Journal of Economic Dynamics & Control 48 (2014) 79–94

Contents lists available at ScienceDirect

Journal of Economic Dynamics & Control

journal homepage: www.elsevier.com/locate/jedc

Comparing the accuracy of multivariate density forecasts in selected regions of the copula support

Cees Diks[a,*], Valentyn Panchenko[b], Oleg Sokolinskiy[c], Dick van Dijk[d]

[a] *Center for Nonlinear Dynamics in Economics and Finance, Faculty of Economics and Business, University of Amsterdam, Roetersstraat 11, NL-1018 WB Amsterdam, The Netherlands*
[b] *School of Economics, Faculty of Business, University of New South Wales, Sydney, NSW 2052, Australia*
[c] *Rutgers Business School: Newark and New Brunswick, Rutgers, The State University of New Jersey, 100 Rockafeller Rd., Rm. 2137 Piscataway, NJ 08854, USA*
[d] *Econometric Institute, Erasmus University Rotterdam, P.O. Box 1738, NL-3000 DR Rotterdam, The Netherlands*

ARTICLE INFO

Article history:
Received 27 September 2013
Received in revised form
19 August 2014
Accepted 20 August 2014
Available online 8 September 2014

JEL classification:
C12
C14
C32
C52
C53

Keywords:
Copula
Density forecast evaluation
Kullback–Leibler Information Criterion
Out-of-sample forecast evaluation

ABSTRACT

This paper develops a testing framework for comparing the predictive accuracy of competing multivariate density forecasts with different predictive copulas, focusing on specific parts of the copula support. The tests are framed in the context of the Kullback–Leibler Information Criterion, using (out-of-sample) conditional likelihood and censored likelihood in order to focus the evaluation on the region of interest. Monte Carlo simulations document that the resulting test statistics have satisfactory size and power properties for realistic sample sizes. In an empirical application to daily changes of yields on government bonds of the G7 countries we obtain insights into why the Student-*t* and Clayton mixture copula outperforms the other copulas considered; mixing in the Clayton copula with the *t*-copula is of particular importance to obtain high forecast accuracy in periods of jointly falling yields.

1. Introduction

In recent years academics and industry experts alike have come to regard the importance of non-linearity in the dependence between financial asset returns. Linear dependence structures characterised by a multivariate Gaussian distribution lead to inadequate measures of trading risk and, consequently, devastating losses. Copulas have emerged as an efficient and popular modeling technique to remedy this situation. However, many different copula specifications are available and the choice of the best copula for forecasting risk is still largely unresolved. Further complicating the task of risk managers, a copula which is best suited for describing moderate financial investment returns may utterly fail to capture the joint dynamics under substantial market instability, such as in periods of common asset depreciations.

To address this issue we propose tests for comparing the predictive accuracy of multivariate density forecasts with different conditional copulas within specific regions of the copula support. The tests are based on average out-of-sample

* Corresponding author.
E-mail addresses: c.g.h.diks@uva.nl (C. Diks), v.panchenko@unsw.edu.au (V. Panchenko), oleg.sokolinskiy@business.rutgers.edu (O. Sokolinskiy), djvandijk@ese.eur.nl (D. van Dijk).

http://dx.doi.org/10.1016/j.jedc.2014.08.021
0165-1889/

scores derived from predicted censored and conditional likelihood. The resulting tests of equal predictive accuracy are sufficiently general to be applied to models with fully-, semi- and non-parametric copulas.

Recognising that the dependence between asset returns is generally non-linear and time-varying, research efforts to accommodate these features have mainly focused on modeling the dynamics of conditional variances and correlations by means of multivariate GARCH and stochastic volatility (SV) models; see the surveys by Silvennoinen and Teräsvirta (2009) and Chib et al. (2009), respectively. Recently, copulas have become an increasingly popular tool for modeling multivariate distributions in finance (Patton, 2009; Genest et al., 2009). The copula approach provides more flexibility than multivariate GARCH and SV models in terms of the types of asymmetric dependence that can be captured. In addition, an attractive property of copulas is that they allow for modeling the marginal distributions and the dependence structure of the asset returns separately.

Many parametric copula families are available, with rather different dependence properties. An important issue in empirical applications therefore is the choice of an appropriate copula specification. In practice, most often this is done by comparing alternative specifications indirectly, subjecting each of them to a battery of goodness-of-fit tests, see Berg (2009) for a detailed review. A direct comparison of alternative copulas from different parametric families has been considered by Chen and Fan (2006) and Patton (2006), adopting the approach based on pseudo likelihood ratio (PLR) tests for model selection originally developed by Vuong (1989) and Rivers and Vuong (2002). These tests compare the candidate copula specifications in terms of their Kullback–Leibler Information Criterion (KLIC), which measures the distance from the true (but unknown) copula. Similar to the goodness-of-fit tests, these PLR tests are based on the in-sample fit of the competing copulas. Diks et al. (2010) approach the copula selection problem from an out-of-sample forecasting perspective. Specifically, the PLR testing approach is extended to compare the predictive accuracy of alternative copula specifications, by using out-of-sample log-likelihood values corresponding with copula density forecasts. An important motivation for considering the (relative) predictive accuracy of copulas is that multivariate density forecasting is one of the main purposes in empirical applications.

