 and 2 in the paper. Next, we report the dimensional study (for the case of $M = 20$) in Tables S7 and S8, which correspond to Tables 3 and 4 in the paper. Finally, we also report the full set of dimensional study (i.e., for $M = 10, 20, 30$) in Tables S9 and S10, which represent the bidirectional counterparts of Tables S3 and S4 in this supplementary document.

We begin with the bivariate study. Table S5 reports the size performance of our bidirectional testing strategy. We find that the size pattern of Q_2 and Q_2^* is very similar to that of Q_1 and Q_1^* reported in Table 1 of the paper. In general, the size of our bidirectional approach is reasonable and it improves as T increases. The size is also stable across the different parameter combinations considered; the four kernel functions studied and their smoothing parameters M. Table S6 reports the power performance of our bidirectional testing approach. We can see that the power of Q_2 and Q_2^* is slightly lower than that of Q_1 and Q_1^* reported in Table 2 of the paper. This minor loss in power is expected because the bidirectional tests check both positive and negative lag order j's for evidence of spillover, whereas the unidirectional tests check only the positive ones. Other than this, the overall power pattern of Q_2 and Q_2^* is very similar to that of Q_1 and Q_1^*.

We now turn to the simulation results of the higher dimensional study of Q_2 and Q_2^* reported in Tables S7, S8, S9 and S10. Because the results in Tables S7 and S8 are embedded, respectively, in Tables S9 and S10, we shall focus on discussing the latter. Table S9 reports the size study of Q_2 and Q_2^* as the portfolio dimension increases. First, we find that the size of Q_2 and Q_2^* increases in d_1, but not overly excessive nor rapid. The size generally improves and stabilizes as T increases. The observed trend is similar to that of the unidirectional tests reported in Table 3 of the paper. We also find the size of Q_2 and Q_2^* to be stable across M for each of the dimension studied, consistent with their unidirectional counterparts reported in Table S3.

Table S10 reports the power study of our bidirectional inferential approach as d_1 increases. We find that Q_2 and Q_2^* have power despite a rather low spillover intensity. We also find that the power increases in d_1 as the spillover evidence becomes stronger due to the increased number of risk recipients. This power trend is consistent with that of the unidirectional tests reported in Table 4 of the paper. We also find that the rejection rates of Q_2 and Q_2^* decrease in M due to the one-period lag spillover, in line with the pattern given by Q_1 and Q_1^* in Table S4. Similar to the bivariate study reported in Table S6, we observe here some minor loss in power of Q_2 and Q_2^* compared with Q_1 and

Table S5: Empirical sizes

T	M	NullA 10	NullA 20	NullA 30	NullB 10	NullB 20	NullB 30	NullC 10	NullC 20	NullC 30
Rejection rates based on asymptotic critical values										
1000	$Q_{2\text{BAR}}$	7.0	6.9	6.8	7.0	6.9	6.7	7.0	7.0	6.8
	$Q_{2\text{DAN}}$	6.9	6.9	6.8	6.8	6.9	6.8	6.8	6.9	6.9
	$Q_{2\text{QS}}$	7.0	6.9	6.9	6.9	6.9	6.8	7.0	6.9	6.9
	$Q_{2\text{TR}}$	7.0	6.8	6.4	7.1	6.8	6.6	6.9	6.8	6.4
1500	$Q_{2\text{BAR}}$	6.7	6.7	6.5	6.7	6.7	6.6	6.7	6.7	6.4
	$Q_{2\text{DAN}}$	6.7	6.6	6.3	6.7	6.8	6.5	6.8	6.5	6.3
	$Q_{2\text{QS}}$	6.7	6.5	6.3	6.8	6.8	6.3	6.8	6.6	6.4
	$Q_{2\text{TR}}$	6.7	6.1	6.4	6.8	6.2	6.5	6.7	6.1	6.4
Rejection rates based on bootstrap critical values										
1000	$Q^*_{2\text{BAR}}$	5.4	5.4	5.4	5.3	5.4	5.4	5.3	5.3	5.4
	$Q^*_{2\text{DAN}}$	5.5	5.4	5.2	5.4	5.5	5.4	5.5	5.5	5.2
	$Q^*_{2\text{QS}}$	5.3	5.3	5.3	5.3	5.5	5.2	5.4	5.5	5.3
	$Q^*_{2\text{TR}}$	5.5	5.1	5.3	5.5	5.3	5.4	5.5	5.3	5.3
1500	$Q^*_{2\text{BAR}}$	4.5	4.7	5.0	4.5	4.6	5.1	4.6	4.7	5.1
	$Q^*_{2\text{DAN}}$	4.6	4.8	5.2	4.5	4.8	5.2	4.7	4.9	5.3
	$Q^*_{2\text{QS}}$	4.6	4.8	5.2	4.6	5.0	5.1	4.6	4.8	5.2
	$Q^*_{2\text{TR}}$	4.9	5.2	5.3	4.8	5.1	5.1	4.9	5.2	5.3

