Specification Testing for Multivariate Time Series Volatility Models

Yongmiao Hong and Yoon-Jin Lee

Department of Economics &
Department of Statistical Science
Cornell University
Ithaca, NY 14853-7601
Emails: yh20@cornell.edu, yl254@cornell.edu

First version, February 2004

We thank Tim Bollerslev for the suggestion of this topics, and Tae-Hwy Lee for helpful comments and discussions.

ABSTRACT

Volatility models have been playing an important role in economics and finance. Using a multivariate generalized spectral approach, we propose a new class of generally applicable omnibus tests for univariate and multivariate volatility models. Both GARCH models and stochastic volatility models are covered. Our tests have a convenient asymptotic null N(0,1) distribution, and can detect a wide range of misspecifications for volatility dynamics. Distinct from the existing tests for volatility models, our tests are robust to higher order time-varying moments of unknown form (e.g., time-varying skewness and kurtosis). Our tests check a large number of lags and are therefore expected to be powerful against neglected volatility dynamics that occurs at higher order lags or display long memory properties. Despite using a large number of lags, our tests do not suffer much from loss of a large number of degrees of freedom, because our approach naturally discounts higher order lags, which is consistent with the stylized fact that economic or financial markets are more affected by the recent past events than by the remote past events. No specific estimation method is required, and parameter estimation uncertainty has no impact on the limit distribution of the test statistics. Moreover, there is no need to formulate an alternative volatility model, and only estimated standardized residuals are needed to implement our tests. We do not have to calculate tedious score functions or derivatives of volatility models with respect to estimated parameters, which are model-specific and are required in some existing popular tests for volatility models. We examine the finite sample performance of the proposed tests. An empirical application to some popular GARCH models for stock returns illustrates our approach.

Key Words: Generalized spectral derivative, Kernel, Multivariate generalized spectrum, Multivariate GARCH models, Nonlinear volatility dynamics, Robustness, Specification testing, Stochastic Volatility Model, Time-varying higher order moments of unknown form.

JEL NO: C4, C2

1. Introduction

Volatility is one of the most important instruments in economics and finance. Volatility modeling and forecasting is important in investment, security valuation, risk management and monetary policy making. As a measure for uncertainty, volatility is a key input to many investment decisions and portfolio creations. And it is crucially important in asset pricing. According to most asset pricing theories, risk premium is determined by the conditional covariance between the future return on the asset and one or more benchmark portfolios (e.g., the market portfolio or the growth rate in consumption). Volatility is also important in pricing derivative securities, where the uncertainty associated with the future price of the underlying asset is the most important determinant for derivative prices. On the other hand, an important source of volatility clustering is information flows arriving in a cluster manner. One can investigate how financial markets interact with each other by examining volatility spillover among different markets. Moreover, the understanding of the volatility transmission mechanism between asset prices and GDP growths is important for policy makers to reduce output volatility. Policy makers often rely on market estimates of volatility as a barometer for the vulnerability of financial markets and economy.

Since Engle's (1982) seminal paper, traditional time series tools such as autoregressive moving average (ARMA) models for the conditional mean have been extended to essentially analogous models for the conditional variance. Autoregressive conditional heteroskedasticity (ARCH) models are now commonly used to capture volatility dynamics of financial time series. This class includes the ARCH and GARCH models of Engle (1982) and Bollerslev (1987), as well as their various nonlinear generalizations (e.g., Bera and Higgins' (1992) nonlinear GARCH models, Nelson's (1991) EGARCH model, Glosten *et al*'s (1993) threshold GARCH model, Sentana's (1995) quadratic GARCH model, Zakoian's (1994) threshold ARCH model, to name just a few). For a survey of ARCH models, see Bollerslev, Chou, and Kroner (1992), Bera and Higgins (1993), Bollerslev, Engle and Nelson (1994) among others.

The empirical success of ARCH models in fitting univariate time series has motivated many researchers to extend these models to multivariate contexts. It is a stylized fact that financial volatilities move together over time across assets and markets. Recognizing this feature through multivariate modeling should lead to more relevant empirical models and deeper insights into financial markets than working with separate univariate models. Apart from possible efficiency gains in parameter estimation, estimation of some financial "coefficients" such as the systematic risk (beta coefficients) and the hedge ratio, requires estimating covariances between relevant variables. The motivation for multivariate GARCH models also stems from the fact that many economic variables react to the same information, and hence, have nonzero covariances conditional on the information set available. From a financial point of view, multivariate GARCH modeling opens the doors to better decision tools in various areas such as asset pricing, portfolio selection, hedging, and Value-at-Risk forecasts. Although there is a huge literature on univariate models for volatility dynamics, asymmetry and fat-tails, much fewer works are concerned with their multivariate extensions. Important examples of multivariate volatility models are diagonal multivariate and vech-representation GARCH models of Bollerslev, Engle and Wooldridge (1988), the constant-correlation multivariate GARCH (CC-MGARCH) models of Bollerslev (1990), the BEKK (named after Baba, Engle, Kraft and Kroner) models of Engle and

Kroner (1995) and the dynamic conditional correlation (DCC) models of Engle (2002) and Tse and Tsui (2002).

A popular class of models alternative to GARCH models in capturing volatility clustering is the stochastic volatility (SV) models introduced by Taylor (1982, 1986). Unlike GARCH models, SV models assume that the volatility process is driven by an unobservable information flow. They are also closely related to continuous-time diffusion processes which are widely used in derivatives pricing and other financial applications (e.g., Hull and White (1987)). See Gyhsels (1996) and Shephard (1996) for excellent surveys on SV models and their applications.

Consistent parameter estimation, optimal volatility forecast, valid hypothesis testing and economic interpretations all require correct specification of volatility models. For example, the systematic risk as measured by the beta coefficient depends on the conditional second moments of asset returns, so does the minimum-variance hedge ratio. Reliable estimates and inference of these quantities depend on well-defined conditional heteroskedasticity models. There have not been rigorously developed specification tests for SV models in the literature. There have been a number of specification tests for GARCH models. Diagnostic tests for GARCH models in the literature can be divided into three categories: portmanteau tests of the Box-Pierce-Ljung type, Lagrange multiplier (LM) tests, and residual-based diagnostics.

The portmanteau tests of Box-Pierce type for the squared standardized residuals of a univariate GARCH model have been used to test adequacy of the GARCH model. They have been also used as the benchmark for detecting inadequacy of multivariate GARCH models. As these test statistics are readily computable from the standardized residuals of a GARCH model, they have been widely used in practice (e.g., Tsay 2001, p.115-118). Often the asymptotic χ^2 distribution is used. However, Li and Mak (1994) showed that the Box-Pierce type tests for volatility models are generally not asymptotically χ^2, because the limit distribution depends on parameter estimation uncertainty in volatility models. In other words, substituting the estimated standardized residuals for the unobserved innovations will change the asymptotic distribution of the test statistic. Li and Mak (1994) modified the Box-Pierce type tests and derived the asymptotic distribution of their modified tests for univariate volatility models. Ling and Li (1997) further extended this work and derived the asymptotic distribution of a modified portmanteau statistic for multivariate volatility models. Ling and Li's test is based on the sum of the squared autocorrelations of suitably transformed residuals. However, Tse and Tsui (1999) pointed out that there is a loss of information in the transformation of the estimated residuals, which may induce severe loss of power.

There have been a number of LM tests for GARCH models, as considered in Bollerslev *et al.* (1988), Engle and Ng (1993), Engle and Kroner (1995) and Lundbergh and Teräsvirta (1998). Lundbergh and Teräsvirta's (1998) LM test is a test of the standardized errors being *i.i.d.* against the alternative that they follow an ARCH model. This test is asymptotically equivalent to Li and Mak's (1994) test. The LM test has an advantage over the portmanteau tests due to its efficiency when the alternative hypothesis is correct. However, it requires the specification of an alternative GARCH model, and the calculation of a LM test statistic depends on the alternative.

Tse (2002) proposed residual-based diagnostic tests for GARCH models. These tests resort to a convenient auxiliary autoregression based on the squared standardized residuals or the cross products of the standardized estimated residuals as dependent variables, while lagged squared standardized residuals or lagged cross products of the standardized residuals are the independent variables. Thus, the form of the regression depends on a particular type of model inadequacy the researcher likes to investigate, which dictates the power of the tests.

From a theoretical point of view, Box-Pierce type tests and residual-based tests for GARCH models can detect many misspecifications in volatility dynamics of practical importance. However, they can only capture linear ARCH alternatives, and may miss important nonlinear volatility dynamics, especially those with zero autocorrelation in standardized residuals. They may overlook certain volatility dynamics, such as asymmetric behaviors in volatility. Asymmetric volatility dynamics are not uncommon in practice. They can be caused by (e.g.) "leverage effects", or by business cycles (Hamilton and Lin (1996)). We note that LM tests can detect some specific nonlinear volatility features, depending on the formulation of the alternative model (see Engle and Ng 1993).

Most existing tests for GARCH models usually check a fixed lag order. Recent empirical studies (e.g., Baillie, Bollerslev and Mikkelsen 1996) find that high-frequency financial time series may display long memory of financial time series in volatility clustering, where volatility depends on a very long past history. Indeed, it is an important feature of a non-Markovian process that volatility may depend on the entire past history rather than only first few lags of it. Thus, it is important to check not only the functional forms of volatility dynamics but also its lag structure.

A volatility model with *i.i.d.* innovations is called a strong form volatility model in the literature (cf. Drost and Nijman 1993). It is possible that a volatility model is correctly specified while the standardized innovation displays higher order dependence possibly of unknown form. Indeed, Drost and Nijman (1993) show that even if the innovation is *i.i.d.* at certain sample frequency, the innovation when aggregated to a lower sample frequency will become serially dependent even if it is an martingale difference sequence (*m.d.s.)*.[1] A volatility model where the innovation is not *i.i.d.* is called the semi-strong or weak form volatility model. Recent studies (e.g., Gallant, Hsieh and Tauchen 1991, Hansen 1994, Harvey and Siddique 1999, 2000, Jondeau and Rockinger 2003) have documented that the conditional skewness and kurtosis of asset returns are time-varying. Indeed, financial time series has been characterized with asymmetric and heavy-tailed non-Gaussian distributions of unknown form. It is therefore important to take into account the impact of other higher order time-varying moments of unknown form when constructing tests for volatility models. All existing tests for volatility models assume *i.i.d.* (possibly non-Gaussian) innovations and are not robust to time-varying higher order dependencies which may generate (e.g.) heavy tails and jumps.

We emphasize that a volatility model is concerned with serial dependence in conditional variance. Thus, tests that check all departures from *i.i.d.* are not suitable to test volatility models. For example, the correlation integral test proposed by Brock *et al.* (1991,1996), popularly known as the BDS test, has been documented to have excellent power against ARCH alternatives. However, the BDS test is not suitable to test validity of GARCH models, because it can lead to a rejection due to the existence

[1] To ignore serial dependence in $\{z_t\}$ by assuming *i.i.d.* will not render inconsistent parameter estimation for mean and variance parameters, although it would complicate the attempts to construct asymptotically efficient semiparametric estimators of the variance parameters (see Gallant and Tauchen 1989, Engle and Granger 1991). Lee and Hansen (1994) explicitly consider QMLE method with *m.d.s.* innovations.

of time-varying higher conditional moments (e.g., skewness and kurtosis) even when a GARCH model is correctly specified. In addition, the limit distribution of the BDS test statistic depends on parameter estimation uncertainty in volatility models (Brock *et al.* 1991, Appendix D).