Comparison of out-of-sample KLIC values for assessing relative predictive accuracy has recently become popular for the evaluation of univariate density forecasts, see Mitchell and Hall (2005), Amisano and Giacomini (2007) and Bao et al. (2007). Amisano and Giacomini (2007) provide an interesting interpretation of the KLIC-based comparison in terms of scoring rules, which are loss functions depending on the density forecast and the actually observed data.

In many applications of density forecasts, we are mostly interested in a particular region of the density. Financial risk management is an example in case. Due to the regulations of the Basel accords, among others, the main concern for banks and other financial institutions is an accurate description of the left tail of the distribution of their portfolio's returns, in order to obtain accurate estimates of Value-at-Risk and related measures of downside risk. Correspondingly, Bao et al. (2004), Amisano and Giacomini (2007) and Diks et al. (2011) consider the problem of evaluating and comparing univariate density forecasts in a specific region of interest. The out-of-sample KLIC values can be adapted to this case by replacing the full likelihood by the conditional likelihood given that the actual observation lies in the region of interest, or by the censored likelihood, with censoring of the observations outside the region of interest. The weighted log-likelihood scoring rules (Amisano and Giacomini, 2007) proposed are not "proper" in that an incorrect density forecast may receive a higher score on average than the actual conditional density, just because it assigns a higher probability to the region of interest (Diks et al., 2011).

This paper extends this to multivariate cases where the aim is to compare the predictive accuracy of two competing density forecasts within a particular region of the copula support. The resulting tests are valid under rather general conditions on the competing copulas, which is achieved by adopting the framework of Giacomini and White (2006). This assumes in particular that any unknown model parameters are estimated using a moving window of fixed size. The finite estimation window essentially allows us to treat competing density forecasts based on different copula specifications, including the time-varying estimated model parameters, as two competing *forecast methods*. Comparing scores for forecast *methods* rather than for *models* simplifies the resulting test procedures considerably, because parameter estimation uncertainty does not play a role; it simply is part of the respective competing forecast methods. In addition, the asymptotic distribution of our test statistic in this case does not depend on whether or not the competing copulas belong to nested families.

We examine the size and power properties of our copula predictive accuracy tests via Monte Carlo simulations, where we employ fully parametric and semi-parametric copula-based multivariate dynamic models. The latter class of models (shortened as SCOMDY) developed by Chen and Fan (2005, 2006) combines parametric specifications for the conditional mean and conditional variance with a semi-parametric specification for the distribution of the (standardized) innovations, consisting of a parametric copula with nonparametric univariate marginal distributions. Our simulation results demonstrate that the predictive accuracy tests have satisfactory size and power properties in realistic sample sizes.

To illustrate the usefulness of the suggested tests, we consider an empirical application to daily changes of bond yields in the G7 economies. The analyzed data covers the period from March 6, 1991, until May 21, 2014. We identify the copula specification with the best forecasting performance and discover what drives its success.

The paper is organised as follows. In Section 2 we develop our predictive accuracy test for copulas based on out-of-sample log-likelihood scores and briefly review copula-based multivariate density models. In Section 3 we investigate its size and power properties by means of Monte Carlo simulations. In Section 4 we illustrate our test with an application to daily bond yield changes in the economies of the G7 countries. We conclude in Section 5.

2. Methodology

The aim of this paper is to develop density forecast evaluation tests that focus on specific regions of the copula domain. This is achieved by introducing weighted likelihood-based scoring rules with weight functions defined on the copula domain.

2.1. Density forecast evaluation using weighted scoring rules

Before introducing the new soring rules we first motivate the use of weighted likelihood-based scoring rules for density forecast evaluation.

Density forecast evaluation: Consider a stochastic process $\{\boldsymbol{Z}_t : \Omega \to \mathbb{R}^{k+d}\}_{t=1}^{T}$, defined on a complete probability space $(\Omega, \mathcal{F}, \mathcal{P})$, and identify $\boldsymbol{Z}_t$ with $(\boldsymbol{Y}_t, \boldsymbol{X}_t)$, where $\boldsymbol{Y}_t : \Omega \to \mathbb{R}^d$ is the real-valued d-dimensional random variable of interest and $\boldsymbol{X}_t : \Omega \to \mathbb{R}^k$ is a vector of exogenous or pre-determined variables. The information set at time t is defined as $\mathcal{F}_t = \sigma(\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_t)$. We consider the case where two competing forecast methods are available, each producing one-step ahead density forecasts, i.e. predictive densities of $\boldsymbol{Y}_{t+1}$, based on $\mathcal{F}_t$.