NOTES: The table reports empirical sizes (in %) of Q_2 under NullA, NullB and NullC at the 5% significance level based on NC-LS modeling. Number of simulations = 10000. $Q_{2\text{BAR}}$, $Q_{2\text{DAN}}$, $Q_{2\text{QS}}$, $Q_{2\text{TR}}$ and $Q^*_{2\text{BAR}}$, $Q^*_{2\text{DAN}}$, $Q^*_{2\text{QS}}$, $Q^*_{2\text{TR}}$ denote the rejection rates of Q_2 using asymptotic and bootstrap critical values, respectively; the subscripts BAR, DAN, QS and TR denote, respectively, the Bartlett kernel, the Daniell kernel, the Quadratic-Spectral kernel and the Truncated kernel. Number of bootstraps = 499. T and M denote the sample size and kernel smoothing parameter, respectively.

Q^*_1. This is again due to the fact that Q_2 and Q^*_2 examine both positive and negative directions for spillover evidence while Q_1 and Q^*_1 check only the positive direction. Apart from this, the overall power pattern of Q_2 and Q^*_2 is very similar to that of Q_1 and Q^*_1, and the power generally improves as T increases.

Table S6: Empirical powers

T	M	AlterA 10	AlterA 20	AlterA 30	AlterB 10	AlterB 20	AlterB 30
Rejection rates based on empirical critical values							
1000	$Q_{2\text{BAR}}$	61.2	52.4	45.7	84.5	76.7	69.9
	$Q_{2\text{DAN}}$	58.7	46.0	38.5	82.5	70.4	61.5
	$Q_{2\text{QS}}$	57.6	44.4	37.3	81.5	68.4	59.9
	$Q_{2\text{TR}}$	37.3	27.5	23.0	60.2	45.7	37.3
1500	$Q_{2\text{BAR}}$	79.3	70.5	63.9	96.7	93.4	89.9
	$Q_{2\text{DAN}}$	76.6	64.0	55.1	95.8	89.9	82.8
	$Q_{2\text{QS}}$	75.4	61.9	53.2	95.3	88.6	81.3
	$Q_{2\text{TR}}$	53.2	40.8	33.6	81.9	67.5	57.1
Rejection rates based on bootstrap critical values							
1000	$Q^*_{2\text{BAR}}$	59.4	49.9	43.8	83.7	75.9	68.3
	$Q^*_{2\text{DAN}}$	56.4	43.4	36.6	81.6	68.5	59.2
	$Q^*_{2\text{QS}}$	54.9	42.2	35.0	80.4	66.6	57.4
	$Q^*_{2\text{TR}}$	35.1	25.0	20.8	57.5	42.4	35.1
1500	$Q^*_{2\text{BAR}}$	77.7	69.4	62.0	95.9	92.2	87.9
	$Q^*_{2\text{DAN}}$	75.5	62.1	52.9	95.0	88.3	81.5
	$Q^*_{2\text{QS}}$	74.1	60.3	51.2	94.6	86.9	80.0
	$Q^*_{2\text{TR}}$	51.7	36.9	30.2	80.3	64.6	53.8

NOTES: The table reports empirical powers (in %) of Q_2 under AlterA and AlterB at the 5% significance level based on NC-LS modeling. Number of simulations = 10000. $Q_{2\text{BAR}}$, $Q_{2\text{DAN}}$, $Q_{2\text{QS}}$, $Q_{2\text{TR}}$ and $Q^*_{2\text{BAR}}$, $Q^*_{2\text{DAN}}$, $Q^*_{2\text{QS}}$, $Q^*_{2\text{TR}}$ denote the rejection rates of Q_2 using empirical and bootstrap critical values, respectively; the subscripts BAR, DAN, QS and TR denote, respectively, the Bartlett kernel, the Daniell kernel, the Quadratic-Spectral kernel and the Truncated kernel. Number of bootstraps = 499. T and M denote the sample size and kernel smoothing parameter, respectively.

