ACADEMIA ECONOMIC PAPERS
38 : 2 (June 2010), 157–210

A Method-of-Moments-Based Synthesis of Estimation and Testing Methods for Financial Time Series Models

Yi-Ting Chen *
Institute of Economics
Academia Sinica

Keywords: CCC/DCC models, Conditional moment test, Copula, Estimating function, GARCH-type models, Generalized residual, Method of moments, ML method, Multiplicative error models, QML method

JEL classification: C12, C13, C22, C32

---

* Correspondence: Yi-Ting Chen, Institute of Economics, Academia Sinica, Taipei 115, Taiwan. Tel: (02) 2782-2791 ext. 622; Fax: (02) 2785-3946; E-mail: ytchen@gate.sinica.edu.tw. I am grateful to two anonymous referees for their valuable comments and suggestions.

ABSTRACT

In this survey paper, we provide a synthesis of parametric (finite-dimensional-moments-based) estimation and testing methods for various classes of financial time series (FTS) models, including partially specified GARCH-type models in the univariate context and their extensions to univariate fully specified models, multivariate partially specified models (constant/dynamic conditional correlation models), multivariate fully specified models (copula-based multivariate dynamic models), and multiplicative error models. This synthesis is based on a unified approach, which is established using the concept of the generalized residual (that encompasses the error terms of various models) and the method of moments (that forms the generalized estimation and testing methods). This approach is systematically applicable to various conditional moment or distribution models. This paper summarizes a number of important FTS models and the associated parametric estimation and testing methods, and highlights some simple but general principles underlying these seemingly different models and methods.

1. INTRODUCTION

The financial time series (FTS) literature has experienced a substantial development since the appearance of Bollerslev's (1986) generalized autoregressive conditional heteroskedasticity (GARCH) model, which is a well-known extension of Engle's (1982) ARCH model. In the univariate context, the GARCH model has been widely accepted as a baseline model for financial volatility and has been extended to a class of "partially specified" GARCH-type models with a variety of conditional mean-and-variance specifications. This class of models has been further extended to a range of "fully specified" GARCH-type models with various conditional distribution (CD) assumptions for various economic and statistical reasons, such as risk management, density forecast, maximum likelihood (ML) estimation, or other higher-order conditional moment (CM) or CD-oriented problems. In the multivariate context, the GARCH model has also been extended to various multivariate GARCH-type models with various conditional covariance matrix specifications, such as the constant/dynamic conditional correlation (CCC/DCC) models, for a set of financial returns. Recently, the latter have been further extended to various copula-based multivariate dynamic (CMD) models as multivariate CD models. These multivariate models are essential for exploring financial problems in a portfolio level.

The aforementioned FTS models have attracted a vast number of applications in empirical finance. In econometric studies, various estimation methods have also been suggested for various classes of FTS models. Representative examples include the Gaussian quasi-ML (QML) method for partially specified GARCH-type models in the univariate context, the ML method for fully specified univariate models, the two-stage methods for the CCC/DCC and CMD models. Accompanying various estimation methods, numerous specification tests have also been introduced for evaluating various FTS models in various directions. The FTS literature has been substantially enriched by these specification, estimation, and testing methods. This development involves voluminous econometric studies. It is therefore essential and useful to have a synthesis of these models and methods for summarizing some important results in this literature and, more importantly, for highlighting some simple but general principles underlying these methods. This survey paper is written for these two main purposes.

Our synthesis is established by a unified approach, which is systematically applicable to various classes of FTS models. Specifically, we first use a generic CM or CD model to unify various FTS models into the same framework. This generic model has a "generalized residual" (or said a generalized error term) which satisfies the martingale-difference condition when it is correctly specified. By combining this generalized residual with a set of "instrumental variables," we can obtain a set of unconditional moment restrictions for estimation under this specification correctness condition. Similarly, we can generate another set of unconditional moment restrictions by combining this generalized residual (or a different but related generalized error term) with another set of instrumental variables for testing the hypothesis of specification correctness (or other related hypotheses). The resulting estimation and testing methods are, respectively, known as the estimating function (EF) method and the CM testing method in the literature. They are both constructed in the spirit of the method of moments (MM). We can use this MM-based unified approach to recover and generate various parametric (finite-dimensional-moments-based) estimation and testing methods for various FTS models by choosing the generalized residual and the instrumental variables in an appropriate way. Thus, by using this unified approach, we do not only provide a survey of a number of important FTS models and methods, but also highlight some common principles underlying various parametric methods in this literature.

It should be addressed that the principles of the generalized residual, the EF method, and the CM testing method are all existing ones in statistics and econometrics. Specifically, the generalized residual is understood as a generalized counterpart of the conventional regression error. This concept is quite powerful for unifying various seemingly different statistical or econometric models; see, e.g., Cox and Snell (1968) for its applications in the context of generalized linear regressions, Gouriéroux et al. (1987) and Pagan and Vella (1989) for microeconometric models, and Cameron and Trivedi (1998, Ch. 5) for count data models, among others. The EF method is also known to be useful for unifying the ML, QML, and other extreme estimation methods; see, e.g., Newey and McFadden (1994), White (1994), Mittelhammer et al. (2000), and Bera and Bilias (2002). The CM testing method was introduced by Newey (1985) and Tauchen (1985) for the static models based on the ML method, and has been substantially extended to various contexts; see, e.g., White (1987), Davidson and MacKinnon (1990, 1993), Wooldridge (1990, 1991), and among many others. Parametric specification tests can often be interpreted as particular CM tests. In this paper, we apply and integrate these principles to show a general picture of the parametric methods for FTS analysis.

It should also be addressed that this paper is not written for providing a more complete survey than existing review studies in this literature. Among the representative ones of the latter, Bollerslev et al. (1992) is a well-cited paper that collects some important univariate and multivariate GARCH-type models developed in the earlier stage of this literature and interprets the specifications and empirical applications of these models. Bollerslev et al. (1994) is a related survey paper that focuses on the theoretical properties of these earlier models. Pagan (1996) is another excellent survey that covers discussions on FTS concepts, summarizes a broad spectrum of studies and issues in financial econometrics, and links FTS models to economic models. Li et al. (2002) provided a review on the theoretical properties, such as the stationarity, ergodicity, and moment conditions, of some GARCH-type models. Bauwens et al. (2006) reviewed and compared multivariate GARCH-type models, including certain recently developed DCC and CMD models that are not discussed by the aforementioned papers; see also Silvennoinen and Teräsvirta (2008) for another recent review on multivariate GARCH-type models. Pacurar (2008) is a useful survey on autoregressive conditional duration (ACD)-type models; this class of FTS models are encompassed by the multiplicative error models (MEMs) that will be discussed later. These review papers include many important discussions but not a unified approach like ours. Therefore, our paper is distinguishable from these papers by highlighting the MM-based unified approach for FTS analysis. This is potentially useful for practitioners who are unfamiliar with, but are interested in getting a heuristic idea about, why and how various seemingly different FTS methods work in a systematic way.

The concepts and literatures discussed in this paper are substantially built on their counterparts that we have considered and collected in Chen (2007, 2008a), Chen and Hsieh (2010), and some related studies that we will mention in proper subsections. The first three papers focus on applying the MM to develop parametric specification tests for CMD models, univariate GARCH-type models, and MEMs, respectively. Unlike these papers, this paper further uses a generalized-residual-based generic model to integrate these and other important FTS models, and shows that the unified MM approach derived for this generic model is applicable to recover the parametric estimation and testing methods for various classes of FTS models. Therefore, this paper shares the basic ideas of the MM with our previous studies, but it extends the latter from particular models to a general family of FTS models. This generalization is not proposed for claiming a new methodological contribution, but is introduced for demonstrating and comparing various FTS methods in an efficient way. Moreover, unlike these previous studies, this paper is completely written as a review to summarize existing results rather than for developing specification tests.

This paper is not intended to be exhaustive either. More specifically, the discussions do not include non-parametric models (see, e.g., Section 2.4 of Silvennoinen and Teräsvirta, 2008) and the continuous unobserved component models such as the stochastic volatility models (see, e.g., Harvey et al., 1994; Kim et al., 1998) and the Tobit-GARCH models (see, e.g., Lee, 1999; Wei, 2002). This is not because these models are inessential but because the MM is incompatible with, or at least not directly applicable to, these contexts. Indeed, the Tobit-GARCH models are essential for studying financial volatility in the presence of price limitations. The conventional MM is not suitable in these cases because the non-parametric methods are based on infinite-dimensional moment conditions and require more complicated asymptotic analyses, and the continuous unobserved component models may need to be estimated by certain simulated ML methods or Bayesian methods. Nonetheless, our discussions still cover a wide range of popular FTS models.

The remainder of this paper is organized as follows. In Section 2, we discuss the MM-based unified approach. In Section 3, we review univariate partially specified GARCH-type models and their extensions to univariate fully specified models, multivariate partially specified models, multivariate fully specified models, and MEMs. We use these classes of models as demonstrative examples to show the general applicability of the MM-based unified approach in FTS analysis. We conclude this paper in Section 4.

2. A UNIFIED APPROACH

In this section, we apply the principles of the generalized residual, the EF method, and the CM testing method to show the MM-based unified approach. As mentioned before, these general principles are all existing ones in statistics and econometrics. Thus, we do not claim a new methodological contribution by deriving this approach. Nonetheless, it is important to sketch this derivation for showing why and how the existing methods reviewed in this study could work in a systematic way.

2.1 Unification: Generic Model and Generalized Residual

Let $\{y_t\}$ be a stationary sequence of n-dimensional continuous financial variables. The dimension could be $n = 1$ in the univariate context or some fixed $n > 1$ in the multivariate context. In most cases, the subscript t represents a time index, and $\{y_t\}$ denotes a sequence of regularly spaced time series data, such as a sequence of daily, weekly, or monthly financial returns. However, the subscript t could also simply represent an ordering index when $\{y_t\}$ is a sequence of irregularly spaced time series data, as in certain MEMs that are discussed in Section 3.5. FTS analysts are interested in explaining or predicting y_t using the history of time series $Y_{t-1} := (y_{t-1}, y_{t-2}, \ldots)$ and a vector of exogenous variables w_t and its history, denoted by $W_t := (w_t, w_{t-1}, \ldots)$. We use $\mathcal{X}_{t-1}$ to represent the information set generated from Y_{t-1} and W_t.

Suppose that $y_t|\mathcal{X}_{t-1}$ has the true, but unknown, CD $F_t(\cdot|\mathcal{X}_{t-1})$ with the conditional probability density function (PDF) $f_t(\cdot|\mathcal{X}_{t-1})$ and the CM:

$$M_t(y_t|\mathcal{X}_{t-1}) := \mathbb{E}[\phi_t(y_t, X_t)|\mathcal{X}_{t-1}] = \int_{\mathbb{R}^n} \phi_t(y, X_t) f_t(y|\mathcal{X}_{t-1}) \mathrm{d}y, \tag{1}$$

for some finite-dimensional moment function ϕ_t, which is dependent on y_t and could also be dependent on X_t which represents a vector of $\mathcal{X}_{t-1}$-measurable explanatory variables such that $\mathbb{E}[X_t|\mathcal{X}_{t-1}] = X_t$. We also let $\mu_t(X_t, \alpha)$ be a generic CM model for the unknown $M_t(y_t|\mathcal{X}_{t-1})$ with the parameter vector α, and $G_t(\cdot|X_t, \beta)$ be a generic CD model for the unknown $F_t(\cdot|\mathcal{X}_{t-1})$ with the conditional PDF $g_t(\cdot|X_t, \beta)$ and the parameter vector β. As shown later, the FTS models of our consideration are either particular $\mu_t(X_t, \alpha)$'s or particular $G_t(\cdot|X_t, \beta)$'s, and hence can be unified into the generic model. To address the fact that any FTS model requires some choice of regressors from the information set $\mathcal{X}_{t-1}$, we express $\mu_t(X_t, \alpha)$ and $G_t(\cdot|X_t, \beta)$ as functions of X_t, rather than functions of $\mathcal{X}_{t-1}$.

The statistical inferences of this generic model are based on the specification correctness condition of this model. The generic CM model is correctly specified when

$$M_t(y_t|\mathcal{X}_{t-1}) = \mu_t(X_t, \alpha_o), \quad \forall\, t, \tag{2}$$

for some α_o in the parameter space of α, and the generic CD model is correctly specified when

$$F_t(\cdot|\mathcal{X}_{t-1}) = G_t(\cdot|X_t, \beta_o), \quad \forall\, t, \tag{3}$$

for some β_o in the parameter spaces of β. Note that condition (2) can be rewritten as:

$$\mathbb{E}\left[\phi_t(y_t, X_t) - \mu_t(X_t, \alpha_o) | \mathcal{X}_{t-1}\right] = 0, \quad \forall\, t. \tag{4}$$

By taking the partial derivative of the probability identity: $\int_{\mathbb{R}^n} g_t(y|X_t, \beta)\mathrm{d}y = 1$ with respect to β and by assuming that the differentiation and integration operators are interchangeable (that is, the Cramér property; see, e.g., Bera and Bilias, 2001, p. 24), we can also write the result:

$$\int_{\mathbb{R}^n} \nabla_\beta g_t(y|X_t, \beta)\mathrm{d}y = \int_{\mathbb{R}^n} \left(\nabla_\beta \ln g_t(y|X_t, \beta)\right) g_t(y|X_t, \beta)\mathrm{d}y = 0. \tag{5}$$

By combining (5) with the PDF counterpart of (3): $f_t(\cdot|\mathcal{X}_{t-1}) = g_t(\cdot|X_t, \beta_o)$ for all the t's, we can transform the specification correctness condition for the generic CD model into:

$$\mathbb{E}\left[\nabla_\beta \ln g_t(y_t|X_t, \beta_o) | \mathcal{X}_{t-1}\right] = 0, \quad \forall\, t, \tag{6}$$

where the conditional expectation operator $\mathbb{E}[\cdot|\mathcal{X}_{t-1}]$ is taken with respect to the true conditional PDF $f_t(\cdot|\mathcal{X}_{t-1})$ as in (1) and (4).

Importantly, conditions (4) and (6) can be rewritten as the martingale-difference condition:

$$\mathbb{E}\left[v_{1,ot}|\mathcal{X}_{t-1}\right] = 0, \quad v_{1,ot} := v_{1t}|_{\theta=\theta_o}, \quad \forall\, t, \tag{7}$$

with the random variable v_{1t} and the parameter vector θ:

$$v_{1t} = \begin{cases} \phi_t(y_t, X_t) - \mu_t(X_t, \alpha), & \text{with } \theta = \alpha, \text{ for the generic CM model,} \\ \nabla_\beta \ln g_t(y_t|X_t, \beta), & \text{with } \theta = \beta, \text{ for the generic CD model.} \end{cases}$$

Such a random variable can be interpreted as an error term of the generic model because it is required to be orthogonal to any $\mathcal{X}_{t-1}$-measurable variables when the generic model is correctly specified. Following the MM literature, we refer to v_{1t} as the generalized "residual," while it is indeed based on the parameter θ rather than an estimator of θ_o. The generalization of v_{1t} is due to the fact that it encompasses various error terms of various FTS models that are encompassed by the generic model. Given this generalized residual, we can apply the EF method to derive a generalized estimator for θ_o under the specification correctness condition (7). Given this estimator, we can further conduct the CM testing method to establish a generalized test for the null hypothesis:

$$H_o : \mathbb{E}\left[v_{2,ot}|\mathcal{X}_{t-1}\right] = 0, \quad v_{2,ot} := v_{2t}|_{\theta=\theta_o}, \quad \forall\, t, \tag{8}$$

in which v_{2t} represents a generalized error term which could be the same as, or different from but related to, the generalized residual v_{1t}. In the case where $v_{1t} = v_{2t}$, this generalized test checks the specification correctness of the generic model. In the case where $v_{1t} \neq v_{2t}$, this test evaluates certain CM restrictions that are implied by, or beyond but related to, condition (7). Practical examples will be given in Sections 3.1.

2.2 The MM: The EF Method and the CM Test

To discuss the EF method and the CM testing method, we let v_{it} be a $r_i \times 1$ function of y_t, X_t, and θ with $\Theta \subseteq \mathbb{R}^{q_1}$ denoting a compact parameter space of θ and $V_{it} := \nabla_{\theta'} v_{it}$ denoting a $r_i \times q_1$ matrix that comprises the partial derivatives of v_{it} taken with respect to θ', where " $'$ " denotes the transpose operator; $i = 1, 2$. Also, we let z_{it} be a $q_i \times r_i$ matrix of $\mathcal{X}_{t-1}$-measurable instrumental variables such that $\mathbb{E}[z_{it}|\mathcal{X}_{t-1}] = z_{it}$ with $i = 1, 2$. Denote $z_{i,ot} := z_{it}|_{\theta=\theta_o}$. By combining v_{it} with z_{it}, we can obtain a set of $q_i \times 1$ unconditional moment restrictions:

$$\mathbb{E}\left[z_{i,ot} v_{i,ot}\right] = 0 \tag{9}$$

with $i = 1$ under the specification correctness condition (7) and with $i = 2$ under the null hypothesis (8).