Following Amisano and Giacomini (2007), by 'forecast method' we mean a given density forecast in terms of past information, resulting from the choices that the forecaster makes at the time of the prediction. These include the variables $\boldsymbol{X}_t$, the econometric model (if any), and the method for estimation of unknown parameters in the model. The only requirement that we impose on the forecast methods is that the density forecasts depend on a finite number R of most recent observations $\boldsymbol{Z}_{t-R+1}, \ldots, \boldsymbol{Z}_t$. Comparing forecast methods rather than forecast models allows for treating parameter estimation uncertainty in a natural way as an integral part of the forecast methods. The use of a fixed (rolling) estimation window R considerably simplifies the asymptotic theory of tests of equal predictive accuracy, as argued by Giacomini and White (2006). It also is convenient in that it enables comparison of density forecasts based on both nested and non-nested models, in contrast to model comparison approaches such as that of West (1996).

Scoring rules: One approach that has been put forward for density forecast evaluation in general is by means of scoring rules, which are commonly used in probability forecast evaluation, see Diebold and Lopez (1996). A scoring rule is a loss function $S^*(\hat{f}_t; \boldsymbol{y}_{t+1})$ depending on the density forecast $\hat{f}_t$ and the actually observed value $\boldsymbol{y}_{t+1}$, such that a density forecast that is 'better' receives a higher score. As argued by Diebold et al. (1998) and Granger and Pesaran (2000), any rational user would prefer the true conditional density p_t of $\boldsymbol{Y}_{t+1}$ over an incorrect density forecast. It is therefore natural to focus on "proper" scoring rules (Gneiting and Raftery, 2007), for which incorrect density forecasts $\hat{f}_t$ do not receive a higher average score than the true conditional density p_t, that is,

$$\mathsf{E}_t(S^*(\hat{f}_t; \boldsymbol{Y}_{t+1})) \le \mathsf{E}_t(S^*(p_t; \boldsymbol{Y}_{t+1})), \quad \text{for all } t.$$

Note that the correct density p_t does not depend on estimated parameters, while density forecasts typically do. This implies that even if the density forecast $\hat{f}_t$ is based on a correctly specified model, if the model includes estimated parameters the average score $\mathsf{E}_t(S^*(\hat{f}_t; \boldsymbol{Y}_{t+1}))$ may not achieve the upper bound $\mathsf{E}_t(S^*(p_t; \boldsymbol{Y}_{t+1}))$, due to non-vanishing estimation uncertainty. As a consequence, a density forecast based on a misspecified model with limited estimation uncertainty may be preferred over a density forecast based on a correctly specified model having larger estimation uncertainty.

Null hypothesis and testing approach: Given a scoring rule of one's choice, there are various ways to construct tests of equal predictive ability. Giacomini and White (2006) distinguish tests of unconditional predictive ability and conditional predictive ability. Here we focus on tests for unconditional predictive ability for clarity of exposition. The suggested approach can be extended to obtain tests of conditional predictive ability in a straightforward manner.

Assume that two competing density forecasts $\hat{f}_{A,t}$ and $\hat{f}_{B,t}$ and corresponding realisations of the variable $\boldsymbol{Y}_{t+1}$ are available for $t = R, R+1, \ldots, T-1$. We may then compare $\hat{f}_{A,t}$ and $\hat{f}_{B,t}$ based on their average scores, by testing formally whether their difference is statistically significantly different from zero on average. Defining the score difference

$$d^*_{t+1} = S^*(\hat{f}_{A,t}; \boldsymbol{Y}_{t+1}) - S^*(\hat{f}_{B,t}; \boldsymbol{Y}_{t+1}),$$

for a given scoring rule S^*, the null hypothesis of equal expected scores is given by

$$H_0 : \mathsf{E}(d^*_{t+1}) = 0, \quad \text{for all } t = R,\ R+1, \ldots, T-1,$$

Let $\overline{d}^*_{R,P}$ denote the sample average of the score differences, that is, $\overline{d}^*_{R,P} = P^{-1} \sum_{t=R}^{T-1} d^*_{t+1}$ with $P = T - R$. To test the null hypothesis, we may use a Diebold and Mariano (1995) type statistic

$$t_{R,P} = \sqrt{P} \frac{\overline{d}^*_{R,P}}{\sqrt{\hat{\sigma}^2_{R,P}}}, \tag{1}$$

where $\hat{\sigma}^2_{R,P}$ is a heteroskedasticity and autocorrelation-consistent (HAC) variance estimator of $\sigma^2_{R,P} = \mathsf{Var}(\sqrt{P}\overline{d}^*_{R,P})$.

Theorem 4 of Giacomini and White (2006) states that, under certain mixing and moment conditions, $t_{R,P}$ in (1) is asymptotically (as $P \to \infty$ with R fixed) standard normally distributed under the null hypothesis.

