

TESTING THE STRUCTURE OF CONDITIONAL CORRELATIONS
IN MULTIVARIATE GARCH MODELS: A GENERALIZED
CROSS-SPECTRUM APPROACH*

BY NADINE McCLOUD AND YONGMIAO HONG¹

*University of the West Indies at Mona, Jamaica; Cornell University, U.S.A.,
and Xiamen University, China*

We introduce a class of generally applicable specification tests for constant and dynamic structures of conditional correlations in multivariate GARCH models. The tests are robust to the presence of time-varying higher-order conditional moments of unknown form and are *pure significance tests*. The tests can identify linear and nonlinear misspecifications in conditional correlations. Our approach does not necessitate a particular parameter estimation method and distributional assumption on the error process. The asymptotic distribution of the tests is invariant to the uncertainty in parameter estimation. We assess the finite sample performance of our tests using simulated and real data.

1. INTRODUCTION

Correlations play a vital role in optimal portfolio diversification and hedge ratio estimation (e.g., Bera and Kim, 2002; Engle, 2002) and are therefore very important in theoretical and empirical economics and finance. Multivariate generalized autoregressive conditional heteroskedasticity (MGARCH) models provide a convenient framework for modeling correlations. These models include the constant conditional correlation MGARCH (CCC-MGARCH) model by Bollerslev (1990), which has been the most widely used model due to its simple variance–covariance matrix decomposition that facilitates theoretical analysis (see, e.g., Jeantheau, 1998; Ling and McAleer, 2003; He and Teräsvirta, 2004; McAleer et al., 2009; Nakatani and Teräsvirta, 2009) and estimation in empirical applications. However, recent empirical evidences suggest that the structure of conditional correlations between certain assets is time-varying (dynamic), which renders the use of the CCC-MGARCH model inappropriate for some empirical applications. To this end, many researchers have become interested in designing MGARCH models that explicitly² accommodate time-varying conditional correlations (see, e.g., Engle, 2002; Cappiello et al., 2006; Pelletier, 2006; McAleer et al., 2008; and others). Despite the importance of specification testing in building and drawing correct inferences

*Manuscript received October 2008; revised January 2010.

¹ We are sincerely grateful to the editor (Frank Schorfheide) and three anonymous referees for their invaluable critiques that helped to craft this article. Also, we thank Wolfgang Härdle, Tae-hwy Lee, Anton Schick, and seminar participants at the 2008 Xiamen University-Humboldt University Joint Workshop on Economics and Finance, 2008 Far Eastern and South Asian Meeting of the Econometric Society, and the Statistics Seminar Series in Department of Mathematical Sciences at State University of New York at Binghamton for their comments and suggestions. Hong thanks the National Science Foundation of China for its support via the Overseas Outstanding Youth Grant. McCloud thanks the Department of Economics and Graduate Student Organization at the State University of New York at Binghamton and the Department of Economics and the Dean of the Faculty of Social Sciences at the University of the West Indies at Mona, and Wang Yanan Institute for Studies in Economics at Xiamen University for travel support. All errors are our own. Please address correspondence to: Yongmiao Hong, Department of Economics, Cornell University, Ithaca, NY 14853. Phone: 607 255 5130; Fax: 607-255-2818; E-mail: yh20@cornell.edu.

² We acknowledge the fact that other MGARCH models such as the VEC model by Bollerslev et al. (1988) and the BEKK model by Engle and Kroner (1995) do allow for time-varying conditional correlations but without assuming a specific functional form for these correlations. That is, the VEC and BEKK formulations model time-varying conditional covariances.

from econometric models, very little effort has been devoted to designing tests for assessing the adequacy of the fit of these time-varying conditional correlation MGARCH models to the data. This article fills this gap in the MGARCH literature by proposing a class of generally applicable tests for constant conditional correlation and parametric specification of time-varying conditional correlations.

There are some tests for constant conditional correlations in MGARCH models. These tests include the Ljung–Box portmanteau test (Bollerslev, 1990), classical tests (e.g., Longin and Solnik, 1995; Tse, 2000; Engle and Sheppard, 2001; Silvennoinen and Teräsvirta, 2005, 2009a), and the Information matrix test (Bera and Kim, 2002); Bauwens et al. (2006) and Silvennoinen and Teräsvirta (2009b) provide an extensive review of the MGARCH literature and most of these tests. Bollerslev (1990) assumes that under the null of constant conditional correlations, the cross products of the standardized residuals are serially uncorrelated and uses the Ljung–Box portmanteau test to investigate the adequacy of the null specification. The absence of serial correlation, however, does not necessarily imply constant conditional correlations (Bera and Kim, 2002). In addition, Li and Mak (1994) find that the portmanteau test statistic is not asymptotically χ^2 -distributed.³ Longin and Solnik (1995) consider three alternative structures—a time trend, a threshold variable, and a linear function of some variables in the information set. Because of the large number of parameters in their model, the adequacy of the null specification relative to each correlation structure is independently assessed using Likelihood Ratio (LR) tests. Thus, their framework does not admit a joint test for several departures from the null hypothesis. In general and under the null specification, an overall LR test is statistically equivalent to the sum of a set of individual LR tests if the individual LR statistics are asymptotically independently distributed.⁴ As such, individual LR tests may fail to detect dependent specification errors that may exist in empirical applications.

Tse (2000) proposes a Lagrange Multiplier (LM) test with lag-1 cross product of standardized residuals alternative. Silvennoinen and Teräsvirta (2005, 2009a) put forward LM tests with Smooth Transition Conditional Correlation GARCH (STCC-GARCH) and Double Smooth Transition Conditional Correlation GARCH (DSTCC-GARCH) alternatives. In the STCC-GARCH and DSTCC-GARCH models, the conditional correlations change smoothly between two extreme states as a function of at most two exogenous or endogenous transition variables. Engle and Sheppard (2001) propose an IID or a Wald test with a p th-order autoregressive alternative.

These classical tests, however, have a common drawback. It is well known that the LM, LR, and Wald tests are asymptotically optimal within a class of contiguous alternatives. In this context, this implies that these classical tests will be consistent against certain forms of time-varying conditional correlations. In particular, in Silvennoinen and Teräsvirta's (2005, 2009a) framework, the outcome of their tests is dependent on the transition variable. Also in Tse's and Engle and Sheppard's works, the use of an exogenous lag order for the alternative model may under- or overutilize the information in the data, thus biasing the power performance of their tests. Thus, in empirical applications where the true structure of conditional correlations is unknown, coupled with the possible lack of empirical or theoretical guidelines to selecting alternative models, the use of the classical tests may be inappropriate. A test for constant conditional correlations that is independent of an alternative specification may be quite useful in these instances.

Bera and Kim (2002) develop an efficient-score form of the information matrix (IM) test for assessing the constancy of the conditional correlation matrix in a bivariate GARCH model. This form of the IM test alleviates the poor size performance in finite samples that is usually exhibited by its outer product gradient counterpart. In contrast to the classical tests, no a priori alternative

³ Accordingly, Li and Mak (1994) introduce a modified portmanteau test statistic that is asymptotically χ^2 -distributed. We thank an anonymous referee for pointing this out.

⁴ In essence, this will require some asymptotic orthogonality conditions on the regressors of the null and alternative specifications.

functional form of conditional correlation is needed to derive their test statistic. However, their IM test is constructed using moment conditions of the bivariate normal distribution. Thus, the greater is the departure from zero excess kurtosis or other forms of nonnormality, the larger is the probability of rejecting the null of constant conditional correlation.⁵

These existing tests for constant conditional correlations are derived under the i.i.d. standardized error vector and are therefore not robust to the presence of time-varying higher-order conditional moments of unknown form. Absent from the existing literature, also, is a test that does not require a particular distributional assumption on the error process. Our generally applicable test for constant conditional correlations avoids these limitations of the existing tests.

Testing for only constant conditional correlation may be insufficient to draw valid inferences from an econometric model, especially in the case where there is evidence against the null specification. For example, King and Wadhvani (1990), Lin et al. (1994), de Santis and Gerard (1997), and Longin and Solnik (2001) find evidence in support of strong correlations between cross-country stock markets during times of financial turbulence but weak or no correlations outside of these events. These empirical findings may be manifestations of time-varying conditional correlations between stock markets and invalidate the frequently imposed constant conditional correlation assumption in some empirical works. The inadequacy of the constant conditional correlation assumption for some data has prompted researchers to design a new class of MGARCH models that admits flexible structures for conditional correlations. The most popular time-varying correlation MGARCH model is by Engle (2002), who extends Bollerslev's (1990) CCC-MGARCH model by incorporating dynamic conditional correlations (the DCC-GARCH model). Engle (2002) imposes a multivariate normal distribution on the innovation process and suggests heterogeneous dynamics, but employs homogeneous dynamics, for conditional correlations. His assumptions have motivated extensions of the DCC-GARCH model that incorporate some stylized facts of financial time series data. For example, Pelagatti and Rondena (2006) retain the dynamic specification of the DCC-GARCH model but consider multivariate, fat-tailed elliptical distributions for the innovation process so as to model excess kurtosis. Hafner and Franses (2009) put forward a generalized dynamic conditional correlation (GDCC) model that allows for all correlations to have different dynamics. Billio et al. (2006) introduce a flexible dynamic conditional correlation (FDCC) model that allows for equal correlation dynamics only within groups of variables. Cappiello et al. (2006) extend the DCC model to accommodate series-specific news, smoothing parameters, and conditional asymmetries in correlation dynamics (the AG-DCC model). Billio and Caporin (2009) formulate a model that nests the DCC, AG-DCC, and FDCC by allowing for constant correlation dynamics only among blocks of assets that are from the same category.

Other forms of time-varying conditional correlations models have been introduced to the literature. Tse and Tsui (2002) introduce the time-varying conditional correlation MGARCH (TVC-MGARCH) model in which the correlation matrix has an autoregressive moving average type structure similar to that of the DCC-GARCH model. Pelletier (2006) proposes a regime switching conditional correlation (RSDC) model that allows for a time-invariant correlation matrix within each regime but possible differences in conditional correlations across regimes, with a latent Markov chain governing the transition between regimes. More recently, McAleer et al. (2008) offer the generalized autoregressive conditional correlation (GARCC) model in which the standardized residuals have a random coefficient vector autoregressive specification that engenders time-varying conditional correlations. Lee and Long (2009) put forward another rich class of MGARCH models for which the higher-order conditional dependence structure is embedded in a copula function.⁶

⁵ Nonnormal distributions yield an actual asymptotic significance level greater than that of its counterpart under the normality assumption. Bera and Kim offer a studentized test statistic as a remedy to the dependency of their statistic on the normality assumption. However, a studentized variant of a test statistic is not robust to all departures from normality (see Koenker, 1981; Wooldridge, 1990).

⁶ We thank an anonymous referee for drawing our attention to Lee and Long's work on copula.

We emphasize that, despite the burgeoning interest in designing time-varying conditional correlation MGARCH models, little interest has been paid to constructing tests for the adequacy of these models. In the existing literature, specification tests for time-varying conditional correlations are of the classical type, for example, that of McAleer et al. (2008) and Silvennoinen and Teräsvirta (2009a). However, as we have already mentioned, the inherent dependence of the power performance of classical tests on the type of alternative models warrants more general specification tests to analyze model adequacy of time-varying conditional correlation models. Moreover, the structures of most time-varying conditional correlation models do not emanate from economic theory. The structures are introduced to mostly fit stylized facts of time series data and ensure positive definiteness of the time-varying conditional correlation matrix in estimation. Thus, choosing an appropriate alternative model may be quite difficult. In addition, the different aforementioned distributions and structures for time-varying conditional correlations highlight the importance of constructing generally applicable specification tests that are consistent against a wide range of time-varying conditional correlation models, not just nested models. This will enable reliable statistical inferences from and more widespread use of time-varying conditional correlation models in empirical applications.

In this article, we develop a class of generally applicable tests for investigating the constancy of conditional correlations and the parametric specification of time-varying conditional correlations.⁷ Our tests are predicated on an extension of Hong's (1999) generalized spectrum approach that is useful for testing univariate time series. Our extension, called the *generalized cross-spectrum*, accommodates multivariate time series. Without modifications, however, the generalized cross-spectrum cannot be employed to assess the structure of conditional correlations. Specifically, this generalized cross-spectral tool can capture cyclical dynamics induced by linear and nonlinear cross dependence in various moments of the standardized error vector but does not permit us to identify the source of these dynamics. To analyze the structure of conditional correlations, we differentiate the generalized cross-spectrum to yield its *generalized cross-spectral derivative*, which is the appropriate device for analyzing various aspects of cross dependence.

Our proposed tests have several attributes. We require no alternative specifications; therefore our tests are *pure significance tests*. That is, the design of our tests does not hinge on an explicit alternative hypothesis. Unlike the existing tests, no distributional assumption on the observations is required for deriving our tests. Moreover, no specific estimation method is required for the parameters; as a result, any \sqrt{T} -consistent estimator is admissible. The asymptotic distribution of the tests is the $N(0, 1)$. Furthermore, our tests are nuisance parameter free in that using the estimated standardized residuals in lieu of the standardized errors has no impact on this asymptotic distribution. In addition, the spectral nature of our tests facilitates the detection of linear and nonlinear misspecifications in conditional correlations.

Our more general test is robust to the presence of time-varying higher-order conditional moments (e.g., skewness and kurtosis) of unknown form in the conditional density of the innovation vector. This distinguishing feature renders the test a nontrivial extension of its i.i.d. counterpart, which is not valid under the null hypothesis of a correctly specified time-varying

⁷ For concreteness, we clarify what we mean by time-varying (dynamic) conditional correlations. A time-varying parameter may be (1) a linear or nonlinear function of variables in the information set at time $t - 1$, (2) a deterministic function of time (see, e.g., Dahlhaus, 1997), or (3) a linear or nonlinear function of variables in the information set at time $t - 1$ and a deterministic function of time. Type 1 can be a stationary process; however, types 2 and 3 are nonstationary or locally stationary processes. An example of a locally stationary model is the local dynamic conditional correlation (LDCC) model proposed by Feng (2006). Feng (2006) allows each conditional variance to have a locally stationary and stationary components and the conditional correlation matrix to be a nonparametric function of the rescaled time variable (the location variable) and past observations. It is well known that stationary and nonstationary processes differ in their implications and consequences for forecasting and asymptotic theory. For example, a model of type 1 can be predicted by an appropriate nonlinear model of variables in its information set whereas a model of type 2 can be predicted by a local moving model. In this article, we abstract from nonstationary processes, and therefore our concept of "time-varying (dynamic) conditional correlations" refers only to stationary processes (type 1).

parametric model for conditional correlations.⁸ Furthermore, the i.i.d. assumption precludes the existence of, say, time-varying conditional skewness and kurtosis. Time-varying higher-order conditional moments in time-series data can arise for many reasons, and their existence cannot be viewed as immaterial. For example, Patton (2006) explains that the monetary policy objectives of central banks and financial decisions of investors can give rise to time-varying higher-order dependence structures between exchange rates. His flexible conditional copula modeling framework confirms the presence of such structure between the Deutsche mark (euro dollar)–U.S. dollar and yen–U.S. dollar daily exchange rates. Patton (2006) also finds nonlinearity in time-varying conditional correlation between these exchange rates. Moreover, the presence of time-varying skewness can affect the time-series properties of lower-order conditional moments (Harvey and Siddique, 1999, 2000). Specification tests for conditional correlations that do not account for time-varying higher-order conditional moments will exhibit poor size performances. The theoretical and empirical relevance of this higher-order moment-robust feature of our tests underscores one of the essential contributions of this article to the existing literature.

Our tests do not impose a priori lag order on the design set; rather, we use an adaptive lag-selection method that allows us to capitalize on the information in the data without sacrificing power. To test constancy of conditional correlations, the estimated standardized residuals and conditional correlation are the only inputs needed to carry out the test. To test a specific time-varying structure for conditional correlations, only the standardized residuals and the vector of the estimated time-varying conditional correlations are required to execute the test.

The layout of the article is as follows. Section 2 formulates the hypotheses of interest for testing the existence of constant conditional correlations and time-varying parametric specification of conditional correlations. Section 3 presents and describes the test statistics and procedures derived from the generalized cross-spectrum. Section 4 establishes the asymptotic theory. Section 5 investigates the finite sample performance of the test for constant conditional correlations. Section 6 provides an application of our tests to a classical asset allocation problem. Section 7 concludes the article. We place a brief outline of the mathematical details in the Appendix. A well-detailed technical appendix and the GAUSS code for executing the tests are available from the authors upon request. Throughout this article, we use C to denote an arbitrary bounded constant, $\|\cdot\|$ the Euclidean norm, and A^* the complex conjugate of A .

2. HYPOTHESIS TESTING

For completeness, we first introduce the constant and time-varying conditional correlation MGARCH models. We then formalize our hypotheses of interest for the structure of conditional correlations.

2.1. The Constant and Dynamic Conditional Correlation MGARCH Models. Let $\{Y_t\}$ with $Y_t = (y_{1t}, y_{2t}, \dots, y_{Nt})'$ be an \mathbb{R}^N -valued process of time series observations that is adapted to a filtration I_{t-1} . Furthermore, suppose

$$(1) \quad Y_t = \mu_t + \Lambda_t z_t,$$

where μ_t and Λ_t are measurable with respect to I_{t-1} . Let $\{z_t\}$ be a N -variate unobservable martingale difference sequence (m.d.s.) innovation vector such that $E(z_t | I_{t-1}) = 0$ a.s. and $E(z_t z_t' | I_{t-1}) = \Phi_t$ a.s., $\Phi_t = [\rho_{ij,t}]$ is the matrix of conditional correlations. This m.d.s. assumption identifies μ_t as the conditional mean vector of Y_t . Assume $\varepsilon_t \equiv \Lambda_t z_t$ is the model error. Note that the m.d.s. property of $\{z_t\}$ implies that $\{\varepsilon_t\}$ is m.d.s. such that $E(\varepsilon_t | I_{t-1}) = 0$ a.s. and $E(\varepsilon_t \varepsilon_t' | I_{t-1}) = \Lambda_t \Phi_t \Lambda_t$ a.s. Then the multiplicatively separable matrix $\Lambda_t \Phi_t \Lambda_t$ is the conditional

⁸ McAleer et al. (2008) circumvent this inherent drawback of assuming i.i.d. innovations by imposing a random coefficient autoregressive structure on the innovations.

variance of Y_t . Also, let $\Lambda_t = \text{diag}(h_{1,t}^{1/2}, h_{2,t}^{1/2}, \dots, h_{N,t}^{1/2})$ be a diagonal matrix of conditional standard deviations and each $h_{i,t}$ has a univariate GARCH (1, 1) specification so that

$$(2) \quad h_{i,t} = \omega_{i0} + \alpha_i \varepsilon_{i,t-1}^2 + \beta_i h_{i,t-1},$$

with $\omega_{i0} > 0, \alpha_i > 0, \beta_i > 0, \alpha_i + \beta_i < 1, \forall i = 1, \dots, N$. We remark that other specifications for Λ_t are admissible in the present context. For example, we could retain Λ_t as a diagonal matrix but permit dynamic dependence between volatility series as in Jeantheau (1998). If Φ_t is a constant matrix, then Y_t follows a constant conditional correlation MGARCH (CCC-MGARCH) model (e.g., Bollerslev, 1990). Otherwise, Y_t follows a dynamic conditional correlation MGARCH model.

The literature offers different functional forms for characterizing the time evolution of the conditional correlation matrix Φ_t (see, e.g., Engle, 2002; Tse and Tsui, 2002; Silvennoinen and Teräsvirta, 2005, 2009a; Pelletier, 2006; McAleer et al., 2008). Under certain parameter restrictions, each of these models nests Bollerslev’s (1990) CCC-MGARCH model. We briefly review some of the specifications of time-varying conditional correlations in MGARCH models.

EXAMPLE 1 (Tse and Tsui, 2002): TVC-MGARCH.

$$(3) \quad \begin{aligned} \rho_{ij,t} &= (1 - \zeta_1 - \zeta_2)\rho_{ij} + \zeta_2\rho_{ij,t-1} + \zeta_1\pi_{ij,t-1}, \\ \pi_{ij,t-1} &= \frac{\sum_{h=1}^M z_{i,t-h}z_{j,t-h}}{\sqrt{\left(\sum_{h=1}^M z_{i,t-h}^2\right)\left(\sum_{h=1}^M z_{j,t-h}^2\right)}}, \end{aligned}$$

with $\zeta_1, \zeta_2 \geq 0$ and $\zeta_1 + \zeta_2 \leq 1$. For this model, $\rho_{ij,t}$ has an autoregressive moving average specification that is the convex combination of $\rho_{ij}, \rho_{ij,t-1}$, and $\pi_{ij,t-1}$, and the parameter ζ_2 represents the degree of inertia in time-varying conditional correlations whereas ζ_1 represents the degree of perturbation to $\rho_{ij,t}$. The matrix $[\pi_{ij,t-1}]$ is a correlation matrix for a subvector of the residuals at time $t - 1$. A necessary condition for the matrix $[\pi_{ij,t-1}]$ to be positive definite is $M \geq N$.

EXAMPLE 2 (Engle, 2002): DCC-GARCH.

$$(4) \quad \begin{aligned} \rho_{ij,t} &= \frac{q_{ij,t}}{\sqrt{q_{ii,t}q_{jj,t}}}, \\ q_{ij,t} &= (1 - \zeta_1 - \zeta_2)\bar{q}_{ij} + \zeta_2q_{ij,t-1} + \zeta_1z_{i,t-1}z_{j,t-1}, \quad \forall i, j, \end{aligned}$$

with $\zeta_1, \zeta_2 \geq 0$ and $\zeta_1 + \zeta_2 \leq 1$. For this model each component of $\rho_{ij,t}, q_{ij,t}, q_{ii,t}$, and $q_{jj,t}$, has an autoregressive moving average specification, and the matrix $[q_{ij,t}]$ is transformed to yield the correlation matrix $[\rho_{ij,t}]$.⁹

⁹ It is important to note that for our purposes, in Examples 1 and 2 we exclude the joint restriction $\zeta_1 = 0$ and $\zeta_2 > 0$ from the quasi-convex set of restrictions given in Tse and Tsui (2002) and Engle (2002). This is because, using Example 1 for illustration,

$$\rho_{ij,t} = (1 - \zeta_2)\rho_{ij} + \zeta_2\rho_{ij,t-1}$$

$$\text{so } (1 - \zeta_2L)\rho_{ij,t} = (1 - \zeta_2)\rho_{ij},$$

where L is the backward shift operator. Then

$$\rho_{ij,t} = (1 - \zeta_2)(1 - \zeta_2L)^{-1}\rho_{ij} = \rho_{ij}.$$

EXAMPLE 3 (Pelletier, 2006): RSDC. Pelletier proposes a general regime switching dynamic correlation matrix of the form

$$\Phi_t = \sum_{i=1}^S \mathbb{I}_{\{F_t=i\}} \Phi_i,$$

where F_t is a latent Markov chain process that is independent of the innovation z_t and can take S possible values, $F_t = 1, 2, \dots, S$, \mathbb{I} is the indicator function, and each Φ_i is an $N \times N$ constant correlation matrix. Pelletier analyzes a simple form of this regime switching model:

$$(5) \quad \Phi_t = \lambda(F_t)\Phi + (1 - \lambda(F_t))I_N,$$

where $\lambda(F_t) \in [0, 1]$ is a univariate process dictated by F_t , Φ is a constant correlation matrix, and I_N is an $N \times N$ identity matrix. For this model, the time-varying conditional correlation matrix at time t is the convex combination of the correlation matrix in each state—in one state the innovations are correlated [$\lambda(F_t) = 1$] and in the other state they are uncorrelated [$\lambda(F_t) = 0$]. The degree of smoothing for Pelletier’s specification is governed by the transition probabilities. Pelletier imposes additional conditions for identification of this model.

These correlation structures may generate quite different correlation dynamics. In fact, in an empirical application with exchange rates, Pelletier finds that the regime switching model generates smoother time-varying correlations than Engle’s DCC model. The disparities in the correlation dynamics of the models point to the importance of investigating the adequacy of the fit of these models to the data. The class of generally applicable specification tests we propose below can detect departures from the null hypotheses of constant conditional correlation and correct specification of time-varying conditional correlations and are robust to misspecification in other aspects of the models.

2.2. *Testing the Structure of Conditional Correlations.* Consider a bivariate version of the model in (1). Let θ be the vector of location, scale, and correlation parameters such that $\theta \in \Theta \subset \mathbb{R}^K$ and Θ is compact and convex. Assume the parametric model for conditional correlations, Ξ_θ , for the stochastic standardized error vector $z_t(\theta)$ is such that

$$(6) \quad \Xi_\theta = \{\rho_t(\theta) : E[z_{1t}(\theta)z_{2t}(\theta) | I_{t-1}] = \rho_t(\theta), \theta \in \Theta \subset \mathbb{R}^K\},$$

where $\rho_t(\theta)$ is measurable with respect to I_{t-1} and the functional form of $\rho_t(\theta)$ is specified up to the unknown finite-dimensional parameter θ .

Let $\theta^0 \in \Theta$ be the true but unknown parameter. Assume $\rho_t(\theta^0)$ characterizes the true but unknown structure of conditional correlations. Furthermore, if the true structure is constant, we let $\rho \equiv \rho_t(\theta^0)$. If the true structure is time-varying, we let $\rho_t \equiv \rho_t(\theta^0)$. Then we say Ξ_θ is correctly specified for time-varying conditional correlations only if $\rho_t \in \Xi_\theta$. We say Ξ_θ is misspecified for time-varying conditional correlations only if $\rho_t \notin \Xi_\theta$. Similarly, Ξ_θ is correctly specified for constant conditional correlations only if $\rho \in \Xi_\theta$, and Ξ_θ is misspecified for constant conditional correlations only if $\rho \notin \Xi_\theta$. We now rewrite these criteria as follows.