The EF method estimates θ_o under condition (7) using the moment restriction set: $\mathbb{E}[z_{1,ot} v_{1,ot}] = 0$, which is "just-identified" in the sense that $\dim(z_{1,t} v_{1,t}) = \dim(\theta) = q_1$. This estimation method is referred to as the EF method because the q_1-dimensional moment function $z_{1t} v_{1t}$ is known as the (elementary) EF of θ_o in the MM literature; see, e.g., Godambe (1991), Bera and Bilias (2002), and Bera et al. (2006), among many others. Let T be the sample size, and $\hat{\theta}_T$ be an estimator of θ_o. Denote $\hat{v}_{it} := v_{it}|_{\theta=\hat{\theta}_T}$

and $\hat{z}_{it} := z_{it}|_{\theta=\hat{\theta}_T}$ with $i = 1, 2$. In the FTS literature, $\hat{\theta}_T$ is typically obtained by some extreme estimation method that maximizes or minimizes a certain objective function of the unknown parameters, such as the ML method or the QML method. We can link an extreme estimation method to the EF method by interpreting the estimating equation:

$$\frac{1}{T}\sum_{t=1}^{T} \hat{z}_{1t}\hat{v}_{1t} = 0, \tag{10}$$

which is the sample counterpart of the restriction $\mathbb{E}[z_{1,ot}v_{1,ot}] = 0$, as the first-order condition of this estimation method. This interpretation also means that, in practical applications, the closed forms of v_{1t} and z_{1t} are determined by the FTS model being estimated and the extreme estimation method being used.

Given the estimator $\hat{\theta}_T$, the CM testing method checks the null hypothesis (8) by testing the moment restriction set: $\mathbb{E}[z_{2,ot}v_{2,ot}] = 0$. Specifically, it checks whether the sample counterpart of $\mathbb{E}[z_{2,ot}v_{2,ot}]$, that is $T^{-1}\sum_{t=1}^{T} \hat{z}_{2t}\hat{v}_{2t}$, is insignificantly different from the $q_2 \times 1$ zero vector as required by the null hypothesis. This type of generalized test is also known as the CM test after Newey (1985) and Tauchen (1985). Importantly, since the restriction $\mathbb{E}[z_{1,ot}v_{1,ot}] = 0$ has been taken as given in the estimation procedure, we should require the elements of $\mathbb{E}[z_{2,ot}v_{2,ot}]$ to be distinct from those of $\mathbb{E}[z_{1,ot}v_{1,ot}]$ for precluding redundant moments in hypothesis testing. In practical applications, the closed forms of v_{2t} and z_{2t} are determined by the null hypothesis being tested and the alternative hypothesis being considered. The elements of z_{2t} are also known as the "misspecification indicators" in testing CM models when $v_{1t} = v_{2t}$; see Wooldridge (1990). This reflects the fact that the choice of z_{2t} is closely related to the power directions that the test is designed to have in testing CM models.

Note that this MM-based unified approach is related to, but different from, the overidentifying generalized MM (GMM) approach of Hansen (1982). Unlike the former, the latter does not distinguish the moment conditions for estimation from those for specification test. In our problem, this GMM approach would estimate and test the generic model simultaneously, using the same overidentifying moment restrictions: $\mathbb{E}[z_{ot}v_{ot}] = 0$ with $v_{ot} = v_{1,ot} = v_{2,ot}$ and $z_{ot} := (z'_{1,ot}, z'_{2,ot})'$. We do not conduct this approach because FTS analysts tend to first estimate their models before testing their estimated models. This statistical inference procedure is particularly useful when the FTS models are complicated. Meanwhile, by relaxing the restriction: $v_{1,ot} = v_{2,ot}$, our approach is also useful for generating more flexible estimation and testing methods.

To complete the MM-based unified approach, we sketch the derivation of the asymptotic distribution of $T^{1/2}(\hat{\theta}_T - \theta_o)$ under condition (7) and that of $T^{-1/2}\sum_{t=1}^{T}\hat{z}_{2t}\hat{v}_{2t}$ under conditions (7) and (8) in a heuristic way by using the delta method. This asymptotic method is quite standard in econometrics. It includes the consistency of $\hat{\theta}_T$ for θ_o, a mean-value expansion, a uniform law of large numbers, and a central limit theorem as four key ingredients. These ingredients require certain technical conditions that we do not discuss here for simplicity, but are widely available in the literature; see, e.g., Davidson (1994), Newey and McFadden (1994), and White (1994), among many others. In our discussions, we maintain these four ingredients as a set of "high-level" assumptions. Specifically, we assume that $\hat{\theta}_T$ is consistent for θ_o under the specification correctness condition (7). Meanwhile, the statistic $T^{-1/2}\sum_{t=1}^{T}\hat{z}_{it}\hat{v}_{it}$ is assumed to have the expansion:

$$\begin{aligned}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\hat{z}_{it}\hat{v}_{it} = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}z_{i,ot}v_{i,ot} &+ \sum_{j=1}^{r_i}\left[\frac{1}{T}\sum_{t=1}^{T}(\nabla_{\theta'}z_{ij,t})v_{ij,t}\right]_{\theta=\theta_T^{\dagger}}\sqrt{T}(\hat{\theta}_T-\theta_o)\\ &+ \left[\frac{1}{T}\sum_{t=1}^{T}z_{it}V_{it}\right]_{\theta=\theta_T^{\dagger}}\sqrt{T}(\hat{\theta}_T-\theta_o), \qquad (11)\end{aligned}$$

for both $i = 1, 2$; $\theta_T^{\dagger} \in \Theta$ is in the line segment between $\hat{\theta}_T$ and θ_o; $v_{ij,t}$ and $z_{ij,t}$ are, respectively, the jth element of v_{it} and the jth column of z_{it} with $j = 1, \ldots, r_i$ and $i = 1, 2$. The sequences $\{(\nabla_{\theta'}z_{ij,t})v_{ij,t}\}$ and $\{z_{it}V_{it}\}$ are also assumed to be stationary and follow a uniform law of large numbers:

$$\frac{1}{T}\sum_{t=1}^{T}(\nabla_{\theta'}z_{ij,t})v_{ij,t} \xrightarrow{p} \mathbb{E}[(\nabla_{\theta'}z_{ij,t})v_{ij,t}] \qquad (12)$$

and

$$\frac{1}{T}\sum_{t=1}^{T}z_{it}V_{it} \xrightarrow{p} \mathbb{E}[z_{it}V_{it}], \qquad (13)$$

uniformly in Θ, for all $j = 1, \ldots, r_i$ and $i = 1, 2$. Finally, the sequence $\{z_{i,ot}v_{i,ot}\}$ is assumed to follow a martingale-difference central limit theorem:

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T} z_{i,ot}v_{i,ot} \xrightarrow{d} N(0, \mathbb{E}[(z_{i,ot}v_{i,ot})(z_{i,ot}v_{i,ot})']) \tag{14}$$

with $i = 1$ under condition (7) and with $i = 2$ under hypothesis (8).

Given these assumptions, we first use the law of iterated expectations to write that

$$\mathbb{E}[(\nabla_{\theta'} z_{ij,t})v_{ij,t}]_{\theta=\theta_o} = \mathbb{E}[(\nabla_{\theta'} z_{ij,t})\mathbb{E}[v_{ij,t}|\mathcal{X}_{t-1}]]_{\theta=\theta_o} = 0$$

with $i = 1$ under (7) and with $i = 2$ under (8). From this result and the consistency of $\hat{\theta}_T$, we can combine (12) and (13) into (11) and obtain

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T} \hat{z}_{it}\hat{v}_{it} = \frac{1}{\sqrt{T}}\sum_{t=1}^{T} z_{i,ot}v_{i,ot} + \mathbb{E}[z_{i,ot}V_{i,ot}]\sqrt{T}(\hat{\theta}_T - \theta_o) + o_p(1), \tag{15}$$

where $V_{i,ot} := V_{it}|_{\theta=\theta_o}$. Note that this derivation requires v_{it} to be continuously differentiable with respect to θ. In the case where v_{it} is a non-smooth function of θ, one may also obtain (15) by replacing the delta method with the stochastic equicontinuity argument and by re-defining V_{it} as $V_{it} = \nabla_{\theta'}\mathbb{E}[v_{it}|\mathcal{X}_{t-1}]$; see Andrews (1994) and Newey and McFadden (1994) for more details of this argument.

2.2.1 The EF estimator

Using the estimating equation (10), we can simplify (15) as the following:

$$\sqrt{T}(\hat{\theta}_T - \theta_o) = -\mathbb{E}[z_{1,ot}V_{1,ot}]^{-1}\frac{1}{\sqrt{T}}\sum_{t=1}^{T} z_{1,ot}v_{1,ot} + o_p(1) \tag{16}$$

in the case where $i = 1$. From (14) and (16), we can obtain the asymptotic normality:

$$\sqrt{T}(\hat{\theta}_T - \theta_o) \xrightarrow{d} N(0, \textstyle\sum_o), \tag{17}$$

where

$$\textstyle\sum_o := \mathbb{E}[z_{1,ot}V_{1,ot}]^{-1}\mathbb{E}[(z_{1,ot}v_{1,ot})(z_{1,ot}v_{1,ot})']\mathbb{E}[V'_{1,ot}z'_{1,ot}]^{-1}. \tag{18}$$

In practical applications, we may estimate $\sum_o$ using its sample counterpart (that is a heteroskedasticity-consistent covariance matrix estimator):

$$\hat{\sum}_T = \left[\frac{1}{T}\sum_{t=1}^{T} \hat{z}_{1t}\hat{V}_{1t}\right]^{-1}\left[\frac{1}{T}\sum_{t=1}^{T}(\hat{z}_{1t}\hat{v}_{1t})(\hat{z}_{1t}\hat{v}_{1t})'\right]\left[\frac{1}{T}\sum_{t=1}^{T}\hat{V}_{1t}'\hat{z}_{1t}'\right]^{-1},$$

where $\hat{V}_{1t} := V_{1t}|_{\theta=\hat{\theta}_T}$. In what follows, we refer to this generalized estimator as the EF estimator.

2.2.2 The CM test

By combining (16) into (15) with $i = 2$, we obtain that

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\hat{z}_{2t}\hat{v}_{2t} = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\xi_{ot} + o_p(1) \tag{19}$$

with $\xi_{ot} := \xi_t|_{\theta=\theta_o}$ and $\xi_t := z_{2t}v_{2t} - \mathbb{E}[z_{2t}V_{2t}]\mathbb{E}[z_{1t}V_{1t}]^{-1}z_{1t}v_{1t}$. Since ξ_t is a linear combination of $z_{1t}v_{1t}$ and $z_{2t}v_{2t}$, we can utilize (14) and (19) to show that

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\hat{z}_{2t}\hat{v}_{2t} \xrightarrow{d} N(0,\Xi_o), \quad \Xi_o := \mathbb{E}[\xi_{ot}\xi_{ot}']. \tag{20}$$

Let $\hat{\Xi}_T$ be a consistent estimator of Ξ_o, such as

$$\hat{\Xi}_T = \frac{1}{T}\sum_{t=1}^{T}\hat{\xi}_t\hat{\xi}_t', \quad \hat{\xi}_t := \hat{z}_{2t}\hat{v}_{2t} - \left[\frac{1}{T}\sum_{t=1}^{T}\hat{z}_{2t}\hat{V}_{2t}\right]\left[\frac{1}{T}\sum_{t=1}^{T}\hat{z}_{1t}\hat{V}_{1t}\right]^{-1}\hat{z}_{1t}\hat{v}_{1t}.$$

By assuming that $\hat{\Xi}_T$ is uniformly positive definite, we can define the test statistic:

$$M_T := \frac{1}{T}\left[\sum_{t=1}^{T}\hat{z}_{2t}\hat{v}_{2t}\right]'\hat{\Xi}_T^{-1}\left[\sum_{t=1}^{T}\hat{z}_{2t}\hat{v}_{2t}\right]. \tag{21}$$

Following (20), this test statistic has the asymptotic null distribution: $M_T \xrightarrow{d} \chi^2(q_2)$ under H_o with $q_2 := \dim(z_{2t}v_{2t})$. Henceforth, we refer to this generalized test as the CM test.

The MM-based unified approach includes the EF estimator and the CM test as the generalized estimation and testing methods for parametric FTS analysis. Indeed, this approach is generally applicable to various CM and CD models, including but not limited to the FTS models of our interest. As mentioned previously, the derivation of this MM-based unified approach is standard in econometrics. Nonetheless, from this derivation, we can summarize a number of properties that have important implications for the use of parametric estimation and testing methods for FTS analysis.

First, the consistency and the asymptotic normality of the EF estimator $\hat{\theta}_T$ mainly result from the specification correctness condition (7). In the presence of model misspecification, condition (7) and hence the asymptotic result in (17) are no longer valid. This means that it is important to check the specification correctness of the FTS model being estimated before concluding the empirical findings from the estimated results. The CM test with $v_{1t} = v_{2t}$ is particularly important in this sense.

Second, the aforementioned point also means that the EF estimator $\hat{\theta}_T$ for the generic CM model is robust to the unknown $F_t(\cdot|\mathcal{X}_{t-1})$ because its consistency and asymptotic normality are derived without requiring any specific form of $F_t(\cdot|\mathcal{X}_{t-1})$. This robustness is an important feature that distinguishes the Gaussian QML method (the Gaussian QML-based CM tests) from the ML method (the ML-based CM tests) in the context of partially specified GARCH-type models, as we will discuss in Section 3.

Third, the asymptotic expansion (15) shows that the statistic $T^{-1/2}\sum_{t=1}^{T}\hat{z}_{2t}\hat{v}_{2t}$ is asymptotically different from its population counterpart $T^{-1/2}\sum_{t=1}^{T} z_{2,ot}v_{2,ot}$, unless $\mathbb{E}[z_{2,ot}V_{2,ot}] = 0$. This discrepancy reflects the effect of parameter estimation uncertainty. This effect is typically model-specific and estimation-method-specific, and has been taken into account by the asymptotic covariance matrix Ξ_o. Since (21) is based on a consistent estimator of Ξ_o, the CM test is robust to the estimation effect. Thus, it is useful for generating the estimation-effect-modified counterparts of some conventional tests, such as the Ljung-Box (1978) and McLeod-Li (1983) tests, that were originally derived for conditionally homoskedastic time series models but are widely applied for evaluating various FTS models; see also Section 3.1 for more discussion on this point.

3. FINANCIAL TIME SERIES MODELS AND METHODS

In this section, we review four classes of FTS models that are spanned from the GARCH model, including univariate partially specified (GARCH-type) models, univariate fully specified models, multivariate partially specified models (CCC/DCC models), and multivariate fully specified models (CMD models). In addition, we also briefly review univariate MEMs that are also closely related to univariate GARCH-type models. These FTS models have attracted a wide range of applications in empirical finance. As we will see later, although these models are proposed for various economic and statistical reasons, they can be systematically unified into the generic CM or CD model. This means that we can show the associated parametric estimation methods (testing methods) as particular examples of the EF method (the CM testing method) by choosing the associated v_{1t} and z_{1t} (v_{2t} and z_{2t}) in an appropriate way. Put differently, the MM-based unified approach includes the generalized residual, the EF method, and the CM testing method as the general principles underlying these seemingly different models and methods.

3.1 Univariate Partially Specified Models

In the context of univariate partially specified GARCH-type models, $\{y_t\}$ is a stationary sequence of financial returns with the time index t and the dimension $n = 1$. This class of models are of the general form:

$$y_t = m_t(X_t, \alpha) + u_t, \quad u_t = h_t(X_t, \alpha)^{\frac{1}{2}} \varepsilon_t, \tag{22}$$

where $m_t := m_t(X_t, \alpha)$ is a specification of $\mathbb{E}[y_t|\mathcal{X}_{t-1}]$, $h_t := h_t(X_t, \alpha)$ is a specification of $\text{var}[y_t|\mathcal{X}_{t-1}]$, u_t is a conditionally heteroskedastic error with $\mathbb{E}[u_t] = 0$, and ε_t represents a standardized error with $\mathbb{E}[\varepsilon_t] = 0$ and $\text{var}[\varepsilon_t] = 1$. In this context, the generic CM model $\mu_t(X_t, \alpha)$ degenerates to a generalized conditional mean-and-variance model for $M_t(y_t|\mathcal{X}_{t-1}) = (\mathbb{E}[y_t|\mathcal{X}_{t-1}], \text{var}[y_t|\mathcal{X}_{t-1}])'$, in which $\mathbb{E}[y_t|\mathcal{X}_{t-1}]$ characterizes the (mean) predictability of financial returns and $\text{var}[y_t|\mathcal{X}_{t-1}]$ measures the volatility of financial returns. Various GARCH-type models provide various approximations to the unknown $\mathbb{E}[y_t|\mathcal{X}_{t-1}]$ and $\text{var}[y_t|\mathcal{X}_{t-1}]$ using various m_t's and h_t's.