In this paper, we will propose a new approach to testing validity of volatility model. Both univariate and multivariate volatility models – GARCH models and SV models are covered. There are many ways to generalize univariate GARCH models to multivariate GARCH models, but the curse of dimensionality quickly becomes a major obstacle because there is a relatively large number of components in the conditional variance-covariance matrix, and each of these components contain several parameters to be estimated. For manageable applications, rather restrictive assumptions usually have to be made, as is the case of Bollerslev's (1990) constant correlation multivariate GARCH models, factor-multivariate GARCH models, and Engle's (2002) time-varying correlation multivariate GARCH model. It is therefore highly desirable to develop a specification test that can check the overall adequacy of a multivariate GARCH model.

Specifically, we propose the multivariate generalized spectral derivative approach by extending Hong's (1999) univariate generalized spectral analysis. Generalized spectrum is a frequency domain nonlinear analytic tool. Because of the use of the characteristic function, our approach can check a variety of linear and nonlinear functional form misspecifications in volatility dynamics. Moreover, our frequency domain approach can check a growing number of lags as the sample size increases without suffering from the curse of dimensionality. Thus, our test is expected to be powerful against long memory volatility processes, such as fractionally integrated GARCH (FIGARCH) model (Baillie and Bollerslev 1996). Usually there would be a loss of power due to the loss of a large number of degrees of freedom, but this is not the case here due to a downward weighting scheme for the lags. Our approach is based on a kernel function and it naturally discounts higher order lags, which is consistent with the stylized fact that economic and financial markets are usually more influenced by the recent events than by the remote past events. The older the information, the less its impacts on the current volatility.

When constructing our tests, we do not require the formulation of any alternative volatility model. Moreover, as an important feature of our tests, they are robust to parameter estimation uncertainty, i.e., the use of *estimated* standardized residuals in place of true unobservable innovations has no impact on the limit distribution of test statistics. Any $\sqrt{T}-$consistent parameter estimator suffices. We do not require a specific method for estimation, and only estimated standardized residuals are needed to implement our tests. In particular, we do not have to compute the tedious case-by-case score functions or derivative of volatility models, unlike some popular tests for volatility models. All these desirable features yield a convenient procedure in practice.

Moreover, when testing multivariate volatility models, our approach is also applicable to test each univariate volatility component and their pairwise correlations. These separate inference procedures can reveal useful information on inadequacy of a multivariate volatility model, which will be useful when reconstructing the volatility model.

Section 2 introduces the volatility models and hypotheses of interest. Section 3 introduces a multivariate generalized spectral analysis and shows how to use the derivatives of the generalized spectral density to test volatility models. Section 4 constructs test statistics. We derive the asymptotic normality distribution of the proposed test statistics in Section 5 and establish their asymptotic power property in Section 6. Section 7 discusses the choice of a data-driven lag order. Section 8 examines the finite sample performance of the tests via Monte Carlo experiments. Section 9 considers an empirical application. Section 10 concludes. All mathematical proofs are given in the appendix. Throughout, we denote C for a generic bounded constant, A^* for the complex conjugate of A, $\mathrm{Re}\, A$ for the real part of A, and $||A||$ for the Euclidean norm of A. All limits are taken as the sample size $T \to \infty$. A GAUSS code to implement our tests is available from the authors upon request.

2. Model and Hypotheses

Consider a stochastic $d \times 1$ vector time series process $\{Y_t\}$, where $d \in \mathbb{N}^+ \equiv \{1, 2, \cdots\}$:

$$\begin{cases} Y_t = \mu_t + \varepsilon_t, \\ \varepsilon_t = H_t^{1/2} z_t, \\ E(z_t|I_{t-1}) = 0 \quad a.s., \\ \mathrm{Var}(z_t) \equiv E(z_t z_t'|I_{t-1}) = \mathbf{I}_d \quad a.s., \end{cases} \tag{2.1}$$

where $\mathbf{I}_d$ is the $d \times d$ identity matrix, $\{z_t\}$ is *a* $d \times 1$ unobservable martingale difference sequence ($m.d.s.$) innovation vector with $\mathrm{var}(z_t|I_{t-1}) = \mathbf{I}_d$. By construction, $\mu_t = E(Y_t|I_{t-1})$ is the $d \times 1$ conditional mean vector of Y_t given the information set available at time $t-1$, I_{t-1}, and $H_t = \mathrm{var}(Y_t|I_{t-1})$ is the $d \times d$ conditional variance-covariance matrix of Y_t given I_{t-1}. Both μ_t and H_t are measurable functions of information set I_{t-1}. Note that I_{t-1} may include not only lagged dependent variables but also exogenous variables and may date back to the infinite remote past. An important feature of most economic and financial time series is that μ_t and H_t may depend on the entire past history of Y_t rather than only a few lags of Y_t, as is the case for ARMA and/or GARCH processes. We note that the conditions of $E(z_t|I_{t-1}) = 0$ and $\mathrm{var}(z_t|I_{t-1}) = \mathbf{I}_d$ ensure that μ_t completely captures the conditional mean dynamics of Y_t, and H_t completely captures the conditional variance and the conditional correlations of Y_t.

In many economic and financial applications, interest has been in modelling the conditional variance-covariance matrix H_t, which characterizes the dynamics in volatility clustering of each time series, as well as the evolution of their conditional correlations. Important examples of univariate volatility models (i.e., when $d = 1$) include Bollerslev's (1986) generalized ARCH (GARCH) model, Nelson's (1991) exponential GARCH (EGARCH) model, Higgins and Bera's (1992) nonlinear ARCH (NARCH) model, Glosten *et al.*'s (1993) asymmetric model (GJR model), Ding *et al.*'s (1993) asymmetric power ARCH (APARCH) model, and Zakoian's (1994) threshold ARCH (TARCH) model. Important examples of multivariate volatility models (i.e., when $d > 1$) include vech-representation form due to Bollerslev *et al.* (1988), the constant correlation multivariate GARCH (CC-MGARCH) model due to Bollerslev (1990), and the BEKK model due to Engle and Kroner (1995), time-varying conditional correlation multivariate GARCH models of Engle (2002) and Tse and Tsui (2002).

We emphasize that process (2.1) also cover both univariate and multivariate stochastic volatility

(SV) models as well. To see this, let us consider a basic univariate SV model as used in Taylor (1986).

$$\begin{cases} Y_t = e^{1/2}\Lambda_t \varepsilon_t, \\ \Lambda_t = \gamma + \delta\Lambda_{t-1} + \nu\eta_t, \\ \{\varepsilon_t\} \sim i.i.d.(0,1), \\ \{\eta_t\} \sim i.i.d.(0,1), \end{cases} \tag{2.2}$$

where the log volatility Λ_t is unobservable. This model is a successful alternative to the class of GARCH models in capturing volatility clustering and heavy tails in financial time series. It is very closely related to continuous-time diffusion models which are widely used in the derivative pricing literature.

Our tests cannot be directly applied to test correct specification of latent volatility model $e^{1/2\Lambda_t}$, which is a latent SV model by considering the observable standardized innovations

$$z_t(\theta) = H_t^{-1/2}(\theta)\left[Y_t - \mu_t(\theta)\right],$$

where $\mu_t(\theta) = E\left(Y_t|I_{t-1}\right), H_t(\theta) =var(Y_t|I_{t-1})$, and $I_{t-1} = \{Y_{t-1}, Y_{t-2}, \cdots\}$ is the observed information set available at time $t-1$. For the SV model in (2.2), we have $\mu_t = 0$ and $H_t(\theta) = E\left[e^{\lambda_t}|I_{t-1}\right]$. The latter can be computed using various filtration techniques, such as Gallant and Tauchen's (1998) reprojections techniques that are based on a projection of the data onto a seminonparametric transition density or Kim *et al*'s (1998) particle filter algorithms, or Liesenfiled and Richard's (2003) EIS filtering method. We will impose our regularity conditions directly on μ_t and $H_t(\theta)$, rather than on the latent volatility model $e^{1/2\lambda_t}$. When the SV model correctly capture the conditional mean μ_t and conditional variance $H_t(\theta)$, the standardized innovations $\{z_t\}$ is a *m.d.s.* with respect to the observable information set I_{t-1} such that $E\left(z_t|I_{t-1}\right) = 0$ and var$(z_t|I_{t-1}) = \mathbf{I}_d$. Therefore we can apply our tests to whether var$(z_t|I_{t-1}) = \mathbf{I}_d$. Note that the distribution of z_t given I_{t-1} under correct specification of the SV model is unknown and can display serial dependence in higher order moments. Therefore, our tests are highly desirable here because they are robust to time-varying higher order moments.

Our interest in this paper is in checking whether a parametric volatility model $H_t(\theta) \equiv H(I_{t-1}, \theta)$ is correctly specified for var$(Y_t|I_{t-1})$, when $\theta \in \Theta$ is a finite dimensional parameter, and Θ is the parameter space. The hypotheses of interest are

$$\mathbb{H}_0 : \Pr[H_t(\theta_0) = \mathrm{Var}(Y_t|I_{t-1})] = 1 \quad \text{for some } \theta_0 \in \Theta$$

versus

$$\mathbb{H}_A : \Pr[H_t(\theta) \neq \mathrm{Var}(Y_t|I_{t-1})] > 0 \quad \forall\ \theta \in \Theta.$$

Because many statistical inferences for economic and financial data are based on model $H_t(\theta)$, a test of the hypothesis $\mathbb{H}_0$ is important from both theoretical and practical points of view. Interest in $\mathbb{H}_0$ is often based on the assumption that the conditional mean $\mu_t(\theta)$ has been correctly specified. Thus, strictly speaking, when $\mathbb{H}_0$ is rejected, it may be due to the misspecification of $H_t(\theta)$ and/or $\mu_t(\theta)$. For high-frequency economic and financial time series, it is believed that there exists mild or little serial dependence in conditional mean. Therefore, the primary focus has been on the modelling of volatility and conditional correlation of $\{Y_t\}$.

