Estimating and Evaluating the Predictive Abilities of Semiparametric Multivariate Models with Application to Risk Management

Valentyn Panchenko*

February 28, 2006

CeNDEF, University of Amsterdam
Roetersstraat 11, 1018WB Amsterdam, The Netherlands
v.panchenko@uva.nl

Abstract

In this paper we propose a new semiparametric procedure for estimating multivariate models with conditioning variables. The semiparametric model is based on the parametric conditional copula of Patton (2005a) and nonparametric conditional marginals. To avoid the curse of dimensionality in the estimation of the latter, we propose a dimension reduction technique. The marginals are estimated using conditional kernel smoothers based on local linear estimator. The semiparametric copula model is compared with the parametric DCC model using predictive likelihood as a criterion. The comparison is based on the recent conditional test for predictive abilities of Giacomini & White (2005). We use various simulations and financial series

*Work in progress: preliminary and incomplete. I express my gratitude to Dr. Cees Diks and Professor Cars Hommes for helpful comments and enthusiastic supervision. I also thank to seminar participants at Tinbergen Institute Amsterdam, University of New South Wales, University of Adelaide and the Quantitative Methods in Finance conference in Sydney. The usual disclaimers apply.

to compare the methods and show when the proposed semiparametric model is expected to be superior to the fully parametric DCC model.

Keywords: risk management, copula, correlation, multivariate time series, nonparametric conditional distribution

JEL classification: G15, C51, C12, C13, C32

1 Introduction

Financial institutions have to address the important problem of quantifying and hedging their exposure to market risk. This risk arises from a volatile market environment, i.e. changing prices of equities, commodities, exchange rates and interest rates. A popular measure for quantifying a market risk as a single number is Value at Risk (VaR). VaR is the maximum expected loss over a period of time, at a given confidence level. Statistically, the VaR of a portfolio is a specific conditional quantile of the distribution function of the portfolio's value over a given future time period (refer to Engle & Manganelli, 2004 for review and recent methodology). In this paper, we reduce the methodological question of the VaR estimation to the prediction of the portfolio value distribution. The latter is derived from the forecast of the multivariate asset return distribution. We specify a flexible semiparametric model for predicting the future multivariate asset return distribution. The major question addressed here is whether the semiparametric model is able to produce a more accurate forecast than a fully parametric multivariate model.

To specify a multivariate model we use the concept of copula. Copulas can fully characterize the dependence structure of a multivariate time series. They have proven to be a useful device in many financial applications; e.g. risk management, portfolio aggregation, spillover effects (for an extensive review of applications see Bouyé et al., 2000, Embrechts et al., 2003). The Basel Committee on Banking and Supervision recognized copulas as one of the important risk management techniques, which is reflected in the New Basel Capital Accord (Basel II).

The copula approach allows for flexible modeling of marginal distributions and their mapping into a joint distribution through a copula function. The marginals may be specified either parametrically (e.g. Patton, 2005b) or nonparametrically (e.g. Genest et al., 1995). The latter approach is commonly used in situations where research is focused on exploiting a dependence structure per se. Copulas themselves can be modeled nonparametrically, i.e. the empirical copula of Deheuvels (1979), however the parametric approach for a copula function is more common. The combination of the nonparametric estimation of marginals and the parametric copula leads to a semiparametric procedure for estimating dependence parameters. This procedure provides a good balance between fully nonparametric and fully parametric models. The former may suffer the curse of dimensionality in multivariate settings, while the latter has to impose distributional assumptions on marginals.

Until recently copulas have been applied to model contemporaneous dependence structure and, therefore, were of limited use in the case of multivariate processes with serial dependence. However, the stylized facts on return series (see e.g. Tsay, 2002) suggest that these processes are of special interest in financial econometrics. Fan & Chen (2002) started to fill this gap by applying copulas to model serial dependence in the univariate time series setting. Later, Patton (2005a) introduced the notion of conditional copula, which allowed for using conditioning variables. This opened the way to an even broader range of copula applications in finance, including prediction.

All known applications of conditional copulas (Patton, 2005a, Patton, 2005b and van den Goorbergh, 2004) apply a parametric specification for their marginal distributions, conditional on past observations. To apply the flexible semiparametric approach in the context of the conditional copulas, we suggest a nonparametric procedure for estimating conditional marginal distributions. The major problem with the application of this procedure to conditional copulas is the requirement for the use of the same set of conditioning variables for all marginals. Conditioning on a large set of variables inevitably will lead to the curse of dimensionality in the context of the nonparametric estimation of conditional distribution function. We suggest several steps to overcome this problem. First, we test nonparametrically whether a specific variable may be excluded from the conditioning set. Then, we apply a dimension reduction technique for the remaining conditioning variables. After reducing the dimension of the conditioning variables to one, we make use of the nonparametric procedure of Hansen (2004) for estimating a smooth conditional cumulative distribution function (CDF). The parametric copula is modeled using a Gaussian specification with time-varying parameters. The semiparametric procedure may be summarized as follows:

1. Granger causality testing
2. If necessary, further dimension reduction to one conditioning variable
3. Nonparametric estimation of conditional marginal CDFs
4. Parameter estimation of a conditional copula
5. Predictive ability evaluation

Once a model is specified and estimated, the next important task is model evaluation and comparison with a benchmark model. The benchmark is specified as a parametric copula model with parametric marginals. We will adopt the recent methodology of Giacomini & White (2005) for testing conditional predictive abilities. This methodology is specifically suitable in our case since it allows for comparison of parametric and semiparametric techniques and acknowledges estimation uncertainty and possible nonstationarities in data.

The rest of the paper is organized as follows. In Section 2 we provide an outline of copula theory. In Section 3 we discuss the specification of conditional marginals, concentrating on the nonparametric estimation and dimension reduction technique. Section 4 describes the estimation procedure for the copula parameters. Further, in Section 5 we discuss the conditional predictive ability test and its application for selecting the best performing model. Section 6 investigates the performance of the suggested semiparametric and parametric models under various data generating processes. In Section 7 we apply the described methodology to the multivariate portfolios of major US stocks, UK stocks and international stock indexes. Section 8 concludes.

2 Copula theory

We start with a general overview of copula theory and provide basic definitions, major theorems and essential properties. For more information we refer the interested reader to the monograph by Nelsen (1999).

2.1 Unconditional copula

The term *copula* was first introduced by Sklar (1959) to describe a function that "couples" univariate marginal distributions to a multivariate joint distribution.