This class of models are quite popular in empirical finance because modelling financial volatility is undoubtedly important for asset pricing, efficient inference of return predictability, and many other problems. Moreover, they are also useful for explaining the distributional and dynamic characteristics of financial return series. It is well-documented that financial returns have the stylized facts of heavy-tails (leptokurtic distributions) and volatility clustering (serially correlated absolute or squared return sequences); see, e.g., Mandelbrot (1963). Stock return series may also have volatility asymmetry (namely, the leverage effect; a negative past return shock causes a greater impact on the current volatility than a positive return shock of the same magnitude); see, e.g., Black (1976).

Note that the conditionally heteroskedastic error u_t in (22) is the product of $h_t^{1/2}$ and ε_t, and hence it has the kurtosis coefficient: $\mathbb{E}[u_t^4]\mathbb{E}[u_t^2]^{-2} = \mathbb{E}[\varepsilon_t^4]\mathbb{E}[h_t^2]\mathbb{E}[h_t]^{-2} = \mathbb{E}[\varepsilon_t^4]\left(1 + \text{var}[h_t]\mathbb{E}[h_t]^{-2}\right)$ when ε_t is independent of $\mathcal{X}_{t-1}$. This kurtosis exceeds $\mathbb{E}[\varepsilon_t^4]$ provided that h_t is a positive random variable. This means that the distribution of u_t tends to be leptokurtic whenever $\mathbb{E}[\varepsilon_t^4] \geq 3$, and shows that GARCH-type models are useful for interpreting (for fitting) the stylized fact of heavy-tails; see, e.g., Engle and González-Rivera (1991, p. 21) for this argument. Meanwhile, these models are also useful for explaining the dynamic features of return series because they include dynamic conditional variance specifications as their key settings. Specifically, Bollerslev's (1986) GARCH$(1,1)$ model has the conditional variance specification:

$$h_t = \alpha_{h0} + \alpha_{h1}h_{t-1} + \alpha_{h2}u_{t-1}^2,$$

where α_{h0}, α_{h1}, and α_{h2} denote unknown parameters, that allows the sequence $\{u_t^2\}$ to be serially correlated and hence is useful for characterizing the stylized fact of volatility clustering. This model includes the IGARCH model as a special case where $\alpha_{h1} + \alpha_{h2} = 1$. In addition, Nelson's (1991) EGARCH$(1,1)$ and Glosten et al.'s (1993) GJR-GARCH$(1,1)$ models are, respectively, of the conditional variance specifications:

$$h_t = \exp\left(\alpha_{h0} + \alpha_{h1}\ln h_{t-1} + \alpha_{h2}|\varepsilon_{t-1}| + \alpha_{h3}\varepsilon_{t-1}\right)$$

and

$$h_t = \alpha_{h0} + \alpha_{h1}h_{t-1} + (\alpha_{h2} + \alpha_{h3}I(u_{t-1} < 0))u_{t-1}^2,$$

with α_{h3} denoting an additional parameter, that allow a negative past return shock to generate a different impact on the current volatility from its positive counterpart when $\alpha_{h_3} \neq 0$. Thus, these "asymmetric" GARCH-type models are useful for explaining not only volatility clustering but also volatility asymmetry. We refer to Hentschel (1995), Duan (1997), and the references therein for many other different h_t's available in the FTS literature and for some nesting families of these h_t's.

It should be addressed that, for any partially specified GARCH-type model, (m_t, h_t) is simply an approximation to the unknown $(\mathbb{E}[y_t|\mathcal{X}_{t-1}], \mathrm{var}[y_t|\mathcal{X}_{t-1}])$, which is not automatically ensured to be correctly specified. The asymptotic properties of its estimators require the assumption of specification correctness, and this assumption should be checked by reasonable tests. In the following, we demonstrate that the EF estimator and the CM test (with $\theta = \alpha$) are useful for these purposes.

Using the fact that model (22) is a particular example of the generic CM model with $\mu_t(X_t, \alpha) = (m_t, h_t)'$, it is straightforward to obtain a particular form of the generalized residual: $v_{1t} = (u_t, u_t^2 - h_t)' = \left(h_t^{1/2}\varepsilon_t, h_t(\varepsilon_t^2 - 1)\right)'$. Since h_t is $\mathcal{X}_{t-1}$-measurable, we can re-define this v_{1t} as

$$v_{1t} = \left(\varepsilon_t, \varepsilon_t^2 - 1\right)' \tag{23}$$

for convenience, without distorting the martingale difference property of (7). Denote $\varepsilon_{ot} := \varepsilon_t|_{\alpha=\alpha_o}$. From (7) and (23), we can see that this generalized model is correctly specified under the conditional mean and variance restrictions: $\mathbb{E}[\varepsilon_{ot}|\mathcal{X}_{t-1}] = 0$ and $\mathbb{E}[\varepsilon_{ot}^2|\mathcal{X}_{t-1}] = 1$. In empirical studies, it is common to estimate partially specified GARCH-type models using the ML method. This indeed requires an additional CD assumption which is unable to be verified under this specification correctness condition, and is hence not robust to the potential failure of the CD assumption. To circumvent this non-robustness problem, Bollerslev and Wooldridge (1992) suggested replacing the ML method with the Gaussian QML method for estimating GARCH-type models. Specifically, this QML method is an extreme estimation method that estimates $\theta_o = \alpha_o$ by maximizing the Gaussian "quasi"-log-likelihood function:

$$L_T(\alpha) = -\frac{1}{2}\ln(2\pi) - \frac{1}{2T}\sum_{t=1}^{T}\ln h_t - \frac{1}{2T}\sum_{t=1}^{T}\frac{(y_t - m_t)^2}{h_t}. \tag{24}$$

This is the same as the log-likelihood function of the Gaussian ML method. However, unlike the Gaussian ML method, the Gaussian QML method derives the asymptotic normality of its estimator without using any specific CD assumption except for the specification correctness condition: $\mathbb{E}[\varepsilon_{ot}|\mathcal{X}_{t-1}] = 0$ and $\mathbb{E}[\varepsilon_{ot}^2|\mathcal{X}_{t-1}] = 1$ (and certain technical conditions). Thus, the Gaussian QML estimator (QMLE) is robust to the unknown CD of $y_t|\mathcal{X}_{t-1}$ in the partially specified context. This property is indeed resulted from the fact that the Gaussian QML method is a particular EF estimator. This point can be seen from the fact that the Gaussian QML method has the estimating equation (the first-order condition):

$$\frac{1}{T}\sum_{t=1}^{T}\left((\nabla_\alpha m_t)h_t^{-\frac{1}{2}}\varepsilon_t + \frac{1}{2}(\nabla_\alpha h_t)h_t^{-1}(\varepsilon_t^2 - 1)\right)\bigg|_{\alpha=\hat{\alpha}_T} = 0, \tag{25}$$

which is a particular case of (10) with the generalized residual v_{1t} shown in (23) and the following $q_1 \times 2$ matrix of instrumental variables:

$$z_{1t} = \left[(\nabla_\alpha m_t)h_t^{-\frac{1}{2}}, \quad \frac{1}{2}(\nabla_\alpha h_t)h_t^{-1}\right]. \tag{26}$$

Put differently, the Gaussian QMLE $\hat{\alpha}_T$ is a special EF estimator with (23) and (26).

Using the EF method, we can further recover the asymptotic normality of the Gaussian QMLE $\hat{\alpha}_T$ under the specification correctness condition by plugging the aforementioned v_{1t} and z_{1t} into (17) and (18) with $\theta = \alpha$. The resulting $\sum_o$ is of the component:

$$\begin{aligned}\mathbb{E}[z_{1,ot}V_{1,ot}] &= \mathbb{E}[z_{1,ot}\mathbb{E}[V_{1,ot}|\mathcal{X}_{t-1}]] \\ &= -\left(\mathbb{E}[(\nabla_\alpha m_t \nabla_{\alpha'} m_t)h_t^{-1}] + \frac{1}{4}\mathbb{E}[(\nabla_\alpha h_t \nabla_{\alpha'} h_t)h_t^{-2}]\right)_{\alpha=\alpha_o},\end{aligned}$$

in which the first equality is due to the law of iterated expectations and the fact that $z_{1,ot}$ is $\mathcal{X}_{t-1}$-measurable; the second equality is due to (7), (23), (26), and the associated $2 \times q_1$ matrix of derivatives:

$$V_{1t} = \begin{bmatrix} -(\nabla_{\alpha'} m_t)h_t^{-\frac{1}{2}} - \frac{1}{2}(\nabla_{\alpha'} h_t)h_t^{-1}\varepsilon_t \\ -(\nabla_{\alpha'} m_t)h_t^{-\frac{1}{2}}\varepsilon_t - \frac{1}{2}(\nabla_{\alpha'} h_t)h_t^{-1}\varepsilon_t^2 \end{bmatrix}. \tag{27}$$

This recovers the asymptotic normality of the Gaussian QMLE, which has been derived by Bollerslev and Wooldridge (1992) in a more general multivariate context.

The choice of z_{1t} is indeed an essential issue in terms of asymptotic efficiency. According to the ML theory, the z_{1t} determined by the Gaussian QML method, namely (26), is asymptotically efficient if the true CD is normal because the EF $z_{1t}v_{1t}$ becomes the true score function and the asymptotic covariance matrix of $\hat{\alpha}_T$ attains the Cramér-Rao lower bound in this case. However, this optimality may not hold in the presence of conditional non-normality. By applying the optimal EF method developed in the statistical literature after Godambe (1960) and Durbin (1960), Li and Turtle (2000, Equation 14) contributed a different choice of z_{1t} which includes (26) as a particular case where the conditional skewness and kurtosis of $\varepsilon_{ot}|\mathcal{X}_{t-1}$ are, respectively, equal to zero and three. Given their z_{1t}, the resulting estimator is asymptotically more efficient, or at least as efficient as, any other EF estimators. Importantly, unlike the ML theory, this optimality is obtained under the specification correctness condition $\mathbb{E}[v_{1,ot}|\mathcal{X}_{t-1}] = 0$ rather than under the entire CD restriction. Nonetheless, because the use of this optimal EF method needs to estimate the unknown conditional skewness and kurtosis (which are the sources of the relative efficiency of this method over the Gaussian QML method), it is still rare to see practical applications of this method in the FTS literature. We refer to Bera et al. (2006) for a recent survey on the optimal EF method.

Following the CM testing method, we can also easily generate a number of Gaussian QML-based tests for partially specified GARCH-type models using the aforementioned v_{1t} and z_{1t}. By setting $v_{2t} = v_{1t}$, the resulting CM test is applicable to testing the specification correctness of model (22) against various types of misspecification by choosing various z_{2t}'s. For example, we may test a GARCH-type model against remaining serial correlation by setting z_{2t} as the following $q \times 2$ matrix:

$$z'_{2t} = \begin{bmatrix} \varepsilon_{t-1}, & \ldots, & \varepsilon_{t-q} \\ 0, & \ldots, & 0 \end{bmatrix},$$

against remaining volatility clustering by setting z_{2t} as another $q \times 2$ matrix:

$$z'_{2t} = \begin{bmatrix} 0, & \ldots, & 0 \\ \varepsilon^2_{t-1} - 1, & \ldots, & \varepsilon^2_{t-q} - 1 \end{bmatrix},$$

or against these two types of misspecification simultaneously by combining these two z_{2t}'s as a new $2q \times 2$ matrix:

$$z_{2t}' = \begin{bmatrix} \varepsilon_{t-1}, & \ldots, & \varepsilon_{t-q}, & 0, & \ldots, & 0 \\ 0, & \ldots, & 0, & \varepsilon_{t-1}^2 - 1, & \ldots, & \varepsilon_{t-q}^2 - 1 \end{bmatrix}.$$

The resulting test statistics are, respectively, of the asymptotic null distributions: $\chi^2(q)$, $\chi^2(q)$, and $\chi^2(2q)$, and have various power directions.

In the FTS literature, Li and Mak (1994) have introduced a Gaussian ML-based volatility clustering test; see also Lundbergh and Teräsvirta (2002) for its score test interpretation and Berkes et al. (2003) for its Gaussian QML counterpart. In addition, Wong and Ling (2005) also considered a Gaussian QML-based omnibus test for serial correlation and volatility clustering. These tests modify the widely used Ljung-Box and McLeod-Li tests by correcting the estimation effect of GARCH-type models ignored by these two popular tests. However, these modified tests are derived under the null hypothesis that ε_{ot} is independent of $\mathcal{X}_{t-1}$. This hypothesis is stricter than the specification correctness condition (7) with (23), and is incompatible with the case where this specification correctness condition is satisfied but ε_{ot} is dependent on $\mathcal{X}_{t-1}$ through certain higher-order CMs. This case includes Hansen's (1994) autoregressive conditional density model as an important example. In comparison, our serial correlation and volatility clustering tests are free of this problem; see also Wooldridge (1990) for a related robustness issue.

The CM test can also check (22) against remaining nonlinearity-in-mean, such as the GARCH-in-mean effect, by choosing

$$z_{2t}' = \begin{bmatrix} \varepsilon_{t-1}^2 - 1, & \ldots, & \varepsilon_{t-q}^2 - 1 \\ 0, & \ldots, & 0 \end{bmatrix}$$

and against remaining volatility asymmetry by using

$$z_{2t}' = \begin{bmatrix} 0, & \ldots, & 0 \\ \varepsilon_{t-1}, & \ldots, & \varepsilon_{t-q} \end{bmatrix};$$

see also Chen (2008a) for more discussions and for a Gaussian QML-based generalized test for the hypothesis of independence. Note that, similar to the score test (the news impact test) of Engle and Ng (1993) and the time irreversibility test of Chen and Kuan (2002), this volatility asymmetry test is useful for testing the GARCH model against asymmetric GARCH-type models. In addition, we can also use the CM test to generate Chen's (2008a) test by allowing v_{2t} to be a finite-dimensional transformation of ε_t, which could be different from (23), and by deriving the resulting Ξ_o using the law of iterated expectations and the hypothesis of independence. If the specification correctness condition (7) with (23) is accepted as satisfied, we may further apply such independence tests to check whether the higher-order CMs of $\varepsilon_{ot}|\mathcal{X}_{t-1}$ are time-varying. This detection may provide useful information for building fully specified GARCH-type models in an appropriate way.

3.2 Univariate Fully Specified Models

Given (22), it is straightforward to establish a univariate fully specified GARCH-type model by making an additional "standardized" CD specification of $\varepsilon_t|\mathcal{X}_{t-1}$, denoted by $G_\varepsilon(\cdot|X_t, \gamma)$, with the zero mean and unit variance, the conditional PDF $g_\varepsilon(\cdot|X_t, \gamma)$, and the parameter vector γ. In this context, the generic CD model $G_t(\cdot|X_t, \beta)$ degenerates to the following form:

$$G_t(y_t|X_t, \beta) = G_\varepsilon\left(\left.\frac{y_t - m_t(X_t, \alpha)}{h_t(X_t, \alpha)^{\frac{1}{2}}}\right| X_t, \gamma\right), \tag{28}$$

with $\beta = (\alpha', \gamma')'$ and the conditional PDF:

$$g_t(y_t|X_t, \beta) = h_t(X_t, \alpha)^{-\frac{1}{2}} g_\varepsilon\left(\left.\frac{y_t - m_t(X_t, \alpha)}{h_t(X_t, \alpha)^{\frac{1}{2}}}\right| X_t, \gamma\right). \tag{29}$$

Given the conditional mean-and-variance specification, various fully specified GARCH-type models provide various approximations to the unknown $F_t(\cdot|\mathcal{X}_{t-1})$ using various $G_\varepsilon(\cdot|X_t, \gamma)$'s.