Under the null hypothesis of a correctly specified functional form for the conditional correlation between z_{1t} and z_{2t} we write

$$(7) \quad \mathbb{H}_0: \Pr\{E[z_{1t}(\theta^0)z_{2t}(\theta^0) | I_{t-1}] = \rho_t(\theta^0)\} = 1 \quad \text{for some } \theta^0 \in \Theta.$$

Thus, the conditional correlation is equal to the unconditional one for all t ; that is, this joint restriction renders ζ_2 an unidentified nuisance parameter under the null hypothesis of constant conditional correlation. We thank an anonymous referee for pointing this out.

Similarly, under the alternative hypothesis of an incomplete characterization by $\rho_t(\theta)$ of the conditional correlation between z_{1t} and z_{2t} we write

$$(8) \quad \mathbb{H}_A: \Pr\{E[z_{1t}(\theta)z_{2t}(\theta) | I_{t-1}] = \rho_t(\theta)\} < 1 \quad \forall \theta \in \Theta.$$

Put

$$(9) \quad m_t(\theta) \equiv z_{1t}(\theta)z_{2t}(\theta) - \rho_t(\theta).$$

Then $z_{1t}(\theta^0)z_{2t}(\theta^0) = \rho_t(\theta^0) + m_t(\theta^0)$ can be viewed as an auxiliary regression function, with $m_t(\theta^0)$ representing an MDS regression standardized error. We therefore label $m_t(\theta)$ as the “generalized-standardized residual.” Using these notations, the corresponding MDS expressions for the null and alternative hypotheses are, respectively,

$$(10) \quad \mathbb{H}_0: E[m_t(\theta^0) | I_{t-1}] = 0 \text{ a.s.},$$

$$(11) \quad \mathbb{H}_A: E[m_t(\theta) | I_{t-1}] \neq 0 \quad \forall \theta \in \Theta.$$

Under the MDS assumption on the innovation vector z_t , the test statistics for constant conditional correlation and time-varying specification of conditional correlations are identical. Thus, to avoid redundancy in our proceeding exposition, we discuss the test statistic and procedures and asymptotic theory only in terms of the specification test for time-varying conditional correlations.

3. TEST STATISTICS AND PROCEDURES

To begin, we say that the real-valued process $m_t(\theta)$ possesses the “ideal” MDS feature if, for some $\theta^0 \in \Theta$, $E[m_t(\theta^0) | I_{t-1}] = 0$. This conditional moment restriction suffers from the “curse of dimensionality” problem since the conditioning set has an infinite dimension. One way of circumventing the problem is by making use of the generalized spectral approach by Hong (1999). However, Hong’s (1999) generalized spectrum approach is univariate, so we extend his framework to accommodate our hypotheses of interest.

Let $z_t \equiv z_t(\theta^0)$ and assume $\{z_t\}$ is a strictly stationary process. Consider a family of conditional distributions associated with z_t given I_{t-1} , namely, $\{G(z_t | I_{t-1}) : G \in \mathcal{G}\}$ and suppose $G^0(z_t | I_{t-1})$ is the null conditional distribution. We define the conditional characteristic function of z_t as

$$E_{\theta^0}(e^{iu'z_t} | I_{t-1}) = \int_{\mathbb{R}^N} e^{iu'z_t} dG^0(z_t | I_{t-1}), \quad u \in \mathbb{R}^N, \quad i = \sqrt{-1},$$

where $E_{\theta^0}(\cdot | I_{t-1})$ is the conditional expectation taken with respect to $G^0(z | I_{t-1})$. The introduction of this conditional characteristic function permits us to investigate the functional form of the conditional correlations. We then define the generalized cross-covariance function as

$$\sigma_j(u, v) \equiv \text{cov}\{e^{iu'z_t(\theta)} - E_{\theta^0}[e^{iu'z_t(\theta)} | I_{t-1}], \quad e^{iv'z_{t-j}(\theta)}\}, \quad u, v \in \mathbb{R}^N.$$

Using this generalized cross-covariance function $\sigma_j(u, v)$, we define the generalized cross-spectrum as

$$f(\omega, u, v) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \sigma_j(u, v)e^{-ij\omega}, \quad \omega \in [-\pi, \pi], \quad u, v \in \mathbb{R}^N.$$

That is, $f(\omega, u, v)$ is the Fourier transform of $\sigma_j(u, v)$, and therefore they contain the same information. In this framework, the moments of z_t need not be finite. This generalized cross-spectrum is “general” in the sense that it can capture all cyclical dynamics induced by linear and nonlinear pairwise dependence in various moments of z_t and, without further modification, does not permit us to identify the source of these dynamics.

Now, let $m = (m_1, m_2, \dots, m_N)'$ be a $N \times 1$ vector of positive integers with $|m| = \sum_{c=1}^N m_c$. The generalized cross-spectrum $f(\omega, u, v)$ can be differentiated to yield its generalized cross-spectral derivative, which is the device for distinguishing various aspects of cross dependence. To this end, define the generalized cross-spectral derivative as

$$(12) \quad f^{(0,m,0)}(\omega, 0, v) = \left. \frac{\partial^{m_1}}{\partial u_1^{m_1}} \dots \frac{\partial^{m_N}}{\partial u_N^{m_N}} f(\omega, u, v) \right|_{u=0} = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \sigma_j^{(m,0)}(0, v) e^{-ij\omega},$$

where the derivative of the generalized cross-covariance function is

$$(13) \quad \sigma_j^{(m,0)}(0, v) = cov \left\{ \prod_{c=1}^N (iz_{ct}(\theta))^{m_c} - E_{\theta^0} [\prod_{c=1}^N (iz_{ct}(\theta))^{m_c} | I_{t-1}], e^{iv'z_{t-|j|}(\theta)} \right\}.$$

For our purposes, where our interest lies in the structure of conditional correlation for the bivariate vector $z_t = (z_{1t}, z_{2t})'$, we set $m = (1, 1)'$. Then using the generalized-standardized residual $m_t(\theta)$, we have

$$(14) \quad \begin{aligned} \sigma_j^{(m,0)}(0, v) &= i^{|m|} cov\{z_{1t}(\theta)z_{2t}(\theta) - E_{\theta^0}[z_{1t}(\theta)z_{2t}(\theta) | I_{t-1}], e^{iv'z_{t-|j|}(\theta)}\} \\ &= i^{|m|} cov[m_t(\theta), e^{iv'z_{t-|j|}(\theta)}] \\ &= (i)^2 E\{m_t(\theta) [e^{iv'z_{t-|j|}(\theta)} - E_{\theta^0}(e^{iv'z_{t-|j|}(\theta)})]\}. \end{aligned}$$

Note that for $j > 0$, $\sigma_j^{(m,0)}(0, v) = 0 \forall v \in \mathbb{R}^2$ if and only if $E[m_t(\theta^0) | z_{1t-j}, z_{2t-j}] = 0$ for some $\theta^0 \in \Theta$ (e.g., Bierens, 1982; Stinchcombe and White, 1998). Clearly, testing this “subordinate” MDS hypothesis of $\{E[m_t(\theta^0) | z_{1t-j}, z_{2t-j}] = 0 \text{ for some } \theta^0 \in \Theta\}$ is not necessarily equivalent to its ideal counterpart of $\{E[m_t(\theta^0) | I_{t-1}] = 0 \text{ for some } \theta^0 \in \Theta\}$. In particular, and by containment, the ideal MDS hypothesis implies the subordinate MDS hypothesis, but the converse is not always true.¹⁰ As we will demonstrate below, we derive our generalized cross-spectral derivative tests from a restricted variant of the aforementioned subordinate hypothesis.

¹⁰ If the null hypothesis for the MDS correlation test is rejected, we can further extend the generalized cross-spectral derivative in (12) to obtain a subclass of specific tests for assessing the type of linear or nonlinear departure from the null specification. To see this in our bivariate model, consider another bivariate vector, $s = (s_1, s_2)'$, of positive integers that is associated with the auxiliary vector v in the (12) and define $|s| = \sum_{c=1}^2 s_c$. Then

$$\sigma_j^{(m,s)}(0, 0) = i^{|m|+|s|} cov \left\{ m_t(\theta), \prod_{c=1}^2 [z_{ct-|j|}(\theta)]^{s_c} \right\}.$$

Case 1: If we set $s = (1, 1)'$, then $\sigma_j^{(m,s)}(0, 0) = cov[m_t(\theta), z_{1t-|j|}(\theta)z_{2t-|j|}(\theta)]$ can be used to test for serial correlation in conditional correlation. This is similar in spirit to Bollerslev’s (1990) correlation test.

Case 2: If we set $s = (2, 2)'$, then $\sigma_j^{(m,s)}(0, 0) = cov[m_t(\theta), z_{1t-|j|}^2(\theta)z_{2t-|j|}^2(\theta)]$ can be used to test for volatility in conditional correlation.

Case 3: If we set $s = (3, 3)'$, then $\sigma_j^{(m,s)}(0, 0) = cov[m_t(\theta), z_{1t-|j|}^3(\theta)z_{2t-|j|}^3(\theta)]$ can be used to test for skewness in conditional correlation.

Case 4: If we set $s = (4, 4)'$, then $\sigma_j^{(m,s)}(0, 0) = cov[m_t(\theta), z_{1t-|j|}^4(\theta)z_{2t-|j|}^4(\theta)]$ can be used to test for kurtosis in conditional correlation.

Under \mathbb{H}_0 , the generalized cross-spectral derivative degenerates to a “flat” spectrum:

$$f_0^{(0,m,0)}(\omega, 0, v) = \frac{1}{2\pi} \sigma_0^{(m,0)}(0, v), \quad \omega \in [-\pi, \pi], \quad v \in \mathbb{R}^2.$$

3.1. *Estimating the Generalized Cross-Spectral Derivative.* Since $\{z_t\}$ is unobservable, the derivative of the generalized cross-covariance function in (13) cannot be implemented in this context. Now suppose we observe a random bivariate sample of size T and \tilde{I}_{t-1} is set of this observed information available at time $t - 1$ such that \tilde{I}_{t-1} contains some starting values and $\tilde{I}_{t-1} \subset I_{t-1}$.¹¹ Also, let $\tilde{\rho}_t(\theta), \tilde{z}_t(\theta), \tilde{\mu}_t(\theta)$, and $\tilde{\Lambda}_t(\theta)$ be measurable with respect to \tilde{I}_{t-1} . Using \tilde{I}_{t-1} we obtain $\hat{\theta}$, a \sqrt{T} -consistent estimator for θ_0 , and $\hat{z}_t \equiv \tilde{\Delta}_t^{-1}(\hat{\theta})\hat{\varepsilon}_t, \hat{\Delta}_t(\hat{\theta}) \equiv \text{diag}(\tilde{h}_{1,t}^{1/2}(\hat{\theta}), \tilde{h}_{2,t}^{1/2}(\hat{\theta}))$, $\hat{\varepsilon}_t \equiv \tilde{\varepsilon}_t(\hat{\theta}) = Y_t - \tilde{\mu}_t(\hat{\theta})$, and $\hat{\rho}_t \equiv \tilde{\rho}_t(\hat{\theta})$. We can consistently estimate the generalized cross-spectral derivative $f^{(0,m,0)}(\omega, 0, v)$ using a kernel estimator

$$(15) \quad \hat{f}^{(0,m,0)}(\omega, 0, v) = \frac{1}{2\pi} \sum_{j=1-T}^{T-1} \left(1 - \frac{|j|}{T}\right)^{1/2} k(j/p) \hat{\sigma}_j^{(m,0)}(0, v) e^{-ij\omega},$$

with $\omega \in [-\pi, \pi], v \in \mathbb{R}^2$, and when

$$(16) \quad \hat{\sigma}_j^{(m,0)}(0, v) = \frac{1}{(T - |j|)} \sum_{t=|j|+1}^T [\hat{z}_{1t}\hat{z}_{2t} - \hat{\rho}_t] \hat{\psi}_{t-j}(v),$$

$\hat{\psi}_{t-j}(v) = e^{iv'\hat{z}_{t-|j|}} - \hat{\varphi}_j(v)$ and $\hat{\varphi}_j(v) = (T - |j|)^{-1} \sum_{t=|j|+1}^T e^{iv'\hat{z}_{t-|j|}}$ is the estimator for the unconditional-marginal characteristic function of $z_{t-|j|}$. Also, $p \equiv p(T)$ is a bandwidth and $k: \mathbb{R} \rightarrow [-1, 1]$ is a symmetric kernel function, for, e.g., the Bartlett kernel

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

or the Parzen kernel

$$k(x) = \begin{cases} 1 - 6x^2 + 6|x|^3, & |x| \leq 0.5, \\ 2(1 - |x|)^3, & 0.5 \leq |x| \leq 1, \\ 0, & \text{otherwise.} \end{cases}$$

The weighting factor $(1 - \frac{|j|}{T})^{1/2}$ is a finite sample correction, which can be normalized to equal one.

Similarly, a consistent estimator for the “flat” generalized cross-spectral derivative $f_0^{(0,m,0)}(\omega, 0, v)$ is

$$(17) \quad \hat{f}_0^{(0,m,0)}(\omega, 0, v) \equiv \frac{1}{2\pi} \hat{\sigma}_0^{(m,0)}(0, v), \quad \omega \in [-\pi, \pi], \quad v \in \mathbb{R}^2.$$

We note that the estimators $\hat{f}^{(0,m,0)}(\omega, 0, v)$ and $\hat{f}_0^{(0,m,0)}(\omega, 0, v)$ converge to the same limit under \mathbb{H}_0 . Thus, our test is based on the divergence between these two estimators. To approximate the divergence between $\hat{f}^{(0,m,0)}(\omega, 0, v)$ and $\hat{f}_0^{(0,m,0)}(\omega, 0, v)$, we use the squared L_2 -norm between

¹¹ The set \tilde{I}_{t-1} can be considered as a truncated version of I_{t-1} .

(15) and (17) so that

(18)

$$L_2^2[\hat{f}^{(0,m,0)}(\omega, 0, v), \hat{f}_0^{(0,m,0)}(\omega, 0, v)] = \frac{\pi T}{2} \int \int_{-\pi}^{\pi} |\hat{f}^{(0,m,0)}(\omega, 0, v) - \hat{f}_0^{(0,m,0)}(\omega, 0, v)|^2 d\omega dW(v) \\ = \sum_{j=1}^{T-1} k^2(j/p)(T-j) \int |\hat{\sigma}_j^{(m,0)}(0, v)|^2 dW(v),$$

where the second equality is by virtue of Parseval’s identity. Moreover, $W : \mathbb{R}^2 \rightarrow \mathbb{R}_+$ is a nondecreasing weighting function that weighs the sets symmetric about the origin equally. Some examples of $W(\cdot)$ are the multivariate independent standard normal cdf or any discrete, symmetric probability distribution.

3.2. *Test Statistics for Conditional Correlations.* To design a specification test for time-varying conditional correlations, we make use of the following points. One, under the null hypothesis of a correctly specified time-varying parametric model for conditional correlations, the assumption of conditional homoskedastic or i.i.d. standardized error vector is invalid. Two, it is well known that most time-series data exhibit time-varying higher-order conditional moments. Furthermore, there is a growing trend to allow for innovations with nonnormal densities (see, e.g., Harvey and Siddique, 1999; Bauwens and Laurent, 2005; Patton, 2006; Pelagatti and Rondena, 2006; Pelletier, 2006). Three, time-varying higher-order conditional moments have been found to have a significant impact on lower-order conditional moments. Our test is designed to accommodate time-varying higher-order conditional moments of unknown structure and is therefore robust in this regard. We note that specification tests for conditional correlations that do not account for time-varying higher-order conditional moments will have poor size performances. Under the *m.d.s.* assumption, the test statistic that accounts for time-varying higher-order conditional moments of unknown structure is given as follows:

$$(19) \quad \hat{Q}_1 = \left[\sum_{j=1}^{T-1} k^2(j/p)(T-j) \int |\hat{\sigma}_j^{(m,0)}(0, v)|^2 dW(v) - \hat{C}_1 \right] / \sqrt{\hat{D}_1},$$

where

$$\hat{C}_1 = \sum_{j=1}^{T-1} k^2(j/p) \frac{1}{(T-j)} \sum_{t=j+1}^T |m_t(\hat{\theta})|^2 \int |\hat{\psi}_{t-j}(v)|^2 dW(v), \\ \hat{D}_1 = 2 \sum_{j=1}^{T-2} \sum_{l=1}^{T-2} k^2(j/p) k^2(l/p) \\ \times \iint \left| \frac{1}{T - \max(j, l)} \cdot \sum_{t=\max(j, l)+1}^T [m_t(\hat{\theta})]^2 \hat{\psi}_{t-j}(u) \hat{\psi}_{t-l}(v) \right|^2 dW(u) dW(v),$$

with $m_t(\hat{\theta}) = (i)^2[\hat{z}_{1t}\hat{z}_{2t} - \hat{\rho}_t]$. Here \hat{C}_1 and \hat{D}_1 , respectively, are the location and scale factors that approximate the mean and variance of (18). This test statistic is also valid for testing constant conditional correlation, by simply replacing $\hat{\rho}_t$ with $\hat{\rho}$.

When testing for constant conditional correlation, imposing the i.i.d. assumption on $\{z_t\}$ may be adequate. This assumption greatly simplifies the preceding test statistic \hat{Q}_1 , which

degenerates to the following:

$$(20) \quad \hat{Q}_2 = \left[\sum_{j=1}^{T-1} k^2(j/p)(T-j) \int |\hat{\sigma}_j^{(m,0)}(0, v)|^2 dW(v) - \hat{C}_2 \right] / \sqrt{\hat{D}_2},$$

where

$$\hat{C}_2 = \hat{\Omega} \int \hat{\sigma}_0(v, -v) dW(v) \sum_{j=1}^{T-2} k^2(j/p),$$

$$\hat{D}_2 = 2\hat{\Omega}^2 \int \int |\hat{\sigma}_0(u, v)|^2 dW(u) dW(v) \sum_{j=1}^{T-2} k^4(j/p),$$

with $\hat{\Omega} = T^{-1} \sum_{t=1}^T (\hat{z}_{1t}\hat{z}_{2t} - \hat{\rho})^2$. We note that \hat{C}_2 and \hat{D}_2 , respectively, are the location and scale factors that approximate the mean and variance of (18) under the i.i.d. assumption. We emphasize that this statistic is only applicable to the test for constant conditional correlation due to the auxiliary i.i.d. assumption.

4. ASYMPTOTIC THEORY

To reiterate, we limit our discussion of the asymptotic theory to the test for misspecification of time-varying conditional correlations, \hat{Q}_1 , given that this test is more general. The following regularity conditions are needed to derive the null asymptotic distribution of the test statistic \hat{Q}_1 .

ASSUMPTION 1. $\{Y_t\}$ is a bivariate GARCH, strictly stationary process as defined in (1) with $E\|Y_t\|^8 \leq C$, $E(z_{it})^4 \leq C$, for $i = 1, 2$ and $E(z_{1t}z_{2t})^4 \leq C$.

ASSUMPTION 2. For each sufficiently large integer q , there exists a strictly stationary process $\{z_{1q,t}z_{2q,t}, \rho_{q,t}\}$ such that $\{z_{1q,t}z_{2q,t} - \rho_{q,t}\}$ is a q -dependent MDS process. Moreover, (i) for $i = 1, 2$, $E(z_{it} - z_{iq,t})^4 \leq Cq^{-\eta}$, (ii) $E|\rho_t - \rho_{q,t}|^2 \leq Cq^{-\eta}$ for some constant $\eta \geq 1$.

ASSUMPTION 3. Let $\rho_t(\theta)$ be a parametric function for conditional correlation where $\theta \in \Theta$ is a parameter in a finite dimensional parameter space that is convex and compact, and for each $\theta \in \Theta$: $\mu_t(\theta)$, $\rho_t(\theta)$, and $\Lambda_t^{-1}(\theta)$ are measurable with respect to I_{t-1} ; $\mu_t(\theta)$, $\rho_t(\theta)$ and $\Lambda_t^{-1}(\theta)$ admit continuous derivatives up to order 2 with respect to $\theta \in \Theta$; (i) $E \sup_{\theta \in \Theta} \|\nabla_{\theta} \rho_t(\theta)\|^2 \leq C$ and $E \sup_{\theta \in \Theta} \|\nabla_{\theta\theta} \rho_t(\theta)\|^2 \leq C$; (ii) $E \sup_{\theta \in \Theta} \|\nabla_{\theta} \mu_t(\theta)\|^4 \leq C$ and $E \sup_{\theta \in \Theta} \|\nabla_{\theta} \Lambda_t^{-1}(\theta)\|^4 \leq C$; and (iii) $E \sup_{\theta \in \Theta} \|\nabla_{\theta\theta} \mu_t(\theta)\|^2 \leq C$ and $E \sup_{\theta \in \Theta} \|\nabla_{\theta\theta} \Lambda_t^{-1}(\theta)\|^2 \leq C$.

ASSUMPTION 4. Let $\tilde{\rho}_t(\theta)$, $\tilde{\mu}_t(\theta)$ and $\tilde{\Lambda}_t(\theta)$ be measurable with respect to \tilde{I}_{t-1} . Then (i) $\lim_{T \rightarrow \infty} \sum_{t=1}^T \{E[\sup_{\theta \in \Theta} |\tilde{\rho}_t(\theta) - \rho_t(\theta)|^2]\} \leq C$; (ii) $\lim_{T \rightarrow \infty} \sum_{t=1}^T \{E[\sup_{\theta \in \Theta} \|\tilde{\mu}_t(\theta) - \mu_t(\theta)\|^4]\} \leq C$; and (iii) $\lim_{T \rightarrow \infty} \sum_{t=1}^T \{E[\sup_{\theta \in \Theta} \|\tilde{\Lambda}_t^{-1}(\theta) - \Lambda_t^{-1}(\theta)\|^4]\} \leq C$.

ASSUMPTION 5. Let $\{z_t, \nabla_{\theta} z_t(\theta^0), \rho_t, \nabla_{\theta} \rho_t(\theta^0)\}$ be a strictly stationary α -mixing process with mixing coefficient $\sum_{j=0}^{\infty} \alpha(j)^{(v-1)/v} \leq C$ for some constant $v > 1$. Moreover, $\Omega_0 \equiv E[z_{1,0}z_{2,0} - \rho_0]^2 < \infty$.

ASSUMPTION 6. $\hat{\theta}$ is an estimator for $\theta^0 \in \Theta$, that is, $\sqrt{T}(\hat{\theta} - \theta^*) = O_P(1)$, where $\theta^* = p \lim_{T \rightarrow \infty} \hat{\theta}$ and $\theta^* = \theta^0$ under \mathbb{H}_0 .

ASSUMPTION 7. $W : \mathbb{R}^2 \rightarrow \mathbb{R}^+$ is a nondecreasing, integrable weighting function that places equal weights on sets that are symmetric about the origin. Also, let $\int \|v\|^4 dW(v) < \infty$.

ASSUMPTION 8. Let (i) $k : \mathbb{R} \rightarrow [-1, 1]$ be a symmetric function that is continuous at zero and all points in \mathbb{R} except a finite number of points; (ii) $k(0) = 1$; (iii) $k(z) \leq c|z|^{-b}$ for some $b > \frac{1}{2}$ as $z \rightarrow \infty$.

