Estimation and Model Selection of Semiparametric Copula-Based Multivariate Dynamic Models Under Copula Misspecification*

Xiaohong Chen† Yanqin Fan‡

First version: February 2004

Abstract

As a response to Granger's (2002) call for flexible modelling of the entire conditional distribution of a multivariate nonlinear time series, Chen and Fan (2003) introduced a new class of semiparametric copula-based multivariate dynamic (SCOMDY) models. A SCOMDY model specifies the conditional mean and the conditional variance of a multivariate time series parametrically, but specifies the multivariate distribution of the standardized innovation semiparametrically as a parametric copula evaluated at nonparametric marginal distributions. In this paper, we first study the large sample properties of the estimators of SCOMDY model parameters proposed in Chen and Fan (2003) under copula misspecification, and then establish pseudo likelihood ratio (PLR) tests for model selection between two SCOMDY models with possibly misspepcified copulas. The tests depend on whether the two models are generalized nonnested or generalized nested, and the limiting distributions of the test statistics are affected by nonparametric estimation of the marginal distributions of the innovations. Finally we consider PLR tests for model selection between more than two SCOMDY models in which one is the benchmark model and the rest are candidate models. Like White (2000), we do not require that all the candidate models are generalized nonnested with the benchmark model, and only assume that at least one is generalized nonnested with the benchmark. Unlike White (2000), our test automatically standardizes the PLR statistic for generalized nonnested models (with the benchmark) and ignores generalized nested models asymptotically. Simple yet novel bootstrap tests are also provided.

JEL Classification: C14; C22; G22

KEY WORDS: Multivariate dynamic models; Misspecified copulas; Multiple model selection; Semiparametric inference

---

*We thank F. Diebold, R. Engle, C. Granger, J. Hamilton, D. Hendry, A. Patton, N. Swanson and participants at the NYU/Stern seminar and the 2004 Predictive Methodology and Application in Economics and Finance Conference in honor of Clive Granger for helpful comments. Chen acknowledges support from the National Science Foundation and the C.V. Starr Center at NYU. Fan acknowledges support from the National Science Foundation.

†The corresponding author: Xiaohong Chen, Department of Economics, New York University, 269 Mercer Street, 7th Floor, New York, NY 10003, USA; E-mail: xiaohong.chen@nyu.edu; Tel: 212-998-8970; Fax: 212-995-4186.

‡Yanqin Fan, Department of Economics, Vanderbilt University, VU Station B 351819, 2301 Vanderbilt Place, Nashville, TN 37235, USA. E-mail: yanqin.fan@vanderbilt.edu

1 Introduction

Economic and financial multivariate time series are typically nonlinear, non-normally distributed, and have nonlinear co-movements beyond the first two conditional moments. Granger (2002) points out that the classical linear multivariate modelling (based on the Gaussian distribution assumption) clearly fails to explain the stylized facts observed in economic and financial time series and that it is highly undesirable to perform various economic policy evaluations, financial forecasts, and risk managements based on the classical conditional (or unconditional) Gaussian modelling. The knowledge of the multivariate conditional distribution (especially the fat-tails, asymmetry, positive or negative dependence) is essential in many important financial applications, including portfolio selection, option pricing, asset pricing models, Value-at-Risk (market risk, credit risk, liquidity risk) calculations and forecasting. Thus the entire conditional distribution of multivariate nonlinear economic and financial time series should be studied, see Granger (2002).

Recently Chen and Fan (2003) introduce a new class of semiparametric copula-based multivariate dynamic (hereafter SCOMDY) models. A SCOMDY model specifies the multivariate conditional mean and conditional variance parametrically, but specifies the distribution of the (standardized) innovations semiparametrically as a parametric copula[1] evaluated at the nonparametric univariate marginals, where the copula function captures the concurrent dependence between the components of the multivariate innovation and the marginal distributions characterize their individual behaviors. Chen and Fan (2003) demonstrate via examples the flexibility of SCOMDY models in capturing a wide range of nonlinear, asymmetric dependence structures and of the marginal behavior of a multivariate time series. In addition, a SCOMDY model allows for the estimation of multivariate conditional distribution semiparametrically, which, according to Granger (2002), is an important feature of a multivariate time series model.

There are three sets of unknown parameters associated with a SCOMDY model: the *dynamic* parameters (i.e., the finite-dimensional parameters of the conditional mean and conditional variance); the *copula dependence* parameters (i.e., the finite-dimensional parameters of the copula function of the standardized multivariate innovation); and the infinite-dimensional *marginal distributions* of each component of the standardized innovation. Chen and Fan (2003) provide simple estimators of the parameters in a correctly specified SCOMDY model and establish their asymptotic properties.

In this paper, we shall first consider estimation of the three sets of parameters associated with a SCOMDY model under a possibly misspecified parametric copula of the standardized innovation. This is motivated by the facts that financial theory and economic theory often shed little light on the specification of a parametric copula and that most of the existing applications have typically used multiple choices of parametric copulas. Under misspecification of the copula, we propose a simple

[1]A copula is simply a multivariate probability distribution function with uniform marginals.

three-step procedure to estimate all the SCOMDY model parameters. While the true unknown dynamic parameters and the marginal distributions are still estimated root-n consistently, (n is the sample size), the estimator of the copula dependence parameters will converge to the pseudo true copula dependence parameters which are defined as the minimizer of the Kullback-Leibler distance between the candidate parametric copula density and the true unknown copula density. Interestingly, the limiting distribution of the estimator of the pseudo true copula dependence parameters is not affected by the estimation of the dynamic parameters, albeit it does depend on the estimation of unknown marginal distributions as if the dynamic parameters were known.

As commonly used parametric copulas such as the Gaussian copula, the Frank copula, and the Clayton copula lead to SCOMDY models that may have very different dependence properties, one important issue in empirical implementation of any SCOMDY model is the choice of an appropriate parametric copula. A number of existing papers have attempted to address this issue for special cases of SCOMDY models. In modelling the dependence structure of multivariate high-frequency data, Breymann, et al. (2003) applied the Akaike information criterion (AIC) to five parametric copulas and chose the Student's t copula. Junker and May (2002) presented a transformed copula to model the dependence structure between risk factors in a portfolio. To compare the transformed copula with the Student's t-copula and Cook-Johnson copula, they applied a χ^2 goodness-of-fit test to each copula and selected the transformed copula on the grounds that it resulted in a smaller value of the test statistic. Granger, et al. (2003) used a conditional version of a special class of SCOMDY models to study the behavior of the conditional dependence between consumption and income over the business cycle. They considered eight alternative conditional copulas including the Gaussian, Clayton, Gumble, among others, and chose the Gumble based on maximizing the log-likelihood value. In Chen, et al. (2003) and Fermanian (2003), they respectively establish tests for the correct specification of a parametric copula for specific members of SCOMDY models. One drawback of these tests is that if the null hypothesis of correct specification is rejected, they provide no guidance as to which copula model to choose.

Admittedly, existing work have taken an important step towards formal statistical model selection in the context of copula-based models. However, there are two issues that need to be addressed. One is related to the statistical uncertainty of the goodness-of-fit criterion being used to select the best model and the other concerns the most appropriate procedure for the comparison of more than two models. In this paper, we attempt to address both issues, while allowing for misspecified parametric copulas and completely unspecified marginals.

In the case with only two models, we extend the likelihood ratio tests for model selection of parametric models in Vuong (1989) to SCOMDY models. Unlike Vuong (1989), the null hypothesis we entertain in this paper is: one copula model performs at least as well as the other in terms of the Kullback-Leibler Information Criterion (hereafter KLIC) which measures the distance between a given distribution and the true distribution, while in Vuong (1989), the null hypothesis is: the two models perform equally well. One may take one copula model as the benchmark model such as the Student's t-copula or Cook-Johnson copula in Junker and May (2002), in which case the benchmark model will be entertained unless there is strong evidence that it is outperformed by the candidate copula model. Our testing procedure is general, allowing both competing parametric copula models to be misspecified under the null and the alternative. Although our testing approach is similar to those in Vuong (1989), Sin and White (1996), Rivers and Vuong (2002), Marcellino (2002) and other work following Vuong (1989), we allow for infinite-dimensional nuisance parameters (marginal distributions) in our model selection criterion. Hence our test is really a pseudo- (or quasi-) likelihood ratio (hereafter PLR) test,[2] and the limiting distributions of our test statistics depend on the estimates of the unknown marginal distributions of the standardized innovations. We distinguish between two cases: generalized non-nested case and generalized nested case. For generalized non-nested models, the test statistic is asymptotically normally distributed and hence easy to implement. For generalized nested models, however, the null limiting distribution is given by that of a weighted sum of independent $\chi^2_{[1]}$ random variables, where the weights depend on the parametric copulas as well as the true data generating process (DGP). As a result, the test in this case is not distribution-free. This motivates us to provide a bootstrap test for generalized nested models. The novelty of our bootstrap test in this case is that it is based on bootstrapping a quadratic form, the limiting distribution of which is the null limiting distribution of the original test statistic; bootstrapping the original test statistic in this case does not work since both copula models can be misspecified. In general, one may not know a priori if the two models are generalized non-nested or nested. A sequential test is thus provided in which one first tests the hypothesis that the two models are generalized nested and then proceed to test model selection based on the result of the pretest.

As we noted earlier, in empirical applications of copulas, it is more common to use several parametric copulas to fit the data and compare the results obtained from different models. To address the model selection issue in this case, we extend the PLR test developed for two competing models to more than two models along the lines of the reality check of White (2000). In this case, the candidate copula models are compared with a benchmark copula model. If no candidate model is closer to the true model (according to the KLIC distance) than the benchmark model, the benchmark model is chosen; otherwise, the candidate model that is closest to the true model will be selected. White (2000) proposes the reality check test for the superior predictive accuracy of at least one candidate model over the benchmark model when at least one of the candidate models is nonnested with the benchmark model. Corradi and Swanson (2003, 2004) and Su and White (2003),

[2]Patton (2002) has applied Vuong (1989)'s likelihood ratio test in his study of purely parametric copula-based dynamic models. Our study differs from his since we do not specify marginal distributions of the standardized innovations.

among others, extend the reality check of White (2000) to different contexts. Hansen (2003) shows via simulation that the power of the reality check of White (2000) can be unduly influenced by certain candidate models and proposes a standardized version of the test. However, the standardized test of Hansen (2003) relies implicitly on the assumption that all candidate models are nonnested with the benchmark model and hence has limited applicability compared with the original test of White (2000). In this paper, we develop a novel test that shares the advantages of both the reality check of White (2000) and the standardized test of Hansen (2003). Our test statistic not only automatically standardizes the PLR statistic associated with generalized nonnested candidate models (with the benchmark model), but also asymptotically removes the effect of generalized nested candidate models (with the benchmark model). Consequently, our test potentially has power gains over the original reality check of White (2000) and is not restricted to the class of nonnested candidate models like the standardized test of Hansen (2003). Although the test is developed for comparison of multiple SCOMDY models, the idea is applicable to other settings such as those in White (2000), Corradi and Swanson (2003, 2004), Su and White (2003) and many other work following White's (2000) approach.