Definition 2.1 (Copula) *The copula function $C(\boldsymbol{u})$ is a joint cumulative distribution function of a probability measure with support in $[0,1]^N$ and uniform marginal distributions.*

The following result is cardinal to the theory of copulas.

Theorem 2.1 (Sklar's theorem) *Given an N-dimensional distribution function F with marginal distribution functions $F_1, F_2, \ldots, F_N$, there exists an N-copula C such that for all $\boldsymbol{x} \in \bar{\mathbb{R}}^N$*

$$F(x_1, x_2, \ldots, x_N) \quad = \quad C(F_1(x_1), F_2(x_2), \ldots, F_N(x_N)).$$

Moreover, if the marginal distributions are continuous the copula C is unique.

For the proof of Theorem 2.1 we refer a reader to Sklar (1996). The inverse of Theorem 2.1 provides one of the methods to construct a copula, i.e. inversion method. This method is used to derive copulas from known parametric multivariate distributions, e.g. the Gaussian copula.

Corollary 2.1 (Inversion method) *Let $F_1^{-1}, F_2^{-1}, \ldots, F_N^{-1}$ denote (quasi) inverse of marginal distribution functions $F_1, F_2, \ldots, F_N$, respectively. Then, for any $\boldsymbol{u} \in [0,1]^N$*

$$C(u_1, u_2, \ldots, u_N) \quad = \quad F(F_1^{-1}(u_1), F_2^{-1}(u_2), \ldots, F_N^{-1}(u_N)).$$

One of the essential properties of copulas is their invariance under strictly increasing transformations of X_n.

Theorem 2.2 (Invariance) *Let $X_1, X_2, \ldots, X_N$ denote continuous random variables with copula C and let $T_1, T_2, \ldots, T_N$ be strictly increasing transformation functions. Then transformed random variables $T_1(X_1), T_2(X_2), \ldots, T_N(X_N)$ have exactly the same copula C.*

The proof of Theorem 2.2 follows from Corollary 2.1 and the properties of the distribution function. As a consequence of Theorem 2.2, the dependence between random variables $X_1, X_2, \ldots, X_N$ is completely captured by the copula regardless of the scale in which each random variable X_n is measured. Therefore any dependence measure may be based directly on the copula.

Definition 2.2 (Copula density) *Suppose $C(u_1, u_2, \ldots, u_N)$ is a continues copula function, then its density $c(u_1, u_2, \ldots, u_N)$ can be defined as follows*

$$c(u_1, u_2, \ldots, u_N) = \frac{\partial^N C(u_1, u_2, \ldots, u_N)}{\partial u_1 \cdots \partial u_N}$$

By taking partial derivatives of the distribution function $F(\boldsymbol{x})$ and the corresponding copula representation (Theorem 2.1) with respect to $x_1 \ldots x_N$, we derive the following decomposition of the multivariate density function:

$$f(\boldsymbol{x}) = c(F_1(x_1), F_2(x_2), \ldots, F_N(x_N)) \prod_{n=1}^{N} f_n(x_n), \tag{1}$$

where $f_n(x_n)$ is a density of the margin $F_n(x_n)$. This decomposition is of a great importance for copula estimation using the maximum likelihood approach.

2.2 Conditional copula

Most of the variables of interest in economics and finance exhibit dependence on past observations. Therefore, it is important to incorporate conditioning on the lagged variables in a quantitative analysis of economic and financial data. Until recently, copulas have been used to model a contemporaneous dependence structure. To extend the application of copulas to the case of conditioning variables, Patton (2005a) introduced the notion of conditional copula.

Definition 2.3 (Conditional Copula) *The conditional copula of random variables* $\boldsymbol{X}|(\boldsymbol{Z}=\boldsymbol{z}) \sim F_{\boldsymbol{X}|\boldsymbol{Z}}(\boldsymbol{x}|\boldsymbol{z})$ *is the conditional multivariate distribution function of uniform random variables* $\boldsymbol{U}$*, such that* $U_i \equiv F_{X_i|\boldsymbol{Z}}(x_i|\boldsymbol{z})$ *given* $\boldsymbol{Z}=\boldsymbol{z}$*.*

Hereafter by referring to U or $\boldsymbol{U}$ we will mean conditionally uniform random variable(s) without explicit remark regarding the set of conditioning variables, when the latter is clear from the context. For each $\boldsymbol{z}$ in the support of $\boldsymbol{Z}$ the conditional copula satisfies the properties of the (unconditional) copula (Definition 2.1). Patton (2005a) showed that Sklar's theorem (Theorem 2.1), its inverse (Corollary 2.1) and the invariance (Theorem 2.2) hold in the case of conditioning. In this case, the unconditional multivariate distributions and unconditional marginals are replaced by conditional multivariate distributions and conditional marginals respectively. It is required that the set of conditioning variables $\boldsymbol{Z}$ is the same for every marginal distribution and the copula and the corresponding joint distribution. Otherwise, Sklar's theorem may fail for conditional copulas. It may be the case that the set of conditioning variables may be reduced for some marginal distribution, i.e. $F_{X_i|\boldsymbol{Z}_1}(x_i|\boldsymbol{z}_1) = F_{X_i|\boldsymbol{Z}_1,\boldsymbol{Z}_2}(x_i|\boldsymbol{z}_1,\boldsymbol{z}_2)$ for all $x_i, \boldsymbol{z}_1, \boldsymbol{z}_2$. In this situation X_i is conditionally independent of $\boldsymbol{Z}_2$ given $\boldsymbol{Z}_1$. To test for such conditional independence, one may consider nonparametric tests for conditional independence (e.g. Su & White, 2005).

For simplicity, we assume that multivariate stationary time series $\{\boldsymbol{X}_t\}$ has a Markovian property. Thus, the conditioning variables $\boldsymbol{Z}_t$ include only the recent past $\boldsymbol{X}_{t-1}$. Empirical applications suggest that this assumption is reasonable for processes characterizing asset returns. Next, we apply the pairwise nonparametric Granger causality test of Diks & Panchenko (2005) to reduce the conditioning set. This test is robust to conditional heteroskedasticity which is likely to be present in financial data. In the case where the conditioning set is not reduced to one variable, we adopt a further dimension reduction technique (refer to Subsection 3.1).