The logarithmic scoring rule: Mitchell and Hall (2005), Amisano and Giacomini (2007), and Bao et al. (2004, 2007) focus on the logarithmic scoring rule:

$$S^l(\hat{f}_t; \boldsymbol{y}_{t+1}) = \log \hat{f}_t(\boldsymbol{y}_{t+1}), \tag{2}$$

such that the score assigned to a density forecast varies positively with the value of $\hat{f}_t$ evaluated at the observation $\boldsymbol{y}_{t+1}$. Defining scores in terms of the realisations $\boldsymbol{y}_{t+1}$ rather than the random variables $\boldsymbol{Y}_{t+1}$ is quite common, and can be helpful in particular if they have a possible interpretation as a likelihood (i.e. given the observed value of $\boldsymbol{Y}_{t+1}$). Since that is the case here, we follow this convention.

Based on the P observations available for evaluation, $\boldsymbol{y}_{R+1}, \dots, \boldsymbol{y}_T$, the density forecasts $\hat{f}_{A,t}$ and $\hat{f}_{B,t}$ can be ranked according to their average scores $P^{-1}\sum_{t=R}^{T-1} \log \hat{f}_{A,t}(\boldsymbol{y}_{t+1})$ and $P^{-1}\sum_{t=R}^{T-1} \log \hat{f}_{B,t}(\boldsymbol{y}_{t+1})$. Obviously, the density forecast yielding the highest average score would be the preferred one. The log score differences $d^l_{t+1} = \log \hat{f}_{A,t}(\boldsymbol{y}_{t+1}) - \log \hat{f}_{B,t}(\boldsymbol{y}_{t+1})$ may be used to test whether the predictive accuracy is significantly different, using the test statistic defined in (1). Note that this coincides with the log-likelihood ratio of the two competing density forecasts.

Weighted likelihood-based scoring rules: Diks et al. (2011) adapt the logarithmic scoring rule for evaluating and comparing density forecasts in a specific region of interest, $M_t \subset \mathbb{R}^d$, say, by replacing the full likelihood in (2) either by the conditional likelihood, given that the observation lies in the region of interest, or by the censored likelihood, with censoring of the observations outside M_t. The conditional likelihood (cl) score function is given by

$$S^{cl}\left(\hat{f}_t; \boldsymbol{y}_{t+1}\right) = \mathrm{I}(\boldsymbol{y}_{t+1} \in M_t) \log\left(\frac{\hat{f}_t(\boldsymbol{y}_{t+1})}{\int_{M_t} \hat{f}_t(\boldsymbol{y})\, \mathrm{d}\boldsymbol{y}}\right), \tag{3}$$

where $\mathrm{I}(\cdot)$ denotes the indicator function, which is one if the event in its argument occurs and zero otherwise, while the censored likelihood (csl) score function is given by

$$S^{csl}(\hat{f}_t; \boldsymbol{y}_{t+1}) = \mathrm{I}(\boldsymbol{y}_{t+1} \in M_t) \log \hat{f}_t(\boldsymbol{y}_{t+1}) + \mathrm{I}(\boldsymbol{y}_{t+1} \in M_t^c) \log\left(\int_{M_t^c} \hat{f}_t(\boldsymbol{y})\, \mathrm{d}\boldsymbol{y}\right), \tag{4}$$

where M_t^c is the complement of M_t.

The conditional and censored likelihood scoring rules introduced above, focus on a sharply defined region of interest M_t. More generally one can introduce a weight function $w_t(\boldsymbol{y}_{t+1})$, leading to the scores

$$S^{cl}\left(\hat{f}_t; \boldsymbol{y}_{t+1}\right) = w_t(\boldsymbol{y}_{t+1}) \log\left(\frac{\hat{f}_t(\boldsymbol{y}_{t+1})}{\int w_t(\boldsymbol{y})\hat{f}_t(\boldsymbol{y})\, \mathrm{d}\boldsymbol{y}}\right) \tag{5}$$

and

$$S^{csl}(\hat{f}_t; \boldsymbol{y}_{t+1}) = w_t(\boldsymbol{y}_{t+1}) \log \hat{f}_t(\boldsymbol{y}_{t+1}) + (1 - w_t(\boldsymbol{y}_{t+1})) \log\left(1 - \int w_t(\boldsymbol{y})\hat{f}_t(\boldsymbol{y})\, \mathrm{d}\boldsymbol{y}\right). \tag{6}$$

We make the following assumptions concerning the density forecasts that are to be compared, and the weight function.

Assumption 1. The density forecasts $\hat{f}_{A,t}$ and $\hat{f}_{B,t}$ satisfy $\mathrm{KLIC}(\hat{f}_{A,t}) < \infty$ and $\mathrm{KLIC}(\hat{f}_{B,t}) < \infty$, where $\mathrm{KLIC}(h_t) = \int p_t(\boldsymbol{y}) \log(p_t(\boldsymbol{y})/h_t(\boldsymbol{y}))\, \mathrm{d}\boldsymbol{y}$ is the Kullback–Leibler divergence between the density forecast h_t and the true conditional density p_t.