The rest of this paper is organized as follows. Section 2 briefly reviews the SCOMDY models. In Section 3, we study the large sample properties of the estimators of the SCOMDY model parameters proposed in Chen and Fan (2003) under possibly misspecified parametric copula. It is very interesting to note that the limiting distribution of the estimate of the copula dependence parameters is not affected by the estimation of the dynamic parameters of the conditional mean and conditional variance, although it does depend on the estimation of the unknown marginal distributions of the standardized innovations. This result is not only important in its own right, but also useful in establishing the asymptotic distribution of the PLR statistic under possibly misspecified copulas. In Section 4, we first present the null hypothesis and the PLR statistic for the model comparison of two SCOMDY models and then provide the limiting distributions of the PLR test statistics. It is again interesting to note that the limiting distribution does not depend on the dynamic parameter estimation, although it does depend on the estimation of marginal distribution of the standardized innovations of the SCOMDY models. Section 5 extends the above results to more than two competing SCOMDY models. Section 6 briefly concludes. All technical proofs are gathered into the Appendix.

2 Semiparametric Copula-based Multivariate Dynamic Models

Let $\{(Y_t', X_t')\}_{t=1}^n$ be a vector stochastic process in which Y_t is of dimension d, and X_t is a vector of predetermined or exogenous variables distinct from the Y's. Let $\mathcal{I}_{t-1}$ denote the information set at time t, which is the sigma-field generated by $\{Y_{t-1}, Y_{t-2}, ...; X_t, X_{t-1}, ...\}$. In Chen and Fan

(2003), they specify the class of SCOMDY models as follows:

$$Y_t = \mu_t(\theta_{o1}) + \sqrt{H_t(\theta_o)}\epsilon_t, \tag{2.1}$$

where

$$\mu_t(\theta_{o1}) = (\mu_{1,t}(\theta_{o1}), ..., \mu_{d,t}(\theta_{o1}))' = E\{Y_t|\mathcal{I}_{t-1}\}$$

is the true conditional mean of Y_t given $\mathcal{I}_{t-1}$, and is correctly parameterized up to a finite-dimensional unknown parameter θ_{o1}; and

$$\begin{aligned} H_t(\theta_o) &= diag.(h_{1,t}(\theta_o), ..., h_{d,t}(\theta_o)), \\ h_{j,t}(\theta_o) &= h_{j,t}(\theta_{o1}, \theta_{o2}) = E[(Y_{jt} - \mu_t(\theta_{o1}))^2|\mathcal{I}_{t-1}], \quad j = 1, ..., d, \end{aligned}$$

is the true conditional variance of Y_{jt} given $\mathcal{I}_{t-1}$, and is correctly parameterized up to a finite-dimensional unknown parameter $\theta_o = (\theta'_{o1}, \theta'_{o2})'$, where θ_{o1} and θ_{o2} do not have common elements. The standardized multivariate innovations $\{\epsilon_t \equiv (\epsilon_{1t}, \ldots, \epsilon_{dt})' : t \geq 1\}$ in (2.1) are independent of $\mathcal{I}_{t-1}$, and are i.i.d. distributed with $E(\epsilon_{jt}) = 0$ and $E(\epsilon_{jt}^2) = 1$ for $j = 1, \ldots, d$. Moreover, $\epsilon_t = (\epsilon_{1t}, \ldots, \epsilon_{dt})'$ has a distribution function $F^o(\epsilon) = C^o(F_1^o(\epsilon_1), \ldots, F_d^o(\epsilon_d); \alpha_o)$, where $F_j^o(\cdot)$ is the true but unknown continuous marginal of ϵ_{jt}, $j = 1, \ldots, d$, and $C^o(u_1, \ldots, u_d; \alpha_o) : [0,1]^d \to [0,1]$ is the copula function which has a continuous copula density function $c^o(u_1, \ldots, u_d; \alpha_o)$ depending on a true but unknown finite dimensional copula parameter α_o.

In Chen and Fan (2003), they provided many examples of SCOMDY models by combining different specifications of $\mu_t(\theta_{o1})$, $H_t(\theta_o)$ and $C^o(u_1, \ldots, u_d; \alpha_o)$. Basically, $\mu_t(\theta_{o1})$ and $H_t(\theta_o)$ can take almost all the commonly used conditional mean and conditional variance specifications such as ARCH, GARCH, VAR, Markov switching, etc., see Granger and Teräsvirta (1993), Hamilton (1994), Tsay (2002), and all the chapters on dependent processes in the Handbook of Econometrics, Vol. 4, edited by Engle and McFadden (1994). Similarly, $C^o(u_1, \ldots, u_d; \alpha_o)$ can be any parametric copula function such as the Normal (or Gaussian) copula, the Student's t-copula, the Frank copula, the Gumble copula, and the Clayton copula, see Joe (1997) and Nelsen (1999) for examples and properties of copulas.

We conclude this section using the following SCOMDY examples as illustration.

Example 1 (GARCH(1,1)+Normal copula): For $j = 1, ..., d$,

$$\begin{aligned} Y_{jt} &= X'_{jt}\delta_j + \sqrt{h_{jt}}\epsilon_{jt}, \\ h_{jt} &= \kappa_j + \beta_j h_{j,t-1} + \gamma_j(Y_{j,t-1} - X'_{j,t-1}\delta_j)^2, \end{aligned} \tag{2.2}$$

where

$$\kappa_j > 0, \ \beta_j \geq 0, \ \gamma_j \geq 0, \text{ and } \beta_j + \gamma_j < 1, \ j = 1, \ldots, d.$$

In terms of our notation, $\theta_1 = (\delta_1, ..., \delta_d)'$, $\theta_2 = (\kappa_1, ..., \kappa_d; \beta_1, ..., \beta_d; \gamma_1, ..., \gamma_d)'$, $\mu_t = (X_{1t}'\delta_1, ..., X_{dt}'\delta_d)$ and $H_t = diag\{h_{1t}, ..., h_{dt}\}$. The standardized multivariate innovations $\{\epsilon_t \equiv (\epsilon_{1t}, \ldots, \epsilon_{dt})' : t \geq 1\}$ are independent of $\mathcal{I}_{t-1}$, and are i.i.d. distributed with $E(\epsilon_{jt}) = 0$ and $E(\epsilon_{jt}^2) = 1$ for $j = 1, \ldots, d$.

The copula of $\epsilon_t = (\epsilon_{1t}, \ldots, \epsilon_{dt})'$ is assumed to be a normal copula with unknown correlation matrix $\alpha = \Sigma$. A $d-$ dimensional normal copula is derived from the $d-$ dimensional Gaussian distribution. Let Φ denote the scalar standard normal distribution, and $\Phi_{\Sigma,d}$ the $d-$dimensional normal distribution with correlation matrix Σ. Then the $d-$dimensional normal copula with correlation matrix Σ is

$$C(\mathbf{u};\Sigma) = \Phi_{\Sigma,d}(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_d)),$$

whose copula density is

$$c(\mathbf{u};\Sigma) = \frac{1}{\sqrt{\det(\Sigma)}} \exp\left\{-\frac{(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_d))' \left(\Sigma^{-1} - I_d\right) \left(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_d)\right)}{2}\right\}.$$

Normal copula with $\Sigma \neq 0$ generates joint symmetric dependence, but there is no tail dependence (i.e., there is no joint extreme events).

Example 2 (GARCH(1,1)+Student's t-copula): the conditional mean and conditional variance of Y_{jt}, $j = 1, ..., d$, are specified in the same way as those in Example 1.

The copula of $\epsilon_t = (\epsilon_{1t}, \ldots, \epsilon_{dt})'$ is assumed to be a Student's t- copula with unknown correlation matrix $\alpha = \Sigma$. A $d-$ dimensional t-copula is derived from the $d-$ dimensional Student's t-distribution. Let T_ν be the scalar standard Student's t distribution with $\nu > 2$ degrees of freedom, and $T_{\Sigma,\nu}$ be the $d-$dimensional Student's t distribution with $\nu > 2$ degrees of freedom and a shape matrix Σ. Then the $d-$dimensional Student's t-copula with correlation matrix Σ is

$$C(\mathbf{u};\Sigma,\nu) = T_{\Sigma,\nu}(T_\nu^{-1}(u_1), ..., T_\nu^{-1}(u_d)).$$

The Student's t copula density is:

$$\begin{aligned} c(\mathbf{u};\Sigma,\nu) &= \frac{\Gamma(\frac{\nu+d}{2})[\Gamma(\frac{\nu}{2})]^{d-1}}{\sqrt{\det(\Sigma)}[\Gamma(\frac{\nu+1}{2})]^d} \left(1 + \frac{\mathbf{x}'\Sigma^{-1}\mathbf{x}}{\nu}\right)^{-\frac{\nu+d}{2}} \prod_{i=1}^{d} \left(1 + \frac{x_i^2}{\nu}\right)^{\frac{\nu+1}{2}}, \\ \text{where } \mathbf{x} &= (x_1, ..., x_d)', \quad x_i = T_\nu^{-1}(u_i). \end{aligned}$$

The Student's t copula with $\Sigma \neq 0$ can generate joint symmetric tail dependence, hence allow for *joint* fat tails (i.e., an increased probability of joint extreme events).

Example 3 (GARCH(1,1)+Clayton copula): the conditional mean and conditional variance of Y_{jt}, $j = 1, ..., d$, are specified in the same way as those in Example 1.

The copula of $\epsilon_t = (\epsilon_{1t}, \ldots, \epsilon_{dt})'$ is assumed to be the Clayton copula:

$$C(u_1, \ldots, u_d; \alpha) = [u_1^{-\alpha} + \ldots + u_d^{-\alpha} - d + 1]^{-1/\alpha}, \quad \text{where } \alpha > 0. \tag{2.3}$$

The copula density of the Clayton copula is given by

$$c(u_1, \ldots, u_d; \alpha) = \{\Pi_{j=1}^d [1 + (j-1)\alpha]\}\{\Pi_{j=1}^d u_j^{-(\alpha+1)}\}[\sum_{j=1}^d u_j^{-\alpha} - d + 1]^{-(\alpha^{-1}+d)}, \quad \text{where } \alpha > 0.$$

Unlike the Gaussian and Student's t copulas, the Clayton copula can generate asymmetric dependence. In particular, the Clayton copula has lower tail dependence, but no upper tail dependence.

All three examples have been applied in empirical finance. For instance, Example 1 has been used in Hull and White (1998) for value-at-risk calculation for asset returns and for exchange rates. It can be regarded as a special case of the DCC model proposed in Engle (2002) and Engle and Sheppard (2001). Examples 2 and 3 have been used in Breymann, et al. (2003) and Junker and May (2002) for joint tail dependence and risk management modelling for multivariate high frequency data. But none of the existing work have considered model selection tests with possibly misspecified copulas.