When $d = 1$, it has been suggested that Box-Pierce type tests for the squared standardized error can be used to test adequacy of a GARCH model $H_t(\theta)$, i.e.,

$$BP_2(p) = T(T+2)\sum_{j=1}^{p}(T-j)^{-1}\hat{\rho}_2^2(j), \quad p \in \mathbb{N}, \tag{2.3}$$

where $\hat{\rho}_2(j)$ is the sample autocorrelation function of the squared standardized residuals $\{z_t^2(\hat{\theta})\}$, $z_t(\hat{\theta}) = H_t(\hat{\theta})^{-1/2}\varepsilon_t(\hat{\theta})$, where $\hat{\theta}$ is an estimator of θ_0. As this test statistic is readily computable from the standardized residuals $\hat{z}_t(\hat{\theta})$, they have been widely used with an asymptotic χ_p^2 distribution in practice (e.g., Hafner 1998 p.112, Tsay 2001, p.115-118). However, Li and Mak (1994) showed that the Box-Pierce type tests are generally not asymptotically χ_p^2.[2] Also, the limit distribution of $BP_2(p)$ depends on parameter estimation uncertainty. In other words, substituting the estimated residuals for the unobserved residuals will change the asymptotic distribution of the test statistic. It is necessary to modify the test statistics to take into account the impact of parameter estimation uncertainty. Li and Mak (1994) propose a modified Box-Pierce type test:

$$Q(p) = n\hat{\rho}_2'\hat{V}^{-1}\hat{\rho}_2, \tag{2.4}$$

where $\hat{\rho}_2 = [\hat{\rho}_2(1), \cdots, \hat{\rho}_2(p)]'$, $\hat{\rho}_2(j)$ is the sample autocorrelation in $\{z_t^2(\hat{\theta})\}$, and $\hat{V}$ is a consistent asymptotic variance estimator which takes into account the impact of parameter estimation uncertainty. The test $Q(p)$ will be asymptotically χ_p^2 under $\mathbb{H}_0$. Ling and Li (1997) further extended this test to the multivariate case. The Ling-Li statistic is based on the squared sample autocorrelation coefficients of a transformed vector of estimated residuals:

$$LL_T(M) = T\sum_{j=1}^{M}\hat{R}^2(j), \; j = 1, \cdots, M,$$

where

$$\hat{R}(j) = \frac{\sum_{t=j+1}^{T}\left[z_t(\hat{\theta})'z_t(\hat{\theta}) - d\right]\left[z_t(\hat{\theta})'z_t(\hat{\theta}) - d\right]}{\sum_{t=1}^{T}\left[z_t(\hat{\theta})'z_t(\hat{\theta}) - d\right]^2}.$$

However, Tse and Tsui (1999) pointed out that there is a loss of information in the transformation of the residual vectors, which may induce a severe loss of power. Tse (2002) proposed a residual-based diagnostic test for GARCH models. These tests can be conveniently implemented by an artificial autoregression procedure with the squared standardized residuals or the cross products of the standardized residuals as dependent variables, and lagged squared standardized residuals or lagged cross products of the standardized residuals as the independent variables. Thus, to a certain extent, the form of the regression depends on a particular type of model inadequacy the researcher wants to investigate.

From a theoretical point of view, Box-Pierce type tests for $z_t^2(\theta)$ can detect many misspecifications

[2] Although the asymptotic distribution of the Box-Pierce statistics has not been firmly established, there have been arguments that the χ^2 distribution may be used as an approximation. Especially, the estimation error is of minor order when testing for serial correlation in the squared residuals (See Bellorslev 1990, footnote 7).

of practical importance. However, they can only capture linear volatility alternatives and may miss important nonlinear volatility alternatives, especially those that render zero autocorrelation in $z_t^2(\theta)$. For example, it can miss an asymmetric dynamic patterns in volatility, such as asymmetric behavior in volatility dynamics and conditional correlations.

On the other hand, most existing tests for volatility models employ a fixed lag order. From a theoretical perspective, such tests can easily miss volatility misspecification that occurs at higher order lags. Moreover, recent empirical studies (e.g., Baillie, Bollerslev and Mikkelsen 1996) find that high-frequency financial time series displays long memory in volatility clustering, where H_t depends on a very long history of Y_t. Some financial theory (e.g., Easley and O'hara 1992) suggests the non-Markovian property of high-frequency asset prices. Indeed, it is an important feature of a non-Markovian time series process that H_t depends on the entire past history of Y_t rather than only first few lags of it. Thus, it may suffer from substantial power loss by using a fixed lag order. In practice, one can employ a large lag order when a large sample size is available. However, the use of a large lag order usually induces loss of a large number of degree of freedoms, causing low power against many alternatives of practical importances. In particular, volatility and conditional correlations will be usually influenced more by the recent market events than by the remote market events. As a consequence, the strength of dependence in $z_t^2(\theta)$ on the past history will decay to zero as lag order increases. Below, we will propose a new generally applicable test for $\mathbb{H}_0$ which avoids the aforementioned undesirable features of the existing tests for volatility models.

3. Multivariate Generalized Spectral Approach

We will propose a unified test for $\mathbb{H}_0$ for both $d = 1$ and $d > 1$, by generalizing the univariate generalized spectral approach proposed in Hong (1999) to a multivariate generalized spectral analysis. The generalized spectrum is a spectral analysis based on the characteristic function. It is a basic frequency domain analytic tool for nonlinear time series, just as the power spectrum is a basic analytic tool for linear time series (e.g., Priestley 1981). Both time domain and frequency domain analytic tools are equally informative on serial dependence of a time series. In some applications, however, the frequency domain analysis is more enlightening and suitable. For example, as will be discussed below, the multivariate generalized spectrum can reveal useful information on cyclical dynamics in volatility clustering and the conditional correlations due to linear or nonlinear dependencies.

Define the standardized error of volatility model $H_t(\theta)$,

$$z_t(\theta) = H_t^{-1/2}(\theta)\varepsilon_t(\theta), \tag{3.1}$$

where $\varepsilon_t(\theta) = Y_t - \mu_t(\theta)$. Then the hypothesis $\mathbb{H}_0$ is equivalent to the hypothesis that

$$\text{var}\left[z_t(\theta_0)|I_{t-1}\right] = \mathbf{I}_d \quad a.s. \text{ for some } \theta_0 \in \Theta, \tag{3.2}$$

where I_{t-1} is the observable information available at time $t-1$. This implies

$$\text{var}\left[z_t(\theta_0)|I_{t-1}^z\right] = \mathbf{I}_d \quad a.s. \text{ for some } \theta_0 \in \Theta, \tag{3.3}$$

where $I_{t-1}^z \equiv \{z_{t-1}(\theta_0), z_{t-2}(\theta_0), \cdots\}$. Thus, to test $\mathbb{H}_0$, we can check if (3.3) holds. This is a standardized residual-based approach. This is quite convenient, because there is no need to compute tedious derivatives as in Li-Mak-Ling tests. Here, we still have the curse of dimensionality problem because I_{t-1}^z has an infinite dimension. Fortunately, the generalized spectral approach provides a sensible way to tackle this difficulty.

Most existing tests are based on the sample autocorrelations in $\{z_t(\theta)\}$. This can only detect misspecifications in volatility model $H_t(\theta)$ that render nonzero autocorrelations in $\{z_t(\theta_0)\}$. Because the autocorrelation is a measure for linear association, it may have lower power against nonlinear volatility alternatives. Nonlinear volatility dynamics is not uncommon in practice. For example, it is well-known that volatility reacts differently to a large price increase and a large price drop. This is a so-called leverage effect. It has also been documented that stock price volatility tends to be higher during the recession and tends to be lower during expansion (e.g., Hamilton and Lin 1996). In multivariate contexts, it is believed that negative shocks tend to transmitted across financial markets or different economies faster than positive shocks, as is the case for asymmetric international spillover of business cycles. Therefore it is highly desirable to develop a test that can check a volatility model against a variety of linear and nonlinear departures.

To be able to detect both linear and nonlinear departures, a sensible approach to testing $\mathbb{H}_0$ is to consider a test based on a smoothed nonparametric regression estimator for $\text{var}[z_t(\theta)|z_{t-j}(\theta)]$ and check whether this estimator is significantly different from zero. Such a test can detect many neglected or misspecified nonlinear volatility dynamics, and is expected to work well when $d \leq 3$. For multivariate cases with $d \geq 4$, this approach will unavoidably encounter the notorious difficulty of "curse of dimensionality". For such a dimension, smoothed nonparametric regressions would require an astronomically large data set even in the *i.i.d.* context (cf. Silverman 1986). On the other hand, this time-domain nonparametric approach does not deal with lag orders. Obviously, the use of a finitely many lags will render the test unable to detect misspecification of volatility models which occurs at higher order lags.

In this paper, we will propose a test for $\mathbb{H}_0$ by using a multivariate generalized spectral approach. This is achieved by generalizing Hong's (1999) univariate generalized spectral analysis to the multivariate time series analysis. Such an extension is useful and important, because it allows us to investigate linear and nonlinear interactions among different time series. There are a number of advantages of our frequency domain approach. First, our test is of nonparametric nature, and therefore is able to detect both linear and nonlinear volatility alternatives. Hence, we avoid the "curse of dimensionality" problem associated with smoothed nonparametric estimation due to a large dimension of d, because there is no need for smoothing at each lag. Second, our frequency domain approach naturally incorporates information from many lags. In other words, we can test a large number of lags without suffering from the "curse of dimensionality". This is particularly appealing in detecting long memory volatility alternatives. Moreover, our nonparametric approach naturally discounts higher order lags, thus alleviating the loss of a large number of degrees freedom due to the use of many lags. As a consequence, our test is expected to be powerful against the alternatives where the dependence in volatility decays to zero as $j \to \infty$. This is consistent with the stylized fact that financial markets are usually more influenced by the recent events than the remote past events. Also, an appealing feature of our approach is that any $\sqrt{T}$-consistent parameter, $\hat{\theta}$ say, does not affect the limit distribution of our test, which is $N(0,1)$ under $\mathbb{H}_0$. One can proceed as if the true parameter θ_0 were known and were equal to $\hat{\theta}$, any $\sqrt{T}$-consistent estimator. This gives a very convenient procedure in practice.

For notational economy, we put $z_t \equiv z_t(\theta^*)$, where $\theta^* = p\lim \hat{\theta}$. Suppose $\{z_t\}$ is a strictly stationary process with marginal characteristic function $\varphi(u) \equiv E(e^{iu'z_t})$ and pairwise joint characteristic function $\varphi_j(u,v) \equiv E(e^{iu'z_t + iv'z_{t-|j|}})$, where $i \equiv \sqrt{-1}$, $(u,v) \in \mathbb{R}^d \times \mathbb{R}^d$, and $j \in \{0, \pm 1, \cdots\}$. Following the basic idea of the generalized spectrum of Hong (1999), who considered a univariate time series, we consider the spectrum of the transformed series $\{e^{iu'z_t}\}$. It is defined as

$$f(\omega,u,v) \equiv \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \sigma_j(u,v) e^{-ij\omega}, \quad \omega \in [-\pi,\pi], \tag{3.4}$$

where ω is the frequency, and $\sigma_j(u,v)$ is the covariance function of the transformed series:

$$\sigma_j(u,v) \equiv \operatorname{cov}(e^{iu'z_t}, e^{iv'z_{t-|j|}}), \quad j \in \{0, \pm 1, \cdots\}. \tag{3.5}$$

Note that $f(\omega,u,v)$ is a complex-valued scalar function, although z_t is a $d \times 1$ vector. Compared to the conventional power spectral density matrix (e.g., Hannan 1970) and higher order spectra (e.g., Brillinger 1980), an appealing feature of $f(\omega,u,v)$ is that no moment condition on $\{z_t\}$ is required. The function $f(\omega,u,v)$ can capture any type of pairwise serial dependence in $\{z_t\}$, i.e., dependence between z_t and z_{t-j} for any nonzero lag j, including that with zero autocorrelation. It may be called the generalized spectrum of $\{z_t\}$ because when $E\|z_t\|^2 < \infty$, it can be differentiated to obtain the conventional power spectral density matrix as a special case:

$$-\frac{\partial^2}{\partial u \partial v} f(\omega,u,v)\,|_{(u,v)=(0,0)} = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \operatorname{cov}(z_t, z_{t-|j|}) e^{-ij\omega}, \quad \omega \in [-\pi,\pi], \tag{3.6}$$

where $\operatorname{cov}(z_t, z_{t-|j|})$ is a $d \times d$ autocovariance matrix of $\{z_t\}$ at lag $|j|$.