3 Marginal distributions

Before specifying a copula model, it is necessary to specify the corresponding marginal distributions. It can be done either parametrically or nonparametrically. The former approach imposes restrictive distributional assumptions, while the latter attempts to deduce the whole distribution entirely from the data. The price to pay for the relaxation of the assumptions is the amount of data necessary for estimation, at a reasonable precision level, and the speed of the estimation procedure. The parametric approach has been widely used in modeling asset returns, e.g. ARCH (Engle, 1982), GARCH (Bollerslev, 1986) models in a univariate case, and dynamic conditional correlations (DCC) of Engle (2002) in a multivariate case. Despite the variety of parametric specifications there is no general agreement on which particular model is to be chosen in practice. To avoid this ambiguity the nonparametric specification of marginals prevailed in the (unconditional) copula approach (see e.g. Genest et al., 1995, Embrechts et al., 2003). However, this advantage of copula modeling has not been exploited yet in the recent conditional copula approach of Patton (2005a). With this in mind, we wish to fill the gap and use the nonparametric specification for conditional marginals in the conditional copula framework.

3.1 Dimension reduction

The major difficulty in implementing a nonparametric specification of marginals for the conditional copula model is the requirement to condition on the same set of conditioning variables for every marginal distribution. If the dimension of the conditioning set is greater than one, the nonparametric approach is practically inapplicable mainly due to the bandwidth choice problem and the curse of dimensionality. Initially we reduced dimensionality using a pairwise nonparametric Granger causality test. However, if Granger causality is not rejected for some variables, we have to apply the dimension reduction technique described below.

Hall & Yao (2005) suggest to approximate conditional distribution function $F_{X|\boldsymbol{Y}}(x|\boldsymbol{y})$ by $F_{X|\beta'\boldsymbol{Y}}(x|\beta'\boldsymbol{y})$ using $\beta'\boldsymbol{y}$ as a one dimensional projection of the K-dimensional conditioning variable $\boldsymbol{Y}$. The parameter β is estimated by minimizing the mean integrated square error over b

$$\int\int\left(\widehat{F}_{X,\beta'\boldsymbol{X}}(x,b'\boldsymbol{y})-\widehat{F}_{X,\boldsymbol{Y}}(x,\boldsymbol{y})\right)^2 w(x,\boldsymbol{y})dyd\boldsymbol{x},$$

where $w(y,\boldsymbol{x})$ is a weighting function. The estimation procedure is based on a local linear regression, which requires bandwidth selection and is computationally intensive.

We suggest an alternative procedure that is based on an ordinary linear regression and therefore does not require bandwidth selection and is computationally fast. The suggested estimator of β enjoys the same properties as the estimator of Hall & Yao (2005), i.e. it is consistent and normally distributed with $\sqrt{T}$-convergence rate. The technique we suggest is closely related to the sufficient dimension reduction of Cook (2004).

Suppose $\{X_t,\boldsymbol{Y}_t\}$ is a weakly dependent time series. In our subsequent developments, we will operate in terms of the standardized predictor $\boldsymbol{S}=\boldsymbol{\Sigma}_{\boldsymbol{Y}}^{-1/2}(\boldsymbol{Y}-E(\boldsymbol{Y}))$, where $\boldsymbol{\Sigma}_{\boldsymbol{Y}}$ is a positive definite variance-covariance matrix. This monotonic transformation does not affect the conditional distribution of interest and is required for technical reasons. To ensure that sample estimate $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{Y}}$ is a positive definite matrix, we employ the eigenvalue method of Rousseeuw & Molenberghs (1993). Similarly to Hall & Yao (2005) our central assumption is that the true conditional distribution function can be approximated by $F_{X|Z}(x|z)$, where $z=\beta'\boldsymbol{s}$ and $Z=\beta'\boldsymbol{S}$ is a one dimensional projection of the standardized K-dimensional conditioning variable $\boldsymbol{S}$. Our goal is to find an estimator of the parameter β. Since $F_{X|Z}$ is unknown, β is identified up to a multiplicative scalar. Suppose $L(x,\zeta)$ is an arbitrary function convex in ζ. Consider the minimization problem:

$$(\alpha^*, \beta^*) \quad = \arg\min_{a,b} \quad E_{X,\boldsymbol{S}}[L(X, a + b'\boldsymbol{S})]. \tag{2}$$

The following theorem due to Li & Duan (1989) establishes an estimation procedure for β.

Theorem 3.1 *Parameter β^* is proportional to parameter β under the following conditions:*

(1) The criterion function $L(x, \zeta)$ is convex in ζ.

(2) There is a proper solution for the minimization problem (2).

(3) The conditional expectation $E(b'\boldsymbol{S}|\beta'\boldsymbol{S})$ exists and is linear in $\beta'\boldsymbol{S}$.

A class of estimators defined by Eq. (2) includes ordinary least squares (OLS), for which assumption (1) is satisfied. Existence condition (2) will hold for OLS if the standardized predictor $\boldsymbol{S}$ is defined (non-multicollinearity). Assumption (3) is extensively discussed in Hall & Li (1993). They claim that low dimensional projections (one dimensional in this case) of high dimensional data are almost linear provided the standardized predictor $\boldsymbol{S}$ is defined. Strong consistency and asymptotic normality with $\sqrt{T}$-convergence rate of a sample estimator of β^* is established by theorem 5.1 of Li & Duan (1989) for the case of IID data. To establish the same results for weakly dependent time series, one has to apply the central limit theorem (Doukhan & Louhichi, 1999) and the law of large numbers (Birkel, 1992) for weakly dependent processes. Relying on the above results we may summarize all conditioning information in one variable $Z = \beta'\boldsymbol{S}$.

3.2 Smooth estimation of conditional distribution function

Hall et al. (1999) present an extensive overview of recent techniques for estimating a conditional distribution function. All the methods described there are based on unsmoothed estimators. Hansen (2004) shows that it is possible (at least asymptotically) to increase the efficiency of the nonparametric procedure by using smoothed estimators. The analytical derivation of this author was supported by simulations for various data generating processes. An analogous conclusion has been drawn by Chen et al. (2004) in the context of unconditional nonparametric marginals. Considering the possible efficiency gains we will follow the smooth method of Hansen (2004). We outline this method and refer the reader to the work of Hansen (2004) and the references therein for a more detailed exposition.

Suppose $\{X_t, Z_t\}$ is a strictly stationary time series with distribution $F_{X,Z}(x, z)$. To make a connection with the previous notation, X_t is one of the elements of multivariate series $\boldsymbol{X}_t$, the conditional marginal distribution $U \equiv F_{X|Z}(x|z)$ of which we are after. We drop the subscripts of F for brevity. Element Z_t summarizes the information about $\boldsymbol{X}_{t-1}$, its precise form was discussed in Subsection 3.1.