Table S7: Empirical sizes

T	d_1	NullD							
		3	4	5	6	7	8	9	10
Rejection rates based on asymptotic critical values									
1000	$Q_{2\text{BAR}}$	7.1	7.0	7.6	7.9	7.3	7.6	7.9	8.3
	$Q_{2\text{DAN}}$	7.0	7.1	7.8	8.0	7.4	8.0	8.3	8.5
	$Q_{2\text{QS}}$	6.9	7.1	7.8	7.9	7.3	7.9	8.3	8.5
	$Q_{2\text{TR}}$	6.9	7.1	7.5	8.6	8.1	8.6	8.7	9.3
1500	$Q_{2\text{BAR}}$	6.3	6.4	6.6	6.8	6.9	7.5	7.4	7.3
	$Q_{2\text{DAN}}$	6.3	6.6	6.8	6.9	6.9	7.4	7.5	7.5
	$Q_{2\text{QS}}$	6.3	6.5	6.6	7.0	6.9	7.6	7.3	7.5
	$Q_{2\text{TR}}$	6.3	6.2	7.0	7.4	7.5	7.4	7.9	7.7
Rejection rates based on bootstrap critical values									
1000	$Q^*_{2\text{BAR}}$	5.3	5.7	5.2	5.4	6.1	6.2	6.0	6.2
	$Q^*_{2\text{DAN}}$	5.6	5.7	5.1	5.3	6.3	6.1	6.0	6.2
	$Q^*_{2\text{QS}}$	5.5	5.6	5.3	5.3	6.2	6.1	6.0	6.2
	$Q^*_{2\text{TR}}$	5.7	5.8	5.5	5.5	6.2	6.3	6.3	6.1
1500	$Q^*_{2\text{BAR}}$	4.6	5.2	4.8	4.6	5.6	5.0	5.7	5.4
	$Q^*_{2\text{DAN}}$	4.9	5.3	4.8	5.1	5.5	5.1	5.8	5.2
	$Q^*_{2\text{QS}}$	5.0	5.3	4.8	5.1	5.3	5.0	5.9	5.3
	$Q^*_{2\text{TR}}$	4.8	5.0	5.1	5.4	5.8	5.5	5.6	5.6

NOTES: The table reports empirical sizes (in %) of Q_2 under NullD at the 5% significance level based on NC-LS modeling. Number of simulations = 10000. $Q_{2\text{BAR}}$, $Q_{2\text{DAN}}$, $Q_{2\text{QS}}$, $Q_{2\text{TR}}$ and $Q^*_{2\text{BAR}}$, $Q^*_{2\text{DAN}}$, $Q^*_{2\text{QS}}$, $Q^*_{2\text{TR}}$ denote the rejection rates of Q_2 using asymptotic and bootstrap critical values, respectively; the subscripts BAR, DAN, QS and TR denote, respectively, the Bartlett kernel, the Daniell kernel, the Quadratic-Spectral kernel and the Truncated kernel. Number of bootstraps = 499. T and d_1 denote the sample size and dimension of portfolio 1, respectively.

Table S8: Empirical powers

T	d_1	AlterC							
		3	4	5	6	7	8	9	10
Rejection rates based on empirical critical values									
1000	$Q_{2\text{BAR}}$	18.6	25.5	34.6	39.7	48.9	55.1	55.9	58.0
	$Q_{2\text{DAN}}$	16.8	21.9	28.4	32.6	41.2	45.8	46.7	48.2
	$Q_{2\text{QS}}$	16.5	21.2	27.3	31.3	39.6	43.9	44.3	45.3
	$Q_{2\text{TR}}$	11.6	14.0	17.2	18.9	22.4	24.5	24.7	25.1
1500	$Q_{2\text{BAR}}$	28.5	42.9	58.0	69.0	78.5	82.7	86.1	89.4
	$Q_{2\text{DAN}}$	23.7	36.0	48.2	59.0	68.0	73.0	77.4	80.5
	$Q_{2\text{QS}}$	23.4	34.8	45.8	56.8	65.4	70.9	75.1	78.3
	$Q_{2\text{TR}}$	14.3	20.3	25.0	31.4	35.9	40.1	42.9	45.4
Rejection rates based on bootstrap critical values									
1000	$Q^*_{2\text{BAR}}$	16.0	23.4	32.1	38.9	46.1	50.7	53.4	57.5
	$Q^*_{2\text{DAN}}$	14.0	20.1	26.3	32.7	37.9	41.6	44.6	47.6
	$Q^*_{2\text{QS}}$	13.8	19.0	25.1	31.2	36.1	39.7	42.5	45.3
	$Q^*_{2\text{TR}}$	10.1	12.0	16.0	18.0	20.8	22.0	23.8	25.1
1500	$Q^*_{2\text{BAR}}$	23.2	38.3	56.0	66.8	75.7	80.7	85.6	87.6
	$Q^*_{2\text{DAN}}$	19.7	31.7	46.7	56.4	65.6	70.7	76.7	79.1
	$Q^*_{2\text{QS}}$	19.3	30.5	44.8	54.0	62.7	68.0	74.0	76.6
	$Q^*_{2\text{TR}}$	13.3	17.2	24.2	29.5	34.0	38.2	41.3	43.9