Assumptions 1 and 2 provide regularity conditions for the DGP. Assumption 2 is needed only under the null hypothesis. This condition states that the m.d.s. $\{z_{1t}z_{2t} - \rho_t\}$ can be approximated by a q -dependent m.d.s. $\{z_{1q,t}z_{2q,t} - \rho_{q,t}\}$ arbitrarily well when q is sufficiently large. In particular the difference between these two processes goes to zero at a geometric rate. In essence, Assumption 2 provides the restrictive conditions on the serial dependence in higher-order moments of $\{z_t\}$. This assumption also admits ergodicity for $\{z_t\}$. We note that condition (i) is derived from the condition $E(z_{1t}z_{2t} - z_{1q,t}z_{2q,t})^2 \leq Cq^{-\eta}$ for some constant $\eta \geq 1$. To understand this assumption, consider the zero-mean time-varying conditional correlation MGARCH model of Tse and Tsui (2002) with specification $Y_t = \varepsilon_t$,

$$\left\{ \begin{aligned} h_{i,t} &= \omega_i + \alpha_i \varepsilon_{i,t-1}^2 + \beta_i h_{i,t-1}, \quad i = 1, 2, \\ \rho_t &= (1 - \zeta_2 - \zeta_2)\rho + \zeta_1 \rho_{t-1} + \zeta_2 \pi_{t-1}, \\ \pi_{t-1} &= \frac{\sum_{h=1}^2 z_{1,t-h} z_{2,t-h}}{\sqrt{\left(\sum_{h=1}^2 z_{1,t-h}^2\right) \left(\sum_{h=1}^2 z_{2,t-h}^2\right)}}, \\ \begin{pmatrix} z_{1t} \\ z_{2t} \end{pmatrix} | I_{t-1} &\sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_t \\ \rho_t & 1 \end{pmatrix} \right], \\ \varepsilon_{i,t} &= \sqrt{h_{i,t}} z_{i,t}, \quad i = 1, 2. \end{aligned} \right.$$

We illustrate the conditions for $i = 1$. With $h_{1,t} = \omega_1 + \alpha_1 \varepsilon_{1,t-1}^2 + \beta_1 h_{1,t-1} = \frac{\omega_1}{1-\beta_1} + \alpha_1 \sum_{k=0}^{\infty} \beta_1^k \varepsilon_{1,t-1-k}^2$, we define $h_{1,q,t} \equiv \frac{\omega_1}{1-\beta_1} + \alpha_1 \sum_{k=0}^q \beta_1^k \varepsilon_{1,t-1-k}^2$ and $z_{1q,t} \equiv \varepsilon_{1,t}/h_{1,q,t}^{1/2}$. Then, we have

$$\begin{aligned} E(z_{1t} - z_{1q,t})^4 &= E \left(\frac{\varepsilon_{1,t}}{h_{1,t}^{1/2}} - \frac{\varepsilon_{1,t}}{h_{1,q,t}^{1/2}} \right)^4 = E \left(\varepsilon_{1,t}^4 \frac{(h_{1,q,t}^{1/2} - h_{1,t}^{1/2})^4}{(h_{1,q,t}^{1/2} h_{1,t}^{1/2})^4} \right) \\ &\leq E \left(\varepsilon_{1,t}^4 \frac{(h_{1,t} - h_{1,q,t})^2}{(h_{1,q,t}^{1/2} h_{1,t}^{1/2})^4} \right) \leq C(E(\varepsilon_{1,t}^8))^{1/2} (E(h_{1,t} - h_{1,q,t})^4)^{1/2} \\ &\leq C(E(h_{1,t} - h_{1,q,t})^4)^{1/2} = C \left(E \left(\alpha_1 \sum_{k=q+1}^{\infty} \beta_1^k \varepsilon_{1,t-1-k}^2 \right)^4 \right)^{1/2} \\ &\leq C \left(\alpha_1 \sum_{k=q+1}^{\infty} \beta_1^k \{E(\varepsilon_{1,t-1-k}^8)\}^{1/4} \right)^2 \leq C \frac{\beta_1^{2q}}{(1-\beta_1)^2}. \end{aligned}$$

Hence, we obtain Assumption 2(i) since $\beta_1 < 1$. The inequalities follow from (a) $h_{1,t}$ and $h_{1,q,t}$ have a lower bound uniformly in all t and parameter vector θ ; (b) $(\sqrt{a} - \sqrt{b})^2 \leq a - b$ for $a, b \geq 0$; (c) Cauchy-Schwarz and Minkowski inequalities; and (d) moment conditions on $\varepsilon_{1,t}$. To show Assumption 2(ii), we write $\rho_t = \frac{1-\zeta_1-\zeta_2}{1-\zeta_2}\rho + \zeta_1 \sum_{k=0}^{\infty} \zeta_2^k \pi_{t-1-k}$, and define $\rho_{q,t} \equiv \frac{1-\zeta_1-\zeta_2}{1-\zeta_2}\rho + \zeta_1 \sum_{k=0}^q \zeta_2^k \pi_{t-1-k}$. Then, $E|\rho_t - \rho_{q,t}|^2 \leq Cq^{-\eta}$ is satisfied provided $E|\pi_{t-1-k}|^2 < \infty$,

which is trivially true since for all t and k , π_{t-1-k} is an entry of a correlation matrix and hence is uniformly bounded in all t .

Assumption 3 imposes regularity conditions on the structure of the dynamic conditional correlations, conditional means, and conditional variances. Conditions (i) and (ii) along with Assumption 1 guarantee the existence of $E \sup_{\theta \in \Theta} \|\nabla_{\theta} z_t(\theta)\|^4$ and $E \sup_{\theta \in \Theta} \|\nabla_{\theta\theta} z_t(\theta)\|^2$.

Assumption 4 imposes regularity conditions on the truncated information set I_{t-1} . This assumption ensures that the limit distribution of \hat{Q}_n is invariant to any use of starting values. To understand this, consider a bivariate variant of Bollerslev’s (1990) CCC-GARCH(1,1) model with specification: $Y_t = \varepsilon_t$ where $\varepsilon_t(\theta) | I_{t-1} \sim N(0, \Lambda_t(\theta)\Phi_t(\theta)\Lambda_t(\theta))$ with $\Lambda_t(\theta) = \text{diag}(h_{1,t}^{1/2}(\theta), h_{2,t}^{1/2}(\theta))$, $h_{i,t}(\theta) = \omega_i + \alpha_i \varepsilon_{i,t-1}^2(\theta) + \beta_i h_{i,t-1}(\theta)$ for $i = 1, 2$. Then $\theta = (\omega_1, \alpha_1, \beta_1, \omega_2, \alpha_2, \beta_2, \rho)'$, $\rho_t(\theta) = \rho$ and $\mu_t(\theta) = 0$. The conditions on $\rho_t(\theta)$ and $\mu_t(\theta)$ are trivial, so we now show that the condition on $\Lambda_t(\theta)$ holds. Assume $\tilde{h}_{1,0} \in \tilde{I}_{t-1}$. First note that

$$\|\tilde{\Lambda}_t^{-1}(\theta) - \Lambda_t^{-1}(\theta)\|^4 = \{|\tilde{h}_{1,t}^{-1/2}(\theta) - h_{1,t}^{-1/2}(\theta)|^2 + |\tilde{h}_{2,t}^{-1/2}(\theta) - h_{2,t}^{-1/2}(\theta)|^2\}^2.$$

Thus, it suffices to show that $\lim_{T \rightarrow \infty} \sum_{t=1}^T \{E[\sup_{\theta \in \Theta} |\tilde{h}_{1,t}^{-1/2}(\theta) - h_{1,t}^{-1/2}(\theta)|^4]\} \leq C$. Note also that

$$|\tilde{h}_{1,t}^{-1/2}(\theta) - h_{1,t}^{-1/2}(\theta)| = \left| \frac{\tilde{h}_{1,t}(\theta) - h_{1,t}(\theta)}{\tilde{h}_{1,t}^{1/2}(\theta) h_{1,t}^{1/2}(\theta) [\tilde{h}_{1,t}^{1/2}(\theta) + h_{1,t}^{1/2}(\theta)]} \right|.$$

By employing recursive substitution, we find that

$$\tilde{h}_{1,t}(\theta) - h_{1,t}(\theta) = \omega_1 + \alpha_1 \sum_{k=0}^{t-2} \beta_1^k \varepsilon_{1,t-1-k}^2 + \alpha \beta^{t-1} \tilde{h}_{1,0} - \omega_1 - \alpha_1 \sum_{k=0}^{t-2} \beta_1^k \varepsilon_{1,t-1-k}^2 - \alpha \beta^{t-1} h_{1,0}(\theta).$$

Then, it follows that

$$\sum_{t=1}^T E \sup_{\theta \in \Theta^0} |\tilde{h}_{1,t}^{-1/2}(\theta) - h_{1,t}^{-1/2}(\theta)|^4 \leq \sum_{t=1}^{\infty} E \sup_{\theta \in \Theta^0} \left| \frac{\alpha_1 \beta_1^{t-1} [\tilde{h}_{1,0} - h_{1,0}(\theta)]}{2\omega_1^{3/2}} \right|^4 \leq C,$$

assuming $\omega_1 > 0$, $0 < \alpha_1, \beta_1 < 1$, $\alpha_1 + \beta_1 < 1$, and $E(h_{1,0}^4)$ exists.

Assumption 5 provides restrictions on the nature of the serial dependence in $\{z_t, \nabla_{\theta} z_t(\theta^0), \rho_t, \nabla_{\theta} \rho_t(\theta^0)\}$. The strictly stationary α -mixing condition is frequently used in the context of nonlinear time series analysis.¹² Assumption 6 states that a \sqrt{T} -consistent estimator, $\hat{\theta}$, of θ^0 will suffice. This assumption therefore accommodates various asymptotic estimators, including asymptotically most efficient estimator, and those obtained via MLE and QMLE. The statistical properties of the QMLE for some of the conditional correlation MGARCH models have been established in the literature. Jeantheau (1998) proposes a set of necessary conditions under which the QMLE of multivariate autoregressive process with conditionally heteroskedastic errors is strongly consistent and verifies these conditions for Bollerslev’s (1990) CCC-GARCH model. Ling and McAleer (2003) prove consistency and asymptotic normality of the QMLE for a class of vector ARMA-GARCH models that nests the CCC-GARCH model. More recently, McAleer et al. (2009) and McAleer et al. (2008) develop sufficient conditions for consistency and asymptotic normality of the QMLE, respectively, for the ARMA-asymmetric GARCH model, which admits constant conditional correlations, and the GARCC model, which admits time-varying conditional correlations.

¹² Note that with the measurability assumption on $\rho_t(\theta)$ in Assumption 3(i), we could also assume that $\{z_t, \nabla_{\theta} z_t(\theta)\}$ is α -mixing since a measurable function of a finite subset of α -mixing processes is also α -mixing and of the same size (see, e.g., White, 2000).

Assumption 7 provides regularity conditions for the weighting function $W(\cdot)$. This assumption is applicable to any CDF with finite fourth moment. Assumption 8 provides the regularity conditions for the kernel function. Imposing continuity at zero and condition (ii) assist in eliminating the bias of the generalized cross-spectral derivative estimator $\hat{f}^{(0,m,0)}(\omega, 0, v)$ as $T \rightarrow \infty$. Condition (iii) dictates the tail behavior of $k(\cdot)$ so that higher-order lags have negligible impact on the statistical properties of $\hat{f}^{(0,m,0)}(\omega, 0, v)$. Some of the most frequently used kernels satisfy this assumption, including the Bartlett and Parzen kernels with $b = \infty$ and the Daniell and Quadratic-spectral kernels $b = 1$ and 2 , respectively.

We now present the asymptotic distribution of $\hat{Q}_a, a = 1, 2$, under \mathbb{H}_0 .

THEOREM 1. *Suppose $p = cT^\lambda$ for $0 < \lambda < (3 + \frac{1}{4b-2})^{-1}$ with $c \in (0, \infty)$. (i) Let Assumptions 1 to 8 hold. Under \mathbb{H}_0 and as $T \rightarrow \infty, \hat{Q}_1 \xrightarrow{d} N(0, 1)$. (ii) Let Assumptions 1, 3, 4, and 6–8 hold. Under \mathbb{H}_0 and as $T \rightarrow \infty$, if $\{z_t | I_{t-1}\}$ is i.i.d. $(0, \Phi)$, then $\hat{Q}_2 \xrightarrow{d} N(0, 1)$.*

A salient feature of $\hat{Q}_a, a = 1, 2$, is that the use of the estimated standardized residuals $\{\hat{z}_t\}$ in lieu of the true standardized residuals $\{z_t\}$ has no impact on the limit distribution of \hat{Q}_a . Hence, one can ignore the fact that the true parameter value θ^0 is unknown and set θ^0 to be equal to $\hat{\theta}$. This substitution is possible because the rate at which the parametric parameter estimator $\hat{\theta}$ converges exceeds that of the nonparametric kernel estimator $\hat{f}^{(0,m,0)}(\omega, 0, v)$ of $f^{(0,m,0)}(\omega, 0, v)$. As such, the limit distribution of \hat{Q}_a is completely governed by $\hat{f}^{(0,m,0)}(\omega, 0, v)$, and using $\hat{\theta}$ in lieu of θ^0 has no impact asymptotically. This ensures that any \sqrt{T} -consistent estimator will suffice.

So far our discussions have been centered around the null hypothesis. We now examine the asymptotic behavior of our test \hat{Q}_a under \mathbb{H}_A .

THEOREM 2. *Suppose $p = cT^\lambda$ for $\lambda \in (0, 1/2)$ and $c \in (0, \infty)$. Then under the conditions in Assumptions 1 and 3 to 8 and for $a = 1, 2$,*

$$\begin{aligned} \frac{p^{1/2}}{T} \hat{Q}_a &\xrightarrow{p} \frac{1}{D^{1/2}} \sum_{j=1}^{\infty} \int |\sigma_j^{(m,0)}(0, v)|^2 dW(v) \\ &= \frac{1}{D^{1/2}} \int \int_{-\pi}^{\pi} |f^{(0,m,0)}(\omega, 0, v) - f_0^{(0,m,0)}(\omega, 0, v)|^2 d\omega dW(v), \end{aligned}$$

where $D = 4\pi \int_0^{\infty} k^4(z) dz \quad \Omega_0^2 \int \int \int_{-\pi}^{\pi} |f(\omega, v, v')|^2 d\omega dW(v) dW(v')$.

Consider the case where $E[m_t(\theta) | z_{1t-j}, z_{2t-j}] \neq 0$ for some $j > 0$. This yields $\int |\sigma_j^{(m,0)}(0, v)|^2 dW(v) > 0$ for any weighting function that satisfies Assumption 7. Consequently, $P[\hat{Q}_a > c(T)] \rightarrow 1$ for any sequence of constants $\{c(T) = o(T/p^{1/2})\}$. Intuitively, this means that \hat{Q}_a has unitary power at any given level of significance whenever $E[m_t(\theta) | z_{1t-j}, z_{2t-j}]$ is nonzero at some lag $j > 0$. This characteristic of \hat{Q}_a suggests that it is sensitive to all forms of model misspecifications that result in $E[m_t(\theta) | z_{1t-j}, z_{2t-j}]$ being nonzero at some lag $j > 0$. To this end, our tests for the structure of conditional correlations may have low power against certain functional forms for time-varying conditional correlations.

5. MONTE CARLO STUDY

In this section, we investigate the empirical size and power of our test for constant conditional correlation and demonstrate how it fares against some existing tests. We choose DGPs with conditional mean normalized to zero to allow us to focus on the main theme of this article. Thus, our DGPs are such that $Y_t = (y_{1t}, y_{2t})' = (\varepsilon_{1t}, \varepsilon_{2t})'$, where $\varepsilon_{it} = \sqrt{h_{i,t}} z_{i,t}$ for $i = 1, 2$. To mitigate

startup effects, for each DGP we discard the first 500 Monte Carlo realizations. We estimate the bivariate CCC-GARCH models using Bollerslev’s (1990) algorithm; thus our estimates are QMLE. Also, we do not impose an upper bound on the unconditional variances.

5.1. *Experimental Design: The Existence of Constant Conditional Correlations.*

5.1.1. *Size.* To investigate the empirical size of the test under \mathbb{H}_0 , we analyze the following data generating processes (DGPs):

DGP1 [CCC-GARCH(1,1)]:

$$\begin{cases} h_{1,t} & = 0.4 + 0.15\varepsilon_{1,t-1}^2 + 0.8h_{1,t-1}, \\ h_{2,t} & = 0.2 + 0.2\varepsilon_{2,t-1}^2 + 0.7h_{2,t-1}, \\ \rho & = 0.2, \\ \begin{pmatrix} z_{1t} \\ z_{2t} \end{pmatrix} | I_{t-1} & \overset{i.i.d.}{\sim} N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right], \end{cases}$$

DGP2 [CCC-GARCH(1,1)]:

$$\begin{cases} h_{1,t} & = 0.4 + 0.15\varepsilon_{1,t-1}^2 + 0.8h_{1,t-1}, \\ h_{2,t} & = 0.2 + 0.2\varepsilon_{2,t-1}^2 + 0.7h_{2,t-1}, \\ \rho & = 0.8, \\ \begin{pmatrix} z_{1t} \\ z_{2t} \end{pmatrix} | I_{t-1} & \overset{i.i.d.}{\sim} N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right], \end{cases}$$

We take DGP1 and DGP2 from Tse (2000). The difference between DGP1 and DGP2 is the degree of constant conditional correlation. Our other null models are two copula-GARCH models, labeled DGP3 and DGP4 later. We first define a few terms and then give the specifications of DGPs 3 and 4. Following Lee and Long (2009), we assume $\varepsilon_t = H_t^{1/2} \Sigma_t^{-1/2} \eta_t$ where we define $H_t \equiv \Delta_t \Phi \Delta_t$ (see Subsection 2.1). The subsequent formulation ensures that $E(\varepsilon_t | I_{t-1}) = 0$ a.s. and $E(\varepsilon_t \varepsilon_t' | I_{t-1}) = H_t$ a.s., and ε_t also has higher-order moments a.s. For $i = 1, 2$, we let $u_i = F_i(\eta_i)$ be the probability integral transform of η_i and F_i be the univariate standard normal cdf; F_{12} be the joint cdf with margins F_1 and F_2 ; $\eta_t | I_{t-1} \sim F_{12}(\eta_1, \eta_2; \delta_t) = C(F_1(\eta_1), F_2(\eta_2); \delta_t) = C(u_1, u_2; \delta_t)$ where $C(\cdot)$ is the implied conditional copula distribution function and δ_t is the time-varying or time-invariant copula parameter that governs the higher-order dependence structure. This association between the joint distribution function F_{12} with margins F_1 and F_2 and the copula function is by virtue of Sklar’s Theorem (see, e.g., Patton, 2006; Lee and Long, 2009, and the references cited therein). In essence, a plethora of well-defined joint distributions with higher-order dependence structures can be generated by combining different margins and copulas. As such, other margins are permissible; our use of univariate standard normal margins simplifies the simulation and estimation procedures given that the parameters of these margins are known. Finally, we suppose $E(\eta_t | I_{t-1}) = 0$ a.s. and $E(\eta_t \eta_t' | I_{t-1}) = \Sigma_t \equiv [\Sigma_{ij,t}(\delta_t)]$ a.s. for identification. The off-diagonal elements of $\Sigma_{ij,t}(\delta_t)$ are extracted from the copula by applying Hoeffding’s Lemma.¹³ Our specifications of DGPs 3 and 4, respectively, are as follows:

¹³ Lee and Long (2009) provide a more detail description of the use of Hoeffding’s Lemma and the construction of copula-based MGARCH models. To simulate and estimate DGPs 3 and 4, as in Lee and Long (2009), we normalize the diagonal elements of Σ_t to be 1; this circumvents identification problems. We thank Tae-Hwy Lee for providing us with his copula codes.

DGP3 [Copula-GARCH(1,1)]:

$$\begin{cases} h_{1,t} &= 0.4 + 0.15\varepsilon_{1,t-1}^2 + 0.8h_{1,t-1}, \\ h_{2,t} &= 0.2 + 0.2\varepsilon_{2,t-1}^2 + 0.7h_{2,t-1}, \\ \rho &= 0.2, \\ \delta_t &= 1 + \exp(0.1 + 0.8\delta_{t-1} + 0.5u_{1,t-1} + 0.5u_{2,t-1}), \\ C(u_{1t}, u_{2t}; \delta_t) &= \exp\{-[(-\ln u_{1t})^{\delta_t} + (-\ln u_{2t})^{\delta_t}]^{1/\delta_t}\}. \end{cases}$$

DGP4 [Copula-GARCH(1,1)]:

$$\begin{cases} h_{1,t} &= 0.4 + 0.15\varepsilon_{1,t-1}^2 + 0.8h_{1,t-1}, \\ h_{2,t} &= 0.2 + 0.2\varepsilon_{2,t-1}^2 + 0.7h_{2,t-1}, \\ \rho &= 0.8, \delta = 2, \\ C(u_{1t}, u_{2t}; \delta) &= (u_{1t}^{-\delta} + u_{2t}^{-\delta} - 1)^{-1/\delta}. \end{cases}$$

DGP3 is a Gumbel-based MGARCH model with constant conditional correlation but time-varying higher-order moments, and DGP4 is a Clayton-based MGARCH model with constant conditional correlation but time-invariant higher-order moments. We introduce DGPs 3 and 4 in our analysis to examine the robustness of our tests to the presence of time-varying higher-order moments and nonnormal, particularly nonelliptical, distributions, which are well-documented features of financial data.¹⁴ We compute \hat{Q}_1 for each of these processes. Note that in the presence of time-invariant higher-order moments, which characterize DGPs 1, 2, and 4, both \hat{Q}_1 and \hat{Q}_2 have suitable asymptotic distributions. We report the empirical levels for 1,000 Monte Carlo realizations from samples of size $T = 500, 1,000, 2,500$.

5.1.2. Power. To analyze the powers of \hat{Q}_1 and \hat{Q}_2 in discriminating the CCC-MGARCH model of DGP1 from alternative models with time-varying conditional correlations, we consider the following DGPs:

DGP5 [TVC-MGARCH(1,1)]:

$$\begin{cases} h_{1,t} &= 0.4 + 0.15\varepsilon_{1,t-1}^2 + 0.8h_{1,t-1}, \\ h_{2,t} &= 0.2 + 0.2\varepsilon_{2,t-1}^2 + 0.7h_{2,t-1}, \\ \rho_t &= 0.07 + 0.8\rho_{t-1} + 0.1\pi_{t-1}, \\ \pi_{t-1} &= \frac{\sum_{h=1}^2 z_{1,t-h}z_{2,t-h}}{\sqrt{\left(\sum_{h=1}^2 z_{1,t-h}^2\right)\left(\sum_{h=1}^2 z_{2,t-h}^2\right)}}, \\ \begin{pmatrix} z_{1t} \\ z_{2t} \end{pmatrix} | I_{t-1} &\sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_t \\ \rho_t & 1 \end{pmatrix}\right]. \end{cases}$$

¹⁴For contour plots of these copula distributions refer to, for example, Patton (2006, p. 532).

DGP6 [Bivariate BEKK(1,1)]:

$$\begin{pmatrix} h_{1,t} & h_{12,t} \\ h_{21,t} & h_{2,t} \end{pmatrix} = \begin{pmatrix} 0.20 & 0.10 \\ 0.10 & 0.20 \end{pmatrix} + \begin{pmatrix} 0.60 & 0.20 \\ 0.20 & 0.60 \end{pmatrix} \begin{pmatrix} h_{1,t-1} & h_{12,t-1} \\ h_{21,t-1} & h_{2,t-1} \end{pmatrix} \begin{pmatrix} 0.60 & 0.20 \\ 0.20 & 0.60 \end{pmatrix}'$$

$$+ \begin{pmatrix} 0.30 & 0.10 \\ 0.10 & 0.30 \end{pmatrix} \begin{pmatrix} \varepsilon_{1,t-1}^2 & \varepsilon_{1,t-1}\varepsilon_{2,t-1} \\ \varepsilon_{1,t-1}\varepsilon_{2,t-1} & \varepsilon_{2,t-1}^2 \end{pmatrix} \begin{pmatrix} 0.30 & 0.10 \\ 0.10 & 0.30 \end{pmatrix}'$$

$$\rho_t = \frac{h_{12,t}}{\sqrt{h_{1,t} h_{2,t}}},$$

$$\begin{pmatrix} z_{1t} \\ z_{2t} \end{pmatrix} | I_{t-1} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_t \\ \rho_t & 1 \end{pmatrix} \right].$$

DGP7 [DCC(1,1)-BGARCH(1,1)]:

$$\begin{cases} h_{1,t} & = 0.4 + 0.15\varepsilon_{1,t-1}^2 + 0.8h_{1,t-1}, \\ h_{2,t} & = 0.2 + 0.2\varepsilon_{2,t-1}^2 + 0.7h_{2,t-1}, \\ \rho_t & = \frac{q_{12t}}{\sqrt{q_{11t} q_{22t}}}, \\ q_{12,t} & = 0.02 + 0.6q_{12,t-1} + 0.3z_{1,t-1}z_{2,t-1}, \\ q_{ii,t} & = 0.1 + 0.6q_{ii,t-1} + 0.3z_{i,t-1}^2, \quad i = 1, 2 \\ \left(\begin{matrix} z_{1t} \\ z_{2t} \end{matrix} \right) | I_{t-1} & \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_t \\ \rho_t & 1 \end{pmatrix} \right]. \end{cases}$$

DGP8 [CCC-DCC Regime Switching]:

$$\begin{aligned} h_{1,t} &= 0.4 + 0.15\varepsilon_{1,t-1}^2 + 0.8h_{1,t-1}, \\ h_{2,t} &= 0.2 + 0.2\varepsilon_{2,t-1}^2 + 0.7h_{2,t-1}, \\ \rho_t &= \begin{cases} \frac{0.02 + 0.5q_{12,t-1} + 0.4z_{1,t-1}z_{2,t-1}}{\sqrt{(0.1 + 0.5q_{11,t-1} + 0.4z_{1,t-1}^2)(0.1 + 0.5q_{22,t-1} + 0.4z_{2,t-1}^2)}}, & \text{if } F_t = 1 \\ 0.2, & \text{if } F_t = 2 \end{cases} \end{aligned}$$

$$f(z_t | I_{t-1}) = \sum_{F_t=1}^2 f_N(z_t | F_t, I_{t-1}) f(F_t | I_{t-1}).$$

DGP9 [CCC-CCC Regime Switching]:

$$\begin{aligned} h_{1,t} &= 0.4 + 0.15\varepsilon_{1,t-1}^2 + 0.8h_{1,t-1}, \\ h_{2,t} &= 0.2 + 0.2\varepsilon_{2,t-1}^2 + 0.7h_{2,t-1}, \end{aligned}$$

$$\rho_t = 0.5 - (F_t - 1), \quad F_t = 1, 2,$$

$$f(z_t | I_{t-1}) = \sum_{F_t=1}^2 f_N(z_t | F_t, I_{t-1}) f(F_t | I_{t-1}).$$

Some of these alternative models yield misspecifications in only conditional correlations whereas the others also generate misspecifications in other conditional moments or distribution. Since our \hat{Q}_1 and \hat{Q}_2 tests are designed for assessing model adequacy of the conditional correlation function, these models allow us to analyze the robustness of \hat{Q}_1 and \hat{Q}_2 to misspecifications in other aspects of the model.