In the context of (unsmooth) conditional distribution estimation, a local linear estimator may achieve lower bias in comparison with Nadaraya-Watson estimator. The same is true in the context of smooth estimation. Therefore, the smooth estimator will be of the form of a modified local linear estimator ensuring that it remains a valid conditional distribution function:

$$\widehat{U} \equiv \widehat{F}_{h,b}(x|z) = \frac{\sum_{t=1}^{T} w_t^* K_h(x - X_t)}{\sum_{t=1}^{T} w_t^*}, \tag{3}$$

where $K_h(s) = K(\frac{s}{h})$ with $K(s) = \int_{-\infty}^{s} \kappa(u)$ is the integrated smooth kernel with bandwidth parameter h, we use Gaussian kernel, and w_t^* denotes modified weights from local linear estimation

$$w_t^* = \begin{cases} 0 & \widehat{\beta}(z - Z_t) > 1 \\ w_t(1 - \widehat{\beta}(z - Z_t)) & \widehat{\beta}(z - Z_t) \leq 1 \end{cases},$$

where $w_t = \frac{1}{b} w(\frac{z - Z_t}{b})$ are kernel weights with kernel w (Gaussian in our application) and bandwidth b and $\widehat{\beta}$ is the slope from the weighted least squares

$$\widehat{\beta} = \left(\sum_{t=1}^{T} w_t (z - Z_t)^2 \right)^{-1} \left(\sum_{t=1}^{T} w_t (z - Z_t) \right).$$

Nonnegative weights w_t^* ensure that $\widehat{F}$ is a proper distribution function.

As in any kernel smoothing problem the bandwidth choice is the crucial part of the analysis. In the present situation the bandwidths (b^*, h^*) are selected by a plug-in method. The plug-in bandwidths minimize the estimated asymptotic mean integrated square error (MISE). The following theorem of Hansen (2004) defined the MISE for the smooth local linear estimator.

Theorem 3.2 *Assume $F(x|z)$ and marginal density of X, $f_X(x)$, are continuously differentiable up to fourth order in both x and z. If $b = cT^{-1/5}$ and $h = O(b)$ as $T \to \infty$, then*

$$\int_{-\infty}^{\infty} E\left(\widehat{F}_{h,b}(x|z) - F(x|z)\right)^2 dx = \frac{R}{f(x)nb}(V - h\psi) + \frac{b^4 V_1}{4} - \frac{h^2 b^2 V_2}{2} + \frac{h^4 V_3}{4} + O(T^{-6/5}),$$

where $R = \int_{-\infty}^{\infty} (w(s))^2 ds$ denotes the roughness of weighting kernel w, $\psi = 2\int_{-\infty}^{\infty} sK(s)\kappa(s)ds > 0$ and

$$\begin{aligned} V &= \int_{-\infty}^{\infty} F(x|z)(1 - F(x|z))dx \\ V_1 &= \int_{-\infty}^{\infty} (F^{(2)}(x|z))^2 dx \\ V_2 &= \int_{-\infty}^{\infty} f(x|z) f^{(2)}(x|z) dx \\ V_3 &= \int_{-\infty}^{\infty} (f'(x|z)) dx. \end{aligned}$$

The notation $g^{(2)}$ refers to second derivative of g. The parameters of the MISE $f(x), V, V_1, V_2, V_3$ are estimated using kernel smoothing and local polynomial regression techniques (see Hansen (2004) for details). Plug-in bandwidths $(\widehat{b}^*, \widehat{h}^*)$ are found by numerical minimization of the estimated MISE.

It can be seen that estimators of the MISE parameters and, thus, bandwidths $(\widehat{b}^*, \widehat{h}^*)$ depend on the conditioning variable Z. Since the bandwidth selection procedure is computationally involved, selection of bandwidths $(\widehat{b}^*, \widehat{h}^*)$ for every Z_t appears to be an infeasible task for reasonable sample sizes $(T \geq 500)$. To reduce the number of computations, the conditioning variables Z_t were grouped according to their ranks into I groups G_i:

$$G_i = \left\{ Z_t : \ \Phi\left(2.5\left[-1 + \frac{2i-2}{I}\right]\right) < \frac{\text{rank}(Z_t)}{T} \leq \Phi\left(2.5\left[-1 + \frac{2i}{I}\right]\right) \right\},$$

where Φ denotes standard Gaussian distribution function. The plug-in bandwidth for each group $(\widehat{b}^*, \widehat{h}^*)_i$ is determined on the basis of observation Z_t with $\text{rank}(Z_t)/T = \Phi(2.5[-1 + (2i-1)/I]$. Intuitively, the proposed grouping puts observations Z_t with similar values into one group. Our simulations showed that for sample size $T = 1000$, $I = 5$ provides good approximation, i.e. its results are hardly distinguishable from the results of the full procedure, when the bandwidths are computed individually for every observation.

Using the procedure described above, we transform each element of the multivariate time series $\{\boldsymbol{X}_t\}$ into corresponding conditional marginal and group them together in the uniform series of conditional marginal transformations $\widehat{\boldsymbol{U}}_t$.

4 Specification and estimation of copula

Similarly to the case of marginals, it is possible to specify a copula either nonparametrically or parametrically. A nonparametric approach will lead to the empirical copula of Deheuvels (1979). For simplicity, here we will pursue the parametric copula approach. As an extension we plan to include nonparametric copula analysis in the future.

There are many copula specifications suggested for the bivariate case (see Nelsen, 1999 for extensive review). However, for most of the copulas, extensions to the higher dimensions are not feasible. Therefore, most of the copulas considered for multivariate modeling are based on the inverses of the

Gaussian distribution or more general Student's t distribution. The former will yield the Gaussian copula with a correlation matrix $\mathbf{\Sigma}$ as a matrix of parameters, while the latter will result in Student's t copula with $\mathbf{\Sigma}$ and an additional parameter, the number of degrees of freedom ν. Several studies (e.g. Chen et al., 2004, Panchenko, 2005) reported a poor fit of the Gaussian copula with constant correlations to the multivariate asset returns data, while Student's t copula provided slightly better fit to the data. In the recent past Engle & Sheppard (2001) and Engle (2002) suggested dynamical conditional correlations (DCC) models for parsimonious specification of time-varying correlation matrix $\mathbf{\Sigma}$ for multivariate GARCH models. Chen et al. (2004) showed that the Gaussian DCC model provided a good fit to the data in the context of the unconditional copula. Relying on these results, we suggest using the DCC model to specify the time evolution of correlation matrix $\mathbf{\Sigma}$. It is important to notice that there are other, possibly better, ways to model the time evolution of $\mathbf{\Sigma}$. Recent work of Pelletier (2005) and Hafner et al. (2005) suggest new promising nonlinear and semiparametric techniques for modeling the correlation dynamics. We adopt the DCC of Engle (2002) for simplicity and plan an extension of this paper to include other correlation dynamics.