NOTES: The table reports empirical powers (in %) of Q_2 under AlterC at the 5% significance level based on NC-LS modeling. Number of simulations = 10000. $Q_{2\text{BAR}}$, $Q_{2\text{DAN}}$, $Q_{2\text{QS}}$, $Q_{2\text{TR}}$ and $Q^*_{2\text{BAR}}$, $Q^*_{2\text{DAN}}$, $Q^*_{2\text{QS}}$, $Q^*_{2\text{TR}}$ denote the rejection rates of Q_2 using empirical and bootstrap critical values, respectively; the subscripts BAR, DAN, QS and TR denote, respectively, the Bartlett kernel, the Daniell kernel, the Quadratic-Spectral kernel and the Truncated kernel. Number of bootstraps = 499. T and d_1 denote the sample size and dimension of portfolio 1, respectively.

Table S9: Empirical sizes

		NullD																							
	d_1	3			4			5			6			7			8			9			10		
T	M	10	20	30	10	20	30	10	20	30	10	20	30	10	20	30	10	20	30	10	20	30	10	20	30
Rejection rates based on asymptotic critical values																									
1000	$Q_{2\text{BAR}}$	7.4	7.1	7.0	6.8	7.0	7.0	7.1	7.6	7.7	7.4	7.9	8.2	7.1	7.3	7.7	7.1	7.6	8.0	7.6	7.9	8.6	7.7	8.3	8.7
	$Q_{2\text{DAN}}$	7.2	7.0	6.9	7.0	7.1	7.1	7.3	7.8	7.7	7.6	8.0	8.4	7.1	7.4	7.7	7.1	8.0	8.5	7.6	8.3	8.8	7.8	8.5	9.0
	$Q_{2\text{QS}}$	7.1	6.9	7.1	7.0	7.1	7.0	7.1	7.8	7.7	7.7	7.9	8.4	7.0	7.3	7.5	7.1	7.9	8.4	7.7	8.3	8.9	7.7	8.5	9.1
	$Q_{2\text{TR}}$	7.1	6.9	7.2	6.9	7.1	7.5	7.5	7.5	7.8	7.9	8.6	8.9	7.3	8.1	8.2	7.8	8.6	9.3	8.2	8.7	10.1	8.1	9.3	9.9
1500	$Q_{2\text{BAR}}$	6.4	6.3	6.2	6.7	6.4	6.3	6.8	6.6	6.6	6.6	6.8	7.1	6.8	6.9	7.2	7.1	7.5	7.6	7.1	7.4	7.6	7.4	7.3	7.6
	$Q_{2\text{DAN}}$	6.4	6.3	6.2	6.6	6.6	6.3	6.9	6.8	6.7	6.5	6.9	7.3	6.6	6.9	7.2	7.2	7.4	7.3	7.4	7.5	7.6	7.4	7.5	7.9
	$Q_{2\text{QS}}$	6.4	6.3	6.2	6.7	6.5	6.2	6.8	6.6	6.8	6.4	7.0	7.1	6.6	6.9	7.1	7.1	7.6	7.4	7.4	7.3	7.6	7.4	7.5	7.8