We take DGP5 from Tse and Tsui (2002). DGP5 is also identical to DGP1 except in the specification of conditional correlation. In DGP5, the conditional correlation at time t is specified as the convex combination of the unconditional correlation, and lag-1 conditional correlation and the sample correlation of $\{z_{t-1}, z_{t-2}\}$. Moreover, there exists a strong dynamic dependence in the time-varying conditional correlations. Using DGP1 to fit the simulated data from DGP5 implies that there will be misspecifications in conditional correlations.¹⁵

The BEKK-parameterization of the conditional variance matrix in DGP6 accommodates dynamic dependence between volatility series and hence has implications for the structure of time-varying conditional correlations. Using DGP1 to fit the simulated data from DGP6 implies that the conditional variances and conditional correlations are misspecified.

DGP7 is Engle’s (2002) DCC model in which the time-varying correlation has three components, with each component having an autoregressive moving average structure. Unlike the BEKK specification, this model does not accommodate dynamic dependence in volatility series. Using DGP1 to fit data from this DCC model will result in misspecifications in conditional correlations.

DGP8 is motivated by Pelletier’s suggestion of an alternative way of introducing a regime switching for the correlations. We allow the parameters ζ_1 and ζ_2 of the correlation function in Engle’s DCC model in (4) to be a function of the regimes. That is,

$$\rho_{ij,t} = \frac{q_{ij,t}}{\sqrt{q_{ii,t} q_{jj,t}}},$$

$$q_{ij,t} = (1 - \zeta_1(F_t) - \zeta_2(F_t))\bar{q}_{ij} + \zeta_2(F_t)q_{ij,t-1} + \zeta_1(F_t)z_{i,t-1}z_{j,t-1}, \quad \forall t, \quad \text{and } i, j = 1, 2,$$

where $F_t = 1, 2$ is the latent Markov chain process, $\zeta_1(1) = 0.4$, $\zeta_1(2) = 0$, $\zeta_2(1) = 0.5$, and $\zeta_2(2) = 0$. In addition, $f(z_t | I_{t-1})$ and $f_N(z_t | F_t, I_{t-1})$ are, respectively, the marginal and conditional normal densities of z_t . Thus, the marginal density is the weighted average of the conditional densities given $F_t = 1$ and $F_t = 2$. These weighting factors are $Pr[F_t = 1 | I_{t-1}]$ and $Pr[F_t = 2 | I_{t-1}]$. We allow for symmetric parameterization of the transition probabilities between regimes 1 and 2 by choosing $p_{11} = p_{22} = 0.9$. The conditional correlations are dynamic in regime 1 and constant in regime 2. Note that the unconditional correlation is equal in both regimes. Using DGP1 to fit data from this CCC-DCC regime switching model will result in misspecifications in conditional correlations and conditional distribution. The nonnormality of the distribution of DGP8 will not affect the powers of \hat{Q}_1 and \hat{Q}_2 because these statistics are robust to distributional assumptions.

DGP9 follows Pelletier’s general specification of a regime switching dynamic correlation model. The correlations are of opposite signs and constant in both regimes. Transitions between regimes is governed by the latent Markov process F_t . We set the transition probabilities to be $p_{11} = p_{22} = 0.9$. Similar to DGP8, using DGP1 to fit data from this CCC-CCC regime switching model will result in misspecifications in conditional correlations and conditional distribution.

We obtain simulated data from DGPs 5, 6, 7, 8, and 9 of sample sizes $T = 500, 1,000, 2,500$. Using 500 Monte Carlo realizations, we use DGP1 to fit each these simulated data sets and

¹⁵ By construction, misspecification in conditional correlations implies misspecification in conditional covariances.

estimate and compare the power of our tests to some existing tests for constant conditional correlations.

5.2. *Competing Test Statistics.* We also consider the tests of constancy of conditional correlations proposed by Engle and Sheppard (2001), Bera and Kim (2002), and Tse (2000).

To compute Bera and Kim’s (2002) bivariate test statistic, BK , we estimate the constant conditional correlation bivariate GARCH model, compute \hat{z}_t and $\hat{\rho} = T^{-1} \sum_{t=1}^T \hat{z}_{1t}\hat{z}_{2t}$, and evaluate BK using

$$BK = \frac{\left[\sum_{t=1}^T (a_{1t}^2 a_{2t}^2 - 1 - 2\hat{\rho}^2) \right]^2}{4T(1 + 4\hat{\rho}^2 + \hat{\rho}^4)},$$

where $a_{1t} = (\hat{z}_{1t} - \hat{\rho}\hat{z}_{2t})/\sqrt{1 - \hat{\rho}^2}$ and $a_{2t} = (\hat{z}_{2t} - \hat{\rho}\hat{z}_{1t})/\sqrt{1 - \hat{\rho}^2}$. Under the assumption that the bivariate vector z_t is normally distributed, $BK \sim \chi_1^2$ asymptotically.

To test of constancy of conditional correlations, Tse (2000) presumes that the time evolution of conditional correlation can be characterized by the function $\rho_t = \rho + \delta\varepsilon_{1,t-1}\varepsilon_{2,t-1}$.¹⁶ Thus, under the null hypothesis, it suffices to show that $\delta = 0$. To employ Tse’s (2000), we estimate the constant conditional correlation bivariate GARCH model and obtain the \sqrt{T} -consistent estimator $\hat{\theta}$. We evaluate the score, $\partial l_t/\partial \theta'$, at $\hat{\theta}$. The corresponding lagrange multiplier statistic under the null hypothesis is obtained using

$$TSE = \iota' \hat{S}(\hat{S}'\hat{S})^{-1} \hat{S}'\iota,$$

where ι is a $T \times 1$ unit vector, \hat{S} is the estimator of the $T \times N$ matrix with rows equal to $\partial l_t/\partial \theta'$, for $t = 1, \dots, T$, evaluated at $\hat{\theta}$, and N is the number of parameters under the alternative model. This statistic is equivalent to TR^2 where R^2 is the uncentered coefficient of determination of the regression ι on \hat{S} . We then compare TSE to a suitable χ_1^2 critical value.

For the Engle and Sheppard (2001) test of constant conditional correlations, the test procedure is executed as follows: (1) Estimate the univariate GARCH processes and standardize the residuals for each series; (2) estimate the correlation of the standardized residuals and jointly standardize the vector of univariate standardized residuals by the symmetric square root decomposition of Φ , the constant correlation matrix; (3) compute $A_t \equiv vechu[(\Phi^{-1/2}z_t)(\Phi^{-1/2}z_t)' - I_2]$ where $vechu$ is the vectorization operator that selects the elements above the main diagonal and $\Phi^{-1/2}z_t$ is a bivariate vector of residuals jointly standardized under the null; (4) estimate the autoregression $A_t = \zeta_0 + \zeta_1 A_{t-1} + \dots + \zeta_p A_{t-p} + \nu_t$. Under the null hypothesis, the intercept and slope coefficients in (4) should be zero. Then the $ES(p)$ test statistic is

$$ES(p) = \frac{\hat{\zeta}'B'\hat{\zeta}}{\hat{\sigma}^2},$$

where $\hat{\zeta} = (\hat{\zeta}_0, \hat{\zeta}_1, \dots, \hat{\zeta}_p)'$ and B is a matrix consisting of the regressors. We compare this $ES(p)$ value to an appropriate $\chi_{(p+1)}^2$ critical value.

5.3. *Practical Implementation of \hat{Q}_1 and \hat{Q}_2 .* To calculate our test statistics \hat{Q}_1 and \hat{Q}_2 , see (19) and (20), we need a weighting function $W(\cdot)$, kernel function $k(\cdot)$, and an estimate of the bandwidth p . Our choice of $W(\cdot)$ is the $N(0, I_2)$, where I_2 is the identity matrix in $\mathbb{R}^{2 \times 2}$. For our choice of kernel function we use the Bartlett kernel, which has bounded support and

¹⁶ Tse also suggests the specification $\rho_t = \rho + \delta z_{1,t-1}z_{2,t-1}$ and mentions that this function cannot be used to obtain analytic derivatives from the likelihood function.

is computationally efficient. The simulation results indicate that the choice of weighting and kernel functions has no qualitative impact on the size and level of our tests.¹⁷

5.3.1. *Choosing a data-driven bandwidth.* We use Hong’s (1999) nonparametric plug-in method to find the adaptive bandwidth \hat{p} . For a sketch of this method, define $f^{(q,m,0)}(\omega, 0, v)$ as

$$(21) \quad f^{(q,m,0)}(\omega, 0, v) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} |j|^q \sigma_j^{(m,0)}(0, v) e^{-ij\omega},$$

and let $\bar{f}^{(0,m,0)}(\omega, 0, v)$ and $\bar{f}^{(q,m,0)}(\omega, 0, v)$ be sample analogues of (12) and (21) evaluated at some initial bandwidth \bar{p} , that is,

$$(22) \quad \bar{f}^{(0,m,0)}(\omega, 0, v) \equiv \frac{1}{2\pi} \sum_{j=1-T}^{T-1} \left(1 - \frac{|j|}{T}\right)^{1/2} \bar{k}(j/\bar{p}) \hat{\sigma}_j^{(m,0)}(0, v) e^{-ij\omega},$$

$$(23) \quad \bar{f}^{(q,m,0)}(\omega, 0, v) \equiv \frac{1}{2\pi} \sum_{j=1-T}^{T-1} |j|^q \left(1 - \frac{|j|}{T}\right)^{1/2} \bar{k}(j/\bar{p}) \hat{\sigma}_j^{(m,0)}(0, v) e^{-ij\omega}.$$

Assume that, for kernel $k(\cdot)$, its Parzen exponent q exists, that is, \exists a finite $q \in \mathbb{R}_+$ such that

$$k^{(q)} \equiv \lim_{|z| \rightarrow 0} \frac{1 - k(z)}{|z|^q} \text{ with } k^{(q)} \in (0, \infty).$$

Intuitively, the Parzen exponent governs the degree of smoothness for $k(\cdot)$ at 0; the larger is q the smoother is $k(\cdot)$ at 0. For example, the Bartlett and Parzen kernels have q equal to 1 and 2, respectively.

Then, following Hong, we can show that the theoretically optimal bandwidth that minimizes the asymptotic integrated mean squared error (IMSE) of the estimator of the generalized cross-spectral derivative $\hat{f}^{(0,m,0)}(\omega, 0, v)$ in (15) is $p_0 = c_0 T^{1/(2q+1)}$ for some tuning constant c_0 .¹⁸ For a workable c_0 , we select its sample counterpart evaluated at \bar{p} , which is

$$(24) \quad \hat{c} = \left\{ \frac{2q(k^{(q)})^2}{\int_{-\infty}^{\infty} k^2(z) dz} \times \frac{\int_{-\pi}^{\pi} \int_{-\pi}^{\pi} |\bar{f}^{(q,m,0)}(\omega, 0, v)|^2 d\omega dW(v)}{\int_{-\pi}^{\pi} \left[\int \bar{f}^{(0,m,0)}(\omega, 0, v) dW(v) \right]^2 d\omega} \right\}^{1/(2q+1)}$$

$$= \left\{ \frac{2q(k^{(q)})^2}{\int_{-\infty}^{\infty} k^2(z) dz} \times \frac{\sum_{j=1-T}^{T-1} |j|^{2q} (T - |j|) \bar{k}^2(j/\bar{p}) \int |\hat{\sigma}_j^{(m,0)}(0, v)|^2 dW(v)}{\sum_{j=1-T}^{T-1} (T - |j|) \bar{k}^2(j/\bar{p}) \hat{M}(j) \int \hat{\sigma}_j(v, -v) dW(v)} \right\}^{1/(2q+1)}$$

with $\hat{M}(j) = (T - |j|)^{-1} \sum_{t=|j|+1}^T m_{t-|j|}(\hat{\theta}) m_t(\hat{\theta})$ and $\bar{k}(\cdot)$ has Parzen exponent q . Note that $\bar{k}(\cdot)$ can be different from the kernel in, say, (20); however, for \hat{c} we use the Bartlett kernel. We select

¹⁷ We also use the Parzen kernel. The results, not reported here, show that there is no asymptotic cost to the choice of kernel.

¹⁸ This p_0 ensures that the optimal convergence rate of $n^{-2q/(2q+1)}$ for the IMSE of $\hat{f}^{(0,m,0)}(\omega, 0, v)$ is achieved.

TABLE 1
EMPIRICAL SIZE OF TEST FOR CONSTANCY OF CONDITIONAL CORRELATIONS

\bar{p}	T	DGP 1: CCC-MGARCH with $\rho = 0.2$						DGP 2: CCC-MGARCH with $\rho = 0.8$					
		500		1,000		2,500		500		1,000		2,500	
		10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%
10	\hat{Q}_1	8.3	5.1	10.0	5.9	11.1	7.0	7.6	3.7	8.8	4.8	8.8	6.1
20	\hat{Q}_1	6.0	2.9	9.3	5.4	10.2	6.9	5.8	2.7	7.4	4.6	9.0	5.7
30	\hat{Q}_1	5.2	2.0	7.8	4.4	9.4	5.6	4.7	2.2	6.0	3.8	8.6	4.6
40	\hat{Q}_1	3.8	1.3	6.5	3.4	8.9	4.7	3.7	1.4	5.4	2.6	8.2	4.1
10	\hat{Q}_2	7.2	3.9	7.2	4.3	7.4	4.1	6.5	3.2	6.6	3.7	7.3	4.0
20	\hat{Q}_2	7.0	3.2	7.9	3.9	7.8	4.2	6.4	3.2	6.9	4.1	7.2	3.7
30	\hat{Q}_2	6.7	3.5	7.5	3.8	6.9	4.1	6.2	3.9	5.9	3.6	7.3	3.1
40	\hat{Q}_2	7.2	3.3	7.0	3.8	6.7	3.7	6.2	3.5	6.3	3.4	7.0	3.7
	<i>TSE</i>	10.1	5.5	10.7	4.5	10.1	4.6	12.5	7.4	11.0	5.5	10.3	5.7
	<i>BK</i>	10.5	5.0	9.7	5.4	8.9	4.4	10.0	4.2	9.9	5.0	9.0	3.7
	<i>ES</i> (5)	4.8	2.0	4.6	2.4	6.4	3.1	3.9	1.8	3.8	2.1	5.0	2.4
	<i>ES</i> (10)	5.2	2.5	6.1	2.9	6.7	3.0	4.5	1.9	5.3	2.2	6.5	2.5

NOTE: We generate 1,000 Monte Carlo realizations for each DGP. \hat{Q}_1 and \hat{Q}_2 are the generalized cross-spectral derivative tests under higher-order conditional moments and i.i.d., respectively, with preliminary bandwidth \bar{p} equal to 10, 20, 30, 40; *BK* represents the Bera and Kim (2002) test statistic; *TSE* represents Tse's (2000) test statistic; *ES*(5) and *ES*(10) represent the Engle and Sheppard (2001) test statistic with the lagged value set to 5 and 10.

the integer-valued $\hat{p} = \lceil \hat{c}T^{1/(2q+1)} \rceil$ as our adaptive bandwidth. The choice of \bar{p} should have an asymptotically negligible impact on \hat{p} and, consequently, \hat{Q}_1 and \hat{Q}_2 . Thus in our simulation we investigate the effect of the initial bandwidth \bar{p} on the size and power of our tests by choosing $\bar{p} \in \{10, 11, \dots, 40\}$. In empirical applications where only a single \hat{Q}_1 or \hat{Q}_2 is desired, we can use a range of \bar{p} to compute a set of \hat{p} and then evaluate \hat{Q}_1 or \hat{Q}_2 at \hat{p}^* for

$$(25) \quad \hat{p}^* = \max \left\{ \max_{\hat{p}} \{ \ln(T), \hat{p} \} \right\}.$$

This choice for \hat{p}^* guarantees that the IMSE of $\hat{f}^{(0,m,0)}(\omega, 0, v)$ is achieved. The choice of the function $\ln(T)$, though ad hoc, guarantees that $\hat{p} \rightarrow \infty$ at a complementary rate.

For each \hat{p} , to evaluate the four-dimensional integral of the variance terms in \hat{Q}_1 and \hat{Q}_2 we randomly draw the auxiliary vectors u and v from a $N(0, I_2)$ distribution and discretize u and v to generate 30 grid points in \mathbb{R}^2 to facilitate Gaussian quadrature.

5.4. *Simulation Results.* Tables 1 and 2 show the empirical sizes of the \hat{Q}_1 , \hat{Q}_2 , *TSE*, *BK*, and *ES* tests for constancy of conditional correlation, assuming nominal sizes of 10% and 5%. We now focus on DGP1. At $T = 500$, both \hat{Q}_1 and \hat{Q}_2 (the tests derived under higher-order conditional moments and i.i.d. respectively) underreject \mathbb{H}_0 but not excessively. The rejection probabilities for \hat{Q}_1 are monotonically decreasing in the preliminary bandwidth \bar{p} . This pattern, however, becomes less pronounced as T increases. For \hat{Q}_2 , the rejection probabilities exhibit a more stable pattern than those of \hat{Q}_1 ; consequently, \hat{Q}_2 has better levels than \hat{Q}_1 the larger is \bar{p} . At $T = 1,000, 2,500$, \hat{Q}_1 overrejects, but not excessively, at lower values of \bar{p} whereas \hat{Q}_2 underrejects, but not excessively.

We now consider DGP2 (the model with the higher degree of conditional correlation). Except in two cases, and for all T , both \hat{Q}_1 and \hat{Q}_2 underreject \mathbb{H}_0 but not excessively. In general, the differences between the empirical and nominal rejection probabilities decline as T increases. At the 10% nominal level and for $T = 500, 1,000$, both \hat{Q}_1 and \hat{Q}_2 have lower rejection probabilities than their counterparts in DGP1.

We now focus on the existing tests for constant conditional correlation. For DGP1, the *TSE* and *BK* tests overreject or underreject \mathbb{H}_0 , but not severely. Nevertheless, for $T = 500, 1,000$

TABLE 2
EMPIRICAL SIZE OF TEST FOR CONSTANCY OF CONDITIONAL CORRELATIONS

		DGP 3: Copula-MGARCH with $\rho = 0.2$ Time-varying Gumbel Copula						DGP 4: Copula-MGARCH with $\rho = 0.8$ Time-invariant Clayton Copula					
\bar{p}	T α	500		1,000		2,500		500		1,000		2,500	
		10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%
10	\hat{Q}_1	7.9	4.8	8.8	5.1	10.9	6.5	7.0	3.9	9.0	5.0	11.6	5.9
20	\hat{Q}_1	6.3	3.0	8.1	4.6	10.1	5.8	4.3	3.0	8.2	4.5	11.1	5.1
30	\hat{Q}_1	5.5	2.3	7.5	4.0	9.6	5.1	3.4	2.5	6.2	3.9	10.5	4.9
40	\hat{Q}_1	4.1	1.8	6.3	3.5	8.8	4.6	3.0	1.9	5.2	2.8	9.7	4.5
10	\hat{Q}_2	36.2	32.5	34.3	30.8	35.1	31.7	7.1	3.4	7.1	3.9	8.6	4.0
20	\hat{Q}_2	37.0	33.4	36.4	34.0	36.9	33.1	7.2	3.6	8.1	4.4	9.0	4.1
30	\hat{Q}_2	36.4	33.0	35.5	32.8	36.1	31.8	7.4	3.9	7.8	4.0	9.3	4.5
40	\hat{Q}_2	37.3	33.6	35.1	32.1	35.9	31.5	8.0	4.1	7.7	3.7	9.3	4.6
	<i>TSE</i>	50.9	45.6	45.1	40.8	36.1	37.2	14.2	8.3	13.6	7.2	12.1	7.4
	<i>BK</i>	70.1	50.3	98.1	95.4	100.0	100.0	0.0	0.0	0.0	0.0	1.9	1.1
	<i>ES(5)</i>	96.3	95.0	100.0	100.0	100.0	100.0	6.1	3.9	7.2	4.1	9.1	4.9
	<i>ES(10)</i>	98.5	97.2	100.0	100.0	100.0	100.0	8.0	3.6	9.5	5.3	11.9	6.0

NOTE: We generate 1,000 Monte Carlo realizations for each DGP. \hat{Q}_1 and \hat{Q}_2 are the generalized cross-spectral derivative tests under higher-order conditional moments and i.i.d., respectively, with preliminary bandwidth \bar{p} equal to 10, 20, 30, 40; *BK* represents the Bera and Kim (2002) test statistic; *TSE* represents Tse's (2000) test statistic; *ES(5)* and *ES(10)* represent the Engle and Sheppard (2001) test statistic with the lagged value set to 5 and 10.

the TSE and BK tests have the best rejection probabilities of all the tests. The TSE test has the best rejection probabilities for $T = 2,500$. At the prespecified lag order of 5 and 10, the ES test underrejects \mathbb{H}_0 more than \hat{Q}_2 and has the least favorable size properties. For DGP2, the TSE test overrejects at $T = 500, 1,000$ but achieves the best rejection probabilities at $T = 2,500$. The BK test has the best rejection probabilities at $T = 500, 1,000$. The ES test continues to underreject \mathbb{H}_0 for all T . Similar to \hat{Q}_1 and \hat{Q}_2 , the TSE test appears to be slightly sensitive to the degree of constant conditional correlations.

Under DGP3 (the model with time-varying higher-order conditional moments) we expect \hat{Q}_1 to have the best size properties for all T . This is confirmed by our simulation results. The differences between the empirical and nominal rejection probabilities associated with \hat{Q}_1 are small, although, for a fixed T , these differences tend to increase with \bar{p} , but for a fixed \bar{p} , these differences tend to decrease with T . The BK and ES tests severely overreject \mathbb{H}_0 . The TSE and \hat{Q}_2 tests overreject \mathbb{H}_0 but not as much as the ES or BK tests. Under DGP4 (the model with time-invariant higher-order conditional moments), the aforementioned trends associated with \hat{Q}_1 also hold. \hat{Q}_2 underrejects \mathbb{H}_0 but not severely. However, the rejection probabilities of \hat{Q}_2 become favorable as T increases. The BK test severely underrejects \mathbb{H}_0 for all T . For the ES test, the differences between the empirical and nominal rejection probabilities decrease in lag order for a given T and decrease in T for a given lag order. The TSE test overrejects but not excessively. Thus for DGP4, \hat{Q}_1 and \hat{Q}_2 have the best size properties. Additional results, not reported here, from DGP4 with a Plackett copula evaluated at shape parameter $\delta = 4$ in lieu of the Clayton copula show that the BK test attains empirical rejection probabilities in excess of 95% for $T \geq 1,000$ whereas all other tests display favorable size properties. All the preceding results suggest that unlike existing tests for constant conditional correlation, the \hat{Q}_1 test is robust to the presence of time-varying higher-order moments whereas both the \hat{Q}_1 and \hat{Q}_2 tests are robust to the presence of time-invariant higher-order moments and nonelliptical distributions.

Tables 3–7 contain the empirical *corrected* and *uncorrected* powers against the time-varying conditional correlation alternatives, DGPs 5–9. We use the empirical critical values obtained under DGP1 to compute these empirical corrected powers. We consider nominal levels of 10% and 5%. We note that our empirical corrected and uncorrected powers are very similar. Thus,

TABLE 3
EMPIRICAL POWER OF TEST FOR CONSTANCY OF CONDITIONAL CORRELATIONS

		DGP 5: TVC-MGARCH											
		$\rho_t = 0.07 + 0.8\rho_{t-1} + 0.1\psi_{t-1}$											
		ACV						ECV					
		500		1,000		2,500		500		1,000		2,500	
\bar{p}	α	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%
10	\hat{Q}_1	85.0	78.0	98.4	97.8	100.0	100.0	84.6	74.0	100.0	100.0	100.0	100.0
20	\hat{Q}_1	83.3	74.8	99.2	98.0	100.0	100.0	86.2	79.8	100.0	100.0	100.0	100.0
30	\hat{Q}_1	79.6	71.0	98.8	97.2	100.0	100.0	86.8	78.0	100.0	100.0	100.0	100.0
40	\hat{Q}_1	76.0	64.4	98.6	96.4	100.0	100.0	85.8	76.0	100.0	100.0	100.0	100.0
10	\hat{Q}_2	8.8	6.0	10.2	6.2	13.3	7.1	13.4	7.0	13.6	7.6	15.0	10.1
20	\hat{Q}_2	9.8	5.0	10.0	5.2	13.0	6.9	14.4	6.8	13.6	6.4	15.2	10.5
30	\hat{Q}_2	10.6	4.6	9.0	5.0	12.5	6.4	13.8	7.0	13.4	6.0	14.9	10.3
40	\hat{Q}_2	8.8	4.4	8.2	4.0	11.1	5.9	12.4	6.4	13.0	6.2	15.1	9.7
	<i>TSE</i>	32.4	22.8	45.0	34.0	51.2	41.6						
	<i>BK</i>	15.8	10.2	18.4	13.2	34.0	26.6						
	<i>ES(5)</i>	12.8	6.6	28.8	20.8	40.3	36.8						
	<i>ES(10)</i>	11.6	6.8	27.2	18.2	35.2	33.1						

NOTE: We generate 500 Monte Carlo realizations for each DGP. \hat{Q}_1 and \hat{Q}_2 are the generalized cross-spectral derivative tests under higher-order conditional moment and i.i.d., respectively, with preliminary bandwidth \bar{p} equal to 10, 20, 30, 40; *BK* represents the Bera and Kim (2002) test statistic; *TSE* represents Tse's (2000) test statistic; *ES(5)* and *ES(10)* represent the Engle and Sheppard (2001) test statistic with the lagged value set to 5 and 10.