Assumption 2. The weight function $w_t(\boldsymbol{y})$ is such that (a) it is determined by the information $\mathcal{F}_t$ available at time t, (b) $0 \le w_t(\boldsymbol{y}) \le 1$, and (c) $\int w_t(\boldsymbol{y}) p_t(\boldsymbol{y})\, \mathrm{d}\boldsymbol{y} > 0$.

Assumption 1 ensures that the expected score differences for the competing density forecasts are finite. Assumption 2(c) is needed to avoid cases where $w_t(\boldsymbol{y})$ takes strictly positive values only outside the support of the data.

Under Assumptions 1 and 2, the generalised conditional likelihood scoring rule given in (5) and the generalised censored likelihood scoring rule given in (6) are proper. A proof for univariate predictive densities has been given in the appendix of Diks et al. (2011). The steps in that proof remain valid if the univariate densities are replaced by multivariate densities, and scalar integration variables by a vector-valued integration variables.

2.2. *Copula comparison with weights on the copula domain*

Patton's (2006) extension of Sklar's (1959) theorem to the time-series case describes how the time-dependent multivariate distribution $F_t(\boldsymbol{y}_{t+1})$ can be decomposed into conditional marginal distributions $F_{j,t}(y_j)$, $j = 1, \dots, d$, and a conditional copula $C_t(\cdot)$, that is

$$F_t(\boldsymbol{y}) = C_t(F_{1,t}(y_1), F_{2,t}(y_2), \dots, F_{d,t}(y_d)), \tag{7}$$

provided that the marginal conditional cumulative distribution functions (CDFs) $F_{i,t}$ are continuous. The subscript t in F_t, $F_{i,t}$ $(i=1,\ldots,d)$ and C_t emphasises that these are conditional on $\mathcal{F}_t$. This decomposition emphasises the attractiveness of the copula approach for modeling multivariate distributions. Given that the marginal distributions $F_{j,t}$, $j=1\ldots,d$, only contain univariate information on the individual variables $Y_{j,t+1}$, their dependence is governed completely by the copula function C_t. As the choice of marginal distributions does not restrict the choice of copula, or vice versa, a wide range of joint distributions can be obtained by combining different marginals with different copulas.

The one-step-ahead predictive log-likelihood associated with $\boldsymbol{y}_{t+1}$ based on (7) is given by

$$\sum_{j=1}^{d} \log f_{j,t}(y_{j,t+1})+\log c_t(F_{1,t}(y_{1,t+1}),F_{2,t}(y_{2,t+1}),\ldots,F_{d,t}(y_{d,t+1})), \tag{8}$$

where $f_{j,t}(y_{j,t+1})$, $j=1,\ldots,d$, are the conditional marginal densities and c_t is the conditional copula density, defined as

$$c_t(u_1,u_2,\ldots,u_d)=\frac{\partial^d}{\partial u_1 \partial u_2 \ldots \partial u_d} C_t(u_1,u_2,\ldots,u_d),$$

which we will assume to exist throughout. Using (8), the conditional likelihood and censored likelihood scoring rule of a density forecast $\hat{f}_t(\boldsymbol{y}_{t+1})$ with marginal predictive densities $\hat{f}_{j,t}$, $j=1,\ldots,d$, and copula density $\hat{c}_t$ can be decomposed as

$$S^{cl}_{t+1}=w_t(\boldsymbol{y}_{t+1})\left(\sum_{j=1}^{d} \log \hat{f}_{j,t}(y_{j,t+1})+\log \hat{c}_t(\hat{\boldsymbol{u}}_{t+1})\right)-w_t(\boldsymbol{y}_{t+1})\log\left(\int w_t(\boldsymbol{y})\hat{f}_t(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}\right), \tag{9}$$

and

$$\begin{aligned} S^{csl}_{t+1}=&\,w_t(\boldsymbol{y}_{t+1})\left(\sum_{j=1}^{d} \log \hat{f}_{j,t}(y_{j,t+1})+\log \hat{c}_t(\hat{\boldsymbol{u}}_{t+1})\right)\\ &+(1-w_t(\boldsymbol{y}_{t+1}))\log\left(1-\int w_t(\boldsymbol{y})\hat{f}_t(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}\right), \end{aligned} \tag{10}$$

where S^*_{t+1} is short-hand notation for $S^*(\hat{f}_t;\boldsymbol{y}_{t+1})$, $\hat{c}_t$ is the conditional copula density associated with the density forecast, and $\hat{\boldsymbol{u}}_{t+1}=(\hat{F}_{1,t}(y_{1,t+1}),\ldots,\hat{F}_{d,t}(y_{d,t+1}))'$ its multivariate conditional probability integral transform (PIT).