	$Q_{2\text{TR}}$	6.3	6.3	6.4	6.5	6.2	6.4	6.7	7.0	6.8	6.7	7.4	7.6	6.9	7.5	7.8	7.4	7.4	8.3	7.2	7.9	8.3	7.4	7.7	8.2
Rejection rates based on bootstrap critical values																									
1000	$Q^*_{2\text{BAR}}$	5.1	5.3	5.7	5.2	5.7	5.8	5.3	5.2	5.2	5.2	5.4	5.4	5.5	6.1	6.1	5.7	6.2	6.1	5.7	6.0	6.0	5.3	6.2	6.2
	$Q^*_{2\text{DAN}}$	5.0	5.6	5.9	5.4	5.7	5.7	5.1	5.1	5.3	5.5	5.3	5.3	5.7	6.3	6.0	6.0	6.1	6.2	5.6	6.0	6.2	5.4	6.2	6.1
	$Q^*_{2\text{QS}}$	5.1	5.5	5.7	5.5	5.6	5.8	5.3	5.3	5.5	5.3	5.3	5.3	5.8	6.2	6.2	6.0	6.1	6.2	5.4	6.0	6.1	5.8	6.2	6.0
	$Q^*_{2\text{TR}}$	5.8	5.7	5.8	5.6	5.8	5.5	5.1	5.5	5.5	5.4	5.5	5.6	6.2	6.2	5.9	6.1	6.3	6.4	5.9	6.3	6.6	6.1	6.1	6.1
1500	$Q^*_{2\text{BAR}}$	4.9	4.6	5.0	5.5	5.2	5.4	5.1	4.8	4.9	4.9	4.6	5.3	5.6	5.6	5.6	5.3	5.0	5.2	5.5	5.7	5.9	5.2	5.4	5.4
	$Q^*_{2\text{DAN}}$	4.9	4.9	4.8	5.2	5.3	5.1	5.2	4.8	4.9	4.9	5.1	5.3	5.8	5.5	5.7	5.3	5.1	5.3	5.3	5.8	6.0	5.2	5.2	5.8
	$Q^*_{2\text{QS}}$	4.7	5.0	4.8	5.3	5.3	5.2	5.3	4.8	4.9	4.8	5.1	5.4	5.7	5.3	5.8	5.3	5.0	5.1	5.5	5.9	6.1	5.2	5.3	5.5
	$Q^*_{2\text{TR}}$	5.0	4.8	4.9	5.3	5.0	5.2	4.5	5.1	4.8	5.1	5.4	5.1	5.8	5.8	5.5	5.0	5.5	5.3	6.0	5.6	5.8	5.3	5.6	5.5

NOTES: The table reports empirical sizes (in %) of Q_2 under NullD at the 5% significance level based on NC-LS modeling. Number of simulations = 10000. $Q_{2\text{BAR}}$, $Q_{2\text{DAN}}$, $Q_{2\text{QS}}$, $Q_{2\text{TR}}$ and $Q^*_{2\text{BAR}}$, $Q^*_{2\text{DAN}}$, $Q^*_{2\text{QS}}$, $Q^*_{2\text{TR}}$ denote the rejection rates of Q_2 using asymptotic and bootstrap critical values, respectively; the subscripts BAR, DAN, QS and TR denote, respectively, the Bartlett kernel, the Daniell kernel, the Quadratic-Spectral kernel and the Truncated kernel. Number of bootstraps = 499. M denotes the kernel smoothing parameter. T and d_1 denote the sample size and dimension of portfolio 1, respectively.