TABLE 4
EMPIRICAL POWER OF TEST FOR CONSTANCY OF CONDITIONAL CORRELATIONS

		DGP 6: BEKK											
		ACV						ECV					
		500		1,000		2,500		500		1,000		2,500	
\bar{p}	α	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%
10	\hat{Q}_1	81.8	75.2	98.6	98.0	99.8	99.8	80.8	71.4	98.4	96.1	99.4	97.8
20	\hat{Q}_1	77.8	69.4	98.4	97.8	99.8	98.4	81.4	74.4	98.1	96.9	99.8	97.6
30	\hat{Q}_1	72.8	61.0	97.4	96.4	99.0	97.8	80.4	71.4	98.0	97.1	99.0	97.6
40	\hat{Q}_1	68.2	54.6	96.8	94.6	98.4	96.6	78.2	68.2	97.3	96.0	98.4	96.6
10	\hat{Q}_2	38.2	31.0	70.8	60.4	99.6	98.4	47.2	33.8	76.0	62.4	99.8	99.0
20	\hat{Q}_2	40.6	31.0	71.8	62.2	99.8	98.4	46.8	36.6	76.8	63.8	100.0	98.8
30	\hat{Q}_2	38.2	30.8	69.0	58.4	99.0	98.4	45.6	33.8	75.4	61.8	99.6	98.6
40	\hat{Q}_2	35.8	27.4	65.0	53.6	98.8	98.0	43.8	32.8	72.2	59.0	99.0	98.4
	<i>TSE</i>	72.4	61.4	89.6	83.6	100.0	99.8						
	<i>BK</i>	10.4	4.2	11.8	6.4	11.4	7.4						
	<i>ES(5)</i>	4.2	2.0	6.8	2.4	11.0	6.0						
	<i>ES(10)</i>	5.4	2.8	6.4	3.4	11.6	5.2						

NOTE: We generate 500 Monte Carlo realizations for each DGP. \hat{Q}_1 and \hat{Q}_2 are the generalized cross-spectral derivative tests under higher-order conditional moment and i.i.d., respectively, with preliminary bandwidth \bar{p} equal to 10, 20, 30, 40; *BK* represents the Bera and Kim (2002) test statistic; *TSE* represents Tse's (2000) test statistic; *ES(5)* and *ES(10)* represent the Engle and Sheppard (2001) test statistic with the lagged value set to 5 and 10.

the rankings of our tests relative to the existing tests are identical regardless of our benchmark empirical power. We discuss only the empirical *corrected* powers in the ensuing analysis.

Under DGP5, the time-varying conditional correlations possess a high degree of inertia, and hence the perturbations to conditional correlations are small. This DGP allows us to assess the sensitivity of our constant correlation tests to detecting minimal time-variation in conditional

TABLE 5
EMPIRICAL POWER OF TEST FOR CONSTANCY OF CONDITIONAL CORRELATIONS

DGP 7: DCC(1,1)-BGARCH(1,1)

$$\rho_t = \frac{q_{12t}}{\sqrt{q_{11t} q_{22t}}}$$

\bar{p}	T α	ACV						ECV					
		500		1,000		2,500		500		1,000		2,500	
		10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%
10	\hat{Q}_1	99.4	98.6	100.0	100.0	100.0	100.0	99.2	98.2	100.0	100.0	100.0	100.0
20	\hat{Q}_1	99.2	98.0	100.0	100.0	100.0	100.0	99.4	98.8	100.0	100.0	100.0	100.0
30	\hat{Q}_1	98.2	96.2	100.0	100.0	100.0	100.0	99.4	98.0	100.0	100.0	100.0	100.0
40	\hat{Q}_1	96.6	93.8	100.0	100.0	100.0	100.0	99.2	96.0	100.0	100.0	100.0	100.0
10	\hat{Q}_2	99.0	98.4	99.8	99.8	100.0	100.0	99.2	98.8	99.8	99.8	100.0	100.0
20	\hat{Q}_2	99.2	98.2	99.8	99.8	100.0	100.0	99.2	98.8	99.8	99.8	100.0	100.0
30	\hat{Q}_2	98.4	97.6	99.8	99.8	100.0	100.0	98.8	98.0	99.8	99.8	100.0	100.0
40	\hat{Q}_2	97.8	97.0	99.8	99.8	100.0	100.0	98.2	97.6	99.8	99.8	100.0	100.0
	<i>TSE</i>	100.0	100.0	100.0	100.0	100.0	100.0						
	<i>BK</i>	92.6	88.8	99.8	99.4	100.0	100.0						
	<i>ES(5)</i>	100.0	100.0	100.0	100.0	100.0	100.0						
	<i>ES(10)</i>	100.0	100.0	100.0	100.0	100.0	100.0						

NOTE: We generate 500 Monte Carlo realizations for each DGP. \hat{Q}_1 and \hat{Q}_2 are the generalized cross-spectral derivative tests under higher-order conditional moment and i.i.d., respectively, with preliminary bandwidth \bar{p} equal to 10, 20, 30, 40; *BK* represents the Bera and Kim (2002) test statistic; *TSE* represents Tse's (2000) test statistic; *ES(5)* and *ES(10)* represent the Engle and Sheppard (2001) test statistic with the lagged value set to 5 and 10.

TABLE 6
EMPIRICAL POWER OF TEST FOR CONSTANCY OF CONDITIONAL CORRELATIONS

DGP 8: CCC-DCC Regime Switching

\bar{p}	T α	ACV						ECV					
		500		1,000		2,500		500		1,000		2,500	
		10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%
10	\hat{Q}_1	72.6	61.2	97.2	94.4	100.0	100.0	78.2	58.0	97.2	93.0	100.0	100.0
20	\hat{Q}_1	65.6	56.0	95.8	91.4	100.0	100.0	71.0	60.8	96.4	91.2	100.0	100.0
30	\hat{Q}_1	57.2	45.8	91.8	88.2	100.0	100.0	68.2	55.0	93.6	89.4	100.0	100.0
40	\hat{Q}_1	49.8	34.2	90.4	82.6	100.0	100.0	65.2	49.8	92.8	87.6	100.0	100.0
10	\hat{Q}_2	70.8	60.8	96.0	92.8	100.0	100.0	75.2	65.0	97.4	94.8	100.0	100.0
20	\hat{Q}_2	69.8	59.2	94.8	91.4	100.0	100.0	74.6	67.2	97.0	92.6	100.0	100.0
30	\hat{Q}_2	65.8	54.8	92.6	90.4	100.0	100.0	73.2	61.4	95.2	91.6	100.0	100.0
40	\hat{Q}_2	60.0	51.0	91.6	88.4	100.0	100.0	68.4	56.8	93.4	90.2	100.0	100.0
	<i>TSE</i>	95.8	94.2	100.0	100.0	100.0	100.0						
	<i>BK</i>	64.8	58.2	89.6	87.2	99.8	99.6						
	<i>ES(5)</i>	95.4	93.2	100.0	100.0	100.0	100.0						
	<i>ES(10)</i>	91.4	87.4	100.0	99.8	100.0	100.0						

NOTE: We generate 500 Monte Carlo realizations for each DGP. \hat{Q}_1 and \hat{Q}_2 are the generalized cross-spectral derivative tests under higher-order conditional moment and i.i.d., respectively, with preliminary bandwidth \bar{p} equal to 10, 20, 30, 40; *BK* represents the Bera and Kim (2002) test statistic; *TSE* represents Tse's (2000) test statistic; *ES(5)* and *ES(10)* represent the Engle and Sheppard (2001) test statistic with the lagged value set to 5 and 10.

correlations. We expect \hat{Q}_1 to be more powerful than \hat{Q}_2 . This is confirmed by our simulation results. \hat{Q}_1 is the most powerful, with power reaching 100% when $T = 1,000$. The TSE test is more powerful than the \hat{Q}_2 , BK, and ES tests. At $T = 2,500$, the powers of the ES test are less than 45% and exceed those of \hat{Q}_2 and BK. For all sample sizes, the powers of \hat{Q}_2 are slightly above the nominal levels, which render \hat{Q}_2 the least powerful test under DGP5.

TABLE 7
EMPIRICAL POWER OF TEST FOR CONSTANCY OF CONDITIONAL CORRELATIONS

DGP 9: CCC-CCC Regime Switching $\rho_1 = 0.5, \rho_2 = -0.5$													
		ACV						ECV					
\bar{p}	T α	500		1,000		2,500		500		1,000		2,500	
		10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%
10	\hat{Q}_1	51.2	36.8	90.0	82.6	100.0	100.0	50.0	33.0	90.0	79.6	100.0	100.0
20	\hat{Q}_1	41.6	30.4	85.6	76.4	100.0	100.0	47.6	36.2	88.0	75.6	100.0	99.8
30	\hat{Q}_1	34.2	22.2	77.8	69.0	100.0	99.8	45.4	31.4	82.2	71.8	100.0	99.8
40	\hat{Q}_1	28.6	16.6	71.8	58.2	99.8	99.4	41.4	28.6	80.0	67.0	99.8	99.6
10	\hat{Q}_2	48.0	35.0	86.4	78.2	100.0	100.0	58.4	39.6	89.0	80.0	100.0	100.0
20	\hat{Q}_2	46.4	33.8	83.6	74.8	100.0	100.0	54.8	40.8	88.4	76.6	100.0	100.0
30	\hat{Q}_2	42.6	31.2	79.4	71.0	100.0	99.8	50.4	36.8	85.2	74.6	100.0	99.8
40	\hat{Q}_2	38.2	30.4	76.2	66.0	99.8	99.4	46.8	34.0	81.0	70.0	99.8	99.6
	TSE	94.6	91.2	99.6	99.2	100.0	100.0						
	BK	100.0	99.2	100.0	100.0	100.0	100.0						
	ES(5)	93.4	89.0	99.6	98.4	100.0	100.0						
	ES(10)	89.8	82.6	99.4	99.0	100.0	100.0						

NOTE: We generate 500 Monte Carlo realizations for each DGP. \hat{Q}_1 and \hat{Q}_2 are the generalized cross-spectral derivative tests under higher-order conditional moment and i.i.d., respectively, with preliminary bandwidth \bar{p} equal to 10, 20, 30, 40; BK represents the Bera and Kim (2002) test statistic; TSE represents Tse's (2000) test statistic; ES(5) and ES(10) represent the Engle and Sheppard (2001) test statistic with the lagged value set to 5 and 10.

DGP6 admits volatility interactions along with time-varying conditional correlations and is misspecified for conditional variances and conditional correlations. \hat{Q}_1 is the most powerful when $T = 500, 1,000$, but the TSE test slightly dominates in power when $T = 2,500$ to attain optimal power at the 10% nominal level. The powers of \hat{Q}_1 and \hat{Q}_2 are very similar when $T = 2,500$. For all sample sizes, the \hat{Q}_2 test dominates the BK and ES tests; the ES test has powers less than 12% whereas the BK test has powers closer to the nominal levels.

We now consider DGP7. The results show that all five tests have excellent power under DGP7. At $T = 500$, we find that the TSE and ES tests attain perfect power whereas, at the 10% level, the \hat{Q}_2 and \hat{Q}_1 , respectively, achieve powers in excess of 99% and 98%, and the BK test achieves power in excess of 93%. At $T = 1,000$, \hat{Q}_1 attains perfect power whereas \hat{Q}_2 and BK achieve identical powers of 99.8% at the 10% level. At $T = 2,500$, all test are equally powerful.

DGP8 is a latent hybrid of a constant conditional correlation model and dynamic conditional correlation model and has a nonnormal conditional distribution. Thus under DGP6, DGP1 is misspecified for conditional correlations and conditional distribution. At $T = 500$, the BK test is the least powerful whereas TSE is the most powerful followed by the ES test. As T increases, the gap in powers between the TSE and ES, \hat{Q}_1 and \hat{Q}_2 tests tapers off. At $T = 1,000$, TSE and ES achieve perfect power, \hat{Q}_1 and \hat{Q}_2 attain power in excess of 90% whereas BK attains power slightly less than 90%. At $T = 2,500$, \hat{Q}_1 and \hat{Q}_2 achieve perfect power whereas BK attains power slightly less than 100%.

DGP9 is also regime switching but the conditional correlations are constant in both regimes. For all $T = 500, 1,000, 2,500$, the BK test is the most powerful and reaches optimal power except for the case when $T = 500$ and the nominal level is 5%. The TSE and ES tests achieve similar powers, greater than 98%, when $T = 1,000$. There is a drastic increase in the powers of \hat{Q}_1 and \hat{Q}_2 as T increases. The \hat{Q}_1 and \hat{Q}_2 tests are the least powerful except when $T = 2,500$ in which case \hat{Q}_1 and \hat{Q}_2 realize their maximum powers of 100% at lower levels of the preliminary bandwidth.

To sum up, we observe:

- The empirical sizes of the i.i.d. test, \hat{Q}_2 , are lower than their nominal counterparts and insensitive to the choice of preliminary bandwidth. The empirical sizes of the higher-order conditional moment test, \hat{Q}_1 , decrease monotonically as the preliminary bandwidth increases, but this pattern becomes less pronounced as the sample size increases. Sample sizes in excess of 1,000 are more desirable for the generalized cross-spectral derivative tests.
- Unlike existing tests, our \hat{Q}_1 test is robust to the presence of time-varying higher-order moments, and both \hat{Q}_1 and \hat{Q}_2 are robust to the presence of time-invariant higher-order moments and nonelliptical distributions.
- The TSE and BK tests have favorable size properties in the presence of a normal error distribution.
- The \hat{Q}_1 test is more powerful than the \hat{Q}_2 test in identifying time-varying conditional correlations even when these variations are small, e.g., the time-varying correlation MGARCH model (DGP5).
- All tests, \hat{Q}_1 , \hat{Q}_2 , TSE, BK, and ES, have good power in discriminating between constant conditional correlation and time-varying conditional correlations that evolve according to Engle's DCC specification. The TSE test is the most powerful in this case.
- The \hat{Q}_1 is not always the most powerful but has good power against *all* time-varying conditional correlation DGPs considered in our simulation study.

6. EMPIRICAL APPLICATION

Engle and Colacito (2006) provide an interesting analysis that quantifies the benefit of knowing the true structure of time-varying conditional correlations within the context of a classical asset allocation framework. Engle and Colacito prove that the infimum of the ratio of the portfolio variances associated with an incorrect estimate of the covariance matrix to that associated with the true covariance matrix is equal to 1. This variance inequality, which holds for an arbitrary vector of expected returns and any required excess return, provides the basis for testing the relative performance of time-varying covariance models. They fit to the data a set of multivariate volatility models, including the DCC and asymmetric DCC (ADCC), and select the model that delivers the lowest estimate of the portfolio variance over a range of expected return vectors. They assume this minimum-variance model is the true model. Hence, holding fixed the expected return vector, the ratio of the estimated portfolio standard deviation of this true model to that of an alternative model is an estimate of the increase in risk from using the incorrect volatility model. By exploiting the symmetrical nature of the asset allocation problem and holding portfolio volatility constant, Engle and Colacito label this increase in risk as the gain in required return from using the true relative to the estimated time-varying covariances.

To value the correlation information, Engle and Colacito simulate a time series of returns using the estimated parameters of the ADCC model that was fitted to the real data but fix the variances of the simulated data to be the unconditional variances of the real data. They approximate the gain in required return that could be demanded by an investor, using the true ADCC model in lieu of the incorrect CCC model, to be at most 23%. Given the dependence of their valuation methods on model adequacy, we use our generalized cross-spectral derivative test for the structure of conditional correlations to assess whether the ADCC model adequately captures the dynamics of the conditional correlations in their data.

The first set of bivariate data in Engle and Colacito consists of daily data of S&P500 (ISPCS00) and 10-year bond (CTYCS00) futures from DataStream for the time span August 26, 1988, to August 26, 2003.¹⁹ The second set consists of daily data on Dow Jones Industrials and S&P500

¹⁹ Silvennoinen and Teräsvirta (2009a, 2009b) also use this data set.

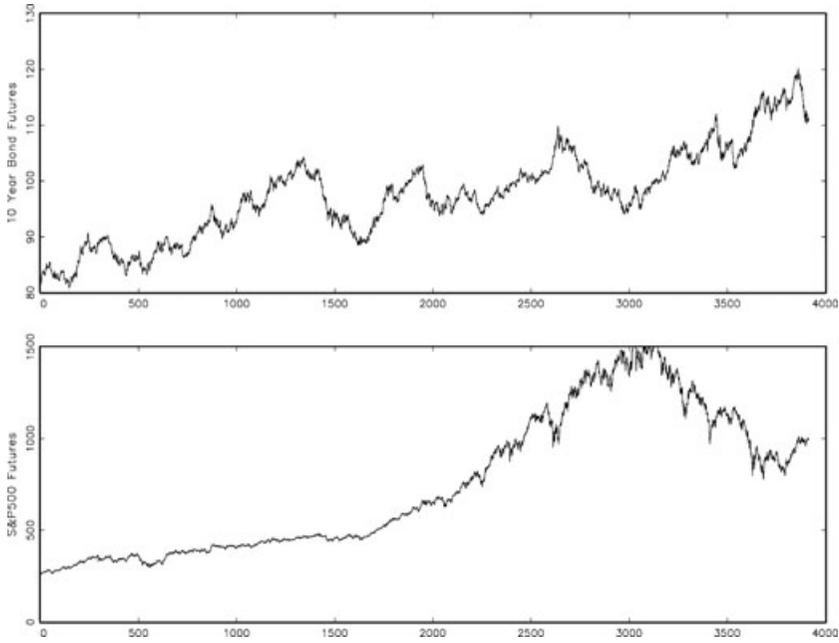


FIGURE 1

TIME SERIES OF S&P500 FUTURES AND 10-YEAR BONDS FUTURES FROM 8/26/88 TO 8/26/03; $N = 3,911$

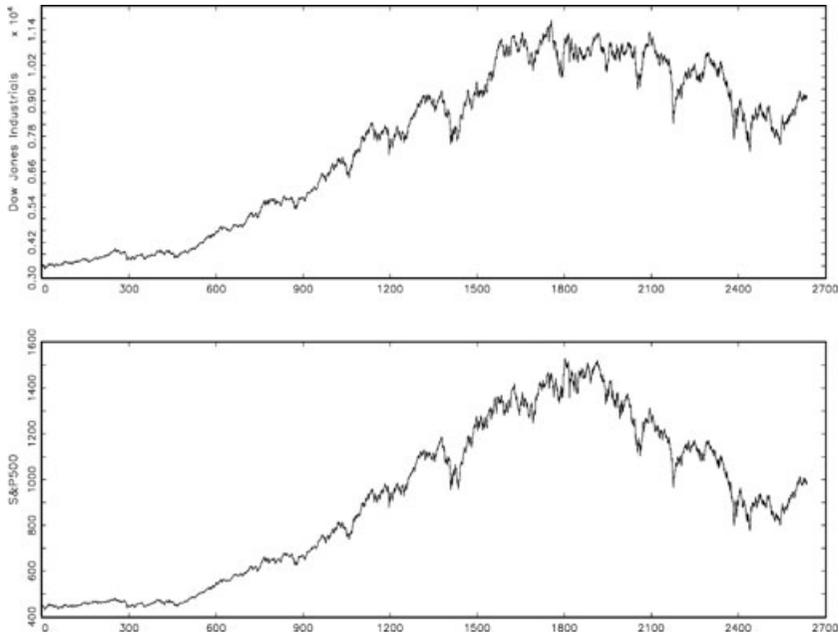


FIGURE 2

TIME SERIES OF S&P500 AND DOW JONES INDUSTRIALS FROM 2/4/93 TO 7/22/03; $N = 2,636$

Indexes from Yahoo! Finance for the time period 2/4/1993 to 7/22/2003.²⁰ In Figures 1 and 2, we replicate Engle and Colacito's time plots for the first and second sets of data (low- and high-correlated data), respectively.

²⁰ Details on these data are in Engle and Colacito (2006).

6.1. *Estimation and Generalized Cross-Spectrum Test Results.* We use our generalized cross-spectrum derivative test, \hat{Q}_1 , that accounts for time-varying higher-order conditional moments. This is the same \hat{Q}_1 we use in our simulation study; we compute the \hat{Q}_1 as outlined in the preceding section. We first test for constant conditional correlation. If we fail to accept the null specification, we then test for adequate specification of time-varying conditional correlations. To execute the first test, we fit a CCC model to the data. As in Engle and Colacito, we choose the GARCH(1,1) specification for the volatility functions of the CCC model. To estimate the DCC and ADCC models, we adopt the exact specifications utilized by Engle and Colacito. Thus, the specification for the ADCC model is $Y_t = H_t^{1/2} \eta_t$, where

$$\begin{aligned}
 H_t &= \begin{pmatrix} h_{1,t}^{1/2} & 0 \\ 0 & h_{2,t}^{1/2} \end{pmatrix} \begin{pmatrix} 1 & \rho_t \\ \rho_t & 1 \end{pmatrix} \begin{pmatrix} h_{1,t}^{1/2} & 0 \\ 0 & h_{2,t}^{1/2} \end{pmatrix}, \\
 h_{1,t} &= \omega_1 + \alpha_1 y_{1,t-1}^2 + \beta_1 h_{1,t-1} + \gamma_1 d_{1,t-1} y_{1,t-1}^2, \\
 h_{2,t} &= \omega_2 + \alpha_2 y_{2,t-1}^2 + \beta_2 h_{2,t-1} + \gamma_2 d_{2,t-1} y_{2,t-1}^2, \\
 \rho_t &= h_{12,t}^* / \sqrt{h_{1,t}^* h_{2,t}^*} \\
 h_{1,t}^* &= (1 - \zeta_1 - \zeta_2 - \zeta_3/2) + \zeta_1 z_{1,t-1}^2 + \zeta_2 h_{1,t-1} + \zeta_3 d_{1,t-1} z_{1,t-1}^2, \\
 h_{2,t}^* &= (1 - \zeta_1 - \zeta_2 - \zeta_3/2) + \zeta_1 z_{2,t-1}^2 + \zeta_2 h_{2,t-1} + \zeta_3 d_{2,t-1} z_{2,t-1}^2, \\
 h_{12,t}^* &= \phi_{12}(1 - \zeta_1 - \zeta_2) - \phi_3 \zeta_3 + \zeta_1 z_{1,t-1} z_{2,t-1} + \zeta_2 h_{12,t-1} + \zeta_3 (d_{1,t-1} z_{1,t-1})(d_{2,t-1} z_{2,t-1}),
 \end{aligned}$$

and $\eta_t \stackrel{iid}{\sim} N(0, I_2)$. Moreover, $d_{1,t}$ and $d_{2,t}$ equal 1 for negative values of $y_{1,t}$ and $y_{2,t}$ and zero otherwise. Also, ϕ_{12} and ϕ_3 are the average sample correlation of returns and the average of the asymmetric component $(d_{1,t-1} z_{1,t-1})(d_{2,t-1} z_{2,t-1})$, and $z_{1,t}$ and $z_{2,t}$ are the standardized residuals. The specification of the DCC model is in the previous section. All the parameter estimates are QMLE.

Table 8 contains the variance and correlation parameter estimates for all three models and both data sets. The qualitative implications of these estimates are parallel to those of Engle and Colacito. For the low-correlated data, Figure 3 displays the results of our analysis, where we plot the \hat{Q}_1 test statistic under various null specifications against an integer sequence of preliminary lag orders. We include two unmarked horizontal demarcations in this figure to represent the standard normal critical values at the 1% and 5% significance levels. The curve labeled CCC confirms Engle and Colacito’s finding that the assumption of constant conditional correlation is inconsistent with the data for S&P500 and 10-year bond futures. We use Equation (25) to obtain the optimal bandwidth, \hat{p}^* , for the DCC and ADCC model, which is 28. The DCC model appears inconsistent with the data at the 5% significance level but consistent with the data at the 1% significance level. However, at both significance levels, the ADCC curve reveals that the ADCC model adequately characterizes the dynamics of conditional correlations in these data. These results suggest Engle and Colacito’s approximation of the gain in expected return that can be achieved by using the true structure of the time-varying conditional correlations is accurate.

For the high-correlated data, Figure 4 shows the results of our \hat{Q}_1 test statistic. These results are all in excess of the standard normal critical value at the 1% significance level; consequently the horizontal demarcations in Figure 3 are not needed in Figure 4. In particular, for Figure 4, we see that the assumption of constant conditional correlation is at odds with the data. The DCC curve indicates that the DCC model is inconsistent with the dynamics of conditional correlations. Also, the ADCC curve suggests that the ADCC model does not provide an accurate fit to the data at the conventional significance levels. Thus, for the high-correlated Dow Jones Industrials and S&P500 data, Engle and Colacito’s estimated value of knowing the true structure of the time-varying conditional correlations may have been underestimated.