The generalized spectrum $f(\omega,u,v)$ itself is not suitable for testing $\mathbb{H}_0$, because it can capture serial dependence not only in mean but also in higher order conditional moments of z_t. An example is that $\{z_t\}$ follows an generalized asymmetric Student t distribution with time-varying skewness and kurtosis (e.g., Hansen 1994). In this case, $\{z_t\}$ is a *m.d.s.* process but is not *i.i.d.* The generalized spectrum $f(\omega,u,v)$ can capture this process, although $\{z_t\}$ is a *m.d.s.* with conditionally homoskedastic errors (i.e., $E(z_t|I_{t-1}^z) = 0$ *a.s.* and $\operatorname{var}(z_t|I_{t-1}^z) = \mathbf{I}_d$ *a.s.*).

However, just as the characteristic function can be differentiated to generate various moments of $\{z_t\}$, $f(\omega,u,v)$ can be differentiated to capture serial dependence in various conditional moments. To check serial dependence in volatility and conditional correlations of Y_t, we can differentiate $f(\omega,u,v)$ and use the following second order generalized spectral derivative

$$f^{(0,2,0)}(\omega,0,v) \equiv \frac{\partial^2}{\partial u \partial u'} f(\omega,u,v)|_{u=0} = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \sigma_j^{(2,0)}(0,v) e^{-ij\omega}, \quad \omega \in [-\pi,\pi], \tag{3.7}$$

where

$$\sigma_j^{(2,0)}(0,v) \equiv \frac{\partial^2}{\partial u \partial u'} \, \sigma_j(u,v)|_{u=0} = -\text{cov}(z_t z_t', e^{iv'z_{t-|j|}}) \tag{3.8}$$

is a $d \times d$ vector. The measure $\sigma_j^{(2,0)}(0,v)$ focuses exclusively on the conditional variance and the correlation dynamics of $\{z_t\}$. It checks whether the autoregression function $\text{var}(z_t|z_{t-j})$ at lag j is constant. Under appropriate conditions, $\sigma_j^{(2,0)}(0,v) = 0$ for all $v \in \mathbb{R}^d$ if and only if $\text{var}(z_t|z_{t-j})$ is a constant matrix.[3] Unlike a smoothed nonparametric estimator for $\text{var}(z_t|z_{t-j})$, $\sigma_j^{(2,0)}(0,v)$ does not involve any smoothed parameter and does not suffer from the "curse of dimensionality". Moreover, the function $f^{(0,2,0)}(\omega,0,v)$ incorporates information on all lags which are difficult to handle using a time domain approach.

It should be noted that the hypothesis of $\text{var}(z_t|I_{t-1}^z) = \mathbf{I}_d$ *a.s.* is not exactly the same as the hypothesis of $\text{var}(z_t|z_{t-j}) = \mathbf{I}_d$ for all $j > 0$. The former implies the latter but not vice versa. There exists a gap between them. This is the price we have to pay to deal with the difficulty of the "curse of dimensionality". Nevertheless, the examples for which $\text{var}(z_t|z_{t-j}) = \mathbf{I}_d$ for all $j > 0$ but $\text{var}(z_t|I_{t-j}^z) \neq \mathbf{I}_d$ may be rare in practice and are thus pathological.[4]

There is another payoff of using $f^{(0,1,0)}(\omega,0,v)$. Define the supremum generalized spectral derivative modulus

$$m(\omega) \equiv \sup_{v \in \mathbb{R}^d} \left\| f^{(0,1,0)}(\omega,0,v) \right\|, \quad \omega \in [-\pi,\pi]. \tag{3.9}$$

This can be viewed as the maximum dependence in mean of $\{z_t\}$ at frequency ω. It can capture cyclical dynamics that is caused by either linear or nonlinear serial dependence in volatility and conditional correlations of Y_t. For example, it has been suggested that volatility tends to be higher during the recession period than the expansion period (e.g., Hamilton and Lin 1996). Similarly, $m(\omega)$ can capture cyclical dynamics in conditional correlations caused by linear or nonlinear dependence. For example, it is often argued that negative shocks tend to be transmitted across markets faster than the positive shock to the markets, as may occur in international spillover of business cycles. Such cyclical patterns in correlation can be easily captured by $m(\omega)$.

Under $\mathbb{H}_0$, the generalized spectral derivative $f^{(0,2,0)}(\omega,0,v)$ becomes

$$f_0^{(0,2,0)}(\omega,0,v) \equiv \frac{1}{2\pi}\sigma_0^{(2,0)}(0,v) = \frac{1}{2\pi}\text{cov}(z_t z_t', e^{iv'z_t}), \quad \omega \in [-\pi,\pi]. \tag{3.10}$$

This is a "flat" second order generalized spectral derivative in the sense that $f_0^{(0,2,0)}(\omega,0,v)$ does not depend on frequency ω; it only depends on v. One can test $\mathbb{H}_0$ by comparing two consistent estimators, one for $f^{(0,2,0)}(\omega,0,v)$, and the other for $f_0^{(0,2,0)}(\omega,0,v)$. Any significant deviation between these estimators will indicate the rejection of $\mathbb{H}_0$. Below, we use a kernel method to develop a new class of tests for $\mathbb{H}_0$.

---

[3] See Bierens (1982) and Stinchcombe and White (1998) for discussion on related issue in an *i.i.d.* context.

[4] This gap can be further narrowed down by using the function $E(z_t z_t'|z_{t-j}, z_{t-l})$, which may be called the bi-autoregression function of r_t at lags (j,l). An equivalent measure is the generalized third order central cumulant function $\sigma_{j,l}^{(2,0)}(0,v) = \text{cov}[z_t z_t', \exp(iv_1'z_{t-j} + iv_2'z_{t-l})]$, where $v = (v_1, v_2)$. This is essentially a generalization of bispectral analysis and still avoids the curse of dimensionality.

4. Test Statistics

Suppose we have a random sample of size T and $\hat{\theta}$ is any $\sqrt{T}$-consistent estimator for θ_0. An example of $\hat{\theta}$ is the quasi-maximum likelihood estimator (e.g., Bollerslev and Wooldridge 1988, Lee and Hansen 1994, Lumsdaine 1996). Put $\hat{z}_t = \hat{H}_t^{-1/2}\hat{\varepsilon}_t, \hat{H}_t = H_t(\hat{\theta})$, and $\hat{\varepsilon}_t = Y_t - \mu_t(\hat{\theta})$. We can estimate the multivariate generalized spectral derivative $f^{(0,2,0)}(\omega, 0, v)$, by the following kernel estimator

$$\hat{f}^{(0,2,0)}(\omega, 0, v) \equiv \frac{1}{2\pi} \sum_{j=1-T}^{T-1} (1 - |j|/T)^{1/2} k(j/p) \hat{\sigma}_j^{(2,0)}(0, v) e^{-ij\omega}, \quad \omega \in [-\pi, \pi], v \in \mathbb{R}^d, \tag{4.1}$$

where

$$\hat{\sigma}_j^{(2,0)}(0, v) = -\frac{1}{T - |j|} \sum_{t=j+1}^{T} \left(\hat{z}_t \hat{z}_t' - d \right) [e^{iv'\hat{z}_{t-j}} - \hat{\varphi}_j(v)],$$

and

$$\hat{\varphi}_j(v) = \frac{1}{T - |j|} \sum_{t=|j|+1}^{T} e^{iv'\hat{z}_{t-|j|}}.$$

Here, $p \equiv p(T)$ is a bandwidth, and $k : \mathbb{R} \to [-1, 1]$ is a symmetric kernel. Examples of $k(\cdot)$ include Bartlett, Daniell, Parzen and Quadratic spectral kernels (e.g., Priestley 1981, p.442). The factor $(1 - |j|/T)^{1/2}$ is a finite-sample correction. It could be replaced by unity. Under certain conditions, $\hat{f}^{(0,2,0)}(\omega, 0, v)$ is consistent for $f^{(0,2,0)}(\omega, 0, v)$. See Theorem 2 below.

On the other hand, the flat generalized spectral derivative $f_0^{(0,2,0)}(\omega, 0, v)$ can be consistently estimated using

$$\hat{f}_0^{(0,2,0)}(\omega, 0, v) \equiv \frac{1}{2\pi} \hat{\sigma}_0^{(2,0)}(0, v), \;\; \omega \in [-\pi, \pi], v \in \mathbb{R}^d. \tag{4.2}$$

Our test will be based on the quadratic form comparing (4.1) and (4.2):

$$\begin{aligned} \hat{Q} &\equiv \int \int_{-\pi}^{\pi} \left\| \text{vech} \left[\hat{f}^{(0,2,0)}(\omega, 0, v) - \hat{f}_0^{(0,2,0)}(\omega, 0, v) \right] \right\|^2 d\omega d\mathcal{W}(v) \\ &= \sum_{j=1}^{T-1} k^2(j/p)(1 - j/T) \int \left\| \text{vech} \left[\hat{\sigma}_j^{(2,0)}(0, v) \right] \right\| d\mathcal{W}(v), \end{aligned} \tag{4.3}$$

where $\mathcal{W}(v) = \Pi_{c=1}^d W_0(v_c)$, $W_0 : \mathbb{R} \to \mathbb{R}^+$ is a nondecreasing weighting function that weighs sets symmetric about zero equally, and the unspecified integrals are taken over the support of $\mathcal{W}(\cdot)$. Examples of $W_0(\cdot)$ include the CDF of any symmetric probability distribution, either discrete or continuous. Note that the second equality follows from Parseval's identity.

4.1 Tests under non-i.i.d. Innovations

A volatility model with *i.i.d.* innovations $\{z_t\}$ in (2.1) is called a strong form volatility model in the literature (cf. Drost and Nijman 1993). It is possible that $H_t(\cdot)$ is correctly specified while the innovation $\{z_t\}$ displays higher order dependence, such as time-varying skewness and kurtosis. Indeed, Drost and Nijman (1993) show that even if $\{z_t\}$ is *i.i.d.* at certain sample frequency, the innovation when aggregated to a lower sample frequency will become serially dependent even if it is an $m.d.s.$[5] A volatility model where $\{z_t\}$ is not $i.i.d.$ is called the semi-strong or weak form volatility model. Recent studies (e.g., Gallant, Hsieh and Tauchen 1991, Hansen 1994, Harvey and Siddique 1999, 2000, Jondeau and Rockinger 2003) find that the conditional skewness and kurtosis of asset returns are time-varying. Indeed, financial time series are characterized by heavy-tailed non-Gaussian distributions of unknown form. For our tests, it is also important to take into account the impact of other higher order time-varying moments which may be displayed in the form of (e.g.) heavy tails and jumps. In light of this, tests assuming $i.i.d.$ innovations for $\{z_t\}$ will not be robust to time-varying conditional moments. They will have incorrect sizes; in particular, they may be likely to incorrectly reject correct GARCH models with time-varying higher order moments. Thus, it is highly desirable to develop tests robust to higher order moments dynamics of unknown form. To our knowledge, there has been no such a test in the earlier literature. All existing tests assumed $i.i.d.$ innovations. Here we provide a test that is robust to time-varying higher order conditional moments of unknown form.