Table S10: Empirical powers

		AlterC																							
	d_1	3			4			5			6			7			8			9			10		
T	M	10	20	30	10	20	30	10	20	30	10	20	30	10	20	30	10	20	30	10	20	30	10	20	30
Rejection rates based on empirical critical values																									
1000	$Q_{2\mathrm{BAR}}$	22.2	18.6	16.5	32.2	25.5	21.9	44.8	34.6	28.7	53.0	39.7	33.3	61.9	48.9	40.9	69.4	55.1	44.8	70.6	55.9	45.9	72.8	58.0	47.1
	$Q_{2\mathrm{DAN}}$	20.9	16.8	15.0	30.0	21.9	18.4	42.1	28.4	23.3	48.8	32.6	26.6	58.0	41.2	33.2	64.8	45.8	35.1	66.3	46.7	36.2	69.1	48.2	37.4
	$Q_{2\mathrm{QS}}$	20.0	16.5	14.4	28.9	21.2	17.6	40.8	27.3	22.6	47.3	31.3	25.9	56.4	39.6	31.8	62.5	43.9	33.6	64.2	44.3	34.2	66.6	45.3	35.6
	$Q_{2\mathrm{TR}}$	14.3	11.6	10.6	18.0	14.0	11.9	22.3	17.2	14.7	26.3	18.9	15.8	32.1	22.4	19.4	35.5	24.5	20.4	36.5	24.7	19.1	37.8	25.1	20.8
1500	$Q_{2\mathrm{BAR}}$	34.8	28.5	24.4	53.9	42.9	36.4	72.4	58.0	48.4	82.5	69.0	58.6	90.3	78.5	67.7	92.9	82.7	72.8	95.2	86.1	77.5	96.6	89.4	80.1
	$Q_{2\mathrm{DAN}}$	33.0	23.7	20.6	49.5	36.0	29.4	67.8	48.2	38.4	79.9	59.0	47.5	87.1	68.0	56.0	90.0	73.0	61.0	92.9	77.4	64.8	94.7	80.5	67.7
	$Q_{2\mathrm{QS}}$	31.4	23.4	19.3	48.1	34.8	28.3	65.6	45.8	36.8	77.9	56.8	45.6	85.3	65.4	53.6	88.7	70.9	58.7	91.4	75.1	61.7	93.8	78.3	64.8
	$Q_{2\mathrm{TR}}$	19.2	14.3	13.2	28.2	20.3	17.0	37.5	25.0	21.1	46.8	31.4	25.3	53.4	35.9	28.4	58.6	40.1	31.0	63.0	42.9	33.0	67.2	45.4	36.8
Rejection rates based on bootstrap critical values																									
1000	$Q^*_{2\mathrm{BAR}}$	19.2	16.0	14.1	29.4	23.4	19.8	41.8	32.1	26.7	51.0	38.9	32.7	59.8	46.1	37.8	64.0	50.7	41.1	67.3	53.4	44.5	70.6	57.5	47.4
	$Q^*_{2\mathrm{DAN}}$	18.2	14.0	12.6	27.5	20.1	15.8	38.3	26.3	21.9	47.3	32.7	26.4	55.5	37.9	30.1	59.8	41.6	32.9	62.3	44.6	35.2	66.6	47.6	37.4
	$Q^*_{2\mathrm{QS}}$	17.6	13.8	12.7	26.6	19.0	15.3	37.4	25.1	21.1	45.4	31.2	24.9	53.4	36.1	29.3	57.6	39.7	31.5	60.5	42.5	34.2	64.5	45.3	36.0
	$Q^*_{2\mathrm{TR}}$	12.2	10.1	9.3	15.6	12.0	10.9	21.1	16.0	13.2	25.0	18.0	15.3	29.6	20.8	17.8	31.0	22.0	18.1	34.8	23.8	19.1	35.5	25.1	19.7
1500	$Q^*_{2\mathrm{BAR}}$	29.8	23.2	20.2	48.9	38.3	32.0	69.6	56.0	46.2	80.9	66.8	56.3	88.5	75.7	65.2	91.8	80.7	70.2	94.5	85.6	75.9	96.2	87.6	78.4
	$Q^*_{2\mathrm{DAN}}$	27.7	19.7	17.1	45.3	31.7	25.9	65.7	46.7	36.9	76.9	56.4	45.7	85.5	65.6	52.1	89.3	70.7	57.8	92.3	76.7	63.0	94.2	79.1	66.1
	$Q^*_{2\mathrm{QS}}$	26.6	19.3	16.7	43.6	30.5	24.8	63.5	44.8	35.2	74.8	54.0	43.5	83.6	62.7	49.7	87.4	68.0	55.2	90.9	74.0	60.6	92.7	76.6	63.2
	$Q^*_{2\mathrm{TR}}$	16.3	13.3	12.0	23.4	17.2	14.2	35.6	24.2	18.6	43.5	29.5	23.6	50.9	34.0	26.9	55.4	38.2	29.3	61.4	41.3	32.3	65.1	43.9	34.6

NOTES: The table reports empirical powers (in %) of Q_2 under AlterC at the 5% significance level based on NC-LS modeling. Number of simulations = 10000. $Q_{2\mathrm{BAR}}$, $Q_{2\mathrm{DAN}}$, $Q_{2\mathrm{QS}}$, $Q_{2\mathrm{TR}}$ and $Q^*_{2\mathrm{BAR}}$, $Q^*_{2\mathrm{DAN}}$, $Q^*_{2\mathrm{QS}}$, $Q^*_{2\mathrm{TR}}$ denote the rejection rates of Q_2 using empirical and bootstrap critical values, respectively; the subscripts BAR, DAN, QS and TR denote, respectively, the Bartlett kernel, the Daniell kernel, the Quadratic-Spectral kernel and the Truncated kernel. Number of bootstraps = 499. M denotes the kernel smoothing parameter. T and d_1 denote the sample size and dimension of portfolio 1, respectively.