TABLE 8
PARAMETER ESTIMATES

Models	Stock variance parameters			Bond variance parameters			Correlations parameters					
	ω_1	α_1	β_1	γ_1	ω_2	α_2	β_2	γ_2	ξ_1	ξ_2	ξ_3	ρ
CCC	0.0004 (0.0001)	0.0278 (0.0030)	0.9594 (0.0043)		0.0014 (0.0002)	0.0472 (0.0035)	0.9476 (0.0039)					0.1569 (0.0122)
DCC	0.0014 (0.0004)	0.0473 (0.0069)	0.9476 (0.0078)		0.0004 (0.0001)	0.0279 (0.0046)	0.9594 (0.0072)		0.0157 (0.0026)	0.9727 (0.0047)		
ADCC	0.0034 (0.0007)	0.0014 (0.0056)	0.9236 (0.0106)	0.1202 (0.0169)	0.0004 (0.0001)	0.0251 (0.0062)	0.9592 (0.0074)	0.0044 (0.0070)	0.0184 (0.0033)	0.9725 (0.0049)	-0.0013 (0.0050)	
Models	S&P500 variance parameters											
CCC	0.0011 (0.0002)	0.0731 (0.0042)	0.9256 (0.0040)		0.0019 (0.0002)	0.0899 (0.0048)	0.9060 (0.0047)					0.9256 (0.0024)
DCC	0.0011 (0.0004)	0.0730 (0.0102)	0.9256 (0.010)		0.0019 (0.0006)	0.0899 (0.0116)	0.9060 (0.0115)		0.0314 (0.0055)	0.9526 (0.0092)		
ADCC	0.0283 (0.0042)	0.0036 (0.0085)	0.9208 (0.0102)	0.1345 (0.0182)	0.0405 (0.0062)	0.0212 (0.010)	0.9101 (0.0115)	0.1135 (0.0178)	0.0441 (0.0073)	0.9590 (0.0080)	0.0055 (0.0061)	
Models	Dow Jones variance parameters											
CCC	0.0011 (0.0002)	0.0731 (0.0042)	0.9256 (0.0040)		0.0019 (0.0002)	0.0899 (0.0048)	0.9060 (0.0047)					0.9256 (0.0024)
DCC	0.0011 (0.0004)	0.0730 (0.0102)	0.9256 (0.010)		0.0019 (0.0006)	0.0899 (0.0116)	0.9060 (0.0115)		0.0314 (0.0055)	0.9526 (0.0092)		
ADCC	0.0283 (0.0042)	0.0036 (0.0085)	0.9208 (0.0102)	0.1345 (0.0182)	0.0405 (0.0062)	0.0212 (0.010)	0.9101 (0.0115)	0.1135 (0.0178)	0.0441 (0.0073)	0.9590 (0.0080)	0.0055 (0.0061)	

NOTE: The numbers in parentheses are standard errors. The top panel corresponds to S&P500 Futures and 10-Year Bonds Futures. The bottom panel corresponds to S&P500 and Dow Jones Industrials data.

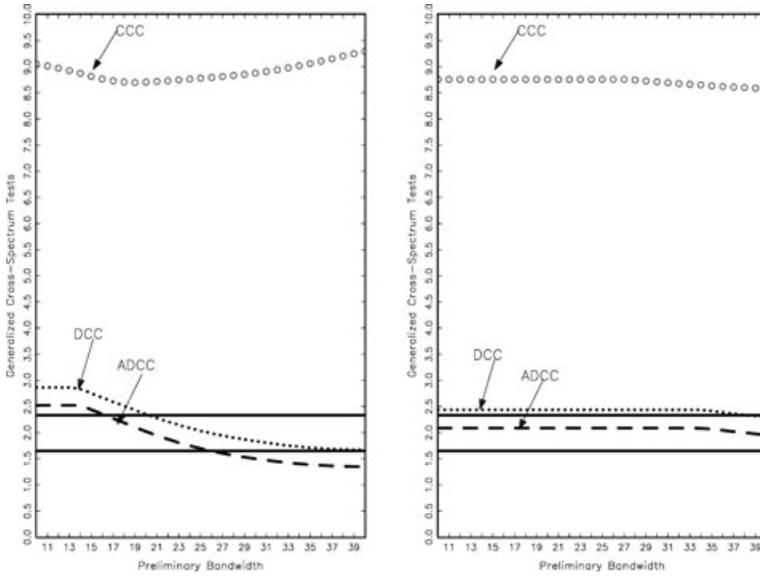


FIGURE 3

GENERALIZED CROSS-SPECTRAL DERIVATIVE TESTS FOR THE S&P500 FUTURES AND 10-YEAR BONDS FUTURES. THE LEFT PANEL USES THE BARTLETT KERNEL AND THE RIGHT PANEL USES THE PARZENS KERNEL. THE TWO UNMARKED HORIZONTAL LINES REPRESENT THE STANDARD NORMAL CRITICAL VALUES AT THE 1% AND 5% SIGNIFICANCE LEVELS.

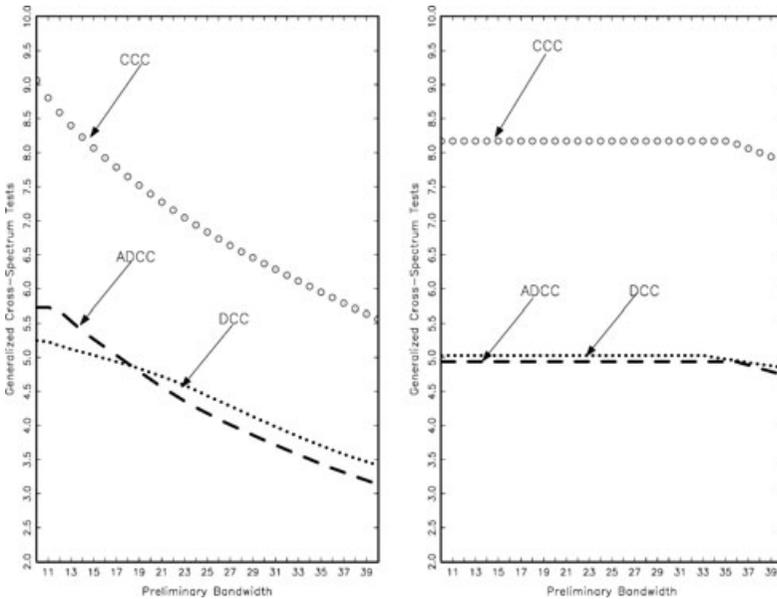


FIGURE 4

GENERALIZED CROSS-SPECTRAL DERIVATIVE TESTS FOR THE S&P500 AND DOW JONES INDUSTRIALS DATA. THE LEFT PANEL USES THE BARTLETT KERNEL AND THE RIGHT PANEL USES THE PARZENS KERNEL.

7. CONCLUSION

Many researchers have echoed the importance of the structure of conditional correlations for numerous types of economic and financial decisions including optimal portfolio diversification and hedging and risk management. The different structures for conditional correlations in multivariate GARCH models that have been put forward warrant general specification tests

to discriminate among competing models and obtain reliable inferences in empirical applications. However, little attention has been paid to general specification tests for the adequacy of these structures, specifically time-varying structures, for conditional correlations. Using a unified framework, we introduce a class of generally applicable tests for assessing the existence of constant conditional correlations and parametric specification of time-varying conditional correlations. Our tests are robust to time-varying higher-order conditional moments, for example, skewness and kurtosis, of unknown form in the conditional density of the innovation vector. Time-varying higher-order conditional moments in time-series data can arise for many reasons, and their existence cannot be viewed as immaterial. It has been argued that monetary policy objectives of central banks and financial decisions of investors can give rise to time-varying higher-order dependence structures in time-series data. Recently, time-varying higher-order conditional moments have been found in equity returns and exchange rate data. Moreover, time-varying higher-order conditional moments have been shown to have a significant impact on time-varying lower-order conditional moments. Specification tests for conditional correlations that are not robust to time-varying higher-order conditional moments will exhibit poor size performances. The theoretical and empirical relevance of this higher-order-moment feature of our tests underscores one of the essential contributions of the present article to the existing literature.

Our tests can identify linear and nonlinear misspecifications in conditional correlations. In addition, our tests do not require an alternative model, a particular estimation method and distributional assumption. Also, the asymptotic distribution of the tests is invariant to the uncertainty in parameter estimation. Our simulation study reveals that our test of constant conditional correlation has good size properties and is consistent against a wide range of structures of time-varying conditional correlations. We illustrate the practicality of our generally applicable tests using real data.

APPENDIX

In this Appendix, we assume $C \in (0, \infty)$ is an arbitrary bounded constant, $\|\cdot\|$ the Euclidean norm, and A^* the complex conjugate of A . We also assume I_{t-1} is the infinite, unobservable information set. We presume that the bivariate data generating process $Y_t = \mu_t + \varepsilon_t$ has conditional variance matrix $H_t = \Lambda_t \Phi_t \Lambda_t$ with $\Lambda_t = \text{diag}(h_{11,t}^{1/2}, h_{22,t}^{1/2})$ and each $h_{ii,t}$ has GARCH(1,1) errors so that $h_{ii,t} = \omega_i^0 + \alpha_i^0 \varepsilon_{it-1}^2 + \beta_i^0 h_{ii,t-1}$, for $i = 1, 2$, with $\omega_i^0, \alpha_i^0, \beta_i^0$ being elements of the finite dimensional parameter vector θ^0 . Also Φ_t is the time-varying conditional correlation matrix with off-diagonal entries equal to ρ_t . For our observed sample, we let \tilde{I}_{t-1} be the observable information set, which contains some initial values and $\tilde{I}_{t-1} \subset I_{t-1}$. We assume that a \sqrt{T} -consistent estimator $\hat{\theta}$ for θ^0 is associated with \tilde{I}_{t-1} and derived from $\tilde{h}_{ii,t}(\theta) = \omega_i + \alpha_i \tilde{\varepsilon}_{it-1}^2(\theta) + \beta_i \tilde{h}_{ii,t-1}(\theta)$, with initial values $\tilde{h}_{ii,t}(\theta) \equiv \tilde{h}_{ii,t} \leq C$ for $t \leq 0$ and $\tilde{\varepsilon}_{it}(\theta) = 0$ for $t \leq 0$. We assume $Q_a, a = 1, 2$, is identical to \hat{Q}_a in (16) and (17) except the unobservable sample $\{z_t \equiv z_t(\theta^0), \rho_t \equiv \rho_t(\theta^0)\}_{t=1}^T$, with $\theta^0 \equiv p \lim \hat{\theta}$, is in lieu of the estimated residual sample $\{\hat{z}_t \equiv \tilde{z}_t(\hat{\theta}), \hat{\rho}_t \equiv \tilde{\rho}_t(\hat{\theta})\}_{t=1}^T$.

PROOF OF THEOREM 1. Here we only consider the proof for \hat{Q}_1 since that of \hat{Q}_2 is less involved. This proof has three main components stated below as Theorems 3 to 5. Intuitively, Theorem 3 states that using the estimated standardized residuals instead of the true standardized error does not affect the limit distribution of \hat{Q}_1 . Theorem 4 states that the use of a sufficiently large subset of the true standardized error, $\{z_{q,t}, \rho_{q,t}\}_{t=1}^T$, does not affect the limit distribution of \hat{Q}_1 .

THEOREM 3. Under the conditions of Theorem 1, $\hat{Q}_1 - Q_1 \xrightarrow{p} 0$.

THEOREM 4. Define Q_{1q} to be Q_1 but with $\{z_{1q,t} z_{2q,t} - \rho_{q,t}\}$ in lieu of $\{z_{1t} z_{2t} - \rho_t\}_{t=1}^T$. Let $q = p^{1 + \frac{1}{4b-2}} (\ln^2 T)^{\frac{1}{2b-1}}$. Under the conditions of Theorem 1, $Q_{1q} - Q_1 \xrightarrow{p} 0$.

THEOREM 5. Let $q = p^{1+\frac{1}{4b-2}}(\ln^2 T)^{\frac{1}{2b-1}}$. Under the conditions of Theorem 1, $Q_{1q} \xrightarrow{d} N(0, 1)$.

PROOF OF THEOREM 3. To proceed, we first establish a bound for the term $T^{-1} \sum_{t=1}^T [\hat{z}_{1t}\hat{z}_{2t} - z_{1t}z_{2t}]^2$. To derive this stochastic bound, we adapt some of the results in Nelson (1990), Andrews (1992), Lee and Hansen (1994), Lumsdaine (1996), and Hong (2001). For $i = 1, 2$, let $\check{z}_{it}(\theta) \equiv \tilde{\varepsilon}_{it}(\theta)/\check{h}_{ii,t}^{1/2}(\theta)$ where

$$(A.1) \quad \check{h}_{ii,t}(\theta) = \omega_i + \alpha_i \varepsilon_{ii,t-1}^2(\theta) + \beta_i \check{h}_{ii,t-1}(\theta) = \frac{\omega_i}{1 - \beta_i} + \alpha_i \sum_{j=0}^{\infty} \beta_i^j \varepsilon_{ii,t-1-j}^2(\theta)$$

is an unobservable strictly stationary process with information set I_{t-1} . Then, we can write $\hat{z}_{it} - z_{it} = [\hat{z}_{it} - \check{z}_{it}(\hat{\theta})] + [\check{z}_{it}(\hat{\theta}) - z_{it}]$ for $i = 1, 2$. We now show that for $i = 1, 2$

$$T^{-1} \sum_{t=1}^T (\hat{z}_{it}(\hat{\theta}) - z_{it})^4 = O_p(T^{-2}) \quad \text{and} \quad T^{-1} \sum_{t=1}^T (\hat{z}_{it} - \check{z}_{it}(\hat{\theta}))^4 = O_p(T^{-1}).$$

We emphasize that although $\check{h}_{ii,t}(\theta^0) = h_{ii,t}$, $\tilde{h}_{ii,t}(\theta^0) \neq h_{ii,t}$ due to the initial value $\tilde{h}_{ii,0}$. This implies that $\check{z}_{it}(\theta^0) = z_{it}$, $\tilde{z}_{it}(\theta^0) \neq z_{it}$. Furthermore, $\check{h}_{ii,t}(\theta) - \tilde{h}_{ii,t}(\theta) = \beta_i^t [\check{h}_{ii,0}(\theta) - \tilde{h}_{ii,0}]$, $\check{z}_{it}(\theta) = \tilde{\varepsilon}_{it}(\theta)/\check{h}_{ii,t}^{1/2}(\theta)$, $\tilde{h}_{ii,t}(\theta) \geq C^{-1}$, and $\tilde{h}_{ii}(\theta) \geq C^{-1}$. Now, $T^{-1} \sum_{t=1}^T (\hat{z}_{it} - \check{z}_{it}(\hat{\theta}))^4 = \sum_{t=1}^T \frac{\varepsilon_{it}^4 [\hat{h}_{ii,t}^{1/2} - \check{h}_{ii,t}^{1/2}(\hat{\theta})]^4}{[\hat{h}_{ii,t}^{1/2} \check{h}_{ii,t}^{1/2}(\hat{\theta})]^4}$,

$$\text{where} \quad [\hat{h}_{ii,t}^{1/2} - \check{h}_{ii,t}^{1/2}(\hat{\theta})]^4 = \frac{[\hat{h}_{ii,t} - \check{h}_{ii,t}(\hat{\theta})]^4}{[\hat{h}_{ii,t}^{1/2} + \check{h}_{ii,t}^{1/2}(\hat{\theta})]^4} \quad \text{and} \quad \hat{h}_{ii,t}^{1/2} + \check{h}_{ii,t}^{1/2}(\hat{\theta}) \geq 2C^{-1/2}.$$

It follows that $T^{-1} \sum_{t=1}^T (\hat{z}_{it} - \check{z}_{it}(\hat{\theta}))^4 \leq \frac{1}{16} C^{-6} T^{-1} [\check{h}_{ii,0}(\hat{\theta}) - \tilde{h}_{ii,0}]^4 \sum_{t=1}^T \beta_i^{4t} \varepsilon_{it}^4$.

Suppose Θ^0 is a convex and compact neighborhood of θ^0 . Assuming $E \sup_{\theta \in \Theta^0} |\check{h}_{ii,0}(\theta)|^4 < \infty$, then by Markov's inequality, $\sup_{\theta \in \Theta^0} \check{h}_{ii,0}^4(\theta) = O_p(1)$. For $0 < p < \frac{1}{4}$, using the C_r and Cauchy Schwarz inequalities, we obtain

$$\begin{aligned} E \left\{ \sup_{\theta \in \Theta^0} |\check{\varepsilon}_{it}^4(\theta)|^p \right\} &= E \sup_{\theta \in \Theta^0} |\varepsilon_{it} + \mu_{it} - \tilde{\mu}_{it}(\theta)|^{4p} \leq E |\varepsilon_{it}|^{4p} + \sup_{\theta \in \Theta^0} |\mu_{it} - \tilde{\mu}_{it}(\theta)|^{4p} \\ &= \{E |z_{it}|^{8p}\}^{1/2} \{E(h_{ii,t}^{8p})\}^{1/2} + \sup_{\theta \in \Theta^0} |\mu_{it} - \tilde{\mu}_{it}(\theta)|^{4p} \leq 2C, \end{aligned}$$

which follows from Assumptions 1 and 4, and also $E(h_{ii,t}^{8p}) \leq C$, which is adapted from Nelson (1990). Consequently, by Markov's inequality we deduce $\sup_{\theta \in \Theta^0} \sum_{t=1}^T \beta_i^{4t} \check{\varepsilon}_{it}^4(\theta) = O_p(1)$ where $0 < \beta_i \leq 1 - \delta < 1$ with $\delta > 0$ and small and $\theta \in \Theta^0$. Hence, $T^{-1} \sum_{t=1}^T (\hat{z}_{it} - \check{z}_{it}(\hat{\theta}))^4 = O_p(T^{-1})$ for $i = 1, 2$.

Employing the Mean Value theorem and the Cauchy Schwarz inequality, we find that

$$(A.2) \quad T^{-1} \sum_{t=1}^T (\check{z}_{it}(\hat{\theta}) - z_{it})^4 \leq \|\hat{\theta} - \theta^0\|^4 \left\{ T^{-1} \sum_{t=1}^T \|\nabla_{\theta} \check{z}_{it}(\bar{\theta})\|^4 \right\} = O_p(T^{-2}).$$

Arguments similar to those of Hong (2001, p. 213) can be used to show that $\sup_{\theta \in \Theta^0} T^{-1} \sum_{t=1}^T \|\nabla_{\theta} \check{z}_{it}(\theta)\|^4 = O_p(1)$.

Then, we have

$$\begin{aligned}
 T^{-1} \sum_{t=1}^T [\hat{z}_{1t} \hat{z}_{2t} - z_{1t} z_{2t}]^2 &= T^{-1} \sum_{t=1}^T [\hat{z}_{1t}(\hat{z}_{2t} - z_{2t}) + z_{2t}(\hat{z}_{1t} - z_{1t})]^2 \\
 &\leq 2T^{-1} \sum_{t=1}^T \hat{z}_{1t}^2 (\hat{z}_{2t} - z_{2t})^2 + 2T^{-1} \sum_{t=1}^T z_{2t}^2 (\hat{z}_{1t} - z_{1t})^2 \\
 &\leq 16 \left\{ T^{-1} \sum_{t=1}^T (\hat{z}_{1t} - \check{z}_{1t}(\hat{\theta}))^4 \right\}^{1/2} \left\{ T^{-1} \sum_{t=1}^T (\hat{z}_{2t} - \check{z}_{2t}(\hat{\theta}))^4 \right\}^{1/2} \\
 &\quad + 16 \left\{ T^{-1} \sum_{t=1}^T (\check{z}_{1t}(\hat{\theta}) - z_{1t})^4 \right\}^{1/2} \left\{ T^{-1} \sum_{t=1}^T (\hat{z}_{2t} - \check{z}_{2t}(\hat{\theta}))^4 \right\}^{1/2} \\
 &\quad + 8 \left\{ T^{-1} \sum_{t=1}^T z_{1t}^4 \right\}^{1/2} \left\{ T^{-1} \sum_{t=1}^T (\hat{z}_{2t} - \check{z}_{2t}(\hat{\theta}))^4 \right\}^{1/2} \\
 &\quad + 16 \left\{ T^{-1} \sum_{t=1}^T (\check{z}_{1t}(\hat{\theta}) - z_{1t})^4 \right\}^{1/2} \left\{ T^{-1} \sum_{t=1}^T (\check{z}_{2t}(\hat{\theta}) - z_{2t})^4 \right\}^{1/2} \\
 &\quad + 16 \left\{ T^{-1} \sum_{t=1}^T (\check{z}_{1t}(\hat{\theta}) - z_{1t})^4 \right\}^{1/2} \left\{ T^{-1} \sum_{t=1}^T (\check{z}_{2t}(\hat{\theta}) - z_{2t})^4 \right\}^{1/2} \\
 &\quad + 8 \left\{ T^{-1} \sum_{t=1}^T z_{1t}^4 \right\}^{1/2} \left\{ T^{-1} \sum_{t=1}^T (\check{z}_{2t}(\hat{\theta}) - z_{2t})^4 \right\}^{1/2} \\
 &\quad + 4 \left\{ T^{-1} \sum_{t=1}^T z_{2t}^4 \right\}^{1/2} \left\{ T^{-1} \sum_{t=1}^T (\hat{z}_{1t} - \check{z}_{1t}(\hat{\theta}))^4 \right\}^{1/2} \\
 &\quad + 4 \left\{ T^{-1} \sum_{t=1}^T z_{2t}^4 \right\}^{1/2} \left\{ T^{-1} \sum_{t=1}^T (\check{z}_{1t}(\hat{\theta}) - z_{1t})^4 \right\}^{1/2} \\
 &= O_p(T^{-1}) + O_p(T^{-3/2}) + O_p(T^{-1/2}) + O_p(T^{-3/2}) \\
 &\quad + O_p(T^{-2}) + O_p(T^{-1}) + O_p(T^{-1/2}) + O_p(T^{-1}) = O_p(T^{-1/2}),
 \end{aligned}$$

where the last decomposition is by virtue of the Cauchy Schwarz inequality and we invoke Markov's inequality to derive $T^{-1} \sum_{t=1}^T z_{it}^4 = O_p(1)$ for $i = 1, 2$.

In a similar manner, we have $T^{-1} \sum_{t=1}^T \|\hat{z}_t - z_t\|^2 = T^{-1} \sum_{t=1}^T (\hat{z}_{1t} - z_{1t})^2 + T^{-1} \sum_{t=1}^T (\hat{z}_{2t} - z_{2t})^2$, where, for $i = 1, 2$, $T^{-1} \sum_{t=1}^T (\hat{z}_{it} - z_{it})^2 \leq 2T^{-1} \sum_{t=1}^T (\hat{z}_{it} - \check{z}_{it}(\hat{\theta}))^2 + 2T^{-1} \sum_{t=1}^T (\check{z}_{it}(\hat{\theta}) - z_{it})^2$. From above, we obtain $T^{-1} \sum_{t=1}^T (\hat{z}_{it} - \check{z}_{it}(\hat{\theta}))^2 = O_p(T^{-1})$ and $T^{-1} \sum_{t=1}^T (\check{z}_{it}(\hat{\theta}) - z_{it})^2 = O_p(T^{-1})$. Therefore, $T^{-1} \sum_{t=1}^T \|\hat{z}_t - z_t\|^2 = O_p(T^{-1})$.

To bound $\sum_{t=1}^T [\hat{\rho}_t - \rho_t]^2$, let $\rho_t(\hat{\theta})$ be associated with I_{t-1} . Then, by Assumption 4

$$(A.3) \quad \sum_{t=1}^T [\hat{\rho}_t - \rho_t(\hat{\theta})]^2 = O_p(1).$$

Similarly, using a first-order Taylor series expansion, $\rho_t(\hat{\theta}) = \rho_t + (\hat{\theta} - \theta^0)' \nabla_{\theta} \rho_t(\theta)$, where $\|\hat{\theta} - \theta^0\|^2 \leq \|\hat{\theta} - \theta^0\|^2$. Utilizing the Cauchy Schwarz inequality and Assumptions 3 and 6 yields

$$(A.4) \quad \sum_{t=1}^T [\rho_t(\hat{\theta}) - \rho_t]^2 \leq T \|\hat{\theta} - \theta^0\|^2 T^{-1} \sum_{t=1}^T \sup_{\theta \in \Theta^0} \|\nabla_{\theta} \rho_t(\theta)\|^2 = O_p(1),$$

where Θ^0 is an ϵ -neighborhood of θ^0 . Therefore,

$$(A.5) \quad \sum_{t=1}^T [\hat{\rho}_t - \rho_t]^2 = O_p(1).$$

Now define $\tilde{\sigma}_j^{(m,0)}(0, v)$ to be $\hat{\sigma}_j^{(m,0)}(0, v)$ but with $m_t(\theta^0)$ in lieu of $m_t(\hat{\theta})$. To prove the theorem we can show (i) $\hat{D}_1^{-1/2} \int \sum_{j=1}^{T-1} k^2(j/p)(T-j)[|\hat{\sigma}_j^{(m,0)}(0, v)|^2 - |\tilde{\sigma}_j^{(m,0)}(0, v)|^2] dW(v) \xrightarrow{p} 0$, (ii) $\hat{C}_1 - \tilde{C}_1 = O_p(T^{-1/2})$, and (iii) $\hat{D}_1 - \tilde{D}_1 \xrightarrow{p} 0$. We note that \hat{D}_1 and \hat{C}_1 grow to infinity at rate p as $p \rightarrow \infty, p/T \rightarrow 0$. Since parts (ii) and (iii) above are straightforward, we will only show the proof for part (i). To begin, we express the integral in (i) as

$$(A.6) \quad \int \sum_{j=1}^{T-1} k^2(j/p)(T-j)[|\hat{\sigma}_j^{(m,0)}(0, v)|^2 - |\tilde{\sigma}_j^{(m,0)}(0, v)|^2] dW(v) = \hat{B}_1 + 2\text{Re}(\hat{B}_2),$$

where

$$\begin{aligned} \hat{B}_1 &= \int \sum_{j=1}^{T-1} k^2(j/p)(T-j) |\hat{\sigma}_j^{(m,0)}(0, v) - \tilde{\sigma}_j^{(m,0)}(0, v)|^2 dW(v), \\ \hat{B}_2 &= \int \sum_{j=1}^{T-1} k^2(j/p)(T-j) [\hat{\sigma}_j^{(m,0)}(0, v) - \tilde{\sigma}_j^{(m,0)}(0, v)] \tilde{\sigma}_j^{(m,0)}(0, v)^* dW(v). \end{aligned}$$

It remains to demonstrate that Propositions 1 and 2 are satisfied. ■

PROPOSITION 1. Under the regularity conditions of Theorem 1, $p^{-1/2} \hat{B}_1 \rightarrow 0$.