3 Estimation of SCOMDY Model Parameters under Copula Misspecification

In this section, we first review the simple estimators of parameters in a SCOMDY model proposed in Chen and Fan (2003) and then establish their large sample properties when the copula is misspecified.

3.1 Estimation of model parameters

Let $\epsilon_t(\theta) \equiv [H_t(\theta)]^{-1/2}(Y_t - \mu_t(\theta_1))$ be the innovation function. The log-likelihood function for the SCOMDY model with a candidate copula function $C(u_1, \ldots, u_d; \alpha)$ is, (up to a constant term)

$$\begin{aligned} L_n(\theta, f; \alpha) &= \frac{1}{n}\sum_{t=1}^n l_t(\theta, f; \alpha) \\ &= \frac{1}{n}\sum_{t=1}^n \left\{ \frac{-\log|H_t(\theta)|}{2} + \sum_{j=1}^d \log f_j(\epsilon_{jt}(\theta)) + \log c(F_1(\epsilon_{1t}(\theta)), \ldots, F_d(\epsilon_{dt}(\theta)); \alpha) \right\} \end{aligned}$$

where $|H_t(\theta)|$ denotes the determinant of $H_t(\theta)$, $c(u_1, \ldots, u_d; \alpha)$ is the copula density function associated with the copula function $C(u_1, \ldots, u_d; \alpha)$, and $f = (f_1, ..., f_d)$ with f_j being the unknown probability density function (pdf) of F_j, $j = 1, ..., d$. As the marginal distributions F_j are completely unspecified, we normalize the mean and variance of the innovation ϵ_{jt} such that $E[\epsilon_{jt}(\theta)] = 0$ and $Var[\epsilon_{jt}(\theta)] = 1$ for $j = 1, \ldots, d$.

We need to estimate three sets of parameters θ_o, $(F_1^o, \ldots, F_d^o)$ and α^*, where α^* is defined as

$$\alpha^* \equiv \arg\max_{\alpha \in \mathcal{A}} E^0[\log c(F_1^o(\epsilon_{1t}(\theta_o)), \ldots, F_d^o(\epsilon_{dt}(\theta_o)); \alpha)]$$

$$= \arg\max_{\alpha\in\mathcal{A}} \int_{[0,1]^d} \log c(u_1,\ldots,u_d;\alpha) c^o(u_1,\ldots,u_d;\alpha_o) du_1\cdots du_d.$$

If the copula density $c(u_1,\ldots,u_d;\alpha)$ correctly specifies the true copula density up to the copula parameter α, then α^* equals the true value α_o. Otherwise, the copula density function $c(u_1,\ldots,u_d;\alpha^*)$ is the closest in the family of parametric copula densities $\{c(u_1,\ldots,u_d;\alpha) : \alpha\in\mathcal{A}\}$ to the true copula density in terms of minimizing the KLIC.

We first consider the estimation of θ_o and $(F_1^o,\ldots,F_d^o)$. The following estimators are proposed in Chen and Fan (2003). The parameter θ_{o1} is estimated by OLS:

$$\widetilde{\theta}_1 = \arg\max_{\theta_1\in\Theta_1}\{\frac{-1}{2n}\sum_{t=1}^{n}[Y_t-\mu_t(\theta_1)]'[Y_t-\mu_t(\theta_1)]\};$$

and the parameter θ_{o2} is estimated by QMLE:

$$\widetilde{\theta}_2 = \arg\max_{\theta_2\in\Theta_2}\frac{-1}{2n}\sum_{t=1}^{n}\sum_{j=1}^{d}\left\{\frac{(Y_{jt}-\mu_t(\widetilde{\theta}_1))^2}{h_{j,t}(\widetilde{\theta}_1,\theta_2)}+\log h_{j,t}(\widetilde{\theta}_1,\theta_2)\right\}.$$

Given the estimator $\tilde{\theta}$, one can estimate F_j^o using the rescaled empirical distribution of $\{\epsilon_{jt}(\tilde{\theta})\}_{t=1}^n$:

$$\widetilde{F}_{nj}(x) = \frac{1}{n+1}\sum_{t=1}^{n}1(\epsilon_{jt}(\tilde{\theta})\le x),\quad j=1,...,d. \tag{3.1}$$

Since the estimators $\tilde{\theta}$ and $(\widetilde{F}_{n1},...,\widetilde{F}_{nd})$ do not depend on the parametric copula specification, their asymptotic properties established in Chen and Fan (2003) still hold. In particular, under mild regularity conditions, $\widetilde{\theta}$ is a $\sqrt{n}-$consistent estimator for θ_o and $(\widetilde{F}_{n1},...,\widetilde{F}_{nd})$ is $\sqrt{n}-$consistent estimator of $(F_1^o,\ldots,F_d^o)$.

Given $(\tilde{\theta},\, \widetilde{F}_{n1},...\widetilde{F}_{nd})$, α^* can be estimated by $\hat{\alpha}$:

$$\hat{\alpha} = \arg\max_{\alpha\in\mathcal{A}}\frac{1}{n}\sum_{t=1}^{n}\log c(\widetilde{F}_{n1}(\epsilon_{1t}(\tilde{\theta})),\ldots,\widetilde{F}_{nd}(\epsilon_{dt}(\tilde{\theta}));\alpha). \tag{3.2}$$

Chen and Fan (2003) establish the asymptotic properties of $\hat{\alpha}$ when the copula is correctly specified. In the next subsection, we will extend their results to the case where the copula is misspecified.

3.2 Asymptotic properties of $\hat{\alpha}$ under copula misspecification

The difficulty in establishing the asymptotic properties of the estimator $\hat{\alpha}$ arises from the fact that for many widely used copula functions including the Gaussian copula, the t-copula and the Clayton copula, the score function and its derivatives blow up to infinity. To handle this difficulty, Chen and Fan (2003) first establish a weighted uniform CLT for the empirical process $\sqrt{n}(\widetilde{F}_{nj}(\cdot)-F_j^o(\cdot))$ based on pseudo-observations $\{\epsilon_{jt}(\widetilde{\theta})\}$ and then used it to prove the $\sqrt{n}-$consistency of $\hat{\alpha}$ under the correct specification of parametric copula. In this section we modify their result to obtain the $\sqrt{n}-$consistent estimation of α^* under misspecified copula.[3]

Let $\mathcal{A}$ be the parameter space, which is a compact subset of $\mathcal{R}^a$. For $\alpha \in \mathcal{A}$, we use $||\alpha - \alpha^*||$ to denote the usual Euclidean metric.

Proposition 3.1 *Under Assumptions D and C stated in the Appendix, we have:* $||\widehat{\alpha}-\alpha^*|| = o_p(1)$.

Proposition 3.1 states that the estimator $\hat{\alpha}$ is a consistent estimator of the pseudo true value α^*. If the parametric copula correctly specifies the true copula in the sense that there exists $\alpha_o \in \mathcal{A}$ such that $C(v_1, \ldots, v_d; \alpha_o) = C^o(v_1, \ldots, v_d)$ for almost all $(v_1, \ldots, v_d) \in (0,1)^d$, then $\alpha^* = \alpha_o$ and $\hat{\alpha}$ consistently estimate α_o.

In the following we denote $l(v_1, \ldots, v_d, \alpha) = \log c(v_1, \ldots, v_d, \alpha)$, $l_\alpha(v_1, \ldots, v_d, \alpha) = \frac{\partial l(v_1,\ldots,v_d,\alpha)}{\partial \alpha}$, $l_j(v_1, \ldots, v_d, \alpha) = \frac{\partial l(v_1,\ldots,v_d,\alpha)}{\partial v_j}$, $l_{\alpha\alpha}(v_1, \ldots, v_d; \alpha) = \frac{\partial^2 l(v_1,\ldots,v_d;\alpha)}{\partial \alpha \partial \alpha'}$ and $l_{\alpha j}(v_1, \ldots, v_d; \alpha) = \frac{\partial^2 l(v_1,\ldots,v_d;\alpha)}{\partial v_j \partial \alpha}$ for $j = 1, \ldots, d$. Define $U_{jt} \equiv F_j^o(\epsilon_{jt}(\theta_o))$ for $j = 1, \ldots, d$ and $U_t = (U_{1t}, \ldots, U_{dt})'$. Denote

$$A_n^* \equiv \frac{1}{n} \sum_{s=1}^{n} \{l_\alpha(U_{1s}, \ldots, U_{ds}, \alpha^*) + \sum_{j=1}^{d} Q_{\alpha j}(U_{js}; \alpha^*)\},$$

where

$$Q_{\alpha j}(U_{js}; \alpha^*) \equiv E^0\{l_{\alpha j}(U_t; \alpha^*)[I\{U_{js} \leq U_{jt}\} - U_{jt}]|U_{js}\}.$$

We also denote $B \equiv -E^0[l_{\alpha\alpha}(U_t; \alpha^*)]$ and $\Sigma \equiv Var^0[l_\alpha(U_s; \alpha^*) + \sum_{j=1}^{d} Q_{\alpha j}(U_{js}; \alpha^*)]$, and assume that both B and Σ are finite, positive definite.

Proposition 3.2 *Let* $\alpha^* \in int(\mathcal{A})$. *Under Assumptions D and N stated in the Appendix, we have: (1)* $\widehat{\alpha} - \alpha^* = B^{-1}A_n^* + o_p(n^{-1/2})$; *(2)* $\sqrt{n}(\widehat{\alpha} - \alpha^*) \to \mathcal{N}(0, B^{-1}\Sigma B^{-1})$ *in distribution.*

The additional terms $Q_{\alpha j}(U_{js}; \alpha^*)$ in A_n^* are introduced by the need to estimate the marginal distribution function $F_j^o(\cdot)$. In the case where the distribution $F_j^o(\cdot)$ is completely known, these terms will disappear from A_n^*. It is interesting to note that the asymptotic variance of $\hat{\alpha}$ does not depend on the functional form of the marginal distribution F_j^o. It is even more interesting to observe that the limiting distribution of $\widehat{\alpha}$ is not affected by the estimation of the dynamic parameters θ_o.
Remark: The asymptotic variance of $\widehat{\alpha}$ can be consistently estimated as $\widehat{B}^- \widehat{\Sigma} \widehat{B}^-$, where $\widehat{B}^-$ is the generalized inverse of

$$\widehat{B} = -\frac{1}{n} \sum_{t=1}^{n} l_{\alpha\alpha}(\tilde{U}_t; \hat{\alpha}),$$

[3] Although White (1982) established the asymptotic properties of the maximum likelihood estimator under misspecified parametric models, his results are not directly applicable here since the estimation of the copula dependence parameter in a SCOMDY model under copula misspecification depends on the estimates of the unknown marginal distributions.

where $\tilde{U}_t = (\tilde{U}_{1t}, \ldots, \tilde{U}_{dt})' = (\tilde{F}_{n1}(\epsilon_{1t}(\widetilde{\theta})), ..., \tilde{F}_{nd}(\epsilon_{1t}(\widetilde{\theta})))'$; and

$$\widehat{\Sigma} = \frac{1}{n}\sum_{t=1}^{n}[l_\alpha(\tilde{U}_t;\hat{\alpha}) + \sum_{j=1}^{d}\widehat{Q}_{\alpha j}(\widetilde{U}_{jt};\hat{\alpha})][l_\alpha(\tilde{U}_t;\hat{\alpha}) + \sum_{j=1}^{d}\widehat{Q}_{\alpha j}(\widetilde{U}_{jt};\hat{\alpha})]',$$

with

$$\widehat{Q}_{\alpha j}(U_{jt};\hat{\alpha}) = \frac{1}{n}\sum_{s=1,s\neq t}^{n}\Big(l_{\alpha j}(U_s;\hat{\alpha})\{I_{\{U_{jt}\leq U_{js}\}} - U_{js}\} + l_{\alpha j}(U_t;\hat{\alpha})\{I_{\{U_{js}\leq U_{jt}\}} - U_{jt}\}\Big).$$

Any inference drawn based on $\hat{\alpha}$ and this variance estimator $\widehat{B}^{-}\widehat{\Sigma}\widehat{B}^{-}$ would still be valid except that it is on the pseudo true value α^* and the estimated parametric copula estimates the closest copula in the parametric family to the true copula in terms of minimizing the KLIC.