The Gaussian DCC (1,1) conditional copula model is defined as follows. Using the inversion method (Corollary 2.1), we derive the Gaussian copula:

$$C_{\Sigma} = \Phi_{\Sigma}(\xi_1, ..., \xi_N), \tag{4}$$

where Φ_{Σ} denotes the multivariate standard normal distribution with linear correlation matrix $\mathbf{\Sigma}$, and $\xi_i = \Phi^{-1}(U_i)$, where Φ^{-1} is the inverse of the univariate standard normal distribution. The time-varying conditional correlation matrix evolves according to the following specification:

$$\mathbf{\Sigma}_t = (1 - \alpha - \beta)\overline{\mathbf{\Sigma}} + \alpha \boldsymbol{\xi}_{t-1}\boldsymbol{\xi}_{t-1}^{'} + \beta \mathbf{\Sigma}_{t-1}, \tag{5}$$

where $\overline{\boldsymbol{\Sigma}}$ is the unconditional correlation matrix. To ensure the positive definiteness of $\boldsymbol{\Sigma}_t$, we require $0 \leq \alpha \leq 1$, $0 \leq \beta \leq 1$ and $\alpha + \beta \leq 1$. The correlation matrix $\overline{\boldsymbol{\Sigma}}$ is computed by compounding pairwise correlation coefficients ρ. It may happen due to numerical problems that $\overline{\boldsymbol{\Sigma}}$ fails to be a positive definite. To avoid this situation, we use the eigenvalue method of Rousseeuw & Molenberghs (1993). We estimate the parameters of the Gaussian DCC copula model using semiparametric maximum likelihood method introduced by Genest et al. (1995). The log likelihood function for the copula and marginals can be derived from the decomposition of the multivariate density (1):

$$L(\boldsymbol{x}) = \log c(F_1(x_1), \dots, F_N(x_N)) + \sum_{n=1}^{N} \log f_n(x_n) = L_c(\boldsymbol{u}) + \sum_{n=1}^{N} L_n(x_n). \quad (6)$$

The estimation procedure consists of two stages. First we perform a conditional marginal CDF transformation and construct the uniform series of transformations $\widehat{\boldsymbol{U}}_t$. In the next stage, we estimate the parameters of the copula by maximizing the log likelihood function of the copula using the marginal transforms obtained from the first stage:

$$(\widehat{\alpha}, \widehat{\beta}) = \arg\max_{a,b} \sum_{t=1}^{T} L_c(\widehat{U}_{1t}, \dots, \widehat{U}_{Nt}, (a, b)).$$

The following theorem, due to Genest et al. (1995), establishes desirable properties of the semiparametric estimator

Theorem 4.1 *Under suitable regularity condition, the semiparametric estimator* $(\widehat{\alpha}, \widehat{\beta})$ *is consistent and* $\sqrt{T}$ *asymptotically normal.*

Chen et al. (2004) indicate that by using the smooth procedure for estimating the conditional marginals, we improve the efficiency of the semiparametric estimator.

A Gaussian copula density is obtained by taking partial derivatives in Eq. 4:

$$c = |\mathbf{\Sigma}|^{-1/2} \exp\left(-\frac{1}{2}\boldsymbol{\xi}'(\mathbf{\Sigma} - \mathbb{I})\boldsymbol{\xi}\right),$$

where $\mathbb{I}$ is the identity matrix. Using the time-varying conditional correlation $\mathbf{\Sigma}_t$, we may specify the log likelihood function:

$$L = -\frac{1}{2}\left(\sum_{t=1}^{T} \log|\mathbf{\Sigma}_t| + \boldsymbol{\xi}_t'(\mathbf{\Sigma}_t - \mathbb{I})\boldsymbol{\xi}_t\right).$$

Since there is no closed form expression for the derivatives of the log-likelihood L, the optimization is conducted numerically. The optimization is performed using the simplex method of Nelder & Mead (1965). This method is suitable for our purposes since it does not require evaluation of the derivatives.

5 Comparing predictive abilities of alternative models

The major question of this paper is whether using the semiparametric approach can improve the accuracy of the forecast of the distribution of future returns. Therefore, we need to specify a similar copula model, but with parametric marginals.

5.1 The benchmark model

We use a slight modification of the commonly used DCC model as a benchmark. This model can be viewed as a copula model with conditional marginals specified parametrically according to the GARCH(1,1) model:

$$X_t = \phi \boldsymbol{Z}_t + \epsilon_t, \qquad \epsilon_t|\epsilon_{t-1} \sim N(0, h_t), \tag{7}$$

$$h_t = \gamma_o + \gamma_1 \epsilon_{t-1}^2 + \gamma_2 h_{t-1}. \tag{8}$$

Variable $\boldsymbol{Z}_t$ compounds all conditioning variables of the copula. Here we have to assume that $\boldsymbol{Z}_t$ enters only the mean equation (7) and does not affect the variance equation (8). Then, the copula takes the form of the Gaussian DCC copula described in the previous section, see Eqs. (4, 5). The model is estimated using a two stage maximum likelihood: first the GARCH(1,1) models for marginals and then, the Gaussian DCC copula model. For the properties of the estimators derived from this two-stage procedure we refer the reader to Engle & Sheppard (2001). The optimization is performed numerically using the Nelder & Mead (1965) algorithm.