In addition to measuring the predictability and volatility and to explaining the stylized facts (such as heavy-tails, volatility clustering, and volatility asymmetry) of financial returns, there are other important financial and statistical problems that are oriented to certain higher-order CMs or even the entire CD of financial returns. Thus, it is essential to extend GARCH-type models from the partially specified conditional mean-and-variance context to the fully specified CD context. For example, Calzolari and Fiorentini (1998) provided an approximated Tobit-GARCH model which is applicable to explore financial under price limitations. Their approximated model replaces the censored lagged return shocks in a GARCH specification with the conditional expectations of these variables. These conditional expectations must be defined using a fully specified CD assumption. It is also important to explore the skewness of $F_t(\cdot|\mathcal{X}_{t-1})$ for studying the skewed-preference investing problems, see, e.g., Harvey and Siddique (1999, 2000) and Patton (2004), and to study the tail properties of $F_t(\cdot|\mathcal{X}_{t-1})$ for risk management. For the latter, we can define the conditional τ-quantile of $y_t|\mathcal{X}_{t-1}$, denoted as $Q_{F_t}(\tau|\mathcal{X}_{t-1})$ for some $\tau \in (0,1)$, using the restriction:

$$\int_{-\infty}^{Q_{F_t}(\tau|\mathcal{X}_{t-1})} f_t(y|\mathcal{X}_{t-1})dy = \int_{\mathbb{R}} I(y \leq Q_{F_t}(\tau|\mathcal{X}_{t-1}))f_t(y|\mathcal{X}_{t-1})dy = \tau. \quad (30)$$

Correspondingly, we can also define the conditional expected shortfall:

$$\mathbb{E}[y_t|y_t < Q_{F_t}(\tau|\mathcal{X}_{t-1}), \mathcal{X}_{t-1}] = \frac{1}{\tau}\int_{-\infty}^{Q_{F_t}(\tau|\mathcal{X}_{t-1})} y f_t(y|\mathcal{X}_{t-1})dy. \quad (31)$$

These two higher-order CMs are widely used Value-at-Risk (VaR) measures in empirical finance; see, e.g., Kupiec (1995), Artzner et al. (1999), and Christoffersen (2003) for related discussions. In addition, the accuracy of an interval forecast also requires the information of $Q_{F_t}(\tau_1|\mathcal{X}_{t-1})$ and $Q_{F_t}(\tau_2|\mathcal{X}_{t-1})$ for some $0 < \tau_1 < \tau_2 < 1$; see, e.g., Christoffersen (1998). For the correct density forecast, we even need the entire $F_t(\cdot|\mathcal{X}_{t-1})$ to maintain the robustness of the forecasting optimality to various loss functions; see, e.g., Diebold et al. (1998). This CD and the associated higher-order CMs are all unknown in practice. Fully specified GARCH-type models are useful for generating various approximations to $F_t(\cdot|\mathcal{X}_{t-1})$ using various forms of (28).

In the FTS literature, researchers have considered a wide range of $G_\varepsilon(\cdot|X_t,\gamma)$ to establish fully specified GARCH-type models, including the static $G_\varepsilon(\cdot|X_t,\gamma)$'s that are free of X_t and the dynamic $G_\varepsilon(\cdot|X_t,\gamma)$'s that are dependent on X_t. Examples of the static distributions include the standard normal distribution of Engle (1982) and Bollerslev (1986) which does not contain the parameter γ, the standardized t distribution of Bollerslev (1987) which includes the degrees of freedom as its γ, the generalized error distribution of Nelson (1991), the exponential generalized Gamma type-II distribution of Wang et al. (2001), the asymmetric exponential power distribution of Komunjer (2007), and the maximum entropy distribution of Park and Bera (2009), among others. Representative examples of dynamic distributions include the skewed t distribution of Hansen (1994), the generalized t distribution of Harvey and Siddique (1999, 2000), and the conditional skewness-kurtosis-based maximum entropy distribution of Rockinger and Jondeau (2002), among others. These dynamic distributions allow their parameters to follow certain X_t-dependent dynamic specifications. For example, Hansen's (1994) autoregressive conditional density model allows the parameters of the skewed t distribution to be dependent on the ε_{t-k}'s. In such a dynamic setting, a fully specified GARCH-type model can permit not only its conditional mean and variance but also its higher-order CMs to be time-varying.

Nonetheless, it should be reminded that, for any fully specified GARCH-type model, $G_\varepsilon\left((y_t - m_t)h_t^{-1/2}|X_t, \gamma\right)$ is simply an approximation to the unknown $F_t(y_t|\mathcal{X}_{t-1})$. The consistency and asymptotic normality of the estimators of this model require the assumption of specification correctness, and this assumption should be evaluated by specification tests. In the following, we demonstrate that the MM-based approach (with $\theta = \beta$) is also useful for unifying existing parametric estimation and testing methods in this context.

Using the fact that model (28) is a particular example of the generic CD model, we can obtain a particular version of the specification correctness condition (7) with the generalized residual:

$$\begin{aligned} v_{1t} &= \ell_{\beta,t} := \nabla_\beta \ln g_t(y_t|X_t, \beta) \\ &= -\frac{1}{2}(\nabla_\beta h_t)h_t^{-1} + \nabla_\beta \ln g_\varepsilon\left((y_t - m_t)h_t^{-\frac{1}{2}}\middle| X_t, \gamma\right), \end{aligned} \tag{32}$$

in which $\nabla_\beta = (\nabla_{\alpha'}, \nabla_{\gamma'})'$, and the last equality is due to (29). Denote $\ell_{\beta,ot} := \ell_{\beta,t}|_{\beta=\beta_o}$. In the context of fully specified models, it is standard to estimate the parameter vector β_o using the ML method. This extreme estimation method has the log-likelihood function:

$$\frac{1}{T}\sum_{t=1}^{T}\ln g_t(y_t|X_t,\beta) \;=\; -\frac{1}{2T}\sum_{t=1}^{T}\ln h_t(X_t,\alpha) \\ -\frac{1}{T}\sum_{t=1}^{T}\ln g_\varepsilon\left(\frac{y-m_t(X_t,\alpha)}{h_t(X_t,\alpha)^{\frac{1}{2}}}\middle|X_t,\gamma\right) \tag{33}$$

and the estimating equation:

$$\frac{1}{T}\sum_{t=1}^{T}\ell_{\beta,t}\Big|_{\beta=\hat{\beta}_T}=0, \tag{34}$$

where $\hat{\beta}_T$ denotes the MLE for β_o. Clearly, this estimating equation is a particular example of (10) with the generalized residual v_{1t} shown in (32) and $z_{1t}=I_{q_1}$, where I_{q_1} denotes a $q_1\times q_1$ identity matrix and $q_1=\dim(\beta)$. This demonstrates that the MLE can also be interpreted as a particular EF estimator.

Note that, unlike the aforementioned models, the regime-switching GARCH model of Hamilton and Susmel (1994) is a different type of fully specified model that includes a finite-state Markov chain in its CD specification. Nonetheless, by weighting out this unobserved component, their model is also of a well-defined likelihood function for estimation and testing. Thus, it can also be unified into the generic model with $v_{1t}=\ell_{\beta,t}$, though the last equality in (32) is not applicable here. We refer to Hamilton (1996) for the derivation of $\ell_{\beta,t}$ of the Markov-switching models.

Following the EF method, we can obtain the asymptotic normality of the MLE under the assumption of specification correctness, by introducing the aforementioned v_{1t} and z_{1t} into (17) with $\theta=\beta$. In this particular example, we have $V_{1t}=\nabla_{\beta'}\ell_{\beta,t}$, and hence we can write the associated $\sum_o$ as:

$$\sum_o=\Big(\mathbb{E}[\nabla_{\beta'}\ell_{\beta,t}]^{-1}\mathbb{E}[\ell_{\beta,t}\ell'_{\beta,t}]\mathbb{E}[\nabla_{\beta'}\ell_{\beta,t}]^{-1}\Big)_{\beta=\beta_o}.$$

Interestingly, $\mathbb{E}[\nabla_{\theta'}\ell_{\beta,t}]_{\beta=\beta_o}$ and $\mathbb{E}[\ell_{\beta,ot}\ell'_{\beta,ot}]$ are, respectively, the expected Hessian matrix and the information matrix. By taking the partial derivative of the specification correctness condition: $\mathbb{E}[\ell_{\beta,ot}|\mathcal{X}_{t-1}]=0$ with respect to β and by assuming that the differentiation and integration operators are interchangeable, we can obtain the CM restriction:

$$\mathbb{E}[\nabla_{\theta'}\ell_{\beta,t}|\mathcal{X}_{t-1}]_{\beta=\beta_o} + \mathbb{E}[\ell_{\beta,ot}\ell'_{\beta,ot}|\mathcal{X}_{t-1}] = 0,$$

which implies the information matrix equality: $\mathbb{E}[\nabla_{\theta'}\ell_{\beta,t}]_{\beta=\beta_o}+\mathbb{E}[\ell_{\beta,ot}\ell'_{\beta,ot}]=0$. Using this equality, we can further simplify the aforementioned $\sum_o$ as $\sum_o=\mathbb{E}[\ell_{\beta,ot}\ell'_{\beta,ot}]^{-1}$. This leads us to the well-known asymptotic covariance matrix of the MLE in the standard ML theory; see, e.g., White (1994).

Following the CM testing method, we can also generate various ML-based specification tests for fully specified GARCH-type models by setting $v_{1t} = \ell_{\beta,t}$ and $z_{1t} = I_{q_1}$. The leading example is the classical score test. Let $\tilde{g}_t(y_t|\tilde{X}_t, \tilde{\beta})$ be a postulated conditional PDF of $y_t|\mathcal{X}_{t-1}$ with $\tilde{X}_t$ denoting a vector of $\mathcal{X}_{t-1}$-measurable variables and $\tilde{\beta} := (\beta', \delta')'$ denoting a parameter vector with $\dim(\delta) = q_2$. Suppose that $\tilde{g}_t(y_t|\tilde{X}_t, \tilde{\beta})$ includes $g_t(y_t|X_t, \beta)$ as a special case where $\delta=0_{q_2\times 1}$. Given this nesting relationship, we can test the null of specification correctness for $g_t(y_t|X_t, \beta)$ against the alternative hypothesis $\tilde{g}_t(y_t|\tilde{X}_t, \tilde{\beta})$ by checking the parameter restriction: $\delta = 0_{q_2\times 1}$ using the score test. Denote $\ell_{\delta,t} := \nabla_\delta \ln \tilde{g}_t(y_t|\tilde{X}_t, \tilde{\beta})$, $\ell_{\delta,ot} := \ell_{\delta,t}|_{\beta=\beta_o,\delta=0_{q_2\times 1}}$, and $\hat{\ell}_{\delta,t} := \ell_{\delta,t}|_{\beta=\hat{\beta}_T,\delta=0_{q_2\times 1}}$. The score test statistic for the parameter restriction $\delta = 0_{q_2\times 1}$ is of the form:

$$\left[\sum_{t=1}^{T}\hat{\ell}_{\delta,t}\right]'\left(\left[\sum_{t=1}^{T}\hat{\ell}_{\delta,t}\hat{\ell}'_{\delta,t}\right]-\left[\sum_{t=1}^{T}\hat{\ell}_{\delta,t}\hat{\ell}'_{\beta,t}\right]\left[\sum_{t=1}^{T}\hat{\ell}_{\beta,t}\hat{\ell}'_{\beta,t}\right]^{-1}\left[\sum_{t=1}^{T}\hat{\ell}_{\beta,t}\hat{\ell}'_{\delta,t}\right]\right)^{-1}\left[\sum_{t=1}^{T}\hat{\ell}_{\delta,t}\right], \tag{35}$$

and has the asymptotic null distribution $\chi^2(q_2)$. This test is indeed a particular CM test with the aforementioned v_{1t} and z_{1t} and the additional setting: $v_{2t} = \ell_{\delta,t}$ and $z_{2t} = I_{q_2}$. To see this point, note that this pair of v_{2t} and z_{2t} generates $T^{-1/2}\sum_{t=1}^{T}\hat{z}_{2t}\hat{v}_{2t} = T^{-1/2}\sum_{t=1}^{T}\hat{\ell}_{\delta,t}$, $V_{2t} = \nabla_{\beta'}\ell_{\delta,t}$, and $\xi_t = \ell_{\delta,t} - \mathbb{E}[\nabla_{\beta'}\ell_{\delta,t}]\mathbb{E}[\nabla_{\beta'}\ell_{\beta,t}]^{-1}\ell_{\beta,t}$. Using the information matrix equality: $\mathbb{E}[\nabla_{\beta'}\ell_{\beta,t}]_{\beta=\beta_o} = -\mathbb{E}[\ell_{\beta,ot}\ell'_{\beta,ot}]$ and the generalized information matrix equality: $\mathbb{E}[\nabla_{\beta'}\ell_{\delta,t}]_{\beta=\beta_o,\delta=0_{q\times 1}} = -\mathbb{E}[\ell_{\delta,ot}\ell'_{\beta,ot}]$, we can write the above ξ_t as

$$\xi_t = \ell_{\delta,t} - \mathbb{E}[\ell_{\delta,t}\ell'_{\delta,t}]\mathbb{E}[\ell_{\beta,t}\ell'_{\beta,t}]^{-1}\ell_{\beta,t}$$

and obtain that

$$\Xi_o = \mathbb{E}[\ell_{\delta,ot}\ell'_{\delta,ot}] - \mathbb{E}[\ell_{\delta,ot}\ell'_{\beta,ot}]\mathbb{E}[\ell_{\beta,ot}\ell'_{\beta,ot}]^{-1}\mathbb{E}[\ell_{\beta,ot}\ell'_{\delta,ot}]. \tag{36}$$

From these results, it is easy to see that (35) is encompassed by (21) when we write $\hat{\Xi}_T$ as the sample counterpart of (36).

In addition to the score test, certain higher-order moment tests are also applicable to testing fully specified GARCH-type models. For instance, we may check these models using the information matrix test which is based on the information matrix equality $\mathbb{E}[\nabla_{\beta'}\ell_{\beta,t}]_{\beta=\beta_o} + \mathbb{E}[\ell_{\beta,ot}\ell'_{\beta,ot}] = 0$. This test can be interpreted as a particular CM test with the choice of $v_{2t} = \text{vech}(\ell_{\beta,t}\ell'_{\beta,t} - \nabla_{\beta'}\ell_{\beta,t})$, $z_{2t} = I_{q_2}$, and $q_2 = \dim(v_{2t})$, where vech$(\cdot)$ stands for the column stacking operator of the lower part of a symmetric matrix. We may also derive a conditional skewness-kurtosis test in this context by setting $v_{2t} = (\varepsilon_t^3 - s_t, \varepsilon_t^4 - k_t)'$, in which s_t and k_t are, respectively, the conditional skewness and kurtosis specifications implied by the model being tested. For instance, we can extend the widely used Jarque-Bera (1980) test from the context of linear regressions to the context of fully specified GARCH-type models by setting $s_t = 0$, $k_t = 3$, and $z_{2t} = I_2$ (a certain X_t-dependent z_{2t}) for testing the conditional normality assumption against static (dynamic) conditional non-normal distributions; see also Kiefer and Salmon (1983) for a simple model-invariant alternative to the Jarque-Bera test, which is directly applicable to the context of GARCH-type models. Bontemps and Meddahi (2005) and Chen (2008b) offer different interpretations of the robustness of the Kiefer-Salmon test to the model-specific estimation effect.

Moreover, it is also important to discuss Rosenblatt's (1952) probability integral transformation (PIT) for testing fully specified GARCH-type models or other CD models. Indeed, this transformation also plays a crucial role in studying the CMD models that are discussed in Section 3.4. Specifically, the true PIT of $y_t|\mathcal{X}_{t-1}$ is defined by $\wp_t^* := F_t(y_t|\mathcal{X}_{t-1})$. This true CD transformation of y_t has an important implication for specification test:

$$P(\wp_t^* < \tau|\mathcal{X}_{t-1}) = P(y_t < Q_{F_t}(\tau|\mathcal{X}_{t-1})|\mathcal{X}_{t-1}) = \tau, \quad \forall\, \tau \in (0,1). \tag{37}$$

This means that the true PIT $\wp_t^*$ is $U(0,1)$-distributed and independent of $\mathcal{X}_{t-1}$. Under the hypothesis of specification correctness, we have $F_t(\cdot|\mathcal{X}_{t-1}) = G_t(\cdot|X_t, \beta_o)$ and hence the modelled PIT $\wp_{ot} := \wp_t|_{\beta=\beta_o}$, with $\wp_t := G_t(y_t|X_t, \beta)$, should be the same as $\wp_t^*$ and hence be $U(0,1)$-distributed and independent of $\mathcal{X}_{t-1}$. This motivates

Diebold et al. (1998), Berkowitz (2001), and Ghosh and Bera (2005), among others to develop the PIT-based moment tests for density forecast evaluation, see also Chen (2010) for their estimation-effect-modified counterparts. These PIT-based moment tests can all be presented as particular examples of the ML-based CM test by setting the associated v_{2t}'s as certain functions of $\wp_t$ and the associated z_{2t} as appropriate identity matrices (X_t-dependent matrices) for testing the uniformity (for testing the independence) of the PIT sequence $\{\wp_{ot}\}$.