Our test statistic that is robust to time-varying higher order conditional moments of unknown form is given as follows:

$$\hat{M}(p) = \left[\sum_{j=1}^{T-1} k^2(j/p)(T-j) \int \left\| \text{vech}\left[\hat{\sigma}_j^{(2,0)}(0,v) \right] \right\|^2 d\mathcal{W}(v) - \hat{C}(p) \right] \Big/ \sqrt{\hat{D}(p)}, \qquad (4.4)$$

where

$$\begin{aligned}
\hat{C}(p) &= \sum_{j=1}^{T-1} k^2(j/p) \frac{1}{T-j} \sum_{t=j+1}^{T} \left[\left\| \text{vech}\left(\hat{z}_t \hat{z}_t' \right) \right\|^2 - d \right] \int \left| \hat{\psi}_{t-j}(v) \right|^2 d\mathcal{W}(v), \\
\hat{D}(p) &= 2 \sum_{j=1}^{T-2} \sum_{l=1}^{T-2} k^2(j/p) k^2(l/p) \sum_{a=1}^{d} \sum_{b=a}^{d} \sum_{a'=1}^{d} \sum_{b'=a'}^{d} \int \int \\
&\quad \times \left| \frac{1}{T-\max(j,l)} \sum_{t=\max(j,l)+1}^{T} [\hat{z}_{at}\hat{z}_{bt}\hat{z}_{a't}\hat{z}_{b't} - \delta_{ab}\delta_{a'b'}] \hat{\psi}_{t-j}(u) \hat{\psi}_{t-j}^*(v) \right|^2 d\mathcal{W}(u) d\mathcal{W}(v),
\end{aligned}$$

and $\hat{\psi}_t(u) = e^{iu'\hat{z}_t} - T^{-1} \sum_{t=1}^{T} e^{iu'\hat{z}_t}$. The centering and scaling factors $\hat{C}(p)$ and $\hat{D}(p)$ are approximately the mean and the variance of the quadratic form $T\hat{Q}$ in (4.3). They have taken into account the impact of time-varying higher order moments of unknown form in $\{z_t\}$, such as time-varying skewness and kurtosis. This ensures a correct level for $\hat{M}(p)$ asymptotically. Note that $\hat{M}(p)$ involves d- and $2d$-dimensional numerical integrations, which can be computationally quite intensive when d is large. In practice, one may choose a finite number of grid points symmetric about zero or generate a finite number of points drawn from a uniform distribution on $[-1, 1]^d$. Our asymptotic theory allows for both discrete and continuous weighting functions $W_0(\cdot)$ which weigh sets symmetric about zero equally. A continuous weighting function for $W_0(\cdot)$ will ensure good power for $\hat{M}(p)$, but there is a trade-off between computational cost and power when choosing a discrete or continuous weighting function

[5] To ignore serial dependence in $\{z_t\}$ by assuming $i.i.d.$ will not render inconsistent parameter estimation for mean and variance parameters, although it would complicate the attempts to construct asymptotically efficient semiparametric estimators of the variance parameters (see Gallant and Tauchen 1989, Engle and Granger 1991). Lee and Hansen (1994) explicitly consider the QMLE method with $m.d.s.$ innovations.

$W_0(\cdot)$. One may expect that the power of $\hat{M}(p)$ will be ensured if sufficiently fine grid points are used.

4.2 Test Statistics Under i.i.d. Innovations

In many applications, practitioners often assume that the innovation $\{z_t\}$ is $i.i.d.(0, \mathbf{I}_d)$. If this assumption is indeed valid, our test statistic can be simplified, by taking into account the implication of the *i.i.d.* properties of the innovations. When $\{z_t\} \sim i.i.d.(0, \mathbf{I}_d)$, and the components of z_t are mutually independent (rather than just uncorrelated), we can simplify our test statistic as follows:

$$\hat{M}^o(p) = \left[\sum_{j=1}^{T-1} k^2(j/p)(T-j) \int \left\| \text{vech} \left[\hat{\sigma}_j^{(2,0)}(0, v) \right] \right\|^2 d\mathcal{W}(v) - \hat{C}^o(p) \right] \Big/ \sqrt{\hat{D}^o(p)}, \tag{4.5}$$

where $\mathcal{W}(v) = \prod_{c=1}^d W_0(v_c)$, and the centering and scaling constants

$$\begin{aligned}
\hat{C}^o(p) &= \left[\sum_{a=1}^d \left(T^{-1} \sum_{t=1}^T \hat{z}_{at}^4 - 1 \right) + \frac{1}{2} d(d-1) \right] \int \hat{\sigma}_0(v, -v) d\mathcal{W}(v) \sum_{j=1}^{T-1} k^2(j/p), \\
\hat{D}^o(p) &= 2 \left[\sum_{a=1}^d \left(T^{-1} \sum_{t=1}^T \hat{z}_{at}^4 - 1 \right)^2 + \frac{1}{2} d(d-1) \right] \int |\hat{\sigma}_0(u, v)|^2 d\mathcal{W}(u) d\mathcal{W}(v) \sum_{j=1}^{T-2} k^4(j/p),
\end{aligned}$$

with $\hat{\sigma}_0(u, v) = \hat{\varphi}(u+v) - \hat{\varphi}(u)\hat{\varphi}(v)$, $\hat{\varphi}(v) = n^{-1} \sum_{t=1}^n e^{iv'\hat{z}_t}$. Both $\hat{C}^o(p)$ and $\hat{D}^o(p)$ have been simplified, because the *i.i.d.* properties of $\{z_t\}$ has been exploited in deriving $\hat{C}^o(p)$ and $\hat{D}^o(p)$.

4.3 Separate Diagnostics

The tests $\hat{M}(p)$ and $\hat{M}^o(p)$ are designed to assess the overall performance of a multivariate volatility model $H_t(\theta)$. When these tests reject $\mathbb{H}_0$, one may like to know what has caused the rejection, that is, the likely source of misspecification. For example, is the rejection due to poor modeling of volatility dynamics in each time series component, or poor modeling of the conditional correlations between different time series components? If it is due to poor modelling of volatility dynamics, which component of Y_t has the poorest fit for its volatility dynamics? Or if it is due to poor modelling of conditional correlation, which pair of Y_t has the poorest fit for their conditional correlation? Information on these patterns will be very useful in reconstructing a more satisfactory multivariate GARCH model.

For this purpose, we can consider a class of separate tests based on each individual component of z_t. When $\{z_t\}$ is $m.d.s.(0, \mathbf{I}_d)$ but not $i.i.d.(0, \mathbf{I}_d)$, we can construct the following individual test statistic. Suppose $r_{ct} = \hat{z}_{a_c t} \hat{z}_{b_c t} - \delta_{a_c b_c}$ where $a_c, b_c \in \{1, \cdots, d\}$ depends on $c \in \{1, \cdots, \frac{1}{2}d(d+1)\}$, we define

$$\hat{M}_c(p) = \left[\sum_{j=1}^{T-1} k^2(j/p)(T-j) \int \left| \hat{\sigma}_{cj}^{(1,0)}(0, v_c) \right|^2 dW_0(v_c) - \hat{C}_c(p) \right] \Big/ \sqrt{\hat{D}_c(p)}, \tag{4.6}$$

where

$$\hat{\sigma}_{cj}^{(1,0)}(0,v_c) = (T-|j|)^{-1} \sum_{t=|j|+1}^{T} r_{ct} \hat{\psi}_{ctj}(v_c),$$

and the centering and scaling factors

$$\begin{aligned}
\hat{C}_c(p) &= \sum_{j=1}^{T-1} k^2(j/p) \frac{1}{T-j} \sum_{t=j+1}^{T} [\hat{z}_{a_c t}^2 \hat{z}_{b_c t}^2 - \delta_{a_c b_c}] \int \left| \hat{\psi}_{ct-j}(v_c) \right|^2 dW_0(v_c), \\
\hat{D}_c(p) &= 2 \sum_{j=1}^{T-2} \sum_{l=1}^{T-2} k^2(j/p) k^2(l/p) \int \int \\
&\quad \times \left| \frac{1}{T-\max(j,l)} \sum_{t=\max(j,l)+1}^{T} [\hat{z}_{a_c t}^2 \hat{z}_{b_c t}^2 - \delta_{a_c b_c}] \hat{\psi}_{ct-j}(u_c) \hat{\psi}_{ct-j}^*(v_c) \right|^2 dW_0(u_c) dW_0(v_c),
\end{aligned}$$

with $\hat{\psi}_{ct}(v) = e^{iv\hat{r}_{ct}} - \hat{\varphi}_c(v)$, and $\hat{\varphi}_c(v) = T^{-1} \sum_{t=1}^{n} e^{iv\hat{r}_{ct}}$.

Note that the components of $\hat{r}_t$ include individual squared standardized residuals and their pairwise cross-products. Thus, $\hat{M}_c(p)$ can reveal useful information about which component has inadequate modelling for volatility dynamics, and which pair has inadequate modeling for their conditional correlation. However, we should emphasize that these tests are generally not asymptotically independent when $\{z_t\}$ is *m.d.s.* but not *i.i.d.* Note also that $\hat{M}(p)$ is not a simple sum of the individual test statistics $\hat{M}_c(p)$. The latter is easier to compute because they only involve one- or two-dimensional numerical integrations.

Suppose $\{z_t\}$ is $i.i.d.(0, \mathbf{I}_d)$, we can have the following simplified individual test statistics: For $r_{ct} = z_{a_c t} z_{b_c t} - \delta_{a_c b_c}$, where $(a_c, b_c) \in \{1, \cdots, d\}$ depends on $c \in \{1, \cdots, \frac{1}{2}d(d+1)\}$, we define

$$\hat{M}_c^o(p) = \left[\sum_{j=1}^{T-1} k^2(j/p)(T-j) \int |\hat{\sigma}_{cj}^{(1,0)}(0,v_c)|^2 dW_0(v_c) - \hat{C}_c^o(p) \right] \Big/ \sqrt{\hat{D}_c^o(p)}, \tag{4.7}$$

where the centering and scaling constants

$$\begin{aligned}
\hat{C}_c^o(p) &= \left[\delta_{a_c b_c} T^{-1} \sum_{t=1}^{T} (z_{a_c t}^4 - 1) + (1 - \delta_{a_c b_c}) \right] \int \hat{\sigma}_{c0}(v_c, -v_c) dW_0(v_c) \sum_{j=1}^{T-1} k^2(j/p), \\
\hat{D}_c^o(p) &= 2 \left[\delta_{a_c b_c} T^{-1} \sum_{t=1}^{T} (z_{a_c t}^4 - 1) + (1 - \delta_{a_c b_c}) \right]^2 \int |\hat{\sigma}_{c0}(u_c, v_c)|^2 dW_0(u_c) dW_0(v_c) \sum_{j=1}^{T-2} k^4(j/p).
\end{aligned}$$

Again, $\hat{M}^o(p)$ is not a simple sum of the individual test statistics $\hat{M}_c^o(p)$. We will use these individual tests in our empirical applications below.