5.2 Conditional test for predictive abilities

Once a benchmark and an alternative model are specified and estimated, we need to determine whether the alternative model is superior to the benchmark. The accuracy of the out-of-sample prediction of the distribution of future returns is used as a criterion. For brevity, we will focus on a one step ahead forecast at time $t+1$, though an extension is possible. The major problem in testing density forecasts arises from the following fact. While producing a forecast of the whole future distribution, we observe only one (possibly multivariate) outcome from the true distribution. Various techniques were suggested in the forecast evaluation literature to overcome this problem (see Diebold et al., 1998 for review). Our attention will focus on so-called predictive log likelihood or scoring rule (Diebold & Lopez, 1996) defined as

$$L_{t+1}(\hat{f}_t, \boldsymbol{X}_{t+1}) \quad = \quad \log \widehat{f_t}(\boldsymbol{X}_{t+1}), \tag{9}$$

where $\hat{f}_t(\cdot)$ is the density forecast of out-of-sample point of evaluation $\boldsymbol{X}_{t+1}$. In the case of the copula, we observe the decomposition of L into copula part L_c and the sum of marginal log likelihoods L_n according to Eq. (1). While the former is a part of the estimation process, the latter is readily available by differentiating smooth conditional distribution function (3) at the point $\boldsymbol{X}_{t+1}$. Intuitively, the model producing the best density forecast will give the highest average score. Predictive log likelihood is especially convenient in the case of multivariate forecast. In principle, any other suitable criterion may be used, e.g. probability integral transform based criteria (Bao et al., 2004, Chen et al., 2004).

Further, we adopt the recent conditional predictive ability test of Giacomini & White (2005). Their test is based on the out-of-sample evaluation using the rolling window scheme. The in-sample size T used for estimation remains constant, while the sample itself and the point(s), at which the forecast is evaluated, move with time. It is assumed that the number of out-of-sample forecasts M tends to infinity while the in-sample size T remains finite. The advantage of the conditional test vs. traditional unconditional methodology of West (1996) is that it can be applied in a more general setting. The test evaluates not only the model itself, but the whole forecasting method, which includes the choice of the in-sample size T. The conditional methodology can be applied for the comparison of a wide range of models, such as parametric, semiparametric, nonparametric and Bayesian models. Non-nested and nested models can be treated in the similar way, while the test of West (1996) is limited to the case of non-nested models. The conditional methodology of Giacomini & White (2005) is specifically applicable for our situation since we want to compare the semiparametric and parametric methodology for a fixed sample size T. In theory we would expect that semiparametric procedure will always beat the parametric one, provided (1) stationarity and (2) infinite sample size T. However, in practice achieving both (1) and (2) is not realistic, and we have to find the sample interval of size T, on which the sample is approximately stationary and T is large enough to ensure precision of the estimates. Although being asymptotically superior, a semiparametric model can actually be beaten by a parametric model for fixed T, since the latter may require less data to achieve the same precision as the former.

Since we are interested in the VaR estimation, we focus on the forecast concerning the left tail of the return density. To incorporate this in the forecast evaluation we use the weighted likelihood ratio test of Amisano & Giacomini (2005). Their test is designed for a univariate model. A slight extension, in the way the weighting function w is defined, allows to it to be used in multivariate settings. Define the weighted likelihood ratio as

$$WLR_{t+1} = w(\boldsymbol{S}_{t+1})(L_{t+1}(\hat{f}_t^{alt}, \boldsymbol{X}_{t+1}) - L_{t+1}(\hat{f}_t^{bm}, \boldsymbol{X}_{t+1})),$$

where $\hat{f}_t^{alt}$ and $\hat{f}_t^{bm}$ are predictive log likelihoods (9) of the alternative model and the benchmark respectively, $\boldsymbol{S}_{t+1}$ is a standardized realization (using unconditional moments), i.e. $\boldsymbol{S} = \boldsymbol{\Sigma}_{\boldsymbol{X}}^{-1/2}(\boldsymbol{X} - E(\boldsymbol{X}))$ with covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{Y}}$, and weighting function w is defined as

$$w(\boldsymbol{S}_{t+1}) = 1 - \Phi\left(\frac{\sum_{n=1}^{N} S_{n,t+1}}{N}\right).$$

The suggested weighting scheme puts higher weights on the likelihood ratio in the case of joint extreme events in the left tail of the distribution. The null hypothesis is stated in terms of the expectation of the weighted likelihood ratio over the number of evaluations M:

$$\begin{aligned} H_0 &: \quad E[WLR_{t+1}] \leq 0, \\ H_A &: \quad E[WLR_{t+1}] > 0. \end{aligned}$$

The null hypothesis states that on average the forecast produced by the alternative method is not better than the forecast produced by the benchmark. The test statistic TS takes the form:

$$TS = \frac{\frac{1}{M}\sum_{t=T}^{T+M} WLR_{t+1}}{\widehat{\sigma}_M/\sqrt{M}},$$

where $\widehat{\sigma}_M^2$ is a heteroskedasticity and autocorrelation consistent Newey & West (1987) estimator of the asymptotic variance:

$$\widehat{\sigma}_M^2 = \frac{1}{M}\sum_{t=T}^{T+M} WLR_{t+1}^2 + 2\left[\sum_{j=1}^{p_M}\left(1 - \frac{j}{p_M+1}\right)\frac{1}{M}\sum_{t=T+j}^{T+M+j} WLR_{t+1}^2 WLR_{t+1-j}^2\right], \tag{10}$$

where p_M is the number of nonzero autocorrelations taken into account, which is a parameter of choice. It is required that $p_M = o(T^{1/4})$ for the consistency of the estimator (10). Using the rule of thumb, we choice $p_M = \text{integer}(T^{1/5})$.

Theorem 5.1 *Under certain regularity conditions and for a fixed estimation window T, test statistic TS converges in distribution to $N(0,1)$ as number of evaluations $M \to \infty$.*

For the proof of Theorem 5.1 we refer to Amisano & Giacomini (2005). This authors also provide an extensive simulation study, which demonstrates good finite sample properties of the test.

Here, we discussed the situation, when only two models are compared. In practice, we may specify several alternative models and, then, test if the best model among them beats the benchmark. To implement this test, we could adapt the "reality check" test of White (2000) to the current conditional methodology.