4 Pseudo Likelihood Ratio Tests for Model Selection between Two SCOMDY Models

In this section we first introduce the appropriate PLR statistic for testing model selection between two SCOMDY models along the lines of Vuong (1989). We then establish the limiting distribution of the PLR statistic.

4.1 The PLR statistic

For each $i = 1, 2$, let $\{C_i(u_1, \ldots, u_d; \alpha_i) : \alpha_i \in \mathcal{A}_i \subset \mathcal{R}^{a_i}\}$ be a class of parametric copulas. Assuming that the conditional mean μ_t and the conditional variance H_t are correctly specified, we are interested in selecting a parametric copula such that the resulting SCOMDY model is closer to the true SCOMDY model from which the multivariate time series $\{Y_t\}_{t=1}^n$ is generated. Let

$$\ell_{t,i}(\alpha) = -\frac{1}{2}\log|H_t(\theta_o)| + \sum_{j=1}^{d}\log f_j^o(\epsilon_{jt}(\theta_o)) + \log c_i(F_1^o(\epsilon_{1t}(\theta_o)), \ldots, F_d^o(\epsilon_{dt}(\theta_o)); \alpha_i),$$

in which $c_i(\cdot;\alpha_i)$ is the density function of the copula $C_i(\cdot;\alpha_i)$, and $f_j^o(\cdot)$ is the density function of the true marginal cdf $F_j^o(\cdot)$ of $\epsilon_{jt}(\theta_o)$, $j = 1, ..., d$. Throughout this paper, we let $E^0[\cdot]$ denote the expectation of $\cdot$ taken with respect to the true distribution $C^o(F_1^o(\cdot), \ldots, F_d^o(\cdot); \alpha_o)$. Denote $\alpha_i^* = \arg\max_{\alpha_i\in\mathcal{A}_i} E^0[\ell_{t,i}(\alpha_i)]$ as the pseudo true value associated with the copula model $i = 1, 2$. Clearly we have

$$\alpha_i^* = \arg\max_{\alpha_i\in\mathcal{A}_i}\int_{[0,1]^d}\log c_i(u_1, \ldots, u_d; \alpha_i)c^o(u_1, \ldots, u_d; \alpha_o)du_1\cdots du_d.$$

Hence the value of α_i^* depends on both the parametric copula $c_i()$ and the true copula $c^o()$.

Following Vuong (1989), we measure the closeness of a SCOMDY model to the true model by the minimum of the KLIC over the distributions in the copula model or equivalently by the maximum of $E^0[\ell_{t,i}(\alpha_i)]$. Since only the third term in the expression for $\ell_{t,i}(\alpha_i)$ depends on the copula, an equivalent measure of the closeness of the i-th copula model to the true copula model is $E^0 \log [c_i(F_1^o(\epsilon_{1t}(\theta_o)), \ldots, F_d^o(\epsilon_{dt}(\theta_o)); \alpha_i^*)]$, $i = 1, 2$; the larger $E^0 \log [c_i(F_1^o(\epsilon_{1t}(\theta_o)), \ldots, F_d^o(\epsilon_{dt}(\theta_o)); \alpha_i^*)]$, the closer is the model to the true model. This motivates the following hypotheses: For pseudo true values α_1^* and α_2^*, the null hypothesis is

$$H_0: \; E^0 \left\{ \log \frac{c_2(F_1^o(\epsilon_{1t}(\theta_o)), \ldots, F_d^o(\epsilon_{dt}(\theta_o)); \alpha_2^*)}{c_1(F_1^o(\epsilon_{1t}(\theta_o)), \ldots, F_d^o(\epsilon_{dt}(\theta_o)); \alpha_1^*)} \right\} \leq 0,$$

meaning that the copula model with the copula $C_1(\cdot; \alpha_1)$ is not worse than the copula model with the copula $C_2(\cdot; \alpha_2)$, and the alternative hypothesis is

$$H_1: \; E^0 \left\{ \log \frac{c_2(F_1^o(\epsilon_{1t}(\theta_o)), \ldots, F_d^o(\epsilon_{dt}(\theta_o)); \alpha_2^*)}{c_1(F_1^o(\epsilon_{1t}(\theta_o)), \ldots, F_d^o(\epsilon_{dt}(\theta_o)); \alpha_1^*)} \right\} > 0,$$

meaning that the copula model with $C_1(\cdot; \alpha_1)$ is worse than the copula model with $C_2(\cdot; \alpha_2)$.

In the above formulation, one can take the copula model $C_1(\cdot; \alpha_1)$ as the benchmark model and the model $C_2(\cdot; \alpha_2)$ as a candidate model. Given the prevalence of the Gaussian distribution in multivariate financial time series modelling, it is natural to take the Gaussian copula model as the benchmark model; the Gaussian copula model will be retained unless the test strongly suggests that the candidate model outperforms the Gaussian copula model. In Junker and May (2002), the benchmark model is either the Student's t-copula or the Cook-Johnson copula.

Define

$$LR_n(\theta_o, F_1^o, \ldots, F_d^o; \alpha_2^*, \alpha_1^*) = \frac{1}{n} \sum_{t=1}^{n} \left\{ \log \frac{c_2(F_1^o(\epsilon_{1t}(\theta_o)), \ldots, F_d^o(\epsilon_{dt}(\theta_o)); \alpha_2^*)}{c_1(F_1^o(\epsilon_{1t}(\theta_o)), \ldots, F_d^o(\epsilon_{dt}(\theta_o)); \alpha_1^*)} \right\}.$$

Let $\hat{\alpha}_i$ denote the two-step estimator of α_i^* for the SCOMDY model with copula $C_i(u_1, \ldots, u_d; \alpha_i)$, $i = 1, 2$. Our tests will be based on the following PLR statistic:

$$LR_n(\tilde{\theta}, \widetilde{F}_{n1}, \ldots, \widetilde{F}_{nd}; \hat{\alpha}_2, \hat{\alpha}_1) = \frac{1}{n} \sum_{t=1}^{n} \left\{ \log \frac{c_2(\widetilde{F}_{n1}(\epsilon_{1t}(\widetilde{\theta})), \ldots, \widetilde{F}_{nd}(\epsilon_{dt}(\widetilde{\theta})); \hat{\alpha}_2)}{c_1(\widetilde{F}_{n1}(\epsilon_{1t}(\widetilde{\theta})), \ldots, \widetilde{F}_{nd}(\epsilon_{dt}(\widetilde{\theta})); \hat{\alpha}_1)} \right\}.$$

4.2 Asymptotic properties of the PLR statistic and PLR tests

As will be shown later, the asymptotic distribution of the PLR statistic takes different form depending on whether the two closest parametric copulas to the true copula are equal. To distinguish between these two cases, we introduce the concept of generalized non-nested and of generalized nested copula models.

Definition 4.1 **(i)** *Two models are generalized non-nested if the set* $\{(v_1, \ldots, v_d) : c_1(v_1, \ldots, v_d; \alpha_1^*) \neq c_2(v_1, \ldots, v_d; \alpha_2^*)\}$ *has positive Lebesgue measure;*

(ii) *Two models are generalized nested if* $c_1(v_1, \ldots, v_d; \alpha_1^*) = c_2(v_1, \ldots, v_d; \alpha_2^*)$ *for almost all* $(v_1, \ldots, v_d) \in (0, 1)^d$.

It is important to note that as the closest copula in a parametric class of copulas depends on the true copula, it is not obvious a priori whether two parametric classes of copulas are generalized non-nested or generalized nested. However commonly used parametric classes of copulas such as the Clayton copula, the Gumble copula, and the Gaussian copula can be shown to be generalized non-nested unless the closest member to the true copula in each class is the independence copula.

We first obtain the probability limit of the PLR statistic:

Proposition 4.1 *Suppose for $i = 1, 2$, the copula model i satisfies assumptions of Proposition 3.1 and C6 in the Appendix. Then:* $LR_n(\widetilde{\theta}, \tilde{F}_{n1}, \ldots, \tilde{F}_{nd}; \hat{\alpha}_2, \hat{\alpha}_1) - E^0\left[\log \frac{c_2(U_{1t},\ldots,U_{dt};\alpha_2^*)}{c_1(U_{1t},\ldots,U_{dt};\alpha_1^*)}\right] = o_p(1)$.