5. Asymptotic Distribution

Because $\{z_t\}$ is not necessarily *i.i.d.* under $\mathbb{H}_0$, the derivation of the asymptotic distribution of $\hat{M}(p)$ and $\hat{M}_c(p)$ is much more challenging than under *i.i.d.*, since we need to take into account the impact of possible time-varying higher order moments of $\{z_t\}$. To derive the null asymptotic distribution of the test statistics $\hat{M}(p)$, we first provide some regularity conditions.

Assumption A.1: $\{Y_t\}$ is a $d \times 1$ strictly stationary process such that $Y_t = \mu_t + H_t^{1/2} z_t$, where $\mu_t \equiv E(Y_t|I_{t-1}), H_t = \text{var}(Y_t|I_{t-1})$, and I_{t-1} is an information set available at time $t-1$ that may contain lagged dependent variables $\{Y_{t-j},\ j > 0\}$, lagged shocks $\{\varepsilon_{t-j} \equiv H_t^{1/2} z_t,\ j > 0\}$, as well as current and lagged exogenous variables $\{X_{t-j},\ j \geq 0\}$, with $E\left\|z_t^8\right\| \leq C$;

Assumption A.2: For a sufficiently large integer q, there exists a strictly stationary process $\{z_{q,t}\}$ measurable with respect to $\{z_{t-1}, z_{t-2}, \cdots, z_{t-q}\}$ such that (a) as $q \to \infty$, $z_{q,t}$ is independent of $\{z_{t-q-1}, z_{t-q-2}, \cdots\}$ for each t, $E(z_{q,t}|I_{t-1}) = 0\ a.s.$, $E(z_{q,t}z_{q,t}') = \Sigma_q\ a.s.$, (b) $E\left\|\text{vech}(z_t z_t') - \text{vech}(z_{qt} z_{qt}')\right\|^4 \leqslant Cq^{-2\kappa}$ for some constant $\kappa \geq 1$; (c) $\Sigma_q \to \mathbf{I}_d$ as $q \to \infty$, and $E(z_{q,t}^{16}) \leqslant C$ for all large q.

Assumption A.3: $\mu(I_{t-1}, \theta)$ and $H(I_{t-1}, \theta)$ are parametric models for μ_t and H_t, where $\theta \in \Theta$ is a finite-dimensional parameter. (a) $\mu(\cdot, \theta)$ and $\sigma(\cdot, \theta)$ are measurable with respect to I_{t-1} for each $\theta \in \Theta$; (b) with probability one, $\mu(I_{t-1}, \cdot)$ and $\sigma(I_{t-1}, \theta)$ are twice differentiable with respect to $\theta \in \Theta$; (c) for $i, j \in \{1, \cdots, d\}$, $E\sup_{\theta\in\Theta} ||\frac{\partial}{\partial\theta}\mu_i(I_{t-1}, \theta)||^{4\nu} \leqslant C$, $E\sup_{\theta\in\Theta} ||\frac{\partial}{\partial\theta}H_{ij}(I_{t-1}, \theta)||^{4\nu} \leqslant C$, $E\sup_{\theta\in\Theta} ||\frac{\partial^2}{\partial\theta\partial\theta'}\mu_i(I_{t-1}, \theta)||^2 \leqslant C$, and $E\sup_{\theta\in\Theta} ||\frac{\partial}{\partial\theta}H_{ij}(I_{t-1}, \theta)||^2 \leqslant C$, where $\nu > 1$.

Assumption A.4: $\hat{\theta} - \theta_0 = O_P(T^{-1/2})$, where $\theta_0 \equiv p\lim(\hat{\theta}) \in \Theta$.

Assumption A.5: Let $\hat{I}_t$ be the observed information set available at period t that may contain some assumed initial values. Then $\lim_{T\to\infty} \sum_{t=1}^{T} \sup_{\theta\in\Theta} |\mu(\hat{I}_{t-1}, \theta) - \mu(I_{t-1}, \theta)| \leqslant C$ and $\lim_{T\to\infty} \sum_{t=1}^{T} \sup_{\theta\in\Theta} |\text{vech}(H(\hat{I}_{t-}$
$\leqslant C$.

Assumption A.6: The kernel $k : \mathbb{R} \longrightarrow [-1, 1]$ is symmetric about 0, and is continuous at 0 and all points except a finite number of points, with $k(0) = 1$, $\int_0^\infty k^2(z)dz < \infty$, and $|k(z)| \leqslant C|z|^{-b}$ as $z \to \infty$ for some $b > \frac{1}{2}$.

Assumption A.7: $W_0 : \mathbb{R} \longrightarrow \mathbb{R}^+$ is nondegenerate, nondecreasing and weighs sets symmetric about zero equally, with $\int_{-\infty}^{\infty} v^2 dW_0(v) < \infty$.

Assumption A.8: The process $\{z_t,\ \frac{\partial}{\partial\theta}\mu_i(I_{t-1}, \theta_0), \frac{\partial}{\partial\theta}H_{ij}(I_{t-1}, \theta_0), i, j = 1, 2, \cdots, d\}$ is an α-mixing process with the α-mixing coefficient satisfying $\sum_{j=-\infty}^{\infty} \alpha(j)^{\frac{\nu-1}{\nu}} \leq C$, where $\nu > 1$ is as in Assumption A.3.

Assumption A.1 is a regularity condition on the data generating process $\{Y_t\}$. Note that $\{Y_t\}$ may not be covariance-stationary. An example is the IGARCH process. Assumption A.2 implies ergodicity for innovations $\{z_t\}$. It holds trivially when $\{z_t\}$ is a Markovian process with an arbitrary but finite order. It also covers many non-Markovian processes for $\{z_t\}$.

Assumption A.3 are standard regularity conditions on the variance model $H(I_{t-1}, \theta)$. We allow for $H(I_{t-1}, \theta)$ to depend on the entire past history I_{t-1}, rather than a vector with fixed dimension. This is a distinct feature from the existing nonparametric test for the conditional variance models (e.g., Li 2001). Assumption A.4 requires a $\sqrt{T}$-consistent $\hat{\theta}$, which need not be asymptotically most efficient. It can be the conditional quasi-maximum likelihood estimator. This is similar in spirit to Wooldridge's (1990, 1991) robust modified moment-based tests for the conditional mean and the variance specifications. Assumption A.5 is a condition on the truncation of information set I_{t-1}, which usually contains information dating back to the remote past and so may not be observable. Because of the truncation, one may have to assume some initial values in estimating volatility model $H(I_{t-1},\theta)$. Assumption A.5 ensures that the use of initial values, if any, has no impact on the limit distribution of $\hat{M}(p)$. For instance, consider ARMA(1,1)-GARCH(1,1) model:

$$\begin{cases} \mu(I_{t-1},\theta) = \alpha Y_{t-1} + \beta\varepsilon_{t-1}, \\ \varepsilon_t = H(I_{t-1},\theta)^{1/2} z_t, \\ H(I_{t-1},\theta) = \gamma + \delta H(I_{t-2},\theta) + \tau\varepsilon_{t-1}^2, \end{cases}$$

where $|\alpha| \leq \bar{\alpha} < 1$ and $|\beta| \leq \bar{\beta} < \infty$. Here $I_{t-1} = \{Y_{t-1}, Y_{t-2}, \cdots\}$ but $\hat{I}_{t-1} = \{Y_{t-1}, Y_{t-2}, \cdots, Y_1, \hat{\varepsilon}_0\}$, where $\hat{\varepsilon}_0$ is an initial value assumed for ε_0. By recursive substitution, we have

$$\begin{aligned} E\sum_{t=1}^{T}\sup_{\theta\in\Theta}\left|\mu(\hat{I}_{t-1},\theta) - \mu(I_{t-1},\theta)\right| &= E\sum_{t=1}^{T}\sup_{\theta\in\Theta}\left|\beta\sum_{j=t-1}^{\infty}\alpha^j\varepsilon_{t-j-1} - \beta\alpha^{t-1}\hat{\varepsilon}_0\right| \\ &\leqslant \bar{\beta}\sum_{t=1}^{T}E\sup_{\alpha}\left|\alpha^{t-1}\left(\sum_{l=0}^{\infty}\alpha^l\varepsilon_{-l} - \hat{\varepsilon}_0\right)\right| \\ &\leqslant 2\bar{\beta}\sum_{t=1}^{T}|\bar{\alpha}|^{t-1}\left[E|\varepsilon_0|\sum_{l=0}^{\infty}\bar{\alpha}^l + E|\hat{\varepsilon}_0|\right] \leqslant C. \end{aligned}$$

We can obtain a similar condition for $H(\cdot,\theta)$ for a GARCH(1,1) model.

Assumption A.6 is a regularity condition on the kernel $k(\cdot)$. It includes all commonly used kernels (see., e.g., Priestley 1981, p.442). For kernels with bounded support, such as the Bartlett and Parzen kernels, $b = \infty$. For the Daniell kernel, $b = 1$, and for the Quadratic-spectral kernel, $b = 2$. These kernels have unbounded support. As a consequence, all $T - 1$ lags contained in the sample is used in the test statistics $\hat{M}(p)$. Assumption A.7 is a condition on the weighting function $W(\cdot)$ for transform parameter v. The CDF of any symmetric continuous distribution with finite variance satisfies this condition. Finally, Assumption A.8 imposes some temporal dependence condition on the related processes. We now state the main result of this section.

Theorem 1: *Suppose Assumptions A.1–A.8 hold, and* $p = cT^{\lambda}$ *for* $\lambda \in (0, (2b-1)/(4b-1))$ *and* $c \in (0,\infty)$. *(i)* $\hat{M}(p) \xrightarrow{d} N(0,1)$ *under* $\mathbb{H}_0$. *(ii) If* $\{z_t\}$ *is i.i.d.*$(0,1)$, *and the components of* z_t *are mutually independent, then* $\hat{M}^o(p) \xrightarrow{d} N(0,1)$.

An important feature of $\hat{M}(p)$ is that the use of *estimated* standardized residuals $\{\hat{\varepsilon}_t\}$ rather than unobservable errors $\{\varepsilon_t\}$ has no impact on the limit distribution of $\hat{M}(p)$. One can proceed as if the true parameter value θ_0 were known and were equal to $\hat{\theta}$. The reason is that the estimator $\hat{\theta}$ converges to θ_0 at the parametric rate $T^{-\frac{1}{2}}$, which is faster than the nonparametric estimator $\hat{f}^{(0,m,0)}(\omega, 0, v)$. Consequently, the limit distribution of $\hat{M}(p)$ is solely determined by $\hat{f}^{(0,m,0)}(\omega, 0, v)$, and replacing $\hat{\theta}$ by θ_0 has no impact on it. This delivers a convenient procedure, because it does not require any specific estimation method and one does not have to be concerned with the impact of parameter estimations uncertainty. Of course, parameter estimation uncertainty in $\hat{\theta}$ may still have nontrivial impact on the small sample distribution of $M(p)$. In this case, one may use a bootstrap procedure similar to that of Hansen (1996) to obtain more accurate levels of the tests in small samples.

6. Asymptotic Power

Our tests are derived without assuming any alternative volatility models. To gain insight into the nature of the alternatives that our tests are able to detect, we now examine the asymptotic behavior of $\hat{M}(p)$ under the alternative to $\mathbb{H}_0$.