4. GARCH misspecification

Based on a correctly specified model, in the paper we find no evidence of volatility spillover between the UK market and the NA market after Brexit. In this section, we investigate the consequence of model misspecification on testing for volatility spillover. To this purpose, we repeat the examination of spillover between UK and NA in the post–Brexit sample with a deliberately misspecified model.

Recall that a well specified model requires order 4 for the conditional variance of the UK series. We first include a control study that is correctly specified. We specify a model with random order 6 for the conditional variance of the UK series. Diagnostic tests reported in Table S11 suggest that the model is well specified. Table S12 reports the volatility spillover test results. Consistent with the findings in the paper, we find no evidence of volatility spillover between the UK market and the NA market. Moreover, the p-values reported in Table S12 are very similar to those reported in the paper. This control study serves two purposes. First, it ensures that our conclusion is not affected by the sensitivity of a correctly specified model lag order. Second, it ensures that any changes in conclusion in the next study is likely to be driven by model misspecification.

To impose model misspecification, we now use order 1 for the variance of the UK market. Diagnostic results reported in Table S13 suggests model misspecification. Table S14 reports the volatility spillover test results. In contrast to the findings based on a correctly specified model, we find that the NA market has a significant spillover effect to the UK market at all M's. Given the control study, this false-positive result is likely to be induced by serial correlations in the event variables as a result of model misspecification. In summary, the exercises in this section highlight the importance of a correctly specified model in testing volatility spillover.

Table S11: Diagnostic tests (UK–NA, Control)

	LB(10)	LB(20)	LB(30)	$LB^2(10)$	$LB^2(20)$	$LB^2(30)$
Post–Brexit (24th June 2016 – 31st December 2019)						
UK	8.342	17.378	24.770	7.035	12.227	15.502
	[0.595]	[0.628]	[0.736]	[0.722]	[0.908]	[0.987]
US	15.436	18.869	23.736	10.546	13.020	17.188
	[0.117]	[0.530]	[0.784]	[0.394]	[0.877]	[0.970]
Canada	6.779	18.437	22.004	4.432	18.823	21.513
	[0.746]	[0.559]	[0.854]	[0.926]	[0.533]	[0.871]

NOTES: The table reports diagnostic analyses for all fitted series. LB(M) and $LB^2(M)$ are the Ljung-Box tests for the null of no serial correlation (up to lag order M) on the standardized and squared standardized residuals, respectively. The values in the squared parentheses are the p-values of the tests.

Table S12: Spillover results (UK–NA, Control)

	Post–Brexit		
M	10	20	30
$Q^*_{1\text{BAR}}$	0.172	0.166	0.236
$Q^*_{-1\text{BAR}}$	0.643	0.599	0.673

NOTES: The table reports bootstrap p-values of the proposed spillover tests. Number of bootstraps = 499. Q^*_{BAR} denotes the one-way test for the null hypothesis of no volatility spillover from the NA market to the UK market. $Q^*_{-1\text{BAR}}$ denotes the one-way test for the null hypothesis of no volatility spillover from the UK market to the NA market. The subscript BAR denotes the Bartlett kernel. M denotes the kernel smoothing parameter.

Table S13: Diagnostic tests (UK–NA, Misspecified)

	LB(10)	LB(20)	LB(30)	$\text{LB}^2(10)$	$\text{LB}^2(20)$	$\text{LB}^2(30)$
Post–Brexit (24th June 2016 – 31st December 2019)						
UK	33.666	49.171	55.323	22.971	38.451	41.231
	[0.000]	[0.000]	[0.003]	[0.011]	[0.008]	[0.083]
US	15.436	18.869	23.736	10.546	13.020	17.188
	[0.117]	[0.530]	[0.784]	[0.394]	[0.877]	[0.970]
Canada	6.779	18.437	22.004	4.432	18.823	21.513
	[0.746]	[0.559]	[0.854]	[0.926]	[0.533]	[0.871]

NOTES: The table reports diagnostic analyses for all fitted series. LB(M) and $\text{LB}^2(M)$ are the Ljung-Box tests for the null of no serial correlation (up to lag order M) on the standardized and squared standardized residuals, respectively. The values in the squared parentheses are the p-values of the tests.