In the following for $i = 1, 2$, we let $l_i(v_1, \ldots, v_d, \alpha_i) \equiv \log c_i(v_1, \ldots, v_d, \alpha_i)$, $B_i \equiv -E^0[l_{i,\alpha\alpha}(U_t; \alpha_i^*)]$ and

$$Q_{i,j}(U_{js}, \alpha_i^*) \equiv E^0\{l_{i,j}(U_t; \alpha_i^*)[I\{U_{js} \leq U_{jt}\} - U_{jt}]|U_{js}\} \quad \text{for } j = 1, ..., d. \tag{4.1}$$

THEOREM 4.2 *Suppose for $i = 1, 2$, the copula model i satisfies assumptions of Proposition 3.2. Then:*

(1) for the generalized non-nested case,

$$n^{1/2}\left\{LR_n(\widetilde{\theta}, \tilde{F}_{n1}, \ldots, \tilde{F}_{nd}; \hat{\alpha}_2, \hat{\alpha}_1) - E^0\left[\log \frac{c_2(U_{1t}, \ldots, U_{dt}; \alpha_2^*)}{c_1(U_{1t}, \ldots, U_{dt}; \alpha_1^*)}\right]\right\} \rightarrow \mathcal{N}(0, \sigma^2),$$

where

$$\sigma^2 = Var^0\left[\log \frac{c_2(U_{1t}, \ldots, U_{dt}; \alpha_2^*)}{c_1(U_{1t}, \ldots, U_{dt}; \alpha_1^*)} + \sum_{j=1}^{d}\{Q_{2,j}(U_{jt}; \alpha_2^*) - Q_{1,j}(U_{jt}; \alpha_1^*)\}\right]. \tag{4.2}$$

(2) for the generalized nested case,

$$\begin{aligned} 2nLR_n(\widetilde{\theta}, \tilde{F}_{n1}, \ldots, \tilde{F}_{nd}; \hat{\alpha}_2, \hat{\alpha}_1) &= n(\alpha_2^* - \hat{\alpha}_2)'B_2(\alpha_2^* - \hat{\alpha}_2) - n(\alpha_1^* - \hat{\alpha}_1)'B_1(\alpha_1^* - \hat{\alpha}_1) + o_p(1) \\ &\rightarrow M_{a_1+a_2}(\cdot; \lambda^*), \end{aligned}$$

where $M_{a_1+a_2}(\cdot; \lambda^*)$ is the distribution of a weighted sum of independent $\chi^2_{[1]}$ random variables with unknown weights $\lambda^* = (\lambda_1^*, \ldots, \lambda_{a_1+a_2}^*)'$, in which the weights depend on the two parametric copulas and the true unknown distribution function.

Compared with Theorem 3.3 in Vuong (1989), the variance of the asymptotic distribution of the PLR statistic for generalized non-nested models has the additional term due to $\sum_{j=1}^{d}\{Q_{2,j}(U_{jt}; \alpha_2^*) - Q_{1,j}(U_{jt}; \alpha_1^*)\}$. All of these additional terms are introduced by the first step estimation of the unknown marginal distributions F_j^o, $j = 1, \ldots, d$. However, it is very interesting to note that the limiting distribution of the PLR statistic does not depend on the functional forms of the unknown marginal distributions F_j^o, $j = 1, \ldots, d$, nor does it depend on the estimation of the dynamic parameters θ_o.

The following proposition shows that $\sigma^2 > 0$ if and only if the two copula models are generalized non-nested. Consequently, for generalized non-nested models, the null limiting distribution of $n^{1/2} LR_n(\widetilde{\theta}, \tilde{F}_{n1}, \ldots, \tilde{F}_{nd}; \hat{\alpha}_2, \hat{\alpha}_1)$ is a normal distribution with a positive variance. This is the basis for the PLR test for model selection in generalized non-nested case developed in this paper.

Proposition 4.3 *Let* $\sigma_a^2 = Var^0[\log \frac{c_2(U_{1t},\ldots,U_{dt};\alpha_2^*)}{c_1(U_{1t},\ldots,U_{dt};\alpha_1^*)}]$ *and* σ^2 *be given by (4.2). Under conditions of Theorem 4.2,* $\sigma^2 = 0$ *if and only if* $\sigma_a^2 = 0$*; and* $\sigma_a^2 = 0$ *if and only if the two copula models under selection are generalized nested.*

4.3 A PLR test for selection of generalized non-nested models

Unlike Vuong (1989), for generalized non-nested models, the null hypothesis in our paper is a composite hypothesis. As a result, the asymptotic distribution of the PLR statistic under the null is not uniquely determined, see Theorem 4.2(1). The usual approach to handling this problem is based on the Least Favorable Configuration (hereafter LFC) which is the point least favorable to the alternative. In our case, the LFC satisfies $E^0\left[\log \frac{c_2(U_{1t},\ldots,U_{dt};\alpha_2^*)}{c_1(U_{1t},\ldots,U_{dt};\alpha_1^*)}\right] = 0$. Under the LFC, Theorem 4.2(1) implies that $n^{1/2} LR_n(\widetilde{\theta}, \tilde{F}_{n1}, \ldots, \tilde{F}_{nd}; \hat{\alpha}_2, \hat{\alpha}_1) \to \mathcal{N}(0, \sigma^2)$. Moreover, $\sigma^2 > 0$ by Proposition 4.3. We now provide a consistent estimator of σ^2.

First, by the definition of $Q_{i,j}(U_{js}; \alpha_i^*)$ in (4.1), we have $E^0\{Q_{i,j}(U_{js}; \alpha_i^*)\} = 0$ for $i = 1, 2$ and $j = 1, ..., d$. Moreover given U_{js}, $Q_{i,j}(U_{js}; \alpha_i^*)$ can be estimated by

$$\widehat{Q}_{i,j}(U_{js}, \hat{\alpha}_i) = \frac{1}{n} \sum_{t=1, t\neq s}^{n} [l_{i,j}(U_t; \hat{\alpha}_i)\{I_{\{U_{js} \leq U_{jt}\}} - U_{jt}\} + l_{i,j}(U_s; \hat{\alpha}_i)\{I_{\{U_{jt} \leq U_{js}\}} - U_{js}\}]. \quad (4.3)$$

Then a consistent estimator of σ^2 is given by $\hat{\sigma}^2 =$

$$\frac{1}{n} \sum_{t=1}^{n} \left[\log \frac{c_2(\tilde{U}_t; \hat{\alpha}_2)}{c_1(\tilde{U}_t; \hat{\alpha}_1)} - \frac{1}{n} \sum_{s=1}^{n} \log \frac{c_2(\tilde{U}_s; \hat{\alpha}_2)}{c_1(\tilde{U}_s; \hat{\alpha}_1)} + \sum_{j=1}^{d} \{\widehat{Q}_{2,j}(\tilde{U}_{jt}; \hat{\alpha}_2) - \widehat{Q}_{1,j}(\tilde{U}_{jt}; \hat{\alpha}_1)\} \right]^2, \quad (4.4)$$

where $\tilde{U}_t = (\tilde{U}_{1t}, \ldots, \tilde{U}_{dt})' = (\tilde{F}_{n1}(\epsilon_{1t}(\widetilde{\theta})), ..., \tilde{F}_{nd}(\epsilon_{1t}(\widetilde{\theta})))'$.

Define the PLR statistic for the selection of generalized non-nested models as

$$T_n^N = \frac{n^{1/2} LR_n(\widetilde{\theta}, \tilde{F}_{n1}, \ldots, \tilde{F}_{nd}; \hat{\alpha}_2, \hat{\alpha}_1)}{\hat{\sigma}}, \quad (4.5)$$

where the superscript "N" in T_n^N is meant for non-nested models and normal limiting distributions.

THEOREM 4.4 *Suppose the conditions of Proposition 4.1 and Theorem 4.2 hold and the two models are generalized non-nested. Then under the LFC,* $T_n^N \to \mathcal{N}(0, 1)$.

Proposition 4.1 and Theorem 4.4 suggest the following directional test: Given a significance level α, reject H_0 in favor of H_1 if $T_n^N > Z_\alpha$, where Z_α is the upper α-percentile of the standard normal distribution (i.e., Z_α is the value of the inverse standard normal distribution evaluated at $1 - \alpha$).

4.4 A PLR test for selection of generalized nested models

We now consider the case where under H_0, the two models are generalized nested. In this case, the null hypothesis becomes a simple hypothesis. Define the test statistic:

$$T_n^Q = 2nLR_n(\widetilde{\theta}, \tilde{F}_{n1}, \ldots, \tilde{F}_{nd}; \hat{\alpha}_2, \hat{\alpha}_1), \tag{4.6}$$

where the superscript "Q" in T_n^Q is meant for nested models and quadratic limiting statistics.

Theorem 4.2(2) implies that in this case the null limiting distribution of the PLR statistic T_n^Q is not distribution-free and is a complicated function of both the parametric copulas and the true distribution function. Moreover, one needs to compute and estimate eigenvalues of a complicated $(a_1 + a_2)$ dimensional matrix in order to estimate the asymptotic critical values.

A typical solution to this problem is provided by the method of bootstrap to approximate the critical values of the test. Typically, in order for a bootstrap test to work, the bootstrap sample must satisfy the null model. In our case, both parametric copulas can be misspecified and hence the null hypothesis does not specify a complete null model. Instead, we will rely on Efron's naive bootstrap, but use the distribution of $2nD_n$ instead of T_n^Q computed on the bootstrap sample, where

$$D_n = \frac{1}{2}(\hat{\alpha}_2 - \alpha_2^*)' B_2 (\hat{\alpha}_2 - \alpha_2^*) - \frac{1}{2}(\hat{\alpha}_1 - \alpha_1^*)' B_1 (\hat{\alpha}_1 - \alpha_1^*). \tag{4.7}$$

This is motivated by the observation that in the generalized nested case, the null limiting distribution of T_n^Q is given by the limiting distribution of $2nD_n$. Specifically,

Step 1: Draw a random sample $\{\tilde{\epsilon}_t^*\}_{t=1}^n$ of size n from the residuals $\{\widetilde{\epsilon}_t \equiv \epsilon_t(\widetilde{\theta})\}_{t=1}^n$ with replacement. This leads to one bootstrap sample.

Step 2: Compute the bootstrap estimates $\widetilde{F}_{nj}^*(x) = \frac{1}{n+1}\sum_{t=1}^n 1(\tilde{\epsilon}_{jt}^* \leq x)$. Let $\widetilde{U}_{jt}^* = \widetilde{F}_{nj}^*(\tilde{\epsilon}_{jt}^*)$, $j = 1, \ldots, d$.

Step 3: Compute the bootstrap estimates $\hat{\alpha}_i^* = \arg\max_{\alpha \in \mathcal{A}_i}[n^{-1}\sum_{t=1}^n \log c_i(\widetilde{U}_{1t}^*, ..., \widetilde{U}_{dt}^*; \alpha)]$ for $i = 1, 2$.

Step 4: Compute $2D_n^* = (\hat{\alpha}_2^* - \hat{\alpha}_2)'\hat{B}_2(\hat{\alpha}_2^* - \hat{\alpha}_2) - (\hat{\alpha}_1^* - \hat{\alpha}_1)'\hat{B}_1(\hat{\alpha}_1^* - \hat{\alpha}_1)$,
where $\hat{B}_i = n^{-1}\sum_{t=1}^n l_{i,\alpha\alpha}(\tilde{F}_{n1}(\widetilde{\epsilon}_{1t}), \ldots, \tilde{F}_{nd}(\widetilde{\epsilon}_{dt}); \hat{\alpha}_i)$ for $i = 1, 2$;

Step 5: Repeat Steps 1-4 a large number of times and use the empirical distribution of the resulting values of $2nD_n^*$ to approximate the null distribution of the test statistic T_n^Q.