3.3 Multivariate Partially Specified Models: CCC/DCC Models

In the context of multivariate partially specified GARCH-type models, $\{y_t\}$ becomes a n-dimensional stationary sequence of financial returns with $y_t = (y_{1t}, \ldots, y_{nt})'$ for some finite $n > 1$, in which y_{it} represents the return of the ith asset at time t. It is also straightforward to extend (22) from the univariate context to the multivariate context by re-defining $m_t = m_t(X_t, \alpha)$ as a $n \times 1$ vector $m_t = (m_{1t}, \ldots, m_{nt})'$ with $m_{it} := m_{it}(X_t, \alpha)$ denoting a specification of $\mathbb{E}[y_{it}|\mathcal{X}_{t-1}]$, $u_t := (u_{1t}, \ldots, u_{nt})'$ as a $n \times 1$ vector of conditionally heteroskedastic errors, $h_t^{1/2} := h_t(X_t, \alpha)^{1/2}$ as a $n \times n$ matrix square root of $h_t = h_t(X_t, \alpha)$, which is a specification of $\text{var}[y_t|\mathcal{X}_{t-1}]$, such that $h_t = h_t^{1/2} h_t^{1/2}$, and $\varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{nt})'$ as a $n \times 1$ vector of standardized errors with $\mathbb{E}[\varepsilon_{it}] = 0$, $\text{var}[\varepsilon_{it}] = 1$, and $\mathbb{E}[\varepsilon_{it}\varepsilon_{jt}] = 0$ when $i \neq j$. In this framework, the generic CM model $\mu_t(X_t, \alpha)$ degenerates to an approximation to $M_t(y_t|\mathcal{X}_{t-1})$, which comprises the $n \times 1$ conditional mean vector $\mathbb{E}[y_t|\mathcal{X}_{t-1}]$ and the $n \times n$ conditional covariance matrix:

$$\text{var}[y_t|\mathcal{X}_{t-1}] = \begin{bmatrix} \text{var}[y_{1t}|\mathcal{X}_{t-1}] & \text{cov}[y_{1t}, y_{2t}|\mathcal{X}_{t-1}] & \ldots & \text{cov}[y_{1t}, y_{nt}|\mathcal{X}_{t-1}] \\ \text{cov}[y_{2t}, y_{1t}|\mathcal{X}_{t-1}] & \text{var}[y_{2t}|\mathcal{X}_{t-1}] & \ldots & \text{cov}[y_{2t}, y_{nt}|\mathcal{X}_{t-1}] \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov}[y_{nt}, y_{1t}|\mathcal{X}_{t-1}] & \text{cov}[y_{nt}, y_{2t}|\mathcal{X}_{t-1}] & \ldots & \text{var}[y_{nt}|\mathcal{X}_{t-1}] \end{bmatrix}; \tag{38}$$

which are, respectively, useful for measuring the predictability and volatility (and cross-sectional relationships) of the returns of a portfolio of assets. Exploring these high-dimensional moments is important for studying the mean-variance optimal portfolio, volatility spillover, financial market co-movements, and other problems.

Similar to univariate GARCH-type models, various multivariate GARCH-type models also include various h_t's as their key specifications. In this approach, representative examples include the VECH model of Bollerslev et al. (1988) and the BEKK model of Engle and Kroner (1995). However, it is known that these models are not necessarily easy to estimate when n becomes moderately large. This difficulty results from the fact that h_t is an approximation to the $n \times n$ conditional covariance matrix, so it is required to be positive definite for each t and could involve a considerable number of parameters in the numerical optimization process for estimation; see, e.g., Engle (2002a) for discussions. In empirical finance, it becomes popular to replace the former models with the CCC/DCC models that are established in a different approach.

Unlike the former models that directly approximate $\text{var}[y_t|\mathcal{X}_{t-1}]$ using various h_t's, the CCC/DCC models indirectly approximate $\text{var}[y_t|\mathcal{X}_{t-1}]$ using the following decomposition:

$$\text{var}[y_t|\mathcal{X}_{t-1}] = D_t R_t D_t, \tag{39}$$

where D_t is a $n \times n$ diagonal matrix that includes $\text{var}[y_{it}|\mathcal{X}_{t-1}]^{1/2}$ as the ith diagonal element, and R_t is a $n \times n$ conditional correlation matrix that includes

$$\text{corr}[y_{it}, y_{jt}|\mathcal{X}_{t-1}] = \mathbb{E}\left[\left(\frac{y_{it} - \mathbb{E}[y_{it}|\mathcal{X}_{t-1}]}{\text{var}[y_{it}|\mathcal{X}_{t-1}]^{\frac{1}{2}}}\right)\left(\frac{y_{jt} - \mathbb{E}[y_{jt}|\mathcal{X}_{t-1}]}{\text{var}[y_{jt}|\mathcal{X}_{t-1}]^{\frac{1}{2}}}\right)\middle|\, \mathcal{X}_{t-1}\right] \tag{40}$$

as the (i, j)th element. Correspondingly, they integrate the counterparts of (22):

$$y_{it} = m_{it}(X_t, \alpha_i) + u_{it}, \quad u_{it} = h_{it}(X_t, \alpha_i)^{\frac{1}{2}} \varepsilon_{it}, \tag{41}$$

with $i = 1, 2, \ldots, n$, to obtain another representation of multivariate GARCH-type models:

$$y_t = m_t(X_t, a) + u_t, \quad u_t = \sigma_t(X_t, a)\varepsilon_t, \quad \varepsilon_t|\mathcal{X}_{t-1} \sim (0, \rho_t(X_t, b)), \tag{42}$$

where $m_t(X_t, a) := (m_{1t}, \ldots, m_{nt})'$ includes $m_{it} := m_{it}(X_t, \alpha_i)$, which is a specification of $\mathbb{E}[y_{it}|\mathcal{X}_{t-1}]$; $\sigma_t := \sigma_t(X_t, a)$ is a $n \times n$ diagonal matrix with ith diagonal element $h_{it}^{1/2} := h_{it}(X_t, \alpha_i)^{1/2}$, which is a specification of var$[y_{it}|\mathcal{X}_{t-1}]$; $\varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{nt})'$ is a $n \times 1$ vector of standardized errors with $\mathbb{E}[\varepsilon_{it}] = 0$ and $\mathbb{E}[\varepsilon_{it}^2] = 1$ for all the i's; m_t and σ_t are of the parameter vector $a := (\alpha_1', \ldots, \alpha_n')'$; $\rho_t := \rho_t(X_t, b)$ is a $n \times n$ matrix denoting a specification of the conditional correlation matrix $\mathbb{E}[\varepsilon_t\varepsilon_t'|\mathcal{X}_{t-1}]$, and has the parameter vector b. Note that the diagonal elements of ρ_t are equal to one, and the ε_t in (42) has a different meaning from its counterpart in the aforementioned multivariate version of (22). Moreover, model (42) has the entire parameter vector $\alpha = (a', b')'$, and it requires the α_i's to be separable in the sense that α_i and α_j do not contain the same elements when $i \neq j$; meanwhile, the parameter vectors a and b are also required to be separable. These parameter separability assumptions allow us to estimate these complicated models using certain two-stage estimation methods, as discussed later.

Clearly, this class of multivariate GARCH-type models approximate D_t and R_t using σ_t and ρ_t, respectively. Given the univariate GARCH-type models in (41), various CCC/DCC models approximate R_t using various ρ_t's. Representative examples include the CCC model of Bollerslev (1990) that specifies ρ_t as a matrix of constants, and the DCC models of Engle (2002a) and Tse and Tsui (2002) that allow their ρ_t's to follow different X_t-dependent dynamic specifications. This class of models have a number of appealing properties. First, following (39), these models approximate var$[y_t|\mathcal{X}_{t-1}]$ using $h_t = \sigma_t\rho_t\sigma_t$ with various ρ_t's. By construction, this form of h_t is ensured to be positive semi-definite and hence is easier to be maintained as a positive-definite matrix; see Engle and Sheppard (2001) and Engle (2002a) for further discussion. Second, it decomposes a complicated conditional covariance matrix specification (the h_t) into a set of univariate GARCH-type models (the h_{it}'s) and a relatively simple conditional correlation model (the ρ_t). This provides a direct linkage between a set of univariate GARCH-type models and a multivariate GARCH-type model. Given the parameter separability assumptions, it also allows practitioners to replace a relatively difficult one-stage estimation method with a simpler two-stage estimation method that first estimates the α_i's separately, and then estimates the parameter vector b based on the first-stage estimators.

In the following, we use a generalized two-stage estimation method to show the applicability of the EF method in this scenario. Specifically, we write θ as $\theta = (\theta_1', \theta_2')'$ with θ_1 denoting the first-stage parameter vector and θ_2 denoting the second-stage parameter vector. We also decompose v_{1t} as $v_{1t} = (v_{1t,1}(\theta_1)', v_{1t,2}(\theta_1, \theta_2)')'$, with $v_{1t,1} := v_{1t,1}(\theta_1)$ denoting the first-stage generalized residual, which is dependent on θ_1 but free of θ_2, and $v_{1t,2} := v_{1t,2}(\theta_1, \theta_2)$ denoting the second-stage generalized residual, which is dependent on both θ_1 and θ_2. Meanwhile, we also consider $z_{1t,1} := z_{1t,1}(\theta_1)$ and $z_{1t,2} := z_{1t,2}(\theta_1, \theta_2)$ as the instrumental variable matrices for the first- and second-stage estimations. The first-stage estimator $\hat{\theta}_{1T}$ is solved from the estimating equation:

$$\frac{1}{T}\sum_{t=1}^{T} z_{1t,1} v_{1t,1}\Big|_{\theta_1=\hat{\theta}_{1T}} = 0, \tag{43}$$

and the second-stage estimator $\hat{\theta}_{2T}$ is solved from the $\hat{\theta}_{1T}$-based estimating equation:

$$\frac{1}{T}\sum_{t=1}^{T} z_{1t,2} v_{1t,2}\Big|_{\theta_1=\hat{\theta}_{1T},\theta_2=\hat{\theta}_{2T}} = 0. \tag{44}$$

By summarizing (43) and (44) as (10) with $v_{1t} = (v_{1t,1}', v_{1t,2}')'$ and

$$z_{1t} = \begin{bmatrix} z_{1t,1} & 0 \\ 0 & z_{1t,2} \end{bmatrix},$$

we can see that $\hat{\theta}_T = (\hat{\theta}_{1T}', \hat{\theta}_{2T}')'$ is a indeed particular EF estimator with this combination of v_{1t} and z_{1t}.

In this scenario, we have

$$V_{1t} = \nabla_{\theta'} v_{1t} = \begin{bmatrix} V_{1t,11} & 0 \\ V_{1t,21} & V_{1t,22} \end{bmatrix} := \begin{bmatrix} \nabla_{\theta_1'} v_{1t,1} & 0 \\ \nabla_{\theta_1'} v_{1t,2} & \nabla_{\theta_2'} v_{1t,2} \end{bmatrix}$$

and

$$\mathbb{E}[z_{1t}V_{1t}] = \begin{bmatrix} \mathbb{E}[z_{1t,1}V_{1t,11}] & 0 \\ \mathbb{E}[z_{1t,2}V_{1t,21}] & \mathbb{E}[z_{1t,2}V_{1t,22}] \end{bmatrix}.$$

Accordingly, we can write $\mathbb{E}[z_{1t}V_{1t}]^{-1}$ as

$$\begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix} := \begin{bmatrix} \mathbb{E}[z_{1t,1}V_{1t,11}]^{-1} & 0 \\ -\mathbb{E}[z_{1t,2}V_{1t,22}]^{-1}\mathbb{E}[z_{1t,2}V_{1t,21}]\mathbb{E}[z_{1t,1}V_{1t,11}]^{-1} & \mathbb{E}[z_{1t,2}V_{1t,22}]^{-1} \end{bmatrix},$$

and express $\mathbb{E}[(z_{1t}v_{1t})(z_{1t}v_{1t})']$ as

$$\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} := \begin{bmatrix} \mathbb{E}[(z_{1t,1}v_{1t,1})(z_{1t,1}v_{1t,1})'] & \mathbb{E}[(z_{1t,1}v_{1t,1})(z_{1t,2}v_{1t,2})'] \\ \mathbb{E}[(z_{1t,2}v_{1t,2})(z_{1t,1}v_{1t,1})'] & \mathbb{E}[(z_{1t,2}v_{1t,2})(z_{1t,2}v_{1t,2})'] \end{bmatrix}.$$

The resulting $\sum_o := \sum |_{\theta=\theta_o}$ is the form:

$$\begin{aligned} \sum = \begin{bmatrix} \sum_{11} & \sum_{12} \\ \sum_{12}' & \sum_{22} \end{bmatrix} :=& \\ &\begin{bmatrix} A_{11}B_{11}A_{11}' & A_{11}B_{11}A_{21}' + A_{11}B_{12}A_{22}' \\ \sum_{12}' & A_{21}B_{11}A_{21}' + A_{21}B_{12}A_{22}' + A_{22}B_{21}A_{21}' + A_{22}B_{22}A_{22}' \end{bmatrix}. \quad (45) \end{aligned}$$

This shows the $\sum_o$ of the two-stage estimator $\hat{\theta}_T = (\hat{\theta}_{1T}', \hat{\theta}_{2T}')'$; see also Newey and McFadden (1994, Theorem 6.1) for a discussion of $\hat{\theta}_{2T}$ and $\sum_{22}$.

In practical applications, it is natural to extend the Gaussian QML method as a two-stage estimation method for estimating the CCC/DCC models. This estimation method can be shown as a special case of the generalized two-stage estimation method where $\theta = \alpha$, $\theta_1 = a = (\alpha_1', \ldots, \alpha_n')'$, and $\theta_2 = b$. In the first-stage, it computes the Gaussian QMLE $\hat{\alpha}_{iT}$ for model (41) using the same estimation method discussed in Section 3.1 for each $i = 1, \ldots, n$. Denote $\hat{a}_T := (\hat{\alpha}_{1T}', \ldots, \hat{\alpha}_{nT}')'$ and $\hat{\varepsilon}_t := \varepsilon_t|_{a=\hat{a}_T}$. In the second stage, it computes the $\hat{a}_T$-based Gaussian QMLE $\hat{b}_T$ of the conditional correlation model $\rho_t(X_t, b)$ by maximizing the multivariate Gaussian quasi-log likelihood function:

$$L_T(\beta|\hat{a}_T) = -\frac{n}{2}\ln 2\pi - \frac{1}{2T}\sum_{t=1}^{T}\det(\rho_t) - \frac{1}{2T}\sum_{t=1}^{T}\hat{\varepsilon}_t'\rho_t^{-1}\hat{\varepsilon}_t. \quad (46)$$

More specifically, following (23) and (26), we can write the estimating equation for the first-stage Gaussian QMLE $\hat{\alpha}_{iT}$ as:

$$\frac{1}{T}\sum_{t=1}^{T} z_{1t,1i} v_{1t,1i}\Big|_{\alpha_i=\hat{\alpha}_{iT}} = 0, \tag{47}$$

with $v_{1t,1i} := (\varepsilon_{it}, \varepsilon_{it}^2 - 1)'$ and $z_{1t,1i} := \left[(\nabla_{\alpha_i} m_{it}) h_{it}^{-1/2}, \quad \frac{1}{2}(\nabla_{\alpha_i} h_{it}) h_{it}^{-1}\right]$ for each $i = 1, \ldots, n$. We can further integrate these n equations as the estimating equation for $\hat{a}_T$, which is a particular example of (43) with $v_{1t,1} = (v'_{1t,11}, v'_{1t,12}, \ldots, v'_{1t,1n})'$ and

$$z_{1t,1} = \begin{bmatrix} z_{1t,11} & 0 & \ldots & 0 \\ 0 & z_{1t,12} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & z_{1t,1n} \end{bmatrix}.$$

Meanwhile, by maximizing (46), we can obtain the $\hat{a}_T$-based estimating equation for $\hat{b}_T$:

$$\frac{1}{T}\sum_{t=1}^{T} (\nabla_b \rho_t)(\rho_t^{-1} \otimes \rho_t^{-1}) \mathrm{vec}(\varepsilon_t \varepsilon_t' - \rho_t)\Big|_{a=\hat{a}_T, b=\hat{b}_T} = 0, \tag{48}$$

where vec$(\cdot)$ denotes the vectorization operator of a matrix, and $\nabla_b \rho_t$ is a $q_{12} \times n^2$ matrix of derivatives with $q_{12} := \dim(b)$; see Bollerslev and Wooldridge (1992) for the counterpart in (48) the one-stage multivariate Gaussian QML method and for the definition of $\nabla_b \rho_t$. Obviously, (48) can be interpreted as a particular example of (44) with $v_{1t,2} = \mathrm{vec}(\varepsilon_t \varepsilon_t' - \rho_t)$ and $z_{1t,2} = (\nabla_b \rho_t)(\rho_t^{-1} \otimes \rho_t^{-1})$. Given these $v_{1t,1}$, $v_{1t,2}$, $z_{1t,1}$, and $z_{1t,2}$, we can obtain the asymptotic normality of $T^{1/2}(\hat{\alpha}_T - \alpha_o)$, with $\hat{\alpha}_T = (\hat{a}_T', \hat{b}_T')'$ and $\alpha_o = (a_o', b_o')'$, under condition (7) with $v_{1t} = (v'_{1t,1}, v'_{1t,2})'$, following the results in (17) and (45).

In empirical studies, practitioners are used to estimating the CCC/DCC models using the two-stage Gaussian ML method, as suggested by Bollerslev (1990) and Engle (2002a). This estimation method is a particular example of the aforementioned two-stage Gaussian QML method with the conditional normality assumption of $y_t|\mathcal{X}_{t-1}$. In comparison, the two-stage Gaussian QML method is derived without using this, or any other specific, multivariate CD assumption, so that it robustifies the two-stage ML method to the unknown multivariate CD in the partially specified context. However, its first-stage estimation requires (m_{it}, h_{it}) to be correctly specified for $(\mathbb{E}[y_{it}|\mathcal{X}_{t-1}], \text{var}[y_{it}|\mathcal{X}_{t-1}])$ for all the i's. Given this assumption, its second-stage estimation further requires the ρ_t being estimated to be correctly specified for the true conditional correlation matrix R_t.