Theorem 2: *Suppose Assumptions A.1 and A.3–A.8 hold, and $p = cT^{\lambda}$ for $\lambda \in (0, \frac{1}{2})$ and $c \in (0, \infty)$.*

$$\frac{p^{1/2}}{T}\hat{M}(p) \xrightarrow{p} \left[2D\int_0^{\infty} k^4(z)dz\right]^{-1/2} \int \left\|\text{vech}\left[f^{(0,2,0)}(\omega,0,v) - f_0^{(0,2,0)}(\omega,0,v)\right]\right\|^2 d\omega d\mathcal{W}(v)$$

$$= \left[2D\int_0^{\infty} k^4(z)dz\right]^{-1/2} \sum_{j=1}^{\infty}\int \left\|\text{vech}\left[\sigma_j^{(2,0)}(0,v)\right]\right\|^2 d\mathcal{W}(v),$$

where

$$D = E\left\|\text{vech}\left(z_1 z_1' - \mathbf{I}_d\right)\right\|^2 \int \left\|\text{vech}\left[f(\omega,0,v) - f_0(\omega,0,v)\right]\right\|^2 d\omega d\mathcal{W}(u) d\mathcal{W}(v)$$

Consequently, for any sequence of nonstochastic constants, $\{C_T = o(T/p^{1/2})\}$,

$$\lim_{T\to\infty} \Pr\left[\hat{M}(p) > C_T\right] = 1,$$

whenever $E(z_t z_t' - \mathbf{I}_d | z_{t-j})$ is a measurable function of z_{t-j} for some $j > 0$.

We thus expect that $\hat{M}(p)$ has relatively omnibus power against a wide variety of linear and nonlinear alternatives with unknown lag structure, as is confirmed in our simulation below. It should be emphasized that the omnibus power property does not mean that the proposed tests are more powerful than any other existing tests against *every* alternative. In fact, just because $\hat{M}(p)$ has to take care of a wide range of possible misspecifications, it may be less powerful against certain specific alternative than a parametric test. Nevertheless, the main advantage of our omnibus test, which is not shared by any other parametric tests, is that $\hat{M}(p)$ can eventually detect all possible model misspecifications that render the autoregression functions $E(z_t z_t' - \mathbf{I}_d | z_{t-j})$ nonzero at any lag $j > 0$. This avoids the blindness of searching for different alternatives when one has no prior information.

The existing tests for $\mu(I_{t-1}, \theta)$ and $H(I_{t-1}, \theta)$ only consider a fixed order lag. They can easily miss alternatives for which misspecification occurs at higher lag orders. Of course, these tests could be used to check a large number of lags when a large sample is available. However, they may not be expected to be powerful when the number of lags is too large. Such power loss, due to the loss of a large number of degrees of freedom, is not shared by our test, thanks to the role played by the kernel $k(\cdot)$. Most non-uniform kernels discount higher order lags (i.e., a higher order lag receives a smaller weight). This enhances good power against stationary processes whose serial dependence decays to zero as lag order j increases. Thus, our generalized spectral approach can check a large number of lags without losing too many degrees of freedom. This feature is not available for popular χ^2-type tests with a large number of lags, which essentially give equal weighting to each lag. Equal weighting is not fully efficient when a large number of lags is considered.

Since $p\lim_{T\to\infty} \hat{Q}(p)$ is positive whenever $\text{var}(z_t | z_{t-j}) \neq 0$ for some lag $j > 0$, $\hat{M}(p)$ is an asymptotically one-sided N(0,1) test. Thus, upper-tailed asymptotic critical values (e.g., 1.645 at the 5%

level) should be used.

7. Data-Driven Lag Order

A practical issue in implementing our test is the choice of lag order or bandwidth p. An advantage of our generalized spectral approach is that it can provide a data-driven method to choose p, which, to some extent, let data themselves speak for a proper p. Before discussing specific data-driven methods, we first justify the use of a data-driven lag order, $\hat{p}$ say. For this purpose, we impose a Lipschitz continuity condition on the kernel $k(\cdot)$. This condition rules out the truncated kernel $k(z) = \mathbf{1}(|z| \leq 1)$, where $\mathbf{1}(\cdot)$ is the indicator function, but it still includes most commonly used kernels.

Assumption A.9: *For any $x, y \in \mathbb{R}$, $|k(x) - k(y)| \leq C|x - y|$ for some constant C.*

Theorem 3: *Suppose Assumptions A.1–A.9 hold, and $\hat{p}$ is a data-driven bandwidth such that $\hat{p}/p = 1 + O_P(p^{-(\frac{3}{2}\beta - 1)})$ for some $\beta > (2b - \frac{1}{2})/(2b - 1)$, where b is as in Assumption A.5, and p is a nonstochastic bandwidth with $p = cT^{\lambda}$ for $\lambda \in (0, (2b-1)/(4b-1))$ and $c \in (0, \infty)$. Then (i) under $\mathbb{H}_0$, $\hat{M}(\hat{p}) - \hat{M}(p) \xrightarrow{p} 0$ and $\hat{M}(\hat{p}) \xrightarrow{d} N(0,1)$. (ii) If $\{z_t\}$ is i.i.d.$(0, \mathbf{I}_d)$, and the components of z_t are mutually independent, then $\hat{M}^o(\hat{p}) - \hat{M}^o(p) \xrightarrow{p} 0$ and $\hat{M}^o(\hat{p}) \xrightarrow{d} N(0,1)$.*

Thus, as long as $\hat{p}$ converges to p sufficiently fast, the use of $\hat{p}$ rather than p has no impact on the limit distribution of $\hat{M}(\hat{p})$. This is an additional "nuisance parameter-free" property.

Theorem 3 allows for a wide range of admissible rates for $\hat{p}$. One plausible choice of $\hat{p}$ is the nonparametric plug-in method considered in Hong (1999). It minimizes an asymptotic integrated mean square error (IMSE) criterion for the estimator $\hat{f}(\omega, 0, v)$. Nonparametric plug-in methods are not uncommon in the literature (e.g., Newey and West 1994, Silverman 1986). Consider some "pilot" generalized spectral derivative estimators based on a preliminary bandwidth $\bar{p}$:

$$\bar{f}^{(0,2,0)}(\omega, 0, v) \equiv \frac{1}{2\pi} \sum_{j=1-T}^{T-1} (1 - |j|/T)^{\frac{1}{2}} \bar{k}(j/\bar{p}) \hat{\sigma}_j^{(2,0)}(0, v) e^{-ij\omega}, \tag{6.1}$$

$$\bar{f}^{(q,2,0)}(\omega, u, v) \equiv \frac{1}{2\pi} \sum_{j=1-T}^{T-1} (1 - |j|/T)^{\frac{1}{2}} \bar{k}(j/\bar{p}) \hat{\sigma}_j^{(2,0)}(0, v) |j|^q e^{-ij\omega}, \tag{6.2}$$

where the kernel $\bar{k}(\cdot)$ need not be the same as the kernel $k(\cdot)$ used in (3.6). For example, $\bar{k}(\cdot)$ can be the Bartlett kernel while $k(\cdot)$ is the Daniell kernel. Note that $\bar{f}(\omega, u, v)$ is an estimator for $f(\omega, u, v)$ and $\bar{f}^{(q,0,0)}(\omega, u, v)$ is an estimator for the generalized spectral derivative

$$f^{(q,2,0)}(\omega, u, v) \equiv \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \sigma_j^{(2,0)}(0, v) |j|^q e^{-ij\omega}. \tag{6.3}$$

Suppose for the kernel $k(\cdot)$, there exists some $q \in (0, \infty)$ such that $0 < k^{(q)} \equiv \lim_{z \to 0} \frac{1 - k(z)}{|z|^q} < \infty$. Then the plug-in bandwidth is defined as

$$\hat{p}_0 \equiv \hat{c}_0 T^{\frac{1}{2q+1}}, \tag{6.4}$$

where the tuning parameter estimator

$$\begin{aligned}\hat{c}_0 &\equiv \left[\frac{2q(k^{(q)})^2}{\int_{-\infty}^{\infty}k^2(z)dz}\frac{\int\int_{-\pi}^{\pi}\left\|\text{vech}\left[\bar{f}^{(q,2,0)}(\omega,0,v)\right]\right\|^2 d\omega d\mathcal{W}(u)}{\text{Re}\int_{-\pi}^{\pi}\int\text{vech}\left[\bar{f}^{(0,2,2)}(\omega,0,0)\right]\bar{f}(\omega,v,-v)dW(v)d\omega}\right]^{\frac{1}{2q+1}}\\ &= \left[\frac{2q(k^{(q)})^2}{\int_{-\infty}^{\infty}k^2(z)dz}\frac{\sum_{j=1-T}^{T-1}(T-|j|)\bar{k}^2(j/\bar{p})|j|^{2q}\int\left\|\text{vech}\left[\hat{\sigma}_j^{(2,0)}(0,v)\right]\right\|^2 dW(u)dW(v)}{\sum_{j=1-T}^{T-1}(T-|j|)\bar{k}^2(j/\bar{p})\text{vech}\left[\hat{R}_2(j)\right]\text{Re}\int\hat{\sigma}_j(v,-v)dW(v)}\right]^{\frac{1}{2q+1}},\end{aligned}$$

with $\hat{R}_2(j)=T^{-1}\sum_{t=j+1}^{T}\hat{z}_t\hat{z}_{t-j}'$. The second equality here follows from Parseval's identity. Note that $\hat{p}_0$ is real-valued. One can take its integer part, and the impact of integer-clipping is expected to be negligible.

The data-driven $\hat{p}_0$ in (6.4) involves the choice of a preliminary bandwidth $\bar{p}$, which can be either fixed or growing with the sample size T. If $\bar{p}$ is fixed, $\hat{p}_0$ generally grows at rate $T^{\frac{1}{2q+1}}$ under $\mathbb{H}_A$, but $\hat{c}_0$ does not converge to the optimal tuning constant that minimizes the IMSE of $\hat{f}(\omega,0,v)$. This is analogous in spirit to a parametric plug-in method. Following Hong (1999), we can show that when $\bar{p}$ grows with T properly, the data-driven bandwidth $\hat{p}_0$ in (6.4) minimizes an asymptotic IMSE of $\hat{f}^{(0,2,0)}(\omega,0,v)$. The choice of $\bar{p}$ is somewhat arbitrary, but we expect that it is of secondary importance. This is confirmed in our simulation below.

From a theoretical point of view, the choice of $\hat{p}$ based on the IMSE criterion may not maximize the power of the test. A more sensible alternative would be to develop a data-driven $\hat{p}$ using a power criterion, or a criterion that trades off level distortion and power loss. This will necessitate higher order asymptotic analysis and is beyond the scope of this paper. We are content with the IMSE criterion here. Our simulation experience suggests that the power of our tests is relatively flat in the neighborhood of the optimal lag order that maximizes the power, and the data-driven $\hat{p}_0$ based on IMSE performs reasonably well in finite samples.