Throughout the paper, we let $P^*(\cdot|\{(Y_t, X_t)\}_{t=1}^n)$ denote the probability law of the resampled series, conditional on the data. Note that the above bootstrap procedure does not rely on any specific forms of $\mu_t(\theta)$ and $H_t(\theta)$, since the estimator $\hat{\alpha}_i$ and the test statistic T_n^Q depends on the data $\{(Y_t, X_t)\}_{t=1}^n$ only through $\{\tilde{\epsilon}_t\}_{t=1}^n$. Moreover, the asymptotic distribution of $\tilde{\alpha}_i$ and the null asymptotic distribution of T_n^Q are the same as the case where θ_o is known. This together with the fact that $\{\epsilon_t(\theta_o)\}_{t=1}^n$ is i.i.d. ensure that the above bootstrap procedure works.

THEOREM 4.5 *Under the conditions of Theorem 4.2(2), $P^*(2nD_n^* \leq x | \{(Y_t, X_t)\}_{t=1}^n)$ converges in probability to $M_{a_1+a_2}(x; \lambda^*)$.*

4.5 A sequential PLR test for model selection

The tests presented in the previous subsections are general in the sense that they apply to cases where both parametric copula models could be misspecified, and none of the competing copula models are required to be correctly specified.[4] However, one needs to know whether the two models are generalized non-nested. As the pseudo-values α_1^* and α_2^* are unknown, it is unknown a priori if this is the case. Vuong (1989) suggests a sequential test in which one first tests if the two models are generalized non-nested and then determines which test to use based on the result of the pre-test.

The null hypothesis of generalized nested models can be tested by testing the null hypothesis $\sigma_a^2 = 0$. A consistent estimator of σ_a^2 is given by

$$\hat{\sigma}_a^2 = \frac{1}{n}\sum_{t=1}^n \left[\log\left\{ \frac{c_2(\tilde{U}_t; \hat{\alpha}_2)}{c_1(\tilde{U}_t; \hat{\alpha}_1)} \right\} - \frac{1}{n}\sum_{s=1}^n \log\left\{ \frac{c_2(\tilde{U}_s; \hat{\alpha}_2)}{c_1(\tilde{U}_s; \hat{\alpha}_1)} \right\} \right]^2, \tag{4.8}$$

where $\tilde{U}_t = (\tilde{U}_{1t}, \ldots, \tilde{U}_{dt})' = (\tilde{F}_{n1}(\epsilon_{1t}(\widetilde{\theta})), ..., \tilde{F}_{nd}(\epsilon_{1t}(\widetilde{\theta})))'$. In the following we denote $\lambda^{*2} = (\lambda_1^{*2}, ..., \lambda_{a_1+a_2}^{*2})'$ as the vector of squares of $\lambda^* = (\lambda_1^*, ..., \lambda_{a_1+a_2}^*)'$, the eigenvalue weights in Theorem 4.2(2).

THEOREM 4.6 *Under the conditions of Proposition 4.1 and Theorem 4.2, we have:*

(1) $\hat{\sigma}_a^2$ given in (4.8) is a consistent estimator of σ_a^2;

(2) When $\sigma_a^2 = 0$, $n\hat{\sigma}_a^2 \to M_{a_1+a_2}(\cdot; \lambda^{*2})$ in distribution.

Theorem 4.6 and Proposition 4.3 suggest that a sequential test can be constructed in our case as well. First, tests the null hypothesis that the two copula models are generalized nested by using the test statistic $n\hat{\sigma}_a^2$; if the pretest suggests that the the two models are generalized nested, then stop; otherwise proceed to use the test T_n^N for H_0.

Like the null limiting distribution of T_n^Q, that of $n\hat{\sigma}_a^2$ is not distribution-free. We propose to use bootstrap to approximate its null distribution. The proof of Theorem 4.6 reveals that the null limiting distribution of $\hat{\sigma}_a^2$ is given by the limiting distribution of V_d, where $V_d =$

$$\begin{bmatrix} \hat{\alpha}_2 - \alpha_2^* \\ \hat{\alpha}_1 - \alpha_1^* \end{bmatrix}' \begin{bmatrix} E^0 l_{2,\alpha}(U_t; \alpha_2^*)' l_{2,\alpha}(U_t; \alpha_2^*) & -E^0 l_{2,\alpha}(U_t; \alpha_2^*)' l_{1,\alpha}(U_t; \alpha_1^*) \\ -E^0 l_{1,\alpha}(U_t; \alpha_1^*)' l_{2,\alpha}(U_t; \alpha_2^*) & E^0 l_{1,\alpha}(U_t; \alpha_1^*)' l_{1,\alpha}(U_t; \alpha_1^*) \end{bmatrix} \begin{bmatrix} \hat{\alpha}_2 - \alpha_2^* \\ \hat{\alpha}_1 - \alpha_1^* \end{bmatrix} \tag{4.9}$$

[4]Our test procedure follows the approach of Vuong (1989) and Rivers and Vuong (2002), which contrasts with Cox's (1962) non-nested testing procedure by not requiring one of the competing models to be correct under the null hypothesis.

This motivates us to bootstrap nV_d as follows. First we follow Steps 1-3 in Section 4.4 and then follow Steps 4-5 below:

Step 4: Compute V_d^* as $V_d^* =$

$$\begin{bmatrix} \hat{\alpha}_2^* - \hat{\alpha}_2 \\ \hat{\alpha}_1^* - \hat{\alpha}_1 \end{bmatrix}' \begin{bmatrix} \frac{1}{n}\sum_{t=1}^n l_{2,\alpha}(\tilde{U}_t;\hat{\alpha}_2)' l_{2,\alpha}(\tilde{U}_t;\hat{\alpha}_2) & \frac{-1}{n}\sum_{t=1}^n l_{2,\alpha}(\tilde{U}_t;\hat{\alpha}_2)' l_{1,\alpha}(\tilde{U}_t;\hat{\alpha}_1) \\ \frac{-1}{n}\sum_{t=1}^n l_{1,\alpha}(\tilde{U}_t;\hat{\alpha}_1)' l_{2,\alpha}(\tilde{U}_t;\hat{\alpha}_2) & \frac{1}{n}\sum_{t=1}^n l_{1,\alpha}(\tilde{U}_t;\hat{\alpha}_1)' l_{1,\alpha}(\tilde{U}_t;\hat{\alpha}_1) \end{bmatrix} \begin{bmatrix} \hat{\alpha}_2^* - \hat{\alpha}_2 \\ \hat{\alpha}_1^* - \hat{\alpha}_1 \end{bmatrix}$$

Step 5: Repeat Steps 1-4 a large number of times and use the empirical distribution of the resulting values of nV_d^* to approximate the null distribution of the test statistic $n\hat{\sigma}_a^2$.

THEOREM 4.7 *Under the conditions of Theorem 4.6, $P^*(nV_d^* \le x | \{(Y_t, X_t)\}_{t=1}^n)$ converges in probability to $M_{a_1+a_2}(x;\lambda^{*2})$.*

5 PLR Tests for Model Selection between Multiple SCOMDY Models

In empirical applications of copulas, several parametric copulas are often used to fit the data and the results from models based on these copulas are then compared, see e.g. Breymann et al. (2003), Junker and May (2002) and Granger, et al. (2003). The PLR tests developed in the previous sections can be extended to the comparison of more than two copulas along the lines of White (2000). In this case, all the candidate copula models are compared with a benchmark copula model. If no candidate model is closer to the true model than the benchmark model according to the KLIC distance, the benchmark model is chosen; otherwise, the candidate model that is closest to the true model will be selected. As mentioned earlier, one natural benchmark model is the Gaussian copula model, although the test applies to any benchmark model.

Let $\{C_i(u_1,\ldots,u_d;\alpha_i) : \alpha_i \in \mathcal{A}_i \subset \mathcal{R}^{a_i}\}$ be a class of parametric copulas with $i = 1, 2, \ldots, M$. As in the previous sections, we are interested in selecting a parametric copula such that the resulting SCOMDY model with copula $C_i(u_1,\ldots,u_d;\alpha_i)$ is closest to the true SCOMDY model with unknown copula $C^o(u_1,\ldots,u_d)$. This can be formulated as follows. Let $C_1(u_1,\ldots,u_d;\alpha_1)$ be the benchmark model and $\{C_i(u_1,\ldots,u_d;\alpha_i)\}_{i=2}^M$ be the candidate models. We are interested in testing if the best candidate model outperforms the benchmark model according to the KLIC distance. Hence, for pseudo true values α_i^*, $i = 1,\ldots,M$, the null hypothesis is

$$H_0^M : \max_{i=2,\ldots,M} E^0 \left\{ \log \frac{c_i(F_1^o(\epsilon_{1t}(\theta_o)),\ldots,F_d^o(\epsilon_{dt}(\theta_o));\alpha_i^*)}{c_1(F_1^o(\epsilon_{1t}(\theta_o)),\ldots,F_d^o(\epsilon_{dt}(\theta_o));\alpha_1^*)} \right\} \le 0,$$

meaning that no candidate copula model is closer to the true model than the benchmark model, and the alternative hypothesis is

$$H_1^M : \max_{i=2,\ldots,M} E^0 \left\{ \log \frac{c_i(F_1^o(\epsilon_{1t}(\theta_o)),\ldots,F_d^o(\epsilon_{dt}(\theta_o));\alpha_i^*)}{c_1(F_1^o(\epsilon_{1t}(\theta_o)),\ldots,F_d^o(\epsilon_{dt}(\theta_o));\alpha_1^*)} \right\} > 0,$$

meaning that there exists a candidate copula model that is closer to the true model than the benchmark model.