The first-stage specification correctness conditions can be checked using the particular CM tests that we have introduced in Section 3.1. In the FTS literature, researchers have also introduced a number of parametric specification tests for multivariate GARCH-type models. For example, Ling and Li (1997) introduced a multivariate extension of the Li-Mak test based on the sum of squared standardized residual autocorrelations, Kroner and Ng (1998) considered the use of Wooldridge's (1990) test based on the discrepancy between the cross-products of residuals and the associated components of h_t, Tse (2000) and Bera and Kim (2002), respectively, proposed a score test and a bivariate information matrix test for the CCC model, and Tse (2002) also derived general residuals-based diagnostic tests. Since these existing tests are all based on certain finite-dimensional moment conditions, we may also recast them as particular examples of the CM test introduced in Section 2. To check the second-stage specification correctness condition of our interest, we may apply the CM test to generate the two-stage Gaussian QML-based counterparts of these existing tests by using the aforementioned v_{1t} and z_{1t}, setting v_{2t} as a sub-vector of $\text{vech}(\varepsilon_t\varepsilon_t' - \rho_t)$, and choosing the associated z_{2t}'s in some appropriate way similar to the examples discussed in Section 3.1.

3.4 Multivariate Fully Specified Models: CMD Models

Given (42), we can also establish a multivariate fully specified GARCH-type model by making an additional multivariate CD assumption of $\varepsilon_t|\mathcal{X}_{t-1}$ that has standardized marginal CDs and the conditional correlation matrix ρ_t. In this context, $F_t(\cdot|\mathcal{X}_{t-1})$ is the true multivariate CD of $y_t|\mathcal{X}_{t-1}$ with $y_t = (y_{1t}, \ldots, y_{nt})'$, and we let $F_{it}(\cdot|\mathcal{X}_{t-1})$ be the univariate CD of $y_{it}|\mathcal{X}_{t-1}$ with $i = 1, 2, \ldots, n$. Modelling $F_t(\cdot|\mathcal{X}_{t-1})$ becomes important when the financial and statistical problems, mentioned in Section 3.2, are extended from the univariate (asset) context to the multivariate (portfolio) context. For example, as considered by Patton (2004), it is important to explore not only the skewness of $F_{it}(\cdot|\mathcal{X}_{t-1})$ but also the cross-dependence asymmetry of $F_t(\cdot|\mathcal{X}_{t-1})$ for studying the skewed-preference investing problem in a portfolio level. Let $Q_{F_{it}}(\tau|\mathcal{X}_{t-1})$ be the conditional τ-quantile of $y_{it}|\mathcal{X}_{t-1}$ for some $\tau \in (0, 0.5)$, some researchers measure the cross-dependence asymmetry (or said the correlation asymmetry) by comparing the down-side and up-side conditional exceedance-correlations:

$$\operatorname{corr}[y_{it}, y_{jt} | y_{it} < Q_{F_{it}}(\tau | \mathcal{X}_{t-1}), y_{jt} < Q_{F_{jt}}(\tau | \mathcal{X}_{t-1}),\ \mathcal{X}_{t-1}]$$

and

$$\operatorname{corr}[y_{it}, y_{jt} | y_{it} > Q_{F_{it}}(1-\tau | \mathcal{X}_{t-1}),\ y_{jt} > Q_{F_{jt}}(1-\tau | \mathcal{X}_{t-1}), \mathcal{X}_{t-1}];$$

see also Longin and Solnik (2001), Ang and Chen (2002), and Campbell-Pownall et al. (2008). Since these multivariate higher-order CMs are closely related to the tail properties of $F_t(\cdot|\mathcal{X}_{t-1})$, they are also important for risk management in a portfolio level. Moreover, the multivariate density forecast even requires the entire multivariate CD to be specified; see, e.g., Diebold et al. (1999) and Poon et al. (2004). Multivariate CD models are useful for approximating the unknown $F_t(\cdot|\mathcal{X}_{t-1})$ and its higher-order CMs.

In this class of models, representative examples include the multivariate conditional normal CCC model of Bollerslev (1990) and DCC models of Engle (2002a) and Tse (2002). These models replace the partial specification in (42), $\varepsilon_t|\mathcal{X}_{t-1} \sim (0, \rho_t(X_t, b))$, with a complete CD specification: $\varepsilon_t|\mathcal{X}_{t-1} \sim N(0, \rho_t(X_t, b))$. This multivariate conditional normal distribution requires $\varepsilon_{it}|\mathcal{X}_{t-1} \sim N(0,1)$ for all the i's. However, since Bollerslev (1987), it is well-documented that the requirement of conditional normality is unlikely to be correct in modelling univariate GARCH-type models. This explains why a considerable number of non-normal univariate CDs have been introduced to the FTS literature, as mentioned in Section 3.2. Recently, there is a fast growing interest in building multivariate CD models using the copula approach, which is quite powerful for integrating various univariate CDs into the multivariate context. The resulting models are referred to as the CMD models.

The CMD models are based on the (conditional version of) Sklar's (1959) theorem. Let $C_o(\cdot|\mathcal{X}_{t-1}) : [0,1]^n \to [0,1]$ be a conditional copula function, which is defined as the multivariate CD of a n-dimensional vector of $U(0,1)$ random variables. This theorem indicates that, there exists such a $C_o(\cdot|\mathcal{X}_{t-1})$ to ensure that

$$F_t(y_t|\mathcal{X}_{t-1}) = C_o(F_{1t}(y_{1t}|\mathcal{X}_{t-1}), \ldots, F_{nt}(y_{nt}|\mathcal{X}_{t-1})|\mathcal{X}_{t-1}); \tag{49}$$

see Patton (2006a, Theorem 1). Importantly, this means that we can decompose the multivariate CD $F_t(\cdot|\mathcal{X}_{t-1})$ into the univariate CDs $F_{it}(\cdot|\mathcal{X}_{t-1})$'s and the conditional copula function $C_o(\cdot|\mathcal{X}_{t-1})$. Since the univariate CDs contain no information about the cross-dependence structures of $y_t|\mathcal{X}_{t-1}$, the information of these structures is fully determined by the conditional copula function. For this reason, the copula function is also known as the (cross-)dependence function.

Corresponding to (49), the CMD models for $F_t(y_t|\mathcal{X}_{t-1})$ are of the general form:

$$G_t(y_t|X_t,\beta) = C(G_{1t}(y_{1t}|X_t,\beta_1),\ldots,G_{nt}(y_{nt}|X_t,\beta_n)|X_t,\beta_c), \tag{50}$$

in which $G_{it}(\cdot|X_t,\beta_i)$ and $C(\cdot|X_t,\beta_c)$ are, respectively, the ith "marginal model" and the conditional copula model for approximating the unknown $F_{it}(\cdot|\mathcal{X}_{t-1})$ and $C_o(\cdot|X_t)$. Model (50) has the entire parameter vector $\beta = (\beta_1',\ldots,\beta_n',\beta_c')'$. Similar to the CCC/DCC models, the CMD models also require β_i to be separable from β_j (and β_c) when $i \neq j$ (for all the i's) to facilitate the two-stage estimation. In addition, the ith marginal model takes the form of (28):

$$G_{it}(y_{it}|X_t,\beta_i) = G_{\varepsilon_i}\left(\left.\frac{y_{it}-m_{it}(X_t,\alpha_i)}{h_{it}(X_t,\alpha_i)^{\frac{1}{2}}}\right| X_t,\gamma_i\right), \tag{51}$$

and it has the conditional PDF:

$$g_{it}(y_{it}|X_t,\beta_i) = h_{it}(X_t,\alpha_i)^{-\frac{1}{2}} g_{\varepsilon_i}\left(\left.\frac{y_{it}-m_{it}(X_t,\alpha_i)}{h_{it}(X_t,\alpha_i)^{\frac{1}{2}}}\right| X_t,\gamma_i\right) \tag{52}$$

and the parameter vector $\beta_i = (\alpha_i',\gamma_i')'$, in which $G_{\varepsilon_i}(\cdot|X_t,\gamma_i)$ and $g_{\varepsilon_i}(\cdot|X_t,\gamma_i)$ are, respectively, the standardized CD specification and the conditional PDF of $\varepsilon_{it}|\mathcal{X}_{t-1}$. The conditional copula model $C(\cdot|X_t,\beta_n)$ could be based on any parametric copula function for specification. Sklar's theorem is also useful for transforming the closed form of a multivariate distribution into the functional form of the associated parametric copula for practical applications, as discussed below.

To demonstrate this point, we let $\Psi : \mathbb{R}^n \to [0,1]$ be a multivariate distribution with the marginal distribution $\Psi_i : \mathbb{R} \to [0,1]$, $i = 1,2,\ldots,n$. This theorem indicates that there exists a copula function C_Ψ ensuring that $\Psi(y_1,\ldots,y_n) = C_\Psi(\Psi_1(y_1),\ldots,\Psi_n(y_n))$ for all $(y_1,\ldots,y_n) \in \mathbb{R}^n$. Denote the PIT $\wp_i := \Psi_i(y_i)$. Correspondingly, we can also write $y_i = \Psi_i^{-1}(\wp_i)$ as a quantile transformation of $\wp_i$.

Using these two transformations, we can rewrite the aforementioned copula function C_Ψ as:

$$C_\Psi(\wp_1, \ldots, \wp_n) = \Psi(\Psi_1^{-1}(\wp_1), \ldots, \Psi_n^{-1}(\wp_n)). \tag{53}$$

Therefore, given the closed form of Ψ, we can obtain the associated formula of C_Ψ using (53). Let ψ and ψ_i be the PDFs of Ψ and Ψ_i, respectively. From (53), we can also show that C_Ψ has the copula density function:

$$\begin{aligned} c_\Psi(\wp_1, \ldots, \wp_n) : &= \frac{\partial^n}{\partial\wp_1 \ldots \partial\wp_n} C_\Psi(\wp_1, \ldots, \wp_n) \\ &= \frac{\psi(\Psi_1^{-1}(\wp_1), \ldots, \Psi_n^{-1}(\wp_n))}{\psi_1(\Psi_1^{-1}(\wp_1)) \ldots \psi_n(\Psi_n^{-1}(\wp_n))}. \end{aligned} \tag{54}$$

The widely used normal, t, and Gumbel copula functions are all derived in this way, and they respectively correspond to the multivariate normal, t, and Gumbel distributions. Note that C_Ψ should contain the same parameter vector as Ψ, such as the correlation coefficients (and the degrees of freedom) of the multivariate normal (multivariate t) distributions. In addition, various C_Ψ's could imply various cross-dependence structures. We refer to Joe (1997), Nelsen (1999, 2002), and Cherubini et al. (2004) for further discussions and for more copulas and their formulas.

In practical applications, it is common to set the conditional copula model $C_t(\cdot| \mathcal{X}_{t-1}, \beta_c)$ as a certain C_Ψ. The former has the same parameter vector as the latter. Similar to the CCC model, this conditional copula model is static when its parameter vector comprises constants. Similar to the DCC models, this conditional copula model could also allow its parameter vector to follow certain X_t-dependent dynamic specifications. Given the $G_{it}(\cdot|X_t, \beta_i)$'s, various $C_t(\cdot|\mathcal{X}_{t-1}, \beta_c)$'s generate various CMD models with various cross-dependence structures. In particular, the aforementioned conditional normal CCC/DCC models are particular CMD models that require all the $G_{it}(\cdot|X_t, \beta_i)$'s and the $C_t(\cdot|X_t, \beta_c)$ to be normal and imply no "extreme" tail-dependence; see Patton (2004, 2006a, 2006b), Hu (2006), Jondeau and Rockinger (2006), and Chen (2007) for more CMD models and empirical applications.

Similar to the CCC/DCC models, the CMD models could also involve a considerable number of parameters and hence may not be easily estimated using the one-stage estimation method. Given the parameter separability assumptions, it is natural to remedy this difficulty using the two-stage estimation method. Meanwhile, since the marginal models and the conditional copula model are all fully specified, it is common to estimate the CMD models using the two-stage ML method, as considered by Patton (2006b). This method can also be shown as a particular example of the generalized two-stage estimation method (and hence a particular EF method) with $\theta = \beta$, $\theta_1 = (\beta_1', \ldots, \beta_n')'$ and $\theta_2 = \beta_c$. Specifically, this two-stage ML method first computes the MLE $\hat{\beta}_{iT}$ for model (51) using the ML method discussed in Section 3.2 for each $i = 1, \ldots, n$. Denote $\hat{\theta}_{1T} = (\hat{\beta}_{1T}', \ldots, \hat{\beta}_{nT}')'$, the PIT $\wp_{it} := G_{it}(y_{it}|X_t, \beta_i)$, and the estimated PIT $\hat{\wp}_{it} := G_{it}(y_{it}|X_t, \hat{\beta}_{iT})$. In the second stage, it computes the $\hat{\theta}_{1T}$-based MLE $\hat{\beta}_{cT}$ by maximizing the copula log-likelihood function:

$$L_T(\beta_c|\hat{\beta}_{1T}, \ldots, \hat{\beta}_{nT}) = \frac{1}{T}\sum_{t=1}^{T} \ln c(\hat{\wp}_{1t}, \ldots, \hat{\wp}_{nt}|X_t, \beta_c), \tag{55}$$

in which $c(\cdot|X_t, \beta_c)$ is the conditional copula density function, and it takes the form of (54) when $C(\cdot|X_t, \beta_c)$ follows the specification of C_Ψ.

Given (52), we write $\ell_{\beta_i,t} := \nabla_{\beta_i} \ln g_{it}(y_{it}|X_t, \beta_i)$. According to (34), we can further write the estimating equation for the first-stage MLE $\hat{\beta}_{iT}$ as:

$$\frac{1}{T}\sum_{t=1}^{T} \ell_{\beta_i,t}\Big|_{\beta_i=\hat{\beta}_{iT}} = 0, \tag{56}$$

which can be represented as $T^{-1}\sum_{t=1}^{T} z_{1t,1i} v_{1t,1i}\Big|_{\beta_i=\hat{\beta}_{iT}} = 0$, by setting $v_{1t,1i} = \ell_{\beta_i,t}$, $z_{1t,1i} = I_{q_{11i}}$, and $q_{11i} = \dim(v_{1t,1i})$. The associated n estimating equations for the $\hat{\beta}_{iT}$'s can be summarized as a particular example of (43) with $v_{1t,1} = (v_{1t,11}', v_{1t,12}', \ldots, v_{1t,1n}')'$, $z_{1t,1} = I_{q_{11}}$, and $q_{11} = \dim(v_{1t,1})$. In addition, the second-stage $\hat{\theta}_{1T}$-based MLE is of the estimating equation:

$$\frac{1}{T}\sum_{t=1}^{T} \ell_{\beta_c,t}\Big|_{\theta_1=\hat{\theta}_{1T}, \beta_c=\hat{\beta}_{cT}} = 0, \tag{57}$$

with $\ell_{\beta_c,t} := \nabla_{\beta_c} \ln c(\wp_{1t}, \ldots, \wp_{nt}|X_t, \beta_c)$, which can be interpreted as a particular example of (44) by setting $v_{1t,2} = \ell_{\beta_c,t}$, $z_{1t,2} = I_{q_{12}}$, and $q_{12} = \dim(v_{1t,2})$. Similar to the case of the CCC/DCC models, we can also obtain the asymptotic normality of $T^{1/2}(\hat{\beta}_T - \beta_o)$ under the specification correctness condition (7) with $v_{1t} = (v'_{1t,1}, v'_{1t,2})'$ by plugging the aforementioned $v_{1t,1}$, $v_{1t,2}$, $z_{1t,1}$, and $z_{1t,2}$ into the general results of (17) and (45).

In this estimation method, the first-stage ML method requires the marginal model $G_{it}(\cdot|X_t, \beta_i)$ to be correctly specified for the true univariate CD $F_{it}(\cdot|\mathcal{X}_{t-1})$ for all the i's. Given this assumption, the second-stage ML method further requires the conditional copula model $C_t(\cdot|X_t, \beta_c)$ to be correctly specified for the true conditional copula function $C_o(\cdot|\mathcal{X}_{t-1})$. The first-stage specification correctness conditions can be evaluated using the methods that were discussed in Sections 3.2. Recently, Chen (2007) introduced a class of moment-based tests, including the Kendall's-tau-based concordance test and a set of more powerful tail-dependence tests, for the functional forms of copula. We can also make this class of tests applicable to the two-stage ML method by applying the CM test to the aforementioned v_{1t} and z_{1t}. This class of tests check the second-stage specification correctness condition by choosing v_{2t} as a certain function of the PITs $\wp_{1t}, \ldots, \wp_{nt}$ and setting $z_{2t} = I_{q_2}$ with $q_2 = \dim(v_{2t})$. For example, the concordance test sets $v_{2t} = \frac{1}{2^{n-1}-1}\left[2^n C(\wp_{1t}, \ldots, \wp_{nt}|X_t; \beta_c) - 1\right] - \tau(X_t; \beta_c)$ with $\tau(X_t; \beta_c)$ representing the Kendall's tau implied by the conditional copula being tested. We may also extend these tests to make them powerful against dynamic misspecifications by replacing the original z_{2t} with certain X_t-dependent z_{2t}'s.