8. Monte Carlo Evidence

[To be completed]

8.1. Size

We examine the size performance of our tests under the following two DGPs:

DGP S.1:

$$\begin{cases} Y_t = 0.2Y_{t-2}+\varepsilon_t,\\ \varepsilon_t=\sqrt{h_t}z_t,\\ h_t=0.2+0.6h_{t-1}+0.2\varepsilon_{t-1}^2,\\ \{z_t\}\sim i.i.d.N(0,1).\end{cases}$$

DGP S.2:

$$\begin{cases} Y_t = 0.2Y_{t-2} + \varepsilon_t, \\ \varepsilon_t = \sqrt{h_t} z_t, \\ h_t = 0.2 + 0.6h_{t-1} + 0.2\varepsilon_{t-1}^2, \\ z_t = \left[a(\varepsilon_{t-1}) + \sqrt{1 - a^2(\varepsilon_{t-1})} \delta_t \right] \xi_t, \\ a(\varepsilon_{t-1}) = [1 + \exp(-\varepsilon_{t-1})]^{-1} \\ \{\xi_t\} \sim i.i.d.N(0,1), \\ \{z_t\} \sim i.i.d.N(0,1). \end{cases}$$

Under DGP S.1, the innovations $\{z_t\}$ is $i.i.d.$(0,1). This is a strong GARCH process, where all conditional higher order moments are time-invarying. Under DGP S.2, the innovations $\{z_t\}$ has the property that $E(z_t|I_{t-1}) = 0$ and $\text{var}(z_t|I_{t-1}) = \mathbf{I}_d$, but $\{z_t\}$ is not $i.i.d.$ It can be shown that $\text{Kurt}(Y_t|I_{t-1}) = 9 - 6a^4(\varepsilon_{t-1})$. This is a semi-strong GARCH process. It was first proposed by Bossaerts, Härdle and Hafner (1996) in its rudimentary form, and extended by Yang and Hafner (1997). Most existing tests are not robust to time-varying conditional higher order moments of $\{z_t\}$.

8.2 Power

We consider the following alternatives:
DGP P.1: Threshold GARCH
DGP P.2: EGARCH
DGP P.3: Stochastic volatility models
DGP P.4: Quadratic GARCH

9. Empirical Application

[To be completed]

10. Conclusion

Volatility models have played an important role in economics and finance such as in the studies of trade-off between return and risk, volatility clustering, and volatility spillover between financial markets or between financial sectors and real sectors. We propose a class of new specification tests for volatility models in time series, where the dimension of the conditioning information set may be infinite. Both univariate and multivariate volatility models are covered. The tests can detect a wide range of model misspecifications in volatility while being robust to higher order time-varying moments of unknown form (e.g., skewness and kurtosis). They can check a large number of lags, and naturally discount higher order lags, which is consistent with the stylized fact that economic or financial markets are more affected by the recent past events than by the remote past events. No specific estimation method is required, and the tests have the appealing "nuisance parameter free" property that parameter estimation uncertainty has no impact on the limit distribution of the test statistics. Only the standardized estimated residuals are needed to implement our tests. There is no need to compute score functions of volatility models as required in some existing tests. We examine the finite sample performance of the proposed tests via a simulation study. An empirical application to GARCH models for stock returns highlights our approach.

References

Baillie, R.T., Bollerslev, T. and H. O. Mikkelsen (1996), "Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics* 74, 3-30.

Bera, A.K. and M.L. Higgins (1993), "ARCH Models: Properties, Estimations and Testing", *Journal of Economic Surveys* 7, 305-362.

Bierens, H. (1982), "Consistent Model Specification Tests", *Journal of Econometrics* 20, 105-134.

Bollerslev, T. (1987), "A Conditionally Heteroscedastic Time Series Model for Speculative Prices and Rates of Return," *The Review of Economics and Statistics* 69, 542-547.

Bollerslev, T. (1990), "Modeling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Model," *Review of Economics and Statistics* 72, 498-505.

Bollerslev, T. and T.R. Chou, and K. Kroner (1988), "ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence," *Journal of Econometrics* 52, 5-59.

Bollerslev, T. and R. F. Engle, and Nelson (1994), "ARCH Models," in Handbook of Econometrics, ed. by Engle, and D. McFadden, chap. 4, 2959-3038.

Bollerslev, T. and R. F. Engle, and J.M. Wooldridge (1988), "A Capital Asset Pricing Model with Time-Varying Covariances," *Journal of Political Economy* 96, 116-131.

Bollerslev, T. and Wooldridge, J. (1992), "Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time Varying Covariances," *Econometric Reviews* 11, 143-172.

Brock, W., D. Hsieh and B. LeBaron (1991), *Nonlinear Dynamics, Chaos, and Instability: Statistical Theory and Economic Evidence*. MIT Press: Cambridge, MA.

Brock, W., D. Dechert, J. Scheinkman and B. LeBaron (1996), "A Test for Independence Based on the Correlation Dimension," *Econometric Reviews* 15, 197-235.

Ding, Z., C.W.J. Granger and R.F. Engle (1993), "A Long Memory Property of Stock Market Returns and a New Model," *Journal of Empirical Finance* 1, 83-106.

Drost and Nijman (1993), "Temporal Aggregation of GARCH processes," *Econometrica* 61, 909-927.

Easley and O'hara (1992), "Time and the Process of Security Price Adjustment," *The Journal of Finance* 47, 577-605.

Engle, R. F. (1982), "Autoregressive Conditional Heteroscedasticity with Estimates of the U.K.Inflation," *Econometrica* 50, 987-1008.

Engle, R. F. (2002), "Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models," *Journal of Business and Economic Statistics* 20, 339-350.

Engle, R. F. and Kroner (1995), "Multivariate Simultaneous Generalized ARCH," *Econometric Theory* 11, 122-150.

Engle, R. F. and V.K. Ng (1991), "Measuring and Testing the Impact of News on Volatility," NBER Working Papers 3681, National Bureau of Economic Research, Inc.

Gallant, A.R., D. Hsieh and G. Tauchen (1991) "On Fitting a Recalcitrant Series: The Pound/Dollar Exchange Rate, 1974-83," in Barnett, William A., James Powell, and George E. Tauchen, eds., Nonparametric and Semiparametric Methods in Econometrics and Statistics, Proceedings of the Fifth

International Symposium in Economic Theory and Econometrics, Cambridge University Press, Cambridge, Chapter 8, 199-240.

Gallant and Tauchen (1989) "Seminonparametric Estimation of Conditionally Constrained Heterogeneous Processes: Asset Pricing Applications," *Econometrica* 57(5), 1091-1120.

Ghysels E., A. Harvey and E. Renault (1996) "Stochastic Volatility,"

Glosten, L.R., R. Jaganathan and D. E. Runkle (1993) "On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks," *Journal of Finance* 48, 1779-1801.

Härdle and C. Hafner (1996)

Hafner, C. (1998) *Nonlinear time series analysis with applications to foreign exchange rate volatility*, Heidelberg; New York : Physica-Verlag.

Hamilton, J. D. (2001), "A Parametric Approach to Flexible Nonlinear Inference," *Econometrica* 69, 537-73.

Hamilton, J. D. and Lin (1996), "," *Econometrica* 69, 537-73

Hansen, B. (1994), "Autoregressive Conditional Density Estimation," *International Economic Review* 35, 705-730.

Harvey, C. and A. Siddique (1999), "Autoregressive Conditional Skewness," *Journal of Financial and Quantitative* Analysis 34, 465-487.

Harvey, C. and A. Siddique (2000), "Conditional skewness in asset pricing tests," *Journal of Finance* 55, 1263-1296.

Higgins M.L. and A.K. Bera (1992), "", *Journal of Economic Surveys* 7, 305-362.

Hong, Y. (1999), "Hypothesis Testing in Time Series via the Empirical Characteristic Function: A Generalized Spectral Density Approach," *Journal of the American Statistical Association* 84, 1201-1220.

Hong, Y. (2001), "A test for volatility spillover with application to exchange rates," *Journal of Econometrics* 103, 183-224.

Hong, Y. and Y.J. Lee (2004), "Generalized Spectral Tests for Conditional Mean Models in Time Series with Conditional Heteroskedasticity of Unknown Form," forthcoming in *Reviews of Economic Studies.*

Hong, Y. and Shehadeh (1999), "A New Test for ARCH Effects and its Finite-Sample Performance," *Journal of Business and Economic Statistics* 17, 91-108.

Hull and H. White (1987), "The Pricing of Options on Assets with Stochastic Volatilities", *Journal of Finance* 42, 281-300.

Jondeau, E. and M. Rockinger (2003), "Conditional Volatility, Skewness, and Kurtosis: Existence, Persistence, and Comovements," *Journal of Economic Dynamics and Control* 27, 1699-1737.

Kim, S., N. Shepherd and S. Chib (1998) "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models ," *The Review of Economics Studies* 65, 361-393.

Lee, S. and B. Hansen (1994), "Asymptotic Theory for the GARCH(1,1) Quasi-Maximum Likelihood Estimator," *Econometric Theory*, 10, 29-52.

Li, W.K. and T. Mak (1994), "On the Squared Autocorrelations in Non-linear Time Series with Conditional Heteroskedasticity," *Journal of Time Series Analysis* 15, 627-636.

Liesenfiled and Richard (2003),

Ling and Li (1997), "Diagnostic Checking of Nonlinear Multivariate Time Series with Multivariate ARCH Errors," *Journal of Time Series Analysis* 18, 447-464.

Lumsdaine, R.L. (1996), "Asymptotic Properties of the Quasi-Maximum Likelihood Estimator in GARCH(1,1) and IGARCH(1,1) Models,"

Lundbergh, S. and T. Teräsvirta (1998), "Evaluating GARCH models," Working Paper Series in Economics and Finance 292, Stockholm School of Economics.

Nelson, D. (1991) "Conditional Heteroskedasticity in Asset Returns: A New Approach," *Econometrica* 59, 347-370.

Priestley, M.B. (1981), *Spectral Analysis and Time Series.* Academic press: London.

Sentana, E. (1995) "Quadratic ARCH models," *Reviews of Economic Studies* 62, 639-661.

Shephard, N. (1996), *Statistical Aspects of ARCH and Stochastic Volatility*, chap.1, 1-67.

Silverman, B. (1986), *Density Estimation for Statistics and Data Analysis*, Chapman and Hall: London.

Stinchcombe M. and H. White (1998), "Consistent Specification Testing with Nuisance Parameters Present only Under the Alternative," *Econometric Theory* 14, 295-324.

Taylor (1986)

Tsay, R.S. (2002) *Analysis of Financial Time Series*, John Wiley and Sons, Inc.

Tse, Y.K. (2002) "Residual-based Diagnostics for Conditional Heteroskedasticity Models," *Econometrics Journal* 5, 358-373.

Tse, Y.K. and A.K. Tsui (1999) "A Note on Diagnosing Multivariate Conditional Heteroskedasticity Models," *Journal of Time Series Analysis* 20, 679-691.

Tse, Y.K. and A.K. Tsui (2002) "A Multivariate Generalized Autoregressive Conditional Heteroskedasticity Model with Time-Varying Correlations," *Journal of Business and Economic Statistics* 20, 351-362.

White, H. (1994), *Estimation, Inference and Specification Analysis*, Cambridge University Press: New York.

Wooldridge, J. (1990), "A Unified Approach to Robust, Regression-Based Specification Tests," *Econometric Theory* 6, 17-43.

Wooldridge, J. (1991), "On the Application of Robust, Regression-Based Diagnostics to Models of Conditional Means and Conditional Variances," *Journal of Econometrics* 47, 5-46.

Yang and C. Hafner (1997,) ","

Zakoian, J.M. (1994), "Threshold Heteroskedasticity Models," *Journal of Economics Dynamincs and Control* 15, 931-955.