Our test will be based on the following PLR statistics ($i = 2, \dots, M$):

$$LR_n(\widetilde{\theta}, \widetilde{F}_{n1}, \dots, \widetilde{F}_{nd}; \hat{\alpha}_i, \hat{\alpha}_1) = \frac{1}{n} \sum_{t=1}^{n} \left\{ \log \frac{c_i(\widetilde{F}_{n1}(\epsilon_{1t}(\widetilde{\theta})), \dots, \widetilde{F}_{nd}(\epsilon_{dt}(\widetilde{\theta})); \hat{\alpha}_i)}{c_1(\widetilde{F}_{n1}(\epsilon_{1t}(\widetilde{\theta})), \dots, \widetilde{F}_{nd}(\epsilon_{dt}(\widetilde{\theta})); \hat{\alpha}_1)} \right\}.$$

In the following we denote $\Omega = (\sigma_{ik})_{i,k=2}^{M}$ in which

$$\begin{aligned} \sigma_{ik} \quad = \quad & \mathrm{Cov}^0 \left[\log \frac{c_i(U_{1t}, \dots, U_{dt}; \alpha_i^*)}{c_1(U_{1t}, \dots, U_{dt}; \alpha_1^*)} + \sum_{j=1}^{d} \{ Q_{i,j}(U_{jt}; \alpha_i^*) - Q_{1,j}(U_{jt}; \alpha_1^*) \}, \right. \\ & \left. \log \frac{c_k(U_{1t}, \dots, U_{dt}; \alpha_k^*)}{c_1(U_{1t}, \dots, U_{dt}; \alpha_1^*)} + \sum_{j=1}^{d} \{ Q_{k,j}(U_{jt}; \alpha_k^*) - Q_{1,j}(U_{jt}; \alpha_1^*) \} \right], \end{aligned}$$

where for $i = 1, 2, ..., M$ and $j = 1, ..., d$,

$$Q_{i,j}(U_{jt}, \alpha_i^*) \equiv E^0 \{ l_{i,j}(U_s; \alpha_i^*) [I\{U_{jt} \le U_{js}\} - U_{js}] | U_{jt} \}.$$

Proposition 5.1 *Suppose that for $i = 1, 2, \dots, M$, the copula model i satisfies conditions of Theorem 4.2. Suppose that Ω is positive semi-definite and its largest eigenvalue is positive. Then jointly*

$$n^{1/2} \left\{ LR_n(\widetilde{\theta}, \tilde{F}_{n1}, \dots, \tilde{F}_{nd}; \hat{\alpha}_i, \hat{\alpha}_1) - E^0 \left[\log \frac{c_i(U_{1t}, \dots, U_{dt}; \alpha_i^*)}{c_1(U_{1t}, \dots, U_{dt}; \alpha_1^*)} \right] \right\}_{i=2,\dots,M} \to (Z_2, \dots, Z_M)',$$

in distribution, where $(Z_2, \dots, Z_M)' \sim \mathcal{N}(0, \Omega)$. Hence

$$\max_{i=2,\dots,M} n^{1/2} \left\{ LR_n(\widetilde{\theta}, \tilde{F}_{n1}, \dots, \tilde{F}_{nd}; \hat{\alpha}_i, \hat{\alpha}_1) - E^0 \left[\log \frac{c_i(U_t; \alpha_i^*)}{c_1(U_t; \alpha_1^*)} \right] \right\} \to \max_{i=2,\dots,M} Z_i \text{ in dist.}$$

Proposition 5.1 implies that under the LFC (i.e., $E^0 \left\{ \log \frac{c_i(U_{1t}, \dots, U_{dt}; \alpha_i^*)}{c_1(U_{1t}, \dots, U_{dt}; \alpha_1^*)} \right\} = 0$ for $i = 2, ..., M$), $\max_{i=2,\dots,M} [n^{1/2} LR_n(\widetilde{\theta}, \tilde{F}_{n1}, \dots, \tilde{F}_{nd}; \hat{\alpha}_i, \hat{\alpha}_1)] \to \max_{i=2,\dots,M} Z_i$ in distribution, which could be used to construct White's (2000) Reality Check (RC) test. However, Hansen (2003) shows via simulation that the power of the reality check of White (2000) can be unduly influenced by large σ_{ii} and a standardized version of the test improves the power of the reality check a great deal. In our context, the standardized test is based on

$$T_{nM} = \max \left[\max_{i=2,\dots,M} \frac{n^{1/2} LR_n(\widetilde{\theta}, \widetilde{F}_{n1}, \dots, \widetilde{F}_{nd}; \hat{\alpha}_i, \hat{\alpha}_1)}{\sqrt{\hat{\sigma}_{ii}}}, 0 \right],$$

where $\hat{\sigma}_{ii}$ is a consistent estimator of σ_{ii}, the i-th diagonal element of Ω, defined as:

$$\hat{\sigma}_{ii} = \frac{1}{n} \sum_{t=1}^{n} \left[\log \frac{c_i(\tilde{U}_t; \hat{\alpha}_i)}{c_1(\tilde{U}_t; \hat{\alpha}_1)} - \frac{1}{n} \sum_{s=1}^{n} \log \frac{c_i(\tilde{U}_s; \hat{\alpha}_i)}{c_1(\tilde{U}_s; \hat{\alpha}_1)} + \sum_{j=1}^{d} \{ \widehat{Q}_{i,j}(\widetilde{U}_{jt}; \hat{\alpha}_i) - \widehat{Q}_{1,j}(\widetilde{U}_{jt}; \hat{\alpha}_1) \} \right]^2,$$

where $\widehat{Q}_{i,j}$ is a consistent estimator of $Q_{i,j}$ and is computed the same way as in (4.3).

It is important to point out that Hansen's (2003) standardized test depends on Ω being positive definite, which requires that none of the candidate models and the benchmark model are generalized nested. This assumption can be restrictive in particular when a large number of candidate models are entertained in a given empirical application. This motivates us to propose the following test statistic:

$$T_{nI} = \max\left[\max_{i=2,\ldots,M}\left\{\frac{n^{1/2}LR_n(\widetilde{\theta},\widetilde{F}_{n1},\ldots,\widetilde{F}_{nd};\hat{\alpha}_i,\hat{\alpha}_1)}{\sqrt{\hat{\sigma}_{ii}}}G_b(\hat{\sigma}_{ii})\right\},0\right],$$

where $b = b_n \to 0$ as $n \to \infty$ and $G_b(\cdot)$ is a smoothed trimming function which trims out small $\hat{\sigma}_{ii}$.

We use the following smoothed trimming that has recently been used by Andrews (1995), Ai (1997) and Linton and Xiao (2001). Let $g(\cdot)$ be a density function that has support $[0,1]$, $g(0) = g(1) = 0$, and let

$$g_b(x) = \frac{1}{b}g(\frac{x}{b}-1),$$

where b is the trimming parameter, then $g_b(x)$ has support on $[b, 2b]$. Letting

$$G_b(x) = \int_{-\infty}^{x} g_b(z)dz,$$

we have

$$G_b(x) = \begin{cases} 0, & x < b \\ \int_{-\infty}^{x} g_b(z)dz, & b \leq x \leq 2b \\ 1, & x > 2b. \end{cases}$$

For example, consider the following Beta density

$$g(z) = B(a+1)^{-1}z^a(1-z)^a,\ 0 \leq z \leq 1,$$

for some positive integer a, where $B(a)$ is the beta function defined by $B(a) = \Gamma(a)^2/\Gamma(2a)$, and $\Gamma(a)$ is the Euler gamma function. Then it can be verified that the function $G_b(x)$ is $(a+1)$-times continuously differentiable on $[0,1]$, see Linton and Xiao (2001). We will suppose that $a \geq 1$.

THEOREM 5.2 *Suppose that the copula model $i = 1, 2, ..., M$ satisfies conditions of Proposition 4.1 and Proposition 5.1. If $b \to 0$ and $nb \to \infty$, then under the null hypothesis we have: $T_{nI} = T_{nM'} + o_p(1)$, where*

$$T_{nM'} = \max\left[\max_{i\in S_N}\left\{\frac{n^{1/2}LR_n(\widetilde{\theta},\widetilde{F}_{n1},\ldots,\widetilde{F}_{nd};\hat{\alpha}_i,\hat{\alpha}_1)}{\sqrt{\hat{\sigma}_{ii}}}\right\},0\right],$$

where $S_N = \{i \in \{2, ..., M\}$: Model i and the benchmark model are generalized nonnested$\}$ and M' is the number of candidate models that are generalized non-nested with the benchmark model.

Theorem 5.2 implies that the null limiting distribution of T_{nI} is the same as that of the standardized test $T_{nM'}$ applied to the set of candidate models that are generalized non-nested with the benchmark model. That is, the trimming by $G_b(\hat{\sigma}_{ii})$ in the test statistic T_{nI} removes the effect of generalized nested models (with the benchmark model) on its limiting distribution. For candidate models that are generalized nonnested with the benchmark model, Proposition 4.3 shows that the corresponding variance-covariance matrix in the limiting distribution of

$$n^{1/2}\left\{LR_n(\widetilde{\theta},\tilde{F}_{n1},\ldots,\tilde{F}_{nd};\hat{\alpha}_i,\hat{\alpha}_1)-E^0\left[\log\frac{c_i(U_{1t},\ldots,U_{dt};\alpha_i^*)}{c_1(U_{1t},\ldots,U_{dt};\alpha_1^*)}\right]\right\}_{i\in S_N}$$

is positive definite. Moreover, Proposition 5.1 implies that the test statistic $T_{nM'}$ satisfies Assumptions 1 and 2 in Hansen (2003) and hence its asymptotic distribution under the null hypothesis depends only on all candidate models $i'\in S_N$ for which

$$E^0\left\{\log\frac{c_{i'}(F_1^o(\epsilon_{1t}(\theta_o)),\ldots,F_d^o(\epsilon_{dt}(\theta_o));\alpha_{i'}^*)}{c_1(F_1^o(\epsilon_{1t}(\theta_o)),\ldots,F_d^o(\epsilon_{dt}(\theta_o));\alpha_1^*)}\right\}=0. \tag{5.1}$$

To sum up, Theorem 1 in Hansen (2003) together with our Proposition 5.1 and Theorem 5.2 imply the following result.

Corollary 5.3 *Under the conditions of Theorem 5.2, the null limiting distribution of* T_{nI} *is given by that of* $\max[\max_{i\in S_{LN}}\frac{Z_i}{\sqrt{\sigma_{ii}}},0]$, *where* $S_{LN}=\{i:\ i\in S_N$ *and model* i *satisfies* $(5.1)\}$.

To identify models in S_N that satisfy (5.1), we follow Hansen (2003) by trimming out small estimates of

$$E^0\left\{\log\frac{c_{i'}(F_1^o(\epsilon_{1t}(\theta_o)),\ldots,F_d^o(\epsilon_{dt}(\theta_o));\alpha_{i'}^*)}{c_1(F_1^o(\epsilon_{1t}(\theta_o)),\ldots,F_d^o(\epsilon_{dt}(\theta_o));\alpha_1^*)}\right\}$$

in the following bootstrap procedure used to approximate the unknown null limiting distribution of T_{nI}.

Let $\tilde{U}_t=(\tilde{U}_{1t},\ldots,\tilde{U}_{dt})'=(\tilde{F}_{n1}(\epsilon_{1t}(\widetilde{\theta})),...,\tilde{F}_{nd}(\epsilon_{1t}(\widetilde{\theta})))'$. Under general conditions, the following bootstrap procedure works. Follow Steps 1-3 defined in section 4.4 and then:

Step 4. Let $\hat{W}_{ti}=\log\frac{c_i(\tilde{U}_{1t},\ldots,\tilde{U}_{dt};\hat{\alpha}_i)}{c_1(\tilde{U}_{1t},\ldots,\tilde{U}_{dt};\hat{\alpha}_1)}$. Calculate its bootstrap value $\hat{W}_{ti}^*=\log\frac{c_i(\widetilde{U}_{1t}^*,\ldots,\widetilde{U}_{dt}^*;\hat{\alpha}_i^*)}{c_1(\widetilde{U}_{1t}^*,\ldots,\widetilde{U}_{dt}^*;\hat{\alpha}_1^*)}$ and define the recentered value as:

$$\hat{W}_{tic}^*=\hat{W}_{ti}^*-[\frac{1}{n}\sum_{t=1}^n\hat{W}_{ti}]I\{\frac{1}{n}\sum_{t=1}^n\hat{W}_{ti}\geq -a_n\},\quad i=2,...,M,$$

where $a_n\to 0$ is a small positive (possibly random) number such that $\sqrt{n}a_n\to\infty$.