Following a referee's suggestion, we also review some domestic empirical studies of multivariate GARCH-type models. A brief sample of these studies includes Gau and Hsieh (2002) that applied a three-dimensional BEKK model with the conditional normality assumption to the VaR evaluation for an exchange rate portfolio. Wang and Chen (2003) used a bivariate GJR-GARCH-in-mean model with a conditional t distribution assumption to examine the return-volatility transmission relationships between the U.S. and Taiwan stock markets before and after the 1997 Asia financial crisis. Wang and Wu (2006) applied a trivariate GARCH-type model with a conditional t distribution assumption to explore the impact of Center Bank regulation on the relationships among the Taiwan spot, non-physical delivery forward, and delivery forward exchange rate markets. These studies all contribute some interesting empirical applications and findings that we do not explore in this survey.

Since these empirical studies are all based on fully specified multivariate GARCH-type models, it is natural to consider the CMD model as an alternative specification for similar empirical problems. Moreover, since the CMD model is built on the univariate and multivariate models discussed in Sections 3.1–3.3, it is also useful to demonstrate some potential advantages of the methods reviewed in this paper by considering this alternative modelling strategy. First, by utilizing the "bottom-up" structure of the CMD model, we can check the partial specifications of this model in a step-by-step way. This is essential for detecting and refining the "lower-order" misspecifications before specifying a fully specified multivariate GARCH-type model. Correspondingly, the estimation and testing methods introduced in the previous subsections also play an essential role in this model building process. Second, this bottom-up structure is also useful for decomposing a complicated multivariate model into a combination of simpler sub-models. Thus, it is easier to estimate the latter than the former, provided that the sub-models satisfy the previously mentioned parameter separability assumptions. This is particularly essential when the dimension of the multivariate CD becomes larger. Third, the CMD model also allows its univariate models to have various CD specifications. This flexibility is not shared by conventional multivariate models. For example, a multivariate conditional normal (t) distribution model requires its univariate marginal models to be conditional normal (t)-distributed, and hence precludes the heavy-tails (asymmetry) of any of these marginal models.

Moreover, the testing methods reviewed in this paper may also be potentially useful for refining some popular diagnostic tests, such as the standardized-residuals-based Ljung-Box and McLeod-Li tests, used in some of these (and many other) empirical studies in this research area. Since these two widely used tests do not account for the model-specific estimation effect in an appropriate way, they could be size-distorted in evaluating GARCH-type models. Moreover, these two tests may tend to accept a misspecified model because they are of no powers against remaining nonlinearity. As discussed in Section 3.1, a number of alternative tests have been proposed to refine these problems, and these refinements are encompassed by the MM-based unified approach.

3.5 Multiplicative Error Models

In addition to the previously mentioned "GARCH-spanned" FTS models, researchers have also proposed various MEMs for modelling various non-negative FTS in the univariate context $(n = 1)$. This is a class of "GARCH-related" models with the general form:

$$y_t = m_t(X_t, \alpha)\varepsilon_t, \tag{58}$$

in which $y_t \geq 0$ is a non-negative dependent variable, $m_t := m_t(X_t, \alpha) \geq 0$ is a specification of $\mathbb{E}[y_t|\mathcal{X}_{t-1}]$, and $\varepsilon_t \geq 0$ is a non-negative multiplicative error with $\mathbb{E}[\varepsilon_t] = 1$. Note that $m_t\varepsilon_t$ has a very similar structure to the squared conditionally heteroskedastic error $u_t^2 = h_t\varepsilon_t^2$ in (22). This reflects the fact that the m_t in (58) resembles the role of the h_t in (22). This class of models include the ACD-type models as an important sub-class of models, in which $\{y_t\}$ is an irregularly spaced sequence of financial durations with t denoting an ordering index (rather than the time index). For example, y_t could be the t-th transaction duration between the $(t-1)$th and the tth transactions in a fixed sampling period. This sub-class of MEMs are known to be essential for exploring market microstructures, market liquidity, and intraday risk management, and they include various m_t's; see, e.g., Engle (2000), Engle and Russell (1998), Bauwens and Giot (2000), and Meitz and Teräsvirta (2006), among many others. In addition, (58) could also be applied to regularly spaced FTS data, such as the daily return range sequence considered by Chou (2005) and the daily trading volume sequence discussed by Manganelli (2005); see Engle (2002b) for a general discussion. Meanwhile, partially specified MEMs can also be extended as fully specified MEMs by adding various unit-mean CD specifications of $\varepsilon_t|\mathcal{X}_{t-1}$ into (58), such as the "standardized" exponential, Weibull, generalized Gamma, and Burr distributions; see, e.g., Engle and Gallo (2006) and Fernandes and Grammig (2006).

In practical applications, many studies have estimated partially specified MEMs using the exponential QML method, as suggested by Engle and Russell (1998). Li and Yu (2003) and Meitz and Teräsvirta (2006) also proposed the exponential ML-based conditional mean tests for MEMs; see also Hautsch (2006) for their exponential-QML-based counterparts and Chen and Hsieh (2010) for further extensions to the independence and distribution tests for $\varepsilon_t|\mathcal{X}_{t-1}$. These estimation and testing methods can also be recasted as particular example of the MM-based unified approach. In particular, the specification correctness condition (7) degenerates to the case where $v_{1t} = \varepsilon_t - 1$ in the context of MEMs, and the exponential QML method can be shown as a particular EF method with this v_{1t} and $z_{1t} = (\nabla_\alpha m_t)m_t^{-1}$. Meanwhile, the exponential QML-based conditional mean tests are particular CM tests with this combination of v_{1t} and z_{1t} and the choice of $v_{2t} = v_{1t}$ and certain z_{2t}'s; the independence and distribution tests are based on more general v_{2t}'s. The ML method and the associated ML-based tests for fully specified MEMs can also be unified into the MM-based approach following the discussions in Section 3.2.

4. CONCLUSIONS

In this survey paper, by integrating the principles of the generalized residual, the EF method, and the CM testing method, we summarize a MM-based unified approach for parametric FTS analysis. As demonstrative examples, we review univariate partially specified (GARCH-type) models, univariate fully specified models, multivariate partially specified models (the CCC/DCC models), multivariate fully specified models (the CMD models), and MEMs, and the associated parametric estimation and testing methods. We illustrate that, although these classes of models are proposed for various economic and statistical reasons, they are all certain CM or CD models. Therefore, the error terms of these FTS models are all encompassed by the generalized residual of the generic CM or CD model, and we can recast their parametric estimation methods (testing methods) as particular examples of the EF method (the CM testing method). In Table 1, we provide a summary list of these specification, estimation, and testing methods. This MM-based synthesis is substantially built on a personal research experience, and is not proposed for claiming a methodological contribution. Nonetheless, it may be useful for summarizing a wide range of FTS models and their statistical inference methods in an efficient way. This synthesis may also be potentially useful for practitioners to derive their own estimation and testing methods when their empirical studies involve some new FTS models that can be fitted into the generic CM or CD model.

Table 1 A Summary of the FTS Specification, Estimation, and Testing Methods

Model	Specification (based on v_{1t})	Estimation (for θ based on z_{1t})	Testing (for $\mathbf{E}[v_{2,ot} \mid \mathcal{X}_{t-1}] = 0$ based on v_{1t}, z_{1t}, and z_{2t})
Generic	$\mathbf{E}[v_{1t} \mid \mathcal{X}_{t-1}]_{\theta=\theta_o} = 0$	$\sqrt{T}(\hat{\theta}_T - \theta_o) \xrightarrow{d} N(0, \Sigma_o)$	$M_T := \left[\Sigma_{t=1}^T \hat{z}_{2t}\hat{v}_{2t}\right]' \left[\Sigma_{t=1}^T \hat{\xi}_t\hat{\xi}_t'\right]^{-1} \left[\Sigma_{t=1}^T \hat{z}_{2t}\hat{v}_{2t}\right]$
		$\Sigma_o := \mathbf{E}[z_{1,ot}V_{1,ot}]^{-1}\mathbf{E}[(z_{1,ot}v_{1,ot})(z_{1,ot}v_{1,ot})']\mathbf{E}[V_{1,ot}'z_{1,ot}']^{-1}$	$\hat{\xi}_t := \hat{z}_{2t}\hat{v}_{2t} - \left[\Sigma_{t=1}^T \hat{z}_{2t}\hat{V}_{2t}\right] \left[\Sigma_{t=1}^T \hat{z}_{1t}\hat{V}_{1t}\right]^{-1} \hat{z}_{1t}\hat{v}_{1t}$
			$M_T \xrightarrow{d} \chi^2(q_2)$, $q_2 := \dim(z_{2t}v_{2t})$
Univariate-	$y_t = m_t + h_t^{1/2}\varepsilon_t$	Gaussian QML: $\theta = \alpha$	Conditional mean-and-variance tests
PS	$v_{1t} = \left(\varepsilon_t, \varepsilon_t^2 - 1\right)'$	$z_{1t} = \left[(\nabla_\alpha m_t)h_t^{-1/2},\ \frac{1}{2}(\nabla_\alpha h_t)h_t^{-1}\right]$	$v_{2t} = v_{1t}$, z_{2t}: a $\mathcal{X}_{t-1}$-measurable matrix
			Conditional distribution tests
			v_{2t}: a higher-order transformation of ε_t
Univariate	$g_t(y_t \mid X_t, \beta) = h_t^{-1/2} \times$	ML: $\theta = \beta$	Score test
FS	$g_\varepsilon\left((y_t - m_t)h_t^{-1/2} \mid X_t, \gamma\right)$	$z_{1t} = I_{q_1}$, $q_1 = \dim(\beta)$, $\beta = (\alpha', \gamma')'$	$v_{2t} = \nabla_\delta \ln \tilde{g}_t(y_t \mid \tilde{X}_t, \beta, \delta)\big\vert_{\delta=0}$
	$v_{1t} = \nabla_\beta \ln g_t(y_t \mid X_t, \beta)$		$\tilde{g}_t(y_t \mid \tilde{X}_t, \beta, 0) = g_t(y_t \mid X_t, \beta)$, $z_{2t} = I_{q_2}$, $q_2 = \dim(\delta)$
CCC/DCC	$y_t = m_t + \sigma_t\varepsilon_t$	Two-stage Gaussian QML: $\theta = (\alpha_1', \ldots, \alpha_n', b')'$	Conditional correlation tests
	$\varepsilon_t \mid \mathcal{X}_{t-1} \sim (0, \rho_t)$	$z_{1t,1i} := \left[(\nabla_{\alpha_i} m_{it})h_{it}^{-1/2},\ \frac{1}{2}(\nabla_{\alpha_i} h_{it})h_{it}^{-1}\right]$	v_{2t}: a sub-vector of $\mathrm{vech}(\varepsilon_t\varepsilon_t' - \rho_t)$
	$y_t = (y_{1t}, \ldots, y_{nt})'$	$z_{1t,1} = \mathrm{diag}(z_{1t,1i})$	z_{2t}: a $\mathcal{X}_{t-1}$-measurable matrix
	$m_t := (m_{1t}, \ldots, m_{nt})'$	$z_{1t,2} = (\nabla_b \rho_t)(\rho_t^{-1} \otimes \rho_t^{-1})$	
	$\sigma_t := \mathrm{diag}(h_{it}^{1/2})$	$z_{1t} = ((z_{1t,1}', 0')', (0', z_{1t,2}')')$	
	$\varepsilon_{it} := (y_{it} - m_{it})h_{it}^{-1/2}$		
	$v_{1t,1i} := (\varepsilon_{it}, \varepsilon_{it}^2 - 1)'$		
	$v_{1t,1} = (v_{1t,11}', \ldots, v_{1t,1n}')'$		
	$v_{1t,2} = \mathrm{vec}(\varepsilon_t\varepsilon_t' - \rho_t)$		
	$v_{1t} = \left(v_{1t,1}', v_{1t,2}'\right)'$		

Table 1 A Summary of the FTS Specification, Estimation, and Testing Methods (continued)

Model	Specification (based on v_{1t})	Estimation (for θ based on z_{1t})	Testing (for $\mathbf{E}[v_{2,ot}\vert\mathcal{X}_{t-1}] = 0$ based on v_{1t}, z_{1t}, and z_{2t})
CMD	$C(\wp_{1t}, \ldots, \wp_{nt}\vert X_t, \beta_c)$	Two-stage ML: $\theta = (\beta_1', \ldots, \beta_n', \beta_c')'$	Copula tests
	$\wp_{it} := G_{it}(y_{1t}\vert X_t, \beta_i)$	$z_{1t,1} = I_{q_{11}}, q_{11} = \dim(v_{1t,1})$	v_{2t}: a transformation of the $\wp_{it}$'s
	$C(\wp_1, \ldots, \wp_n\vert X_t, \beta_c) = \int_0^{\wp_1} \cdots \int_0^{\wp_n}$	$z_{1t,2} = I_{q_{12}}, q_{12} = \dim(v_{1t,2})$	implied by certain moment restrictions of C
	$c(\wp_1, \ldots, \wp_n\vert X_t, \beta_c)\mathrm{d}\wp_n \ldots \mathrm{d}\wp_1$	$z_{1t} = ((z_{1t,1}', 0')', (0', z_{1t,2}')')$	$z_{2t} = I_{q2}, q_2 = \dim(v_{2t})$
	$G_{it}(y_i\vert X_t, \beta_i) = \int_{-\infty}^{y_i} g_{it}(y\vert X_t, \beta_i)\mathrm{d}y$		
	$g_{it}(y_{it}\vert X_t, \beta_i) = h_{it}^{-1/2} \times$		
	$g_{\varepsilon_i}\left((y_{it} - m_{it})h_{it}^{-1/2}\vert X_t, \gamma_i\right)$		
	$v_{1t,1i} = \nabla_{\beta_i} \ln g_{it}(y_{it}\vert X_t, \beta_i)$		
	$v_{1t,1} = (v_{1t,11}', \ldots, v_{1t,1n}')'$		
	$v_{1t,2} = \nabla_{\beta_c} \ln c(\wp_{1t}, \ldots, \wp_{nt}\vert X_t, \beta_c)$		
	$v_{1t} = (v_{1t,1}', v_{1t,2}')'$		
MEMs	$y_t = m_t\varepsilon_t$	Exponential QML: $\theta = \alpha$	Conditional mean tests
	$v_{1t} = \varepsilon_t - 1$	$z_{1t} = (\nabla_\alpha m_t)m_t^{-1}$	$v_{2t} = v_{1t}$, z_{2t}: a $\mathcal{X}_{t-1}$-measurable matrix
			Conditional distribution tests
			v_{2t}: a higher-order transformation of ε_t

Note: "Univariate-PS (FS)" stands for univariate partially (fully) specified models. $\mathrm{diag}(\cdot)$ represents a block-diagonal matrix. The exact form of z_{2t} is dependent on the type of test and the direction of powers being considered. For CM tests, various z_{2t} generate various testing power directions; see Section 3.1 for examples. For unconditional distribution (or copula) tests, z_{2t} is an identity matrix. See the main text for the definition of notations and detailed discussions.

REFERENCES

Andrews, D. W. K. (1994), "Empirical Process Methods in Econometrics," in R. F. Engle and D. L. McFadden (eds), *Handbook of Econometrics*, Vol. 4, 2247–2294, Amsterdam: Elsevier.

Ang, A. and J. Chen (2002), "Asymmetric Correlations of Equity Portfolios," *Review of Financial Studies*, 63, 443–494.

Artzner, P., F. Delbaen, J. Eber, and D. Heath (1999), "Coherent Measures of Risk," *Mathematical Finance*, 9, 203–228.

Bauwens, L. and P. Giot (2000), "The Logarithm ACD Model: An Application to the Bid-Ask Quote Process of Three NYSE Stocks," *Annales dÉconomie et de Statistique*, 60, 117–150.

Bauwens, L., S. Laurent, and J. V. K. Rombouts (2006), "Multivariate GARCH Models: A Survey," *Journal of Applied Econometrics*, 21, 79–109.

Bera, A. K. and Y. Bilias (2001), "Rao's Score, Neyman's $C(\alpha)$ and Silvey's LM Tests: An Essay on Historical Developments and Some New Results," *Journal of Statistical Planning and Inference*, 97, 9–44.

Bera, A. K. and Y. Bilias (2002), "The MM, ME, ML, EL, EF and GMM Approaches to Estimation: A Synthesis," *Journal of Econometrics*, 107, 51–86.

Bera, A. K., Y. Bilias, and P. Simlai (2006), "Estimating Functions and Equations: An Essay on Historical Developments with Applications to Econometrics," in T. C. Mills and K. Patterson (eds), *Palgrave Handbook of Econometrics*, Vol. 1, 427–476, London: Palgrave Macmillan.