6 Simulations

We next investigate the performance of the proposed semiparametric model and the parametric DCC model using various data generating processes (DGPs). Since we apply the same specification for the conditional correlation structure in both models, we will focus on the multivariate DGPs with the same correlation structure, but various specifications for the marginals. The correlation matrix is specified as in Eq. 5 with $\alpha = 0.1$ and $\beta = 0.6$. First, we sample uniform random variates u_{ti} from the Gaussian copula (4) and transform them into i.i.d. (in terms of marginals, but not jointly!) $N(0,1)$ marginals $\{\epsilon_{ti}\}$. For brevity, the same specification is applied for each marginal and, therefore, subscript index i is dropped. Then, we use the following DGPs for specifying conditional variance h_t of the marginals:

DGP 0. $\quad h_t = 2 \times 10^{-6} + 0.091\epsilon_{t-1}^2 + 0.899h_{t-1}.$

DGP 1. $\quad h_t = 2 \times 10^{-6} + 0.029\epsilon_{t-1}^2 + 0.076\epsilon_{t-1}^2\mathbb{I}(\epsilon_{t-1} < 0) + 0.931h_{t-1}.$

DGP 2. $\quad \ln(h_t) = -1.001 - 0.0794\epsilon_{t-1} + 0.2647|\epsilon_{t-1}| + 0.899\ln(h_{t-1}).$

DGP	DGP 0	DGP 1	DGP 2
TS	-2.32	1.83	2.11
p-value	0.99	0.03	0.02

Table 1: *Test statistic TS and the corresponding p-values of the predictive abilities test for simulated DGPs.*

DGP 0 is the GARCH(1,1) model of Bollerslev (1986), DGP 1 is the threshold (T)ARCH(1,1) model of Zakoian (1994) and DGP 2 is the exponential (E)GARCH(1,0) model of Nelson (1991). DGP 0 is symmetric in lagged returns, while DGP 1 and 2 allow for asymmetries. The the parameter values used are typical for the S&P 500 index (see e.g. Tsay, 2002, Engle, 2003). We set number of components $N = 3$, in-sample size $T = 1000$ and number of out-of-sample evaluations $M = 5000$. The values of the test statistic TS and the corresponding p-values are reported in Table 1.

As anticipated, the null hypothesis of no better performance than a benchmark is not rejected for DGP 0. In this case, the functional form and the distribution of marginals are equivalent to the parametric DCC model. Under these conditions the parametric DCC will dominate the semiparametric model because of the higher efficiency of the parameter estimators. However, this is not the case for DGP 1 and 2, where the null hypothesis is rejected. The major reason for the rejection is the bias resulting from an incorrect marginal specification in the parametric DCC model. A more detailed analysis of the results in the case of DGP 1 and 2 showed that the parametric DCC model misspecified not only the marginals, but also the parameters of the conditional correlation matrix equation (5). In practice the true DGP is usually unknown. Therefore, we expect that a semiparametric model will provide better forecasts in settings, where the marginals strongly deviate from the benchmark GARCH(1,1). For brevity, we considered DGPs with conditioning only on the lagged variable of the corresponding marginal, which is to be extended in the future.

Portfolio	5 US stocks	5 UK stocks	4 Int Indexes
TS	-0.45	-0.93	2.29
p-value	0.67	0.82	0.01

Table 2: *Test statistic TS and the corresponding p-values of the predictive abilities test on asset return data.*

7 Application to asset portfolios

Next, we evaluate the predictive abilities of the semiparametric model on daily log-return series of three asset portfolios. Log-returns are defined as $X_t = \ln(P_t/P_{t-1})$, where P_t is the dividend-adjusted closing price (except for international indexes) on day t. The whole sample covers the period 08/1990–07/2005. For the sake of stationarity a few dates with extremely abnormal returns, e.g. September 11, 2003, were eliminated from the sample. The first portfolio consists of five major US traded stocks from different sectors (INTERNATIONAL BUS.MACH., CITIGROUP, WAL MART STORES, CHEVRON, VERIZON COMMS.); the second is comprised of five major UK traded stocks (BARCLAYS, BP, BRITISH AIRWAYS, REUTERS GROUP, VODAFONE GROUP); and the third consists of four international indexes (S&P 500, FTSE 100, DAX and CAC 40) measured simultaneously at 16:00 hours London time. This avoids the non-synchronization problem (Martens & Poon, 2001) for the third portfolio. The in-sample size is set to $T = 1000$. A lower sample size is generally not feasible either for the parametric or for the semiparametric model, while a higher sample size might provide worse forecasts due to nonstationarities. Given the available data-set and in-sample size T, we set number of the out-of-sample evaluations to $M = 2900$. Bivariate non-parametric Granger causality testing (Diks & Panchenko, 2005) on an in-sample set and the consequent dimension reduction (Subsection 3.1) was embedded in the procedure. Table 2 reports the values of the test statistic TS and the corresponding p-values.

The test fails to reject the null of no better predictive performance of the semiparametric model compared to the parametric DCC model for the portfolios of US and UK stocks, while the test on the portfolio of four international indexes clearly rejects the null. A possible explanation is that GARCH(1,1) model can adequately explain the dynamics of the considered individual US and UK stocks, however is not fully applicable to the dynamics of international stock indexes. Though, more evidence is needed to draw stronger conclusions about the applicability of semiparametric models, we may expect them to predict better in settings, where a portfolio includes a range of various assets with qualitatively different dynamics, e.g. stocks, bonds, foreign currencies, stock indexes.

8 Conclusions

In line with current developments in quantitative risk management we suggested a flexible multivariate semiparametric model to forecast future asset returns. The procedure heavily relies on the concept of conditional copula, which was shortly reviewed. Motivated by the fact that until now conditional copulas had been used in a fully parametric setting, our aim was to construct a semiparametric copula model with nonparametric marginals. We used conditional kernel smoothers based on local linear estimators for conditional marginal CDFs. The major pitfalls in using a conditional copula is the requirement of the same conditioning set for each marginal. This would inevitably lead to the curse of dimensionality in the estimation of the conditional marginals. To avoid this situation we suggested a dimension reduction technique. For simplicity, we applied the Gaussian copula with a time-varying correlation matrix specified as in the DCC model. The copula specifications may be extended to include other multivariate copulas and reacher correlation dynamics. Moreover, a nonparametric copula can be used as an option.

To assess the predictive abilities of the semiparametric model we confronted it with widely used fully parametric DCC model. We used the conditional predictive abilities framework in conjunction with a predictive likelihood as a loss function.

Simulations on various DGPs suggested better predictive performance of the semiparametric model, when the marginals are misspecified relatively to the parametric model.

In an application to three asset portfolios, we found that the suggested semiparametric model outperformed the fully parametric DCC model in the case of the portfolio of four international stock indexes and did not show evidence of better performance in the cases of the portfolio of five US stocks and the portfolio of five UK stocks. Although, jointly with the simulations, this shed some light on the setting where the semiparametric model is expected to forecast better, further investigation is necessary to draw stronger conclusion regarding its practical applicability.

References

Amisano, G. & R. Giacomini (2005) Comparing Density Forecasts via Weighted Likelihood Ratio Tests. Mimeo, University of California, Los Angeles.