Table S14: Spillover results (UK–NA, Misspecified)

	Post–Brexit		
M	10	20	30
$Q^*_{1\text{BAR}}$	0.078	0.048	0.060
$Q^*_{-1\text{BAR}}$	0.721	0.607	0.687

NOTES: The table reports bootstrap p-values of the proposed spillover tests. Number of bootstraps = 499. Q^*_{BAR} denotes the one-way test for the null hypothesis of no volatility spillover from the NA market to the UK market. $Q^*_{-1\text{BAR}}$ denotes the one-way test for the null hypothesis of no volatility spillover from the UK market to the NA market. The subscript BAR denotes the Bartlett kernel. M denotes the kernel smoothing parameter.

5. Computational efficiency study

In this section, we conduct a computational efficiency experiment to highlight the speed advantage of our adjusted least-squares estimation over the QMLE. We simulate the following simple ARCH(1) process with standard Gaussian error

$$\epsilon_t = \sqrt{h_t}\xi_t, \quad \xi_t \overset{\text{iid}}{\sim} \mathrm{N}(0,1),$$
$$h_t = 0.1 + 0.2\epsilon_{t-1}^2.$$

Next, we fit the simulated ARCH process h_t using our adjusted least-squares estimation as well as QMLE. During estimation, we measure the machine elapsed time using the Matlab `tic` and `toc` functions. For improved accuracy, we repeat the procedure 10000 times and we compute the mean elapsed time. We find that the average machine time required by our estimation method is about 0.0012 (machine unit) per simulation, while that required by QMLE is about 0.0266 per simulation. In other words, our method requires only about 4.3% of the the computational time of QMLE.

6. Robustness analysis

In this section, we conduct a robustness analysis for the empirical results reported in the paper. In particular, we repeat all of the volatility spillover examinations reported in the paper, with 8 observations of the estimated event variables curtailed before and after the Brexit referendum. Results for the UK–NA (EU–NA) nexus is reported in Table S15 (S16). We can see that the p-values reported in Table S15 (S16) are very similar to those reported in Table 6 (9) of the paper. This highlights the robustness of the paper's empirical findings.

Table S15: Spillover results (UK–NA, Robustness)

	Pre–Brexit			*Post–Brexit*		
M	10	20	30	10	20	30
$Q^*_{1\mathrm{BAR}}$	0.026	0.044	0.040	0.152	0.124	0.170
$Q^*_{-1\mathrm{BAR}}$	0.086	0.054	0.062	0.727	0.691	0.790

NOTES: The table reports bootstrap p-values of the proposed spillover tests, with 8 observations curtailed before and after the Brexit referendum date. Number of bootstraps = 499. $Q^*_{1\mathrm{BAR}}$ denotes the one-way test for the null hypothesis of no volatility spillover from the NA market to the UK market. $Q^*_{-1\mathrm{BAR}}$ denotes the one-way test for the null hypothesis of no volatility spillover from the UK market to the NA market. The subscript BAR denotes the Bartlett kernel. M denotes the kernel smoothing parameter.

Table S16: Spillover results (EU–NA, Robustness)

	Pre–Brexit			*Post–Brexit*		
M	10	20	30	10	20	30
$Q^*_{1\mathrm{BAR}}$	0.066	0.154	0.098	0.082	0.142	0.160
$Q^*_{-1\mathrm{BAR}}$	0.210	0.150	0.086	0.144	0.078	0.072

NOTES: The table reports bootstrap p-values of the proposed spillover tests, with 8 observations curtailed before and after the Brexit referendum date. Number of bootstraps = 499. $Q^*_{1\mathrm{BAR}}$ denotes the one-way test for the null hypothesis of no volatility spillover from the NA market to the EU market. $Q^*_{-1\mathrm{BAR}}$ denotes the one-way test for the null hypothesis of no volatility spillover from the EU market to the NA market. The subscript BAR denotes the Bartlett kernel. M denotes the kernel smoothing parameter.