PROPOSITION 2. Under the regularity conditions of Theorem 1, $p^{-1/2} \hat{B}_2 \rightarrow 0$.

PROOF OF PROPOSITION 1. Let $\hat{\delta}_t(v) \equiv e^{iv' \hat{z}_t} - e^{iv' z_t}$, $\psi_{t-j}(v) \equiv e^{iv' z_{t-j}} - \varphi_j(v)$, where $\varphi_j(v) \equiv E(e^{iv' z_{t-j}})$, $T_j \equiv T - j$, and suppose $j > 0$. It is easy to show that

$$(A.7) \quad \begin{aligned} \hat{\sigma}_j^{(m,0)}(0, v) - \tilde{\sigma}_j^{(m,0)}(0, v) &= T_j^{-1} \sum_{t=j+1}^T [\hat{z}_{1t} \hat{z}_{2t} - z_{1t} z_{2t}] \hat{\delta}_{t-j}(v) + T_j^{-1} \sum_{t=j+1}^T [\hat{z}_{1t} \hat{z}_{2t} - z_{1t} z_{2t}] \psi_{t-j}(v) \\ &\quad - \left(T_j^{-1} \sum_{t=j+1}^T [\hat{z}_{1t} \hat{z}_{2t} - z_{1t} z_{2t}] \right) \left(T_j^{-1} \sum_{t=j+1}^T \hat{\delta}_{t-j}(v) \right) \\ &\quad - \left(T_j^{-1} \sum_{t=j+1}^T [\hat{z}_{1t} \hat{z}_{2t} - z_{1t} z_{2t}] \right) \left(T_j^{-1} \sum_{t=j+1}^T \psi_{t-j}(v) \right) \end{aligned}$$

$$\begin{aligned}
 &+ T_j^{-1} \sum_{t=j+1}^T z_{1t} z_{2t} \hat{\delta}_{t-j}(v) - \left(T_j^{-1} \sum_{t=j+1}^T z_{1t} z_{2t} \right) \left(T_j^{-1} \sum_{t=j+1}^T \hat{\delta}_{t-j}(v) \right) \\
 &- T_j^{-1} \sum_{t=j+1}^T [\hat{\rho}_t - \rho_t] \hat{\delta}_{t-j}(v) - T_j^{-1} \sum_{t=j+1}^T [\hat{\rho}_t - \rho_t] \psi_{t-j}(v) \\
 &+ \left(T_j^{-1} \sum_{t=j+1}^T [\hat{\rho}_t - \rho_t] \right) \left(T_j^{-1} \sum_{t=j+1}^T \hat{\delta}_{t-j}(v) \right) \\
 &+ \left(T_j^{-1} \sum_{t=j+1}^T [\hat{\rho}_t - \rho_t] \right) \left(T_j^{-1} \sum_{t=j+1}^T \psi_{t-j}(v) \right) - T_j^{-1} \sum_{t=j+1}^T \rho_t \hat{\delta}_{t-j}(v) \\
 &+ \left(T_j^{-1} \sum_{t=j+1}^T \rho_t \right) \left(T_j^{-1} \sum_{t=j+1}^T \hat{\delta}_{t-j}(v) \right) = \sum_{b=1}^{12} \hat{E}_{bj}(v).
 \end{aligned}$$

Therefore, to complete the proof of Proposition 1, it suffices to prove the following lemma. ■

LEMMA 1. For $b = 1, \dots, 12$, let $\hat{E}_{bj}(v)$ be as defined above. Then

- (1) $\sum_{j=1}^{T-1} k^2(j/p) T_j \int |\hat{E}_{1j}(v)|^2 dW(v) = O_p(\frac{p}{T^{1/2}})$.
- (2) $\sum_{j=1}^{T-1} k^2(j/p) T_j \int |\hat{E}_{2j}(v)|^2 dW(v) = O_p(1)$.
- (3) $\sum_{j=1}^{T-1} k^2(j/p) T_j \int |\hat{E}_{3j}(v)|^2 dW(v) = O_p(\frac{p}{T^{1/2}})$.
- (4) $\sum_{j=1}^{T-1} k^2(j/p) T_j \int |\hat{E}_{4j}(v)|^2 dW(v) = O_p(\frac{p}{T^{1/2}})$.
- (5) $\sum_{j=1}^{T-1} k^2(j/p) T_j \int |\hat{E}_{5j}(v)|^2 dW(v) = O_p(\frac{p}{T})$.
- (6) $\sum_{j=1}^{T-1} k^2(j/p) T_j \int |\hat{E}_{6j}(v)|^2 dW(v) = O_p(\frac{p}{T})$.
- (7) $\sum_{j=1}^{T-1} k^2(j/p) T_j \int |\hat{E}_{7j}(v)|^2 dW(v) = O_p(\frac{p}{T})$.
- (8) $\sum_{j=1}^{T-1} k^2(j/p) T_j \int |\hat{E}_{8j}(v)|^2 dW(v) = O_p(1)$.
- (9) $\sum_{j=1}^{T-1} k^2(j/p) T_j \int |\hat{E}_{9j}(v)|^2 dW(v) = O_p(\frac{p}{T})$.
- (10) $\sum_{j=1}^{T-1} k^2(j/p) T_j \int |\hat{E}_{10j}(v)|^2 dW(v) = O_p(\frac{p}{T})$.
- (11) $\sum_{j=1}^{T-1} k^2(j/p) T_j \int |\hat{E}_{11j}(v)|^2 dW(v) = O_p(\frac{p}{T})$.
- (12) $\sum_{j=1}^{T-1} k^2(j/p) T_j \int |\hat{E}_{12j}(v)|^2 dW(v) = O_p(\frac{p}{T})$.

PROOF OF PROPOSITION 2. To begin, note that

$$|[\hat{\sigma}_j^{(m,0)}(0, v) - \tilde{\sigma}_j^{(m,0)}(0, v)] \tilde{\sigma}_j^{(m,0)}(0, v)^*| \leq \sum_{b=1}^{12} |\hat{E}_{bj}(v)| |\tilde{\sigma}_j^{(m,0)}(0, v)|,$$

where $\hat{E}_{bj}(v)$ is as previously defined. For ease of exposition, let

$$\tilde{E}_{bj}(v) \equiv \sum_{j=1}^{T-1} k^2(j/p) T_j \int |\hat{E}_{bj}(v)| |\tilde{\sigma}_j^{(m,0)}(0, v)| dW(v).$$

By the Cauchy-Schwarz inequality and for each $b \in \{1, \dots, 12\}$ but with $b \neq 2$ or 8 , we write

$$\tilde{E}_{bj} \leq \left\{ \sum_{j=1}^{T-1} k^2(j/p) T_j \int |\hat{E}_{bj}(v)|^2 dW(v) \right\}^{1/2} \left\{ \sum_{j=1}^{T-1} k^2(j/p) T_j \int |\tilde{\sigma}_j^{(m,0)}(0, v)|^2 dW(v) \right\}^{1/2}.$$

Under the null hypothesis, $\sup_{v \in \mathbb{R}^2} E|\tilde{\sigma}_j^{(m,0)}(0, v)|^2 \leq CT_j^{-1}$. Moreover, by Markov's inequality and the m.d.s. property of $\{z_{1t}z_{2t} - \rho_t\}$,

$$\sum_{j=1}^{T-1} k^2(j/p) T_j \int |\tilde{\sigma}_j^{(m,0)}(0, v)|^2 dW(v) = O_p(p).$$

Then, using Lemma 1, we find that for $b = 1, 3, 4$, $\tilde{E}_{bj} = O_p(p^{1/2}/T^{1/4})O_p(p^{1/2}) = o_p(p^{1/2})$, for $p = cT^\lambda, 0 < \lambda < (3 + \frac{1}{4b-2})^{-1}, p \rightarrow \infty, T \rightarrow \infty, p/T \rightarrow 0$; when $b = 5, 6, 7, 9, 10, 11, 12$, $\tilde{E}_{bj} = O_p(p^{1/2}/T^{1/2})O_p(p^{1/2}) = o_p(p^{1/2})$. To complete this proof, we need to establish bounds for \tilde{E}_{2j} and \tilde{E}_{8j} . Since their derivations are quite similar, we only show that $\tilde{E}_{8j} = o_p(p^{1/2})$. Now,

$$\begin{aligned} \text{(A.8)} \quad \tilde{E}_{8j}(v) &\equiv \sum_{j=1}^{T-1} k^2(j/p) T_j \int |\hat{E}_{8j}(v)| |\tilde{\sigma}_j^{(m,0)}(0, v)| dW(v) \\ &\leq \sum_{j=1}^{T-1} k^2(j/p) T_j \int |\hat{E}_{8j1}(v)| |\tilde{\sigma}_j^{(m,0)}(0, v)| dW(v) \\ &\quad + \sum_{j=1}^{T-1} k^2(j/p) T_j \int |\hat{E}_{8j2}(v)| |\tilde{\sigma}_j^{(m,0)}(0, v)| dW(v). \end{aligned}$$

Then, for the first term in (A.8), we have

$$\begin{aligned} \sum_{j=1}^{T-1} k^2(j/p) T_j \int |\hat{E}_{8j1}(v)| |\tilde{\sigma}_j^{(m,0)}(0, v)| dW(v) &\leq 2 \left[\sum_{t=1}^T \sup_{\theta \in \Theta^0} |\tilde{\rho}_t(\theta) - \rho_t(\theta)| \right] \\ &\quad \times \sum_{j=1}^{T-1} k^2(j/p) \int |\tilde{\sigma}_j^{(m,0)}(0, v)| dW(v) \\ &= O_p(p/T^{1/2}), \end{aligned}$$

by virtue of the kernel bound, Assumption 4, and the m.d.s property of $\{z_{1t}z_{2t} - \rho_t\}$.

For the second term in (A.8), we have

$$\begin{aligned} \text{(A.9)} \quad &\sum_{j=1}^{T-1} k^2(j/p) T_j \int |\hat{E}_{8j2}(v)| |\tilde{\sigma}_j^{(m,0)}(0, v)| dW(v) \\ &\leq \|\hat{\theta} - \theta^0\| \sum_{j=1}^{T-1} k^2(j/p) T_j \int \left\| T_j^{-1} \sum_{t=j+1}^T \nabla_{\theta} \rho_t(\theta^0) \psi_{t-j}(v) \right\| |\tilde{\sigma}_j^{(m,0)}(0, v)| dW(v) \end{aligned}$$

$$\begin{aligned}
 & + \|\sqrt{T}(\hat{\theta} - \theta^0)\|^2 \left[T^{-1} \sum_{t=1}^T \sup_{\theta \in \Theta^0} \|\nabla_{\theta\theta} \rho_t(\theta)\| \right] \sum_{j=1}^{T-1} k^2(j/p) \int |\tilde{\sigma}_j^{(m,0)}(0, v)| dW(v) \\
 & = O_p \left(1 + \frac{p}{T^{1/2}} \right) + O_p \left(\frac{p}{T^{1/2}} \right) = o_p(p^{1/2}).
 \end{aligned}$$

Note that for the first term in (A.9) we have made use of the Cauchy Schwarz, α -mixing condition on $\{\nabla_{\theta} \rho_t(\theta^0) \psi_{t-j}(v)\}$, Assumptions 3 and 6, and the m.d.s property of $\{z_{1t} z_{2t} - \rho_t\}$. In a similar manner, we obtain $\tilde{E}_{2j} = o_p(p^{1/2})$. This concludes the proof. ■

PROOF OF THEOREM 4. This proof has a structure similar to that of Theorem 3. Let \hat{B}_{1q} and \hat{B}_{2q} be as in \hat{B}_1 and \hat{B}_2 , respectively, but with $\{z_{1q,t} z_{2q,t}, \rho_{q,t}\}_{t=1}^T$ in lieu of $\{z_{1t} z_{2t}, \rho_t\}_{t=1}^T$. We will now show that $p^{-1/2} \hat{B}_{1q} \xrightarrow{p} 0$ and $p^{-1/2} \hat{B}_{2q} \xrightarrow{p} 0$.

To advance, we set $\delta_{q,t} \equiv e^{iv'z_t} - e^{iv'z_{q,t}}$ and $\psi_{q,t} \equiv e^{iv'z_{q,t}} - \varphi_q(v)$ with $\varphi_q(v) \equiv E(e^{iv'z_{q,t}})$. Also, let $\tilde{\sigma}_{q,j}^{(m,0)}(0, v)$ be as in $\tilde{\sigma}_j^{(m,0)}(0, v)$ but with $\{z_{1q,t}, z_{2q,t}, \rho_{q,t}\}_{t=1}^T$ in lieu of $\{z_{1t}, z_{2t}, \rho_t\}_{t=1}^T$. As in (A.7), we obtain the following decomposition:

$$\tilde{\sigma}_j^{(m,0)}(0, v) - \tilde{\sigma}_{q,j}^{(m,0)}(0, v) = \sum_{b=1}^{12} \hat{E}_{bjq}(v).$$

Repeating the steps in the proof of Theorem 3 we find that as $q/p \rightarrow 0$ and for $\eta \geq 1$

$$p^{-\frac{1}{2}} \hat{B}_{1q} \leq \sum_{b=1}^{12} \sum_{j=1}^{T-1} k^2(j/p) T_j \int |\hat{E}_{bjq}(v)|^2 dW(v) = O_p(p^{1/2}/q^{\frac{\eta}{2}}) = o_p(1)$$

$$p^{-\frac{1}{2}} \hat{B}_{2q} = 2p^{-\frac{1}{2}} \sum_{b=1}^{12} \sum_{j=1}^{T-1} k^2(j/p) T_j \operatorname{Re} \int \hat{E}_{bjq}(v) \tilde{\sigma}_{q,j}^{(m,0)}(0, v)^* dW(v) = O_p(p^{1/2}/q^{\frac{\eta}{2}}) = o_p(1),$$

where and we have made use of Assumption 2 and the m.d.s property of $\{z_{1t} z_{2t} - \rho_t\}$ and $\{z_{1q,t} z_{2q,t} - \rho_{q,t}\}$. ■

PROOF OF THEOREM 5. Let $\tilde{\sigma}_{q,j}^{(m,0)}(0, v)$ be as $\tilde{\sigma}_j^{(m,0)}(0, v)$, \tilde{C}_{1q} be as \tilde{C}_1 , and \tilde{D}_{1q} be as \tilde{D}_1 but with $\{z_{q,t}\}_{t=1}^T$ and $\{\rho_{q,t}\}_{t=1}^T$ both in lieu of $\{z_t\}_{t=1}^T$ and $\{\rho_t\}_{t=1}^T$. In what follows, we prove the following propositions:

PROPOSITION 3. Under the conditions of Theorem 1,

$$p^{-1/2} \sum_{j=1}^{T-1} k^2(j/p) T_j \int |\tilde{\sigma}_{q,j}^{(m,0)}(0, v)|^2 dW(v) = p^{-1/2} \tilde{C}_{1q} + p^{-1/2} \tilde{V}_q + o_p(1),$$

where

$$\tilde{V}_q = \sum_{t=2q+2}^T [z_{1q,t} z_{2q,t} - \rho_{q,t}] \sum_{j=1}^q a_T(j) \int \psi_{q,t-j}(v) \left[\sum_{s=1}^{t-2q-1} [z_{1q,s} z_{2q,s} - \rho_{q,s}] \psi_{q,s-j}(v)^* \right] dW(v)$$

and $a_T(j) \equiv k^2(j/p) T_j^{-1}$.

PROPOSITION 4. $\tilde{D}_{1q}^{-1/2} \tilde{V}_q \xrightarrow{d} N(0, 1)$.

PROOF OF PROPOSITION 3. Since $\tilde{\sigma}_{q,j}^{(m,0)}(0, v) = T_j^{-1} \sum_{t=j+1}^T [z_{1q,t} z_{2q,t} - \rho_{q,t}] \psi_{q,t-j}(v)$, where $\psi_{q,t-j}(v) = e^{iv' z_{q,t-j}} - \varphi_{q,j}(v)$ with $\varphi_{q,j}(v) = E(e^{iv' z_{q,t-j}})$, then

(A.10)

$$\begin{aligned} \sum_{j=1}^{T-1} k^2(j/p) T_j \int |\tilde{\sigma}_{q,j}^{(m,0)}(0, v)|^2 dW(v) &= \sum_{j=1}^{T-1} a_T(j) \int \left| \sum_{t=1}^T [z_{1q,t} z_{2q,t} - \rho_{q,t}] \psi_{q,t-j}(v) \right|^2 dW(v) \\ &+ \sum_{j=1}^{T-1} a_T(j) \int \left| \sum_{t=1}^j [z_{1q,t} z_{2q,t} - \rho_{q,t}] \psi_{q,t-j}(v) \right|^2 dW(v) \\ &- 2\text{Re} \sum_{j=1}^{T-1} a_T(j) \int \left[\sum_{t=1}^T [z_{1q,t} z_{2q,t} - \rho_{q,t}] \psi_{q,t-j}(v) \right] \\ &\times \left[\sum_{t=1}^j [z_{1q,t} z_{2q,t} - \rho_{q,t}] \psi_{q,t-j}(v) \right]^* dW(v) \\ &\equiv \tilde{R}_{0q} + \tilde{R}_{1q} - 2\text{Re}(\tilde{R}_{2q}). \end{aligned}$$

We continue the decomposition through the term \tilde{R}_{0q} :

(A.11)

$$\begin{aligned} \tilde{R}_{0q} &= \sum_{j=1}^{T-1} a_T(j) \int \sum_{t=1}^T [z_{1q,t} z_{2q,t} - \rho_{q,t}]^2 |\psi_{q,t-j}(v)|^2 dW(v) \\ &+ 2\text{Re} \sum_{j=1}^{T-1} a_T(j) \int \sum_{t=2}^T \sum_{s=1}^{t-1} [z_{1q,t} z_{2q,t} - \rho_{q,t}] [z_{1q,s} z_{2q,s} - \rho_{q,s}] \\ &\times \psi_{q,t-j}(v) \psi_{q,s-j}(v)^* dW(v) \\ &\equiv \tilde{C}_q + 2\text{Re}(\tilde{U}_q). \end{aligned}$$

Also, we have

(A.12)

$$\begin{aligned} \tilde{U}_q &= \sum_{t=2q+2}^T [z_{1q,t} z_{2q,t} - \rho_{q,t}] \int \sum_{j=1}^{T-1} a_T(j) \psi_{q,t-j}(v) \\ &\times \sum_{s=1}^{t-2q-1} [z_{1q,s} z_{2q,s} - \rho_{q,s}] \psi_{q,s-j}(v)^* dW(v) \\ &+ \sum_{t=2}^T [z_{1q,t} z_{2q,t} - \rho_{q,t}] \int \sum_{j=1}^{T-1} a_T(j) \psi_{q,t-j}(v) \\ &\times \sum_{s=\max(1, t-2q)}^{t-1} [z_{1q,s} z_{2q,s} - \rho_{q,s}] \psi_{q,s-j}(v)^* dW(v) \\ &\equiv \tilde{U}_{1q} + \tilde{R}_{3q}. \end{aligned}$$

This simplification allows for the processes $\{z_{q,t}, \rho_{q,t}, \psi_{q,t-j}(v)\}_{j=1}^q$ and $\{z_{q,s}, \rho_{q,s}, \psi_{q,s-j}(v)\}_{j=1}^q$ in \tilde{U}_{1q} to be independent since $s < t - 2q$. In \tilde{R}_{3q} $s \geq t - 2q$. Furthermore, we have

$$\begin{aligned}
 \text{(A.13)} \quad \tilde{U}_{1q} &= \sum_{t=2q+2}^T [z_{1q,t}z_{2q,t} - \rho_{q,t}] \int \sum_{j=1}^{T-1} a_T(j) \psi_{q,t-j}(v) \\
 &\quad \times \sum_{s=1}^{t-2q-1} [z_{1q,s}z_{2q,s} - \rho_{q,s}] \psi_{q,s-j}(v)^* dW(v) \\
 &\quad + \sum_{t=2q+2}^T [z_{1q,t}z_{2q,t} - \rho_{q,t}] \int \sum_{j=1}^{T-1} a_T(j) \psi_{q,t-j}(v) \\
 &\quad \times \sum_{s=1}^{t-2q-1} [z_{1q,s}z_{2q,s} - \rho_{q,s}] \psi_{q,s-j}(v)^* dW(v) \\
 &\equiv \tilde{V}_q + \tilde{R}_{4q}.
 \end{aligned}$$

It follows that

$$p^{-1/2} \sum_{j=1}^{T-1} k^2(j/p) T_j \int |\tilde{\sigma}_{q,j}^{(m,0)}(0, v)|^2 dW(v) = \tilde{C}_q + 2\text{Re}(\tilde{V}_q) + \tilde{R}_{1q} - 2\text{Re}(\tilde{R}_{2q} - \tilde{R}_{3q} - \tilde{R}_{4q}).$$

Assuming $q = p^{1+\frac{1}{4b-2}} (\ln^2 T)^{\frac{1}{2b-1}}$ and $p = CT^\lambda$ for $0 < \lambda < (3 + \frac{1}{4b-2})^{-1}$, to complete the proof we show Lemmas 2 to 6 and conclude $p^{-1/2}[\tilde{C}_q - \tilde{C}_{1q}] = o_p(1)$ and $p^{-1/2}\tilde{R}_{nq} = o_p(1)$ for $n = 1, 2, 3, 4$. ■

LEMMA 2. For \tilde{C}_q as in A.11, $\tilde{C}_q - \tilde{C}_{1q} = O_p(p^2/T)$.

LEMMA 3. For \tilde{R}_{1q} as in A.10, $\tilde{R}_{1q} = O_p(p^2/T)$.

LEMMA 4. For \tilde{R}_{2q} as in A.10, $\tilde{R}_{2q} = O_p(p^{3/2}/T^{1/2})$.

LEMMA 5. For \tilde{R}_{3q} as in A.12, $\tilde{R}_{3q} = O_p(q^{1/2}p/T^{1/2})$.

LEMMA 6. For \tilde{R}_{4q} as in A.13, $\tilde{R}_{4q} = O_p(p^{2b} \ln(T)/q^{2b-1})$.