As in Diks et al. (2010) we assume that the two competing multivariate density forecasts differ only in their copula specifications and have identical predictive marginal densities $\hat{f}_{j,t}$, $j=1,\ldots,d$. The two competing copula specifications are assumed to have well-defined densities $\hat{c}_{A,t}$ and $\hat{c}_{B,t}$. The null hypothesis of equal predictive ability is

$$H_0\colon \mathsf{E}(S^*_{A,t+1})=\mathsf{E}(S^*_{B,t+1}),$$

where '$*$' stands for either '*cl*' or '*csl*', and where we have, with slight abuse of notation, used the same notation for the random variables $S^*_{\cdot,t+1}$ and their realisations. Since the conditional marginals are identically specified under both density forecasts, the logarithms of the marginal densities in (9) and (10) cancel out, so that an equivalent formulation of the null hypothesis is

$$H_0\colon \mathsf{E}(\mathcal{S}^*_{A,t+1})=\mathsf{E}(\mathcal{S}^*_{B,t+1}),$$

based on the simpler scores (note the absence of the contributions from the predictive marginals)

$$\mathcal{S}^{cl}_{t+1}=w_t(\boldsymbol{y}_{t+1})\left(\log(\hat{c}(\hat{\boldsymbol{u}}_{t+1}))-\log\int w_t(\boldsymbol{y})\hat{f}_t(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}\right)$$

and

$$\mathcal{S}^{csl}_{t+1}=w_t(\boldsymbol{y}_{t+1})\log \hat{c}_t(\hat{\boldsymbol{u}}_{t+1})+(1-w_t(\boldsymbol{y}_{t+1}))\log\left(1-\int w_t(\boldsymbol{y})\hat{f}_t(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}\right).$$

We use the weight function to focus on specific regions of the copula support. This can be achieved by taking weight functions of the form:

$$w_t(\boldsymbol{y}_{t+1})=\tilde{w}(\hat{u}_{1,t+1}(y_{1,t+1}),\ldots,\hat{u}_{d,t+1}(y_{d,t+1})),$$

where $\tilde{w}(u_1,\ldots,u_d)$ is a weight function defined on the copula support. Note that

$$\begin{aligned} \int w_t(\boldsymbol{y})\hat{f}_t(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}&=\int w_t(\boldsymbol{y})\hat{c}_t(\hat{\boldsymbol{u}}(\boldsymbol{y}))\hat{f}_{1,t}(y_1)\times\cdots\times\hat{f}_{d,t}(y_d)\,\mathrm{d}\boldsymbol{y}\\ &=\int \tilde{w}_t(\hat{\boldsymbol{u}})\hat{c}_t(\hat{\boldsymbol{u}})\,\mathrm{d}\hat{\boldsymbol{u}}. \end{aligned}$$

This allows us to rewrite the scores $\mathcal{S}^*_{t+1}$ as

$$\mathcal{S}^{cl}_{t+1}=\tilde{w}_t(\hat{\boldsymbol{u}}_{t+1})\left(\log \hat{c}_t(\hat{\boldsymbol{u}}_{t+1})-\log\int \tilde{w}_t(\boldsymbol{u})\hat{c}_t(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\right) \tag{11}$$

and

$$\mathcal{S}^{csl}_{t+1} = \tilde{w}_t(\hat{\boldsymbol{u}}_{t+1})(\log \hat{c}_t(\hat{\boldsymbol{u}}_{t+1})) + (1-\tilde{w}_t(\hat{\boldsymbol{u}}_{t+1}))\log\left(1-\int \tilde{w}_t(\boldsymbol{u})\hat{c}_t(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\right). \tag{12}$$

By choosing the weight function $\tilde{w}_t(\boldsymbol{u}) = 1$ one retrieves the log scoring rule:

$$\mathcal{S}^{l}_{t+1} = \log \hat{c}_t(\hat{\boldsymbol{u}}_{t+1}) \tag{13}$$

considered in Diks et al. (2010) for density forecast comparison on the full copula domain.

Note that the 'reduced' scoring rules (11) and (12) take the same form as the weighted likelihood-based scoring rules (5) and (6) derived before, but now involve only the density forecast copula instead of the full density forecast and the observed conditional PITs $\hat{\boldsymbol{u}}_{t+1}$ instead of the variable $\boldsymbol{y}_{t+1}$. This points out an important difference with Diks et al. (2010), where no weight function was used on the copula, and emphasis was placed on comparing density forecasts with the same predictive marginals merely to obtain simplified scoring rules. Here the fact the forecasts have identical predictive marginals also serves another purpose, namely of guaranteeing that the two competing density forecasts agree on the conditional PITs $\hat{\boldsymbol{u}}_{t+1}$ on which the weight function is defined.

In the cases considered in this paper, $\tilde{w}_t$ will be time independent, and will take the form of an indicator function of a given fixed subset of the copula support. This allows for a simplification of the scoring rules, since the integrals occurring in (11) and (12) can be expressed in term of the copula $\hat{C}_t$.