Step 5. Compute the bootstrap value $\hat{\sigma}_{ii}^*$ of $\hat{\sigma}_{ii}$, $i=2,...,M$, and define the bootstrap value of T_{nI} as

$$T_{nI}^*=\max\left[\max_{i=2,\ldots,M}\frac{n^{-1/2}\sum_{t=1}^n\hat{W}_{tic}^*}{\sqrt{\hat{\sigma}_{ii}^*}}G_b(\hat{\sigma}_{ii}^*),0\right].$$

Step 6. Repeat Steps 1-5 for a large number of times and use the empirical distribution function of the resulting values T^*_{nI} to approximate the null distribution of T_{nI}.

It is worthwhile to summarize the roles played by the two trimming functions: $G_b(\cdot)$ used in defining the test statistic T_{nI} and the trimming of $n^{-1}\sum_{t=1}^{n}\hat{W}_{ti}$ from below by $-a_n$ in Step 4. of the bootstrap procedure. The trimming function $G_b(\cdot)$ removes the effect of candidate models that are generalized nested with the benchmark model on the limiting distribution of T_{nI}, and the second trimming identifies among the class of generalized nonnested candidate models (with the benchmark model) the ones that are not strictly dominated by the benchmark model under the null hypothesis.

THEOREM 5.4 *Under the conditions of Theorem 5.2, the conditional distribution of T^*_{nI} given $\{(Y_t, X_t)\}_{t=1}^{n}$ converges in probability to the null limiting distribution of T_{nI}.*

6 Conclusion

Recently Chen and Fan (2003) proposed a class of SCOMDY models which specify the conditional mean and the main diagonal of the conditional covariance of a multivariate time series parametrically, but specify the multivariate distribution of the standardized innovation semiparametrically as a parametric multivariate copula evaluated at nonparametric marginal distributions. They demonstrated that the class of SCOMDY models can capture the entire conditional distribution of a multivariate nonlinear time series flexibly, and studied the estimation of a SCOMDY model under the correct specification of the model. Specific members of the class of SCOMDY models have been applied in empirical finance and insurance, but all the existing applications have used multiple copula specifications without taking into account the statistical uncertainty involved in copula model selection.

In this paper, we first extend the large sample properties of the estimators of SCOMDY model parameters proposed in Chen and Fan (2003) under copula misspecification. Interestingly enough, the limiting distribution of the estimator of the pseudo true copula dependence parameter is not affected by the estimation of the dynamic parameters, albeit it does depend on the estimation of unknown marginal distributions as if the dynamic parameters were known. Therefore, the common practice in empirical finance of ignoring the estimation error of the dynamic parameters is theoretically justified according to our first order large sample theory. Nevertheless, our results show that the statistical uncertainty of the goodness-of-fit model selection criterion cannot be ignored. We establish PLR tests for model selection of two SCOMDY models with possibly misspecified parametric copulas for both generalized nonnested copulas and generalized nested copulas. Finally we consider the PLR test for model selection between more than two SCOMDY models in which one is the benchmark model and the rest are candidate models. Here we assume that the benchmark and at least one of the candidate models are generalized nonnested; our test has advantages over both the reality check of White (2000) and the standardized test of Hansen (2003).

We are currently working on several extensions of the results reported in this paper. First, instead of the in-sample PLR model comparison, we could consider out-of-sample PLR model comparison. Second, we could follow the encompassing approach to perform in-sample and/or out-of-sample multiple SCOMDY model comparison. See Hendry and Richard (1982), Mizon and Richard (1986), Diebold (1989), White (1994), Clements and Hendry (1998), West (2001) and many others for the encompassing tests of model comparison. Third, since some of the copula applications are in terms of option pricing and forecasting, we could consider alternative loss functions instead of the KLIC, see e.g. Machina and Granger (2000), Elliott and Timmermann (2002) and Su and White (2003). Finally, we could consider the misspecification and model comparison of conditional mean, conditional variance, and copula specifications jointly. This will be related to the model comparison of semiparametric multivariate conditional distributions. The ideas in Diebold, et al. (1999), Giacomini and White (2003), and Corradi and Swanson (2004) might be useful here.

References

[1] Ai, C. (1997), "A Semiparametric Maximum Likelihood Estimator," *Econometrica* 65, 933-963.

[2] Andrews, D. W. K. (1995), "Nonparametric Kernel Estimation for Semiparametric Models," *Econometric Theory* 11, 560-596.

[3] Breymann, W. , A. Dias, and Embrechts, P. (2003), "Dependence Structures for Multivariate High-Frequency Data in Finance," *Quantitative Finance* 3(1), 1-16.

[4] Chen, X. and Y. Fan (2003), "Specification and Estimation of A New Class of Semiparametric Copula-based Multivariate Dynamic Models," Manuscript, New York University and Vanderbilt University.

[5] Chen, X., Y. Fan and A. Patton (2003), "Simple Tests for Models of Dependence Between Multiple Financial Time Series: with Applications to U.S. Equity Returns and Exchange Rates," Manuscript, New York University, Vanderbilt University and London School of Economics.

[6] Clements, M. and D. Hendry (1998), *Forecasting Economic Time Series*, Cambridge University Press.

[7] Corradi, V. and N. Swanson (2003), "A Test for Comparing Multiple Misspecified Conditional Distributions," Manuscript, Rutgers University.

[8] Corradi, V. and N. Swanson (2004), "Predictive Density Accuracy Tests," Manuscript, Rutgers University.

[9] Cox, D. (1962), "Further Results on Tests of Separate Families of Hypotheses," *J. R. Statist. Soc. B* 24, 406-424.

[10] Diebold, F.X. (1989), "Forecast Combination and Encompassing: Reconciling Two Divergent Literature," *International Journal of Forecasting* 5, 589-592.

[11] Diebold, F.X., J. Hahn and A. Tay (1999), "Multivariate Density Forecast Evaluation and Calibration in Financial Risk Management: High Frequency Returns on Foreign Exchange," *Review of Economics and Statistics* 81, 661-673.

[12] Elliott, G. and A. Timmermann (2002), "Optimal Forecast Combinations under General Loss Functions and Forecast Error Distributions," Manuscript, UCSD.

[13] Engle, R. (2002), "Dynamic Conditional Correlation-A Simple Class of Multivariate GARCH Models," Forthcoming in *Journal of Business and Economic Statistics.*

[14] Engle, R. and D. McFadden (1994), *Handbook of Econometrics, Vol. 4*, Elsevier: New York.

[15] Engle, R. and K. Sheppard (2001), "Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH," Working paper, UCSD and New York University.

[16] Fermanian, J.-D. (2003), "Goodness of Fit Tests for Copulas," Mimeo, CREST.

[17] Genest, C. , K. Ghoudi, and L.-P. Rivest (1995), "A Semiparametric Estimation Procedure of Dependence Parameters in Multivariate Families of Distributions," *Biometrika* 82 (3), 543-552.

[18] Giacomini, R. and H. White (2003), "Tests of Conditional Predictive Ability," Working paper, UCSD.

[19] Gonçalves, S. and H. White (2004), "Maximum Likelihood and the Bootstrap for Nonlinear Dynamic Models," forthcoming in *Journal of Econometrics*.

[20] Granger, C.W.J. (2002), "Time Series Concept for Conditional Distributions," Manuscript, UCSD.

[21] Granger, C.W.J. and T. Teräsvirta (1993), *Modelling Nonlinear Economic Relationships*. Oxford: Oxford University Press.

[22] Granger, C.W.J., T. Teräsvirta and A. Patton (2003), "Common Factors in Conditional Distributions," Forthcoming in *Journal of Econometrics*.

[23] Hamilton, J. (1994), *Time Series Analysis*. Princeton: Princeton University Press.

[24] Hansen, R. P. (2003), "A Test for Superior Predictive Ability," Manuscript, Brown University.

[25] Hendry, D. and J. Richard (1982), "On the Formulation of Empirical Models in Dynamic Econometrics," *Journal of Econometrics* 20, 3-33.

[26] Hull, J. and A. White (1998): "Value at Risk when Daily Changes in Market Variables are not Normally Distributed," *Journal of Derivatives* 5, 9-19.

[27] Joe, H. (1997), *Multivariate Models and Dependence Concepts*, Chapman & Hall/CRC.

[28] Junker, M. and A. May (2002), "Measurement of Aggregate Risk With Copulas," Preprint, Caesar.

[29] Linton, O. and Z. Xiao (2001), "A Nonparametric Regression Estimator that Adapts to Error Distribution of Unknown Form," Manuscript, London School of Economics.

[30] Machina, M. and C. Granger (2000), "Forecast Comparison Without the Use of Cost Functions," Working paper, UCSD.

[31] Marcellino, M. (2002), "Model Selection for Nonlinear Dynamic Models," Manuscript.

[32] Mizon, G. and J. Richard (1986), "The Encompassing Principle and its Application to Testing Non-nested Hypotheses," *Econometrica* 54, 657-678.

[33] Nelsen, R. B. (1999), *An Introduction to Copulas*, Springer.

[34] Patton, A.J. (2002), "Modeling Time-Varying Exchange Rate Dependence Using the Conditional Copula," Working Paper 01-09, University of California, San Diego.

[35] Rivers, D. and Q. H. Vuong (2002), "Model Selection Tests for Nonlinear Dynamic Models," *Econometric Review*.

[36] Sin, C. and H. White (1996), "Information Criteria for Selecting Possibly Misspecified Parametric Models," *Journal of Econometrics* 71, 207-225.

[37] Su, L. and H. White (2003), "Multiple Forecast Comparison under General Loss Functions: $\overline{\mathcal{X}}^2$ and Resampling Methods," Working paper, UCSD.

[38] Tsay, R. (2002), *Analysis of Financial Time Series*, Wiley Interscience: New York.

[39] Vuong, Q. H. (1989), "Likelihood Ratio Test for Model Selection and Non-nested Hypotheses," *Econometrica* 57, 307-333.

[40] West, K. (2001), "Encompassing Tests When no Model is Encompassing," *Journal of Econometrics* 105, 287-308.

[41] White, H. (1982), "Maximum Likelihood Estimation of Misspecified Models," *Econometrica* 50, 143-161.

[42] White, H. (1994), *Estimation, Inference and Specification Analysis*, Cambridge University Press, Cambridge, UK.

[43] White, H. (2000), "A Reality Check for Data Snooping," *Econometrica* 68, 1097-1126.