Bera, A. K. and S. Kim (2002), "Testing Constancy of Correlation and Other Specifications of the BGARCH Model with an Application to International Equity Returns," *Journal of Empirical Finance*, 9, 171–195.

Berkes, I., L. Horváth, and P. Kokoszka (2003), "Asymptotics for GARCH Squared Residual Correlations," *Econometric Theory*, 19, 515–540.

Berkowitz, J. (2001), "Testing Density Forecasts, Applications to Risk Management," *Journal* *of Business and Economic Statistics*, 19, 465–474.

Black, F. (1976), "Studies of Stock Price Volatility Changes," *1976 Proceedings of the American Statistical Association*, Business and Economic Statistics Section, 177–181.

Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 307–327.

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

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

Bollerslev, T., R. Y. Chou, and K. F. Kroner (1992), "ARCH Modelling in Finance," *Journal of Econometrics*, 52, 5–59.

Bollerslev, T., R. Engle, and D. Nelson (1994), "ARCH Models," in R. Engle and D. McFadden (eds), *Handbook of Econometrics*, Vol. 4, 2959–3038, Amsterdam: North-Holland.

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

Bontemps, C. and N. Meddahi (2005), "Testing Normality: A GMM Approach," *Journal of Econometrics*, 124, 149–186.

Calzolari, G. and G. Fiorentini (1998), "A Tobit Model with GARCH Errors," *Econometric Review*, 17, 85–104.

Cameron, A. C. and P. K. Trivedi (1998), *Regression Analysis of Count Data*, New York: Cambridge University Press.

Campbell-Pownall, R. A. J., C. S. Forbes, C. G. Koedijk, and P. Kofman (2008), "Increasing Correlations or Just Fat Tails?" *Journal of Empirical Finance*, 15, 287–309.

Chen, Y.-T. (2007), "Moment-Based Copula Tests for Financial Returns," *Journal of Business and Economic Statistics*, 25, 377–397.

Chen, Y.-T. (2008a), "A Unified Approach to Standardized-Residuals-Based Correlation Tests for GARCH-Type Models," *Journal of Applied Econometrics*, 23, 111–133.

Chen, Y.-T. (2008b), "Moment Tests for Standardized Error Distributions: A Simple Robust

Approach," Working Paper, Institute of Economics, Academia Sinica.

Chen, Y.-T. (2010), "Moment Tests for Density Forecast Evaluation in the Presence of Parameter Estimation Uncertainty," *Journal of Forecasting*, forthcoming.

Chen, Y.-T. and C.-S. Hsieh (2010), "Generalized Moment Tests for Autoregressive Conditional Duration Models," *Journal of Financial Econometrics*, forthcoming.

Chen, Y.-T. and C.-M. Kuan (2002), "Time Irreversibility and EGARCH Effects in U.S. Stock Index Returns," *Journal of Applied Econometrics*, 17, 565–578.

Cherubini, U., E. Luciano, and W. Vecchiato (2004), *Copula Methods in Finance*, New York: John Wiley & Sons.

Chou, R. (2005), "Forecasting Financial Volatilities with Extreme Values: The Conditional Autoregressive Range (CARR) Model," *Journal of Money, Credit, and Banking*, 37, 561–582.

Christoffersen, P. F. (1998), "Evaluating Interval Forecasts," *International Economic Review*, 39, 841–862.

Christoffersen, P. F. (2003), *Elements of Financial Risk Management*, San Diego: Academic Press.

Cox, D. and E. Snell (1968), "A General Definition of Residuals," *Journal of the Royal Statistical Society*, Series B, 30, 248–265.

Davidson, J. (1994), *Stochastic Limit Theory*, New York: Oxford University Press.

Davidson, R. and J. G. MacKinnon (1990), "Specification Tests Based on Artificial Regressions," *Journal of the American Statistical Association*, 85, 220–227.

Davidson, R. and J. G. MacKinnon (1993), *Estimation and Inference in Econometrics*, New York: Oxford University Press.

Diebold, F. X., T. A. Gunther, and A. S. Tay (1998), "Evaluating Density Forecasts with Applications to Financial Risk Management," *International Economic Review*, 39, 863–883.

Diebold, F. X., J. Hahn, and A. S. Tay (1999), "Multivariate Density Forecast Evaluation and Calibration in Financial Risk Management," *Review of Economics and Statistics*, 81, 661–673.

Duan, J.-C. (1997), "Augmented GARCH (p, q) Process and Its Diffusion Limit," *Journal of Econometrics*, 79, 97–127.

Durbin, J. (1960), "Estimation of Parameters in Time-Series Regression Models," *Journal of* *the Royal Statistical Society*, Series B, 22, 139–153.

Engle, R. F. (1982), "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50, 987–1007.

Engle, R. F. (2000), "The Econometrics of Ultra-High Frequency Data," *Econometrica*, 68, 1–22.

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

Engle, R. F. (2002b), "New Frontiers for ARCH Models," *Journal of Applied Econometrics*, 17, 425–446.

Engle, R. F. and G. M. Gallo (2006), "A Multiple Indicators Model for Volatility Using Intra-Daily Data," *Journal of Econometrics*, 131, 3–27.

Engle, R. F. and G. González-Rivera (1991), "Semiparametric ARCH Models," *Journal of Business and Economic Statistics*, 9, 345–359.

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

Engle, R. F. and V. K. Ng (1993), "Measuring and Testing the Impact of News on Volatility," *Journal of Finance*, 48, 1749–1778.

Engle, R. F. and J. R. Russell (1998), "Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data," *Econometrica*, 66, 1127–1162.

Engle, R. F. and K. Sheppard (2001), "Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH," Working Paper, National Bureau of Economic Research.

Fernandes, M. and J. Grammig (2006), "A Family of Autoregressive Conditional Duration Models," *Journal of Econometrics*, 130, 1–23.

Gau, Y.-F. and J.-H. Hsieh (2002), "Value at Risk and Multivariate GARCH Models," *Taiwan Economic Review*, 30, 273–312. (in Chinese)

Ghosh, A. and A. K. Bera (2005), "Smooth Test for Density Forecast Evaluation," Working Paper, University of Illinois at Urbana-Champaign.

Glosten, L. R., R. Jagannathan, 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*, 5, 1779–1801.

Godambe, V. P. (1960), "An Optimum Property of Regular Maximum Likelihood Estimation," *Annals of Mathematical Statistics*, 31, 1208–1211.

Godambe, V. P. (1991), *Estimating Functions*, Oxford: Oxford University Press.

Gouriéroux, C., A. Monfort, E. Renault, and A. Trognon (1987), "Generalised Residuals," *Journal of Econometrics*, 34, 5–32.

Hamilton, J. D. (1996), "Specification Testing in Markov-Switching Time Series Models," *Journal of Econometrics*, 70, 127–157.

Hamilton, J. D. and R. Susmel (1994), "Autoregressive Conditional Heteroskecasticity and Changes in Regime," *Journal of Econometrics*, 64, 307–333.

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

Hansen, L. P. (1982), "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029–1054.

Harvey, A., E. Ruiz, and N. Shephard (1994), "Multivariate Stochastic Variance Models," *Review of Economic Studies*, 61, 247–264.

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

Harvey, C. R. and A. Siddique (2000), "Conditional Skewness in Asset Pricing Tests," *Journal of Finance*, 55, 1263–1295.

Hautsch, N. (2006), "Testing the Conditional Mean Function of Autoregressive Conditional Duration Models," Working Paper, University of Copenhagen.

Hentschel, L. (1995), "All in the Family Nesting Symmetric and Asymmetric GARCH Models," *Journal of Financial Economics*, 39, 71–104.

Hu, L. (2006), "Dependence Patterns Across Financial Markets: A Mixed Copula Approach," *Applied Financial Economics*, 16, 717–729.

Jarque, C. M. and A. K. Bera (1980), "Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals," *Economics Letters*, 6, 255–259.

Joe, H. (1997), *Multivariate Models and Dependence Concepts*, London: Chapman & Hall.

Jondeau, E. and M. Rockinger (2006), "The Copula-GARCH Model of Conditional Dependencies: An International Stock Market Application," *Journal of International Money and Finance*, 25, 827–853.

Kiefer, N. M. and M. Salmon (1983), "Testing Normality in Econometric Models," *Economics Letters*, 11, 123–127.

Kim, S., N. Shephard, and S. Chib (1998), "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models," *Review of Economic Studies*, 65, 361–393.

Komunjer, I. (2007), "Asymmetric Power Distribution: Theory and Applications to Risk Measurement," *Journal of Applied Econometrics*, 22, 891–921.

Kroner, K. F. and V. K. Ng (1998), "Modeling Asymmetric Comovements of Asset Returns," *Review of Financial Studies*, 11, 817–844.

Kupiec, P. H. (1995), "Techniques for Verifying the Accuracy of Risk Measurement Models," *Journal of Derivatives*, 3, 73–84.

Lee, L.-F. (1999), "Estimation of Dynamic and ARCH Tobit Models," *Journal of Econometrics*, 92, 355–390.

Li, D. X. and H. J. Turtle (2000), "Semiparametric ARCH Models: An Estimating Function Approach," *Journal of Business and Economic Statistics*, 18, 174–186.

Li, W. K., S. Ling, and M. McAleer (2002), "Recent Theoretical Results for Time Series Models with GARCH Errors," *Journal of Economic Surveys*, 16, 245–269.

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

Li, W. K. and P. L. H. Yu (2003), "On the Residual Autocorrelation of the Autoregressive Conditional Duration Model," *Economics Letters*, 79, 169–175.

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

Ljung, G. M. and G. E. P. Box (1978), "On a Measure of Lack of Fit in Time Series Models," *Biometrika*, 65, 297–303.

Longin, F. and B. Solnik (2001), "Extreme Correlation of International Equity Markets," *Journal of Finance*, 56, 649–676.

Lundbergh, S. and T. Teräsvirta (2002), "Evaluating GARCH Models," *Journal of Econometrics*, 110, 417–435.

Mandelbrot, B. (1963), "The Variation of Certain Speculative Prices," *Journal of Business*, 36, 394–419.

Manganelli, S. (2005), "Duration, Volume and Volatility Impact of Trades," *Journal of Financial Markets*, 8, 377–399.

McLeod, A. I. and W. K. Li (1983), "Diagnostic Checking ARMA Time Series Models Using Squared-Residual Autocorrelations," *Journal of Time Series Analysis*, 4, 269–273.

Meitz, M. and T. Teräsvirta (2006), "Evaluating Models of Autoregressive Conditional Duration," *Journal of Business and Economic Statistics*, 24, 104–124.

Mittelhammer, R., G. Judge, and D. Miller (2000), *Econometric Foundations*, New York: Cambridge University Press.

Nelsen, R. B. (1999), *An Introduction to Copulas*, New York: Springer-Verlag.

Nelsen, R. B. (2002), "Concordance and Copulas: A Survey," in C. M. Cuadras, J. Fortiana, and J. A. Rodriguez-Lallena (eds), *Distributions with Given Marginals and Statistical Modelling*, 169–177, Dordrecht: Kluwer Academic Publishers.

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

Newey, W. K. (1985), "Maximum Likelihood Specification Testing and Conditional Moment Tests," *Econometrica*, 53, 1047–1070.

Newey, W. K. and D. L. McFadden (1994), "Large Sample Estimation and Hypothesis Testing," in R. F. Engle and D. L. McFadden (eds), *Handbook of Econometrics*, Vol. 4, 2111–2245, Amsterdam: Elsevier.

Pacurar, M. (2008), "Autoregressive Conditional Duration (ACD) Models in Finance: A Survey of the Theoretical and Empirical Literature," *Journal of Economic Surveys*, 22, 711–751.

Pagan, A. (1996), "The Econometrics of Financial Markets," *Journal of Empirical Finance*, 3, 15–102.

Pagan, A. and F. Vella (1989), "Diagnostic Tests for Models Based on Individual Data: A Survey," *Journal of Applied Econometrics*, 4, S29–S59.

Park, S. Y. and A. K. Bera (2009), "Maximum Entropy Autoregressive Conditional Heteroskedasticity Model," *Journal of Econometrics*, 150, 219–230.

Patton, A. (2004), "On the Out-of-Sample Importance of Skewness and Asymmetric Dependence for Asset Allocation," *Journal of Financial Econometrics*, 2, 130–168.

Patton, A. (2006a), "Modelling Asymmetric Exchange Rate Dependence," *International Eco-*

nomic Review, 47, 527–556.

Patton, A. (2006b), “Estimation of Multivariate Models for Time Series of Possibly Different Lengths,” *Journal of Applied Econometrics*, 21, 147–173.

Poon, S.-H., M. Rockinger, and J. Tawn (2004), “Extreme Value Dependence in Financial Markets: Diagnostics, Models, and Financial Implications,” *Review of Financial Studies*, 17, 581–610.

Rockinger, M. and E. Jondeau (2002), “Entropy Densities with an Application to Autoregressive Conditional Skewness and Kurtosis,” *Journal of Econometrics*, 106, 119–142.

Rosenblatt, M. (1952), “Remarks on a Multivariate Transformation,” *Annals of Mathematical Statistics*, 23, 470–472.

Silvennoinen, A. and T. Teräsvirta (2008), “Multivariate GARCH Models,” in T. G. Andersen, R. A. Davis, J.-P. Kriess, and T. Mikosch (eds), *Handbook of Financial Time Series*, 1–25, New York: Springer.

Sklar, A. (1959), “Fonctions de Réparition à n Dimensions et Leurs Marges,” *Publications de l′Institut de Statistique de l′Université de Paris*, 8, 229–231.

Tauchen, G. (1985), “Diagnostic Testing and Evaluation of Maximum Likelihood Models,” *Journal of Econometrics*, 30, 415–443.

Tse, Y. K. (2000), “A Test for Constant Correlations in a Multivariate GARCH Model,” *Journal of Econometrics*, 98, 107–127.

Tse, Y. K. (2002), “Residual-Based Diagnostics for Conditional Heteroscedasticity Models,” *Econometrics Journal*, 5, 358–373.

Tse, Y. K. and A. K. C. Tsui (2002), “A Multivariate GARCH Model with Time-Varying Correlations,” *Journal of Business and Economic Statistics*, 20, 351–362.

Wang, K.-L. and M.-L. Chen (2003), “The Dynamic Linkages of U.S. and Taiwan Stock Market Indices in the Pre- and Post-Asian Financial Crises Periods,” *Taiwan Economic Review*, 31, 191–252. (in Chinese)

Wang, K.-L., C. Fawson, C. B. Barrett, and J. B. McDonald (2001), “A Flexible Parametric GARCH Model with an Application to Exchange Rates,” *Journal of Applied Econometrics*, 16, 521–536.

Wang, K.-L. and C.-F. Wu (2006), “Dynamic Modeling of Spot, NDF, and DF Markets: An Investigation of Taiwan NDF Restriction Policy,” *Academia Economic Papers*, 34, 93–

126. (in Chinese)

Wei, S. X. (2002), "A Censored-GARCH Model of Asset Returns with Price Limits," *Journal of Empirical Finance*, 9, 197–223.

White, H. (1987), "Specification Testing in Dynamic Models," in T. Bewley (ed.), *Advances in Econometrics - Fifth World Congress*, Vol. I, 1–58, New York: Cambridge University Press.

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

Wong, H. and S. Ling (2005), "Mixed Portmanteau Tests for Time-Series Models," *Journal of Time Series Analysis*, 26, 569–579.

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

Wooldridge, J. M. (1991), "Specification Testing and Quasi-Maximum-Likelihood Estimation," *Journal of Econometrics*, 48, 29–55.

財務時間序列模型之估計與檢定方法
－以動差法為基礎的綜合回顧

陳宜廷 *

中央研究院經濟研究所

關鍵詞: 固定/動態條件相關係數模型、條件動差檢定、關聯函數、估計函數、一般化自我迴歸條件異質性模型、一般化殘差、動差法、最大概似法、乘式誤差模型、準最大概似法

JEL 分類代號: C12, C13, C22, C32

* 聯繫作者: 陳宜廷, 中央研究院經濟研究所, 台北市 115 南港區研究院路二段 128 號。電話: (02) 2782-2791 分機 622; 傳真: (02) 2785-3946; E-mail: ytchen@gate.sinica.edu.tw。

摘　要

本文針對以有限維度動差限制式建構之財務時間序列模型估計與檢定方法,提供一般化的綜合回顧。這些模型包含單變量部分設定架構下的一般化自我迴歸條件異質性模型,及其在單變量完整設定、多變量部分設定(固定/動態條件相關係數模型)、多變量完整設定(以關聯函數為基礎之多變量動態模型)等架構下的延伸模型,與乘式誤差模型。這個綜合回顧的一般性,係來自於一般化殘差的觀念(用以涵蓋不同模型的誤差項)以及動差法(用以得出一般化的估計與檢定)。利用這些基本觀念所整合出的方法,可應用於分析不同的條件動差與條件分配模型。本文藉此整合財務時間序列文獻中一些重要的模型以及對應的參數化估計與檢定方法,並強調這些不同的設定、估計與檢定方法背後具有簡單的通則,可供實證應用者參考。