Above, we have assumed that the two competing multivariate density forecasts differ only in their copula specifications and have identical predictive marginal densities $\hat{f}_{j,t}$, $j=1,\ldots,d$. Implicitly this assumes that the parameters in the marginals and the copula can be separated from each other, so that they can be estimated in a multi-stage procedure for a given dataset. No other restrictions are put on the marginals. In particular, they may be specified parametrically, nonparametrically, or semi-parametrically.

2.3. Copula-based multivariate dynamic models

In the Monte Carlo simulations and the empirical application in subsequent sections we use copula-based multivariate dynamic models (see Patton, 2012, 2013 for reviews). To keep this paper self-contained we briefly describe this class of models here. The general copula-based multivariate dynamic model is specified as

$$\boldsymbol{Y}_t = \boldsymbol{\mu}_t(\boldsymbol{\theta}_1) + \sqrt{H_t(\boldsymbol{\theta})}\boldsymbol{\varepsilon}_t, \tag{14}$$

where

$$\boldsymbol{\mu}_t(\boldsymbol{\theta}_1) = (\mu_{1,t}(\boldsymbol{\theta}_1), \ldots, \mu_{d,t}(\boldsymbol{\theta}_1))' = \mathsf{E}[\boldsymbol{Y}_t|\mathcal{F}_{t-1}]$$

is a specification of the conditional mean, parameterised by a finite dimensional vector of parameters $\boldsymbol{\theta}_1$, and

$$H_t(\boldsymbol{\theta}) = \mathrm{diag}(h_{1,t}(\boldsymbol{\theta}), \ldots, h_{d,t}(\boldsymbol{\theta})),$$

where

$$h_{j,t}(\boldsymbol{\theta}) = h_{j,t}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = \mathsf{E}[(Y_{j,t} - \mu_{j,t}(\boldsymbol{\theta}_1))^2|\mathcal{F}_{t-1}], \quad j = 1, \ldots, d,$$

is the conditional variance of $Y_{j,t}$ given $\mathcal{F}_{t-1}$, parameterised by a finite-dimensional vector of parameters $\boldsymbol{\theta}_2$, where $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ do not have common elements. The innovations $\boldsymbol{\varepsilon}_t = (\varepsilon_{1,t}, \ldots, \varepsilon_{d,t})'$ are independent of $\mathcal{F}_{t-1}$ and independent and identically distributed (i.i.d.) with $\mathsf{E}(\varepsilon_{j,t}) = 0$ and $\mathsf{E}(\varepsilon^2_{j,t}) = 1$ for $j = 1, \ldots, d$. Applying Sklar's theorem, the joint distribution function $F(\boldsymbol{\varepsilon})$ of $\boldsymbol{\varepsilon}_t$ can be written as

$$F(\boldsymbol{\varepsilon}) = C(F_1(\varepsilon_1), \ldots, F_d(\varepsilon_d); \boldsymbol{\alpha}) \equiv C(u_1, \ldots, u_d; \boldsymbol{\alpha}), \tag{15}$$

where $C(u_1, \ldots, u_d; \boldsymbol{\alpha})$: $[0,1]^d \to [0,1]$ is a member of a parametric family of copula functions with finite dimensional parameter vector $\boldsymbol{\alpha}$.

The univariate marginal distributions of the standardised innovations $F_j(\cdot)$, $j = 1, \ldots, d$ are estimated either parametrically or nonparametrically. The models with parametrically specified marginal distributions are typically estimated in two stages. First, for a given marginal distribution univariate quasi-maximum likelihood is used to estimate the parameters of the marginals $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ separately for each for each $j = 1, \ldots, d$. Second, the parameters $\boldsymbol{\alpha}$ of the given copula specification are estimated by maximising the corresponding copula log-likelihood and conditioning on the parameters of the marginals obtained in the first stage. While this procedure is less efficient than full MLE, it is much simpler, and the loss of efficiency is generally not substantial (Patton, 2006).

Chen and Fan (2006) propose estimating the marginal distributions of the standardised innovations nonparametrically and introduce the class of semi-parametric copula-based multivariate dynamic (SCOMDY) models. They suggest a three-stage procedure to estimate the SCOMDY model parameters. The first stage coincides with the first stage of the parametric estimation under the assumptions of normality of the standardised innovations $\varepsilon_{j,t}$. The second stage estimates the marginal distributions $F_j(\cdot)$ by means of the empirical CDF (ECDF) transformation of the residuals $\hat{\varepsilon}_{j,t} \equiv (y_{j,t} - \mu_{j,t}(\hat{\boldsymbol{\theta}}_1))/\sqrt{h_{j,t}(\hat{\boldsymbol{\theta}})}$. Finally, the parameters of a given parametric copula are estimated using a parametric procedure conditioning on the estimates of the marginal CDFs obtained in the second stage.