PROOF OF PROPOSITION 4. Let $\tilde{V}_q = \sum_{t=2q+2}^T V_q(t)$, where

$$V_q(t) \equiv [z_{1q,t}z_{2q,t} - \rho_{q,t}] \int \sum_{j=1}^{T-1} a_T(j) \psi_{q,t-j}(v) G_{j,t-2p-1}(v) dW(v),$$

and $G_{j,t-2p-1}(v) \equiv \sum_{s=1}^{t-2q-1} [z_{1q,s}z_{2q,s} - \rho_{q,s}] \psi_{q,s-j}(v)^*$. To derive asymptotic normality, we employ Brown's (1971) central limit theorem for martingale processes, which guarantees that $\text{var}(2\text{Re}\tilde{V}_q)^{-1/2} 2\text{Re}\tilde{V}_q \xrightarrow{d} N(0, 1)$ under

Condition 1: $\text{var}(2\text{Re}\tilde{V}_q)^{-1} \sum_{t=1}^T [2\text{Re}\tilde{V}_q]^2 \mathbb{I}[|2\text{Re}\tilde{V}_q| > \nu \text{var}(2\text{Re}\tilde{V}_q)^{1/2}] \rightarrow 0 \forall \nu > 0,$

Condition 2: $\text{var}(2\text{Re}\tilde{V}_q)^{-1} \sum_{t=1}^T E[2\text{Re}\tilde{V}_q^2 | I_{t-1}] \xrightarrow{p} 1.$

We first establish the existence of $\text{var}(2\text{Re}\tilde{V}_q)$. Using the m.d.s property of $\{z_{1q,t}z_{2q,t} - \rho_{q,t}, I_{t-1}\}$ under the null hypothesis and the q -dependent assumption on $\{z_{1q,t}z_{2q,t} - \rho_{q,t}\}$, we obtain

$$\begin{aligned} E(\tilde{V}_q^2) &= \sum_{t=2q+2}^T E \left[[z_{1q,t}z_{2q,t} - \rho_{q,t}]^2 \right. \\ &\quad \times \left. \left(\int \sum_{j=1}^q a_T(j) \psi_{q,t-j}(v) \sum_{s=1}^{t-2q-1} [z_{1q,s}z_{2q,s} - \rho_{q,s}] \psi_{q,s-j}(v)^* dW(v) \right)^2 \right] \\ &= \sum_{j=1}^q \sum_{l=1}^q a_T(j) a_T(l) \int \int \sum_{t=2q+2}^T \sum_{s=1}^{t-2q-1} E\{[z_{1q,t}z_{2q,t} - \rho_{q,t}]^2 \cdot \psi_{q,t-j}(v) \psi_{q,t-l}(v')\} \\ &\quad \times E\{[z_{1q,s}z_{2q,s} - \rho_{q,s}]^2 \cdot \psi_{q,s-j}^*(v) \psi_{q,s-l}^*(v')\} dW(v) dW(v') \\ &= \frac{1}{2} \sum_{j=1}^q \sum_{l=1}^q k^2(j/p) k^2(l/p) \int \int |E\{[z_{1q,0}z_{2q,0} - \rho_{q,0}]^2 \psi_{q,-j}(v) \psi_{q,-l}(v')\}|^2 \\ &\quad dW(v) dW(v') [1 + o(1)]. \end{aligned}$$

In the same manner, we deduce

$$\begin{aligned} E(\tilde{V}_q^*)^2 &= \frac{1}{2} \sum_{j=1}^q \sum_{l=1}^q k^2(j/p) k^2(l/p) \\ &\quad \times \int \int |E\{[z_{1q,0}z_{2q,0} - \rho_{q,0}]^2 \psi_{q,-j}(v) \psi_{q,-l}(v')\}|^2 dW(v) dW(v') [1 + o(1)], \\ E|\tilde{V}_q|^2 &= \frac{1}{2} \sum_{j=1}^q \sum_{l=1}^q k^2(j/p) k^2(l/p) \\ &\quad \times \int \int |E\{[z_{1q,0}z_{2q,0} - \rho_{q,0}]^2 \psi_{q,-j}(v) \psi_{q,-l}(v')\}|^2 dW(v) dW(v') [1 + o(1)]. \end{aligned}$$

Under Assumption 2, we obtain $E(\tilde{V}_q^2) = E(\tilde{V}_q^*)^2 = E|\tilde{V}_q|^2$. Then

$$\begin{aligned} \text{var}(2\text{Re}\tilde{V}_q) &= E(\tilde{V}_q^2) + E(\tilde{V}_q^*)^2 + 2E|\tilde{V}_q|^2 \\ &= 2 \sum_{j=1}^q \sum_{l=1}^q k^2(j/p) k^2(l/p) \\ &\quad \times \int \int |E\{[z_{1,0}z_{2,0} - \rho_0]^2 \psi_{-j}(v) \psi_{-l}(v')\}|^2 dW(v) dW(v') [1 + o(1)]. \end{aligned}$$

Note the convergence $E\{[z_{1,0}z_{2,0} - \rho_{q,0}]^2 \psi_{q,-j}(v) \psi_{q,-l}(v')\} \rightarrow E\{[z_{1,0}z_{2,0} - \rho_0]^2 \psi_{-j}(v) \psi_{-l}(v')\}$ as $q \rightarrow \infty$ follows from Assumption 2. For further simplification, we apply the α -mixing

condition in Assumption 5 to $\{[z_{1,0}z_{2,0} - \rho_0]^2 \psi_{-j}(v)\psi_{-l}(v')\}$. First, let $\Omega_0 \equiv E[z_{1,0}z_{2,0} - \rho_0]^2$ and set $\eta_{j,l}(v) \equiv E\{[(z_{1,0}z_{2,0} - \rho_0)^2 - \Omega_0] \psi_{-j}(v)\psi_{-l}(v')\}$. Then by virtue of Assumption 5 we have $|\eta_{j,l}(v)|^2 \leq C\alpha(l-j)^{v-1/v}$, $\sum_{j,l=1}^q |\eta_{j,l}(v)|^2 \leq C$. Under $|k(\cdot)| \leq 1$, a change of variable, and the convergence criterion $p^{-1} \sum_{j=m+1}^q k^2(j/p)k^2((j-m)/p) \rightarrow \int_0^\infty k^4(z)dz$ with $p \rightarrow \infty$ and $q/p \rightarrow 0$, we deduce the following:

$$\begin{aligned} \text{var}(2\text{Re}\tilde{V}_q) &= 2 \sum_{j=1}^q \sum_{l=1}^q k^2(j/p)k^2(l/p) \int \int \Omega_0^2 |\sigma_{l-j}(v, v')|^2 dW(v) dW(v') [1 + o(1)] \\ &= 2p \sum_{m=1-q}^{q-1} \left[p^{-1} \sum_{j=m+1}^q k^2(j/p)k^2((j-m)/p) \right] \\ &\quad \times \Omega_0^2 \int \int |\sigma_m(v, v')|^2 dW(v) dW(v') [1 + o(1)] \\ &= 4\pi p \int_0^\infty k^4(z) dz \Omega_0^2 \int \int \int_{-\pi}^\pi |f(\omega, v, v')|^2 d\omega dW(v) dW(v') [1 + o(1)]. \end{aligned}$$

We now show that condition 1 holds. Using the m.d.s. property of $\{z_{1q,t}, z_{2q,t} - \rho_{q,t}, I_{t-1}\}$ and Rosenthal’s inequality, we find $E |G_{j,t-2p-1}(v)|^4 \leq Ct^2$ for $1 \leq j \leq q$. By virtue of Minkowski’s inequality, we deduce

$$\begin{aligned} E|V_q(t)|^4 &\leq \left[\sum_{j=1}^q a_T(j) \int \left(E|[z_{1q,t}z_{2q,t} - \rho_{q,t}] \psi_{q,t-j}(v)G_{j,t-2p-1}(v)|^4 \right)^{1/4} dW(v) \right]^4 \\ &\leq Ct^2 \left[\sum_{j=1}^q a_T(j) \right]^4 = O(p^4 t^2 / T^4). \end{aligned}$$

Then for $p^2/T \rightarrow 0$, we have $\sum_{t=2q+2}^T E|\tilde{V}_q^2(t)|^4 = o(p^2)$. This shows that condition 1 of Brown’s (1971) theorem is satisfied.

It remains to show condition 2 is also satisfied. Define $\rho_{z,q,t} \equiv E\{[z_{1q,t}z_{2q,t} - \rho_{q,t}]^2 | I_{t-1}\}$ and $\tilde{H}_{q,t}^{j,l}(v, v') \equiv \rho_{z,q,t} \psi_{q,t-j}(v)\psi_{q,t-l}(v') - E[\rho_{z,q,t} \psi_{q,t-j}(v)\psi_{q,t-l}(v')]$. Then

$$\begin{aligned} \text{(A.14)} \quad E[\tilde{V}_q^2(t)] &= \rho_{z,q,t} \left[\sum_{j=1}^q a_T(j) \int \psi_{q,t-j}(v)G_{j,t-2p-1}(v) \right]^2 \\ &= \sum_{j=1}^q \sum_{l=1}^q a_T(j)a_T(l) \int \int \rho_{z,q,t} \psi_{q,t-j}(v)\psi_{q,t-l}(v') \\ &\quad \times G_{j,t-2p-1}(v)G_{l,t-2p-1}(v') dW(v) dW(v') \end{aligned}$$

$$\begin{aligned}
 &= \sum_{j=1}^q \sum_{l=1}^q a_T(j)a_T(l) \int \int E[\rho_{z,q,t}\psi_{q,t-j}(v)\psi_{q,t-l}(v')] \\
 &\quad \times G_{j,t-2p-1}(v)G_{l,t-2p-1}(v') dW(v) dW(v') \\
 &\quad + \sum_{j=1}^q \sum_{l=1}^q a_T(j)a_T(l) \int \int \tilde{H}_{q,t}^{j,l}(v, v')G_{j,t-2p-1}(v)G_{l,t-2p-1}(v') dW(v) dW(v') \\
 &\equiv S_{1q}(t) + V_{1q}(t).
 \end{aligned}$$

In a similar manner, we obtain

$$\begin{aligned}
 S_{1q}(t) &= \sum_{j=1}^q \sum_{l=1}^q a_T(j)a_T(l) \int \int E[\rho_{z,q,t}\psi_{q,t-j}(v)\psi_{q,t-l}(v')] \\
 &\quad \times E[G_{j,t-2p-1}(v)G_{l,t-2p-1}(v')]dW(v) dW(v') \\
 &\quad + \sum_{j=1}^q \sum_{l=1}^q a_T(j)a_T(l) \int \int E[\rho_{z,q,t}\psi_{q,t-j}(v)\psi_{q,t-l}(v')] \\
 &\quad \times \{G_{j,t-2p-1}(v)G_{l,t-2p-1}(v') - E[G_{j,t-2p-1}(v)G_{l,t-2p-1}(v')]\}dW(v) dW(v') \\
 &\equiv E[V_q^2(t)] + S_{2q}(t).
 \end{aligned}$$

We note that $E[V_q^2(t)]$ can be simplified to give

$$E[V_q^2(t)] = \sum_{j=1}^q \sum_{l=1}^q (t - q - 1)a_T(j)a_T(l) \int |E[\rho_{z,q,t}\psi_{q,t-j}(v)\psi_{q,t-l}(v')]|dW(v) dW(v').$$

Let $H_{q,s}^{j,l}(v, v') \equiv (z_{1q,s}z_{2q,s} - \rho_{q,s})^2\psi_{q,s-j}(v)\psi_{q,s-l}(v') - E[(z_{1q,s}z_{2q,s} - \rho_{q,s})^2\psi_{q,s-j}(v)\psi_{q,s-l}(v')]$. We decompose $S_{2q}(t)$ as follows:

(A.15)

$$\begin{aligned}
 S_{2q}(t) &= \sum_{j=1}^q \sum_{l=1}^q a_T(j)a_T(l) \int E[(z_{1q,t}z_{2q,t} - \rho_{q,t})^2\psi_{q,t-j}(v)\psi_{q,t-l}(v')] \\
 &\quad \times \sum_{s=1}^{t-2q-1} H_{q,s}^{j,l}(v, v') dW(v) dW(v') \\
 &\quad + \sum_{j=1}^q \sum_{l=1}^q a_T(j)a_T(l) \int E[(z_{1q,t}z_{2q,t} - \rho_{q,t})^2\psi_{q,t-j}(v)\psi_{q,t-l}(v')] \\
 &\quad \times \sum_{s=2}^{t-2q-1} \sum_{\tau=1}^{s-1} (z_{1q,s}z_{2q,s} - \rho_{q,s})\psi_{q,s-j}(v)(z_{1q,\tau}z_{2q,\tau} - \rho_{q,\tau})\psi_{q,\tau-l}(v') dW(v) dW(v') \\
 &\equiv V_{2q}(t) + S_{3q}(t).
 \end{aligned}$$

Finally, we have

$$\begin{aligned}
 \text{(A.16)} \quad S_{3q}(t) &= \sum_{j=1}^q \sum_{l=1}^q a_T(j) a_T(l) \int E[(z_{1q,t} z_{2q,t} - \rho_{q,t})^2 \psi_{q,t-j}(v) \psi_{q,t-l}(v')] \\
 &\quad \times \sum_{0 < s-\tau \leq 2q} \sum_{0 < s-\tau \leq 2q} (z_{1q,s} z_{2q,s} - \rho_{q,s}) \psi_{q,s-j}(v) (z_{1q,\tau} z_{2q,\tau} - \rho_{q,\tau}) \psi_{q,\tau-l}(v') dW(v) dW(v') \\
 &\quad + \sum_{j=1}^q \sum_{l=1}^q a_T(j) a_T(l) \int E[(z_{1q,t} z_{2q,t} - \rho_{q,t})^2 \psi_{q,t-j}(v) \psi_{q,t-l}(v')] \\
 &\quad \times \sum_{s-\tau > 2q} \sum_{s-\tau > 2q} (z_{1q,s} z_{2q,s} - \rho_{q,s}) \psi_{q,s-j}(v) (z_{1q,\tau} z_{2q,\tau} - \rho_{q,\tau}) \psi_{q,\tau-l}(v') dW(v) dW(v') \\
 &\equiv V_{3q}(t) + V_{4q}(t).
 \end{aligned}$$

Combining the above equations, we see that $\sum_{t=2q+2}^T \{E[V_q^2(t) | I_{t-1}] - E[V_q^2(t)]\} = \sum_{b=1}^4 \sum_{s=1}^{t-2q-1} V_{bq}(t)$. To complete the proof, we show that the conditions of the foregoing Lemmas 7 to 10 are satisfied. These naturally produce the result $E |\sum_{t=2q+2}^T \{E[V_q^2(t) | I_{t-1}] - E[V_q^2(t)]\}|^2 = o(p^2)$ for $q = p^{1+\frac{1}{4b-2}} (\ln^2 T)^{\frac{1}{2b-1}}$ and $p = CT^\lambda$ where $0 < \lambda < (3 + \frac{1}{4b-2})^{-1}$. Consequently, condition 2 of Brown's (1971) central limit theorem holds. ■

LEMMA 7. Given $V_{1q}(t)$ as in (A.14), then $E |\sum_{t=2q+2}^T V_{1q}(t)|^2 = O(qp^4/T)$.

LEMMA 8. Given $V_{2q}(t)$ as in (A.15), then $E |\sum_{t=2q+2}^T V_{2q}(t)|^2 = O(qp^4/T)$.

LEMMA 9. Given $V_{3q}(t)$ as in (A.16), then $E |\sum_{t=2q+2}^T V_{3q}(t)|^2 = O(qp^4/T)$.

LEMMA 10. Given $V_{4q}(t)$ as in (A.16), then $E |\sum_{t=2q+2}^T V_{4q}(t)|^2 = O(p)$.

REMARK 1. We emphasize the following relations and bounds that will be used to prove the above lemmas. For $j, l \in \{1, 2, \dots, q\}$ and an arbitrary constant C ,

- (1) $\tilde{H}_{q,t}^{j,l}(v, v')$ is independent of $\{G_{j,t-2p-1}(v) G_{l,t-2p-1}(v')\}$,
- (2) $\tilde{H}_{q,t}^{j,l}(v, v')$ is independent of $\tilde{H}_{q,\tau}^{j,l}(v, v')$ for $|t - \tau| > 2q$,
- (3) $E |G_{j,t-2p-1}(v)|^4 \leq Ct^2$.
- (4) $H_{q,s}^{j,l}(v, v')$ is independent of $H_{q,\tau}^{j,l}(v, v')$ for $|s - \tau| > 2q$,
- (5) $E |\sum_{s=1}^{t-q-1} H_{q,s}^{j,l}(v, v')|^2 = \sum_{|s-\tau| \leq 2q} \sum_{|s-\tau| \leq 2q} E[H_{q,s}^{j,l}(v, v') H_{q,\tau}^{j,l}(v, v')] \leq Ctq$,
- (6) $[\sum_{j=1}^q a_T(j)]^4 = O(p^4/T^4)$.

PROOF OF THEOREM 2. This proof is separated into Theorems 6 and 7, which we state and prove later. ■

THEOREM 6. Under the regularity conditions of Theorem 2, $p^{1/2}/T [\hat{Q}_1 - Q_1] \xrightarrow{p} 0$.

THEOREM 7. Under the regularity conditions of Theorem 2,

$$p^{1/2}/T Q_1 \xrightarrow{p} 0 \frac{1}{D_1^{1/2}} \int \int_{-\pi}^{\pi} |f^{(0,m,0)}(\omega, 0, v) - f_0^{(0,m,0)}(\omega, 0, v)|^2 d\omega dW(v).$$

PROOF OF THEOREM 6. The three main components of this proof are

$$(A.17) \quad (i) \ T^{-1} \int \sum_{j=1}^{T-1} k^2(j/p) T_j [|\hat{\sigma}_j^{(m,0)}(0, v)|^2 - |\tilde{\sigma}_j^{(m,0)}(0, v)|^2] dW(v) \xrightarrow{p} 0,$$

(ii) $p^{-1}[\hat{C}_1 - \tilde{C}_1] = O_p(1)$ and (iii) $p^{-1}[\hat{D}_1 - \tilde{D}_1] \xrightarrow{p} 0$, where \tilde{C}_1 and \tilde{D}_1 are as in \hat{C}_1 and \hat{D}_1 but with $\{z_t, \rho_t\}_{t=1}^T$ in lieu of $\{\hat{z}_t, \hat{\rho}_t\}_{t=1}^T$. Here, we will only prove (A.17), since parts (ii) and (iii) of this proof are straightforward. Following the proof of Theorem 1, we decompose the left-hand side of (A.17) so that

$$T^{-1} \int \sum_{j=1}^{T-1} k^2(j/p) T_j [|\hat{\sigma}_j^{(m,0)}(0, v)|^2 - |\tilde{\sigma}_j^{(m,0)}(0, v)|^2] dW(v) = \hat{B}_1 + 2Re\hat{B}_2,$$

where \hat{B}_1 and \hat{B}_2 are as defined in (A.6). From Theorem 7, we see that under the alternative hypothesis, $T^{-1} \int \sum_{j=1}^{T-1} k^2(j/p) T_j |\hat{\sigma}_j^{(m,0)}(0, v)|^2 dW(v) = O_p(1)$. Then, applying the Cauchy-Schwarz inequality to \hat{B}_2 , we see that showing $p^{-1}\hat{B}_1 \xrightarrow{p} 0$ will be sufficient. To begin, we retain our previous decomposition of \hat{B}_1 in (A.7). The steps for this proof are identical to those employed in Lemma 1, with the exception that although $\{z_t\}$ is not m.d.s. under the alternative we still have $E(\sum_{t=j+1}^T z_{1t} z_{2t}) \leq CT_j$. Thus, it is straightforward to show the condition $T^{-1} \int \sum_{j=1}^{T-1} k^2(j/p) T_j |\hat{E}_{bj}|^2 dW(v) \xrightarrow{p} 0$ for $b \in \{1, 2, \dots, 12\}$. ■

PROOF OF THEOREM 7. With slight modifications, this proof is easily obtained from Hong (1999, Proof of Theorem 5). ■

REFERENCES

ANDREWS, D., "Generic Uniform Convergence," *Econometric Theory* 8 (1992), 214–57.
 BAUWENS, L., AND S. LAURENT, "A New Class of Multivariate Skew Densities, with Application to Generalized Autoregressive Conditional Heteroscedasticity Models," *Journal of Business and Economic Statistics* 23 (2005), 346–54.
 —, —, AND J. V. K. ROMBOUTS, "Multivariate GARCH Models: A Survey," *Journal of Applied Econometrics* 21 (2006), 79–109.
 BERA, A. K., AND S. KIM, "Testing Constancy of Correlation and Other Specifications of the BGARCH Model with an Application to International Equity Returns," *Journal of Empirical Finance* 9 (2002), 171–95.
 BIERENS, H. J., "Consistent Model Specification Tests," *Journal of Econometrics* 20 (1982), 105–34.
 BILLIO, M., AND M. CAPORIN, "A Generalized Dynamic Conditional Correlation Model for Portfolio Risk Evaluation," *Mathematics and Computers in Simulation* 79 (2009), 2566–78.
 —, —, AND M. GOBBO, "Flexible Dynamic Conditional Correlation Multivariate GARCH Models for Asset Allocation," *Applied Financial Economics Letters* 2 (2006), 123–30.
 BOLLERSLEV, T., "Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized Arch Model," *The Review of Economics and Statistics* 72 (1990), 498–505.
 —, R. F. ENGLE, AND J. M. WOOLDRIDGE, "A Capital Asset Pricing Model with Time-Varying Covariances," *The Journal of Political Economy* 96 (1988), 116–31.
 BROWN, B., "Martingale Central Limit Theorems," *The Annals of Mathematical Statistics* 42 (1971), 59–66.
 CAPPIELLO, L., R. F. ENGLE, AND K. SHEPPARD, "Asymmetric Dynamics in the Correlations of Global Equity and Bond Returns," *Journal of Financial Econometrics* 4 (2006), 537–72.
 DAHLHAUS, R., "Fitting Time Series Models to Nonstationary Processes," *The Annals of Statistics* 25 (1997), 1–37.
 DE SANTIS, G., AND B. GERARD, "International Asset Pricing and Portfolio Diversification with Time-Varying Risk," *Journal of Finance* 52 (1997), 1881–912.
 ENGLE, R. F., "Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models," *Journal of Business and Economic Statistics* 20 (2002), 339–50.

- , AND R. COLACITO, “Testing and Valuing Dynamic Correlations for Asset Allocation,” *Journal of Business and Economic Statistics* 24 (2006), 238–53.
- , AND K. F. KRONER, “Multivariate Simultaneous Generalized ARCH,” *Econometric Theory* 11 (1995), 122–50.
- , AND K. SHEPPARD, “Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH,” Working Paper, University of California, San Diego, 2001.
- FENG, Y., “A Local Dynamic Conditional Correlation Model,” Munich Personal RePEc Archive, Working Paper No. 1592, Maxwell Institute for Mathematical Sciences, Heriot-Watt University, 2006.
- HAFNER, C. M., AND P. H. FRANSES, “A Generalised Dynamic Conditional Correlation Model: Simulation and Application to Many Assets,” *Econometric Reviews* 28 (2009), 612–31.
- HARVEY, C., AND A. SIDDIQUE, “Autoregressive Conditional Skewness,” *The Journal of Financial and Quantitative Analysis* 34 (1999), 465–87.
- AND —, “Conditional Skewness in Asset Pricing Tests,” *The Journal of Finance* 55 (2000), 1263–95.
- HE, C., AND T. TERÄSVIRTA, “An Extended Constant Conditional Correlation GARCH model and its Fourth-Moment Structure,” *Econometric Theory* 20 (2004), 904–26.
- HONG, Y., “Hypothesis Testing in Time Series via the Empirical Characteristic Spectral Density Approach,” *Journal of the American Statistical Association* 94 (1999), 1201–20.
- , “A Test for Volatility Spillover with Application to Exchange Rates,” *Journal of Econometrics* 103 (2001), 183–224.
- JEANTHEAU, T., “Strong Consistency of Estimations of Multivariate ARCH Model,” *Econometric Theory* 14 (1998), 70–86.
- KING, M. A., AND S. WADHWANI, “Transmission of Volatility between Stock Markets,” *The Review of Financial Studies* 3 (1990), 5–33.
- KOENKER R., “A Note on Studentizing a Test for Heteroscedasticity,” *Journal of Econometrics* 17 (1981), 107–12.
- LEE, S-W., AND B. HANSEN, “Asymptotic Theory for the GARCH(1,1) Quasi-Maximum Likelihood Estimator,” *Econometric Theory* 10 (1994), 29–52.
- LEE, T-H., AND X. LONG, “Copula Based Multivariate GARCH Model with Uncorrelated Dependent Errors,” *Journal of Econometrics* 150 (2009), 207–18.
- LI, W. K., AND T. K. MAK, “On the Squared Residual Autocorrelations in Non-linear Time Series with Conditional Heteroscedasticity,” *Journal of Time Series Analysis* 15 (1994), 627–36.
- LIN, W.-L., R. F. ENGLE, AND T. ITO, “Do Bulls and Bears Move across Borders? International Transmission of Stock Returns and Volatility,” *The Review of Financial Studies* 7 (1994), 507–38.
- LING, T., AND M. MCALEER, “Asymptotic Theory for a Vector ARMA-GARCH Model,” *Econometric Theory* 19 (2003), 280–310.
- LONGIN, F., AND B. SOLNIK, “Is the Correlation in International Equity Returns Constant: 1960-1990?” *Journal of International Money and Finance* 14 (1995), 3–26.
- , AND —, “Extreme Correlation of International Equity Markets,” *Journal of Finance* 56 (2001), 649–76.
- LUMSDAINE, R., “Consistency and Asymptotic Normality of the Quasi-Maximum Likelihood Estimator in IGARCH(1,1) and Covariance Stationary GARCH(1,1) Models,” *Econometrica* 64 (1996), 575–96.
- MCALEER, M., F. CHAN, S. HOTI, AND O. LIEBERMAN, “Structure and Asymptotic Theory for Multivariate Asymmetric Conditional Volatility,” *Econometric Reviews* 28 (2009), 422–40.
- , —, AND —, “Generalized Autoregressive Conditional Correlation,” *Econometric Theory* 24 (2008), 1554–83.
- NAKATANI, T., AND T. TERÄSVIRTA, “Testing for Volatility Interactions in the Constant Conditional Correlation GARCH Model,” *Econometrics Journal* 12 (2009), 147–63.
- NELSON, D., “Stationary and Persistence in the GARCH(1,1) Model,” *Econometric Theory* 6 (1990), 318–34.
- PATTON, A., “Modelling Asymmetric Exchange Rate Dependence,” *International Economic Review* 47 (2006), 527–56.
- PELAGATTI, M., AND S. RONDENA, “Dynamic Conditional Correlation with Elliptical Distributions,” Working Papers 20060508, Università degli Studi di Milano-Bicocca, Dipartimento di Statistica, 2006.
- PELLETIER, D., “Regime Switching for Dynamic Correlations,” *Journal of Econometrics* 131 (2006), 445–73.
- SILVENNÖINEN, A., AND T. TERÄSVIRTA, “Multivariate Autoregressive Conditional Heteroskedasticity with Smooth Transitions in Conditional Correlations,” SSE/EFI Working Paper Series in Economics and Finance No. 577, Stockholm School of Economics, 2005.
- , AND —, “Modelling Multivariate Autoregressive Conditional Heteroskedasticity with the Double Smooth Transition Conditional Correlation GARCH Model,” *Journal of Financial Econometrics* 7 (2009a), 373–411.
- , AND —, “Multivariate GARCH Models,” in T.G. Andersen, R.A. Davis, J.-P. Kreiss, and T. Mikosch, eds., *Handbook of Financial Time Series* (New York: Springer, 2009b).

- STINCHCOMBE, M., AND H. WHITE, "Consistent Specification Testing with Nuisance Parameters Present Only Under the Alternative," *Econometric Theory* 14 (1998), 295–325.
- TSE, Y. K., "A Test for Constant Correlations in a Multivariate GARCH Model," *Journal of Econometrics* 98 (2000), 107–27.
- TSE, Y., AND A. TSUI, "A Multivariate GARCH Model with Time-Varying Correlations," *Journal of Business and Economic Statistics* 20 (2002), 351–62.
- WHITE, H., *Asymptotic Theory for Econometricians: Revised Edition* (San Diego, CA: Academic Press, 2000).
- WOOLDRIDGE, J., "A Unified Approach to Robust, Regression-based Specification Tests," *Econometric Theory* 6 (1990), 17–43.