Bao, Y., T.-H. Lee, & B. Saltoglu (2004) A Test for Density Forecast Comparison with Applications to Risk Management. Mimeo, Department of Economics, University of California, Riverside.

Birkel, T. (1992) Laws of large numbers under dependence assumptions. *Statistics and Probability Letters* **14**, 355–362.

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

Bouyé, E., V. Durrleman, A. Nikeghbali, G. Riboulet, & T. Roncalli (2000) Copulas for Finance: A Reading Guide and Some Applications. Working paper, Groupe de Recherche Operationnelle, Credit Lyonnais.

Chen, X., Y. Fan, & A. Patton (2004) Simple Tests for Models of Dependence Between Multiple Financial Time Series, with Applications to U.S. Equity Returns and Exchange Rates. Discussion Paper 483, Financial Markets Group, London School of Economics.

Chen, X., Y. Fan, & V. Tsyrennikov (2004) Efficient Estimation of Semiparametric Multivariate Copula Models. Working paper 04-W20, Department of Economics, Vandebilt university.

Cook, D. R. (2004) Testing predictor contributions in sufficient dimension reduction. *Annals of Statistics* **32**, 1062–1099.

Deheuvels, P. (1979) La fonction de dépendance empirique et ses propriétés. *Acad. Roy. Belg. Bull. Cl. Sci.* **5**, 274–292.

Diebold, F. X., T. A. Gunther, & A. S. Tay (1998) Evaluating density forecasts with applications to financial risk management. *International Economic Review* **39**, 863–83.

Diebold, F. X. & J. A. Lopez (1996) Forecast Evaluation and Combination. In G. S. Maddala & C. R. Rao (eds.), *Handbook of Statistics*, Volume 14 of *Statistical Methods in Finance*, pp. 241–268. North Holland, Amsterdam.

Diks, C. & V. Panchenko (2005) A new statistic and practical guidelines for nonparametric Granger causality testing. *Journal of Economic Dynamics and Control*. Forthcoming.

Doukhan, P. & S. Louhichi (1999) A new weak dependence condition and applications to moment inequalities. *Stochastic Processes and their Applications* **84**, 313–342.

Embrechts, P., F. Lindskog, & A. McNeil (2003) Modelling Dependence with Copulas and Applications to Risk Management. In S. Rachev (ed.), *Handbook of Heavy Tailed Distributions in Finance*, Handbooks in Finance, 1, pp. 329–384. Elsevier.

Engle, R. (1982) Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica* **50**, 987–1008.

Engle, R. (2002) Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics* **20**, 339–350.

Engle, R. F. (2003) Risk and Volatility: Econometric Models and Financial Practice. Nobel lecture, New York University, Department of Finance.

Engle, R. F. & S. Manganelli (2004) CAViaR: conditional value at risk by regression quantiles. *Journal of Business & Economic Statistics* **22**, 367–381.

Engle, R. F. & K. Sheppard (2001) Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH. Working paper 2001-15, University of California, San Diego.

Fan, Y. & X. Chen (2002). Estimation of Copula-Based Semiparametric Time Series Models. . Department of Economics, Vanderbilt University, Working paper, 02-W26R.

Genest, C., K. Ghoudi, & L.-P. Rivest (1995) A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika* **82**, 543–552.

Giacomini, R. & H. White (2005) Tests of Conditional Predictive Ability. mimeo, University of California, San Diego.

Hafner, C., D. van Dijk, & P. Franses (2005) Semi-parametric modelling of correlation dynamics. Econometric Institute Research Report 2005-26, Erasmus University Rotterdam.

Hall, P. & K.-C. Li (1993) On almost linearity of low dimensional projections from high dimensional data. *The Annals of Statistics* **21**, 867–889.

Hall, P., R. C. L. Wolff, & Q. Yao (1999) Methods for estimating a conditional distribution function. *Journal of American Statistical Association* **94**, 154–163.

Hall, P. & Q. Yao (2005) Approximating conditional distribution functions using dimension reduction. *The Annals of Statistics* **33**, 1404–1421.

Hansen, B. E. (2004) Nonparametric Estimation of Smooth Conditional Distributions. Working paper, University of Wisconsin-Madison.

Li, K.-C. & N. Duan (1989) Regression analysis under link violation. *The Annals of Statistics* **17**, 1009–1052.

Martens, M. & S.-H. Poon (2001) Returns synchronization and daily correlation dynamics between international stock markets. *Journal of Banking & Finance* **25**, 1805–1827.

Nelder, J. A. & R. Mead (1965) A simplex method for function minimization. *Computer Journal* **7**, 308–315.

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

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

Newey, W. K. & K. D. West (1987) A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica* **55**, 703–708.

Panchenko, V. (2005) Goodness-of-fit test for copula. *Physica A* **355**, 176–182.

Patton, A. (2005) Estimation of Multivariate Models for Time Series of Possibly Different Lengths. *Journal of Applied Econometrics*. Forthcoming.

Patton, A. (2005) Modelling Asymmetric Exchange Rate Dependence. *International Economic Review*. Forthcoming.

Pelletier, D. (2005) Regime switching for dynamical correlations. *Journal of Econometrics*. Forthcoming.

Rousseeuw, P. & G. Molenberghs (1993) Transformation of non positive semidefinite correlations matrices. *Communications in statistics: theory and methods* **22**, 965–984.

Sklar, A. (1959) Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris* **8**, 229–231.

Sklar, A. (1996) Random Variables, Distribution Functions, and Copulas - a Personal Look Backward and Forward. In L. Rüschendorf, B. Schweizer, & M. D. Taylor (eds.), *Distributions with Fixed Marginals and Related Topics*, pp. 1–14. Hayward, CA: Institute of Mathematical Statistics.

Su, L. & H. White (2005) A Nonparametric Hellinger Metric Test for Conditional Independence. Mimeo, University of California, San Diego.

Tsay, R. S. (2002) *Analysis of Financial Time Series*Wiley series in probability and statistics. Wiley-Interscience.

van den Goorbergh, R. (2004) A Copula-Based Autoregressive Conditional Dependence Model of International Stock Markets. Working Paper 022, Research Department, Netherlands Central Bank.

West, K. D. (1996) Asymptotic Inference about Predictive Ability. *Econometrica* **64**, 1067–1084.

White, H. (2000) A Reality Check For Data Snooping. *Econometrica* **68**, 1097–1127.

Zakoian, J. M. (1994) Threshold heteroskedastik models. *Journal of Economic Dynamics and Control* **18**, 931–955.