3. Monte Carlo simulations

In this section we use Monte Carlo simulation to examine the finite-sample behaviour of our predictive accuracy tests for comparing alternative copula specifications in specific regions of the copula support.

In all experiments, the true data generating process (DGP) is based on an AR(1) specification for the conditional mean with zero constant term and AR coefficient set to 0.1 and a GARCH(1,1) specification for the conditional variance, with coefficients that are typical for financial applications, that is, the constant term, the ARCH and GARCH coefficients are set to $0.1, 0.05$ and 0.85, respectively.

The vector-valued innovations $\boldsymbol{\varepsilon_t}$ are i.i.d. with mean zero and variance one, but the elements $\varepsilon_{j,t}, j \in 1, \ldots, d$ for given t are not independent. Specifically, for given t, each of the variables $\varepsilon_{j,t}$ are marginally standard normally distributed, while their true dependence is described by the Student-t copula. The alternative copulas used in the simulations are the Gaussian copula, the Clayton copula and the non-exchangeable Clayton copula. For a detailed description of these copulas and their applications in financial econometrics we refer to Nelsen (2006) and the reviews of Patton (2009, 2012, 2013).

Non-exchangeable copulas are not invariant under permutation of their arguments (see Liebscher, 2008 for their construction and examples). For size simulations we employ a non-exchangeable copula based on the Clayton copula family, that is,

$$C^{\mathrm{nonexCl}}(\boldsymbol{u}; \alpha) = \prod_{j=1}^{d} u_j^{1-\theta_j} \left(\sum_{j=1}^{d} u_j^{-\alpha\theta_j} - d + 1 \right)^{-1/\alpha} \quad \text{with } \alpha > 0, \quad 0 \le \theta_j \le 1.$$

The parameter α determines the joint strength of dependence for all variables in the copula and the parameter θ_j tilts each dimension away from the independence copula.

All models are estimated using maximum likelihood. For models involving marginal distributions we estimate the parameters of the marginal distributions first and after obtaining the standardised innovations we transform them into PITs, either using the ECDF or the corresponding parametric CDF. These are then used to estimate the copula parameters.

In the simulation experiments we consider copulas of various dimensions, namely, $d=2$, $d=5$ and $d=10$. We set the number of observations for the moving in-sample window to $R=1000$ and compare the results for two different out-of-sample forecasting periods $P=1000$ and $P=5000$. The asymptotic distribution results for the considered tests assume that P tends to infinity with R fixed, but in practice it is not always possible to have large P. This motivates us to consider the finite sample properties of the tests for the more feasible situation $R=P=1000$. The number of replications in each experiment is set to $B=1000$.

Table 1
Size test of equal performance of the two non-exchangeable Clayton copulas for the DGP with the Student-t copula.

Simulation parameters		$r=0.10$		$r=0.15$		$r=0.25$		$r=0.30$	
		cens	cond	cens	cond	cens	cond	cens	cond
$P=1000$									
$d=2$	$\nu=5$	0.025	0.025	0.020	0.019	0.023	0.022	0.024	0.026
	$\nu=10$	0.030	0.031	0.017	0.017	0.023	0.022	0.026	0.025
	$\nu=20$	0.040	0.039	0.024	0.024	0.026	0.024	0.026	0.026
$d=5$	$\nu=5$	0.025	0.024	0.043	0.041	0.054	0.049	0.056	0.048
	$\nu=10$	0.023	0.020	0.040	0.040	0.049	0.047	0.056	0.047
	$\nu=20$	0.019	0.018	0.026	0.027	0.037	0.037	0.051	0.046
$d=10$	$\nu=5$	0.012	0.006	0.022	0.018	0.041	0.039	0.049	0.041
	$\nu=10$	0.007	0.000	0.019	0.016	0.041	0.038	0.048	0.043
	$\nu=20$	0.006	0.002	0.019	0.016	0.053	0.054	0.041	0.036
$P=5000$									
$d=2$	$\nu=5$	0.010	0.010	0.011	0.009	0.011	0.009	0.013	0.013
	$\nu=10$	0.009	0.009	0.009	0.009	0.010	0.010	0.013	0.013
	$\nu=20$	0.010	0.013	0.011	0.009	0.012	0.012	0.007	0.006
$d=5$	$\nu=5$	0.044	0.046	0.055	0.052	0.074	0.062	0.079	0.064
	$\nu=10$	0.043	0.046	0.056	0.054	0.051	0.044	0.068	0.060
	$\nu=20$	0.051	0.048	0.051	0.050	0.055	0.055	0.065	0.049
$d=10$	$\nu=5$	0.043	0.038	0.047	0.042	0.056	0.053	0.051	0.046
	$\nu=10$	0.044	0.041	0.050	0.047	0.053	0.047	0.050	0.044
	$\nu=20$	0.032	0.034	0.058	0.057	0.052	0.052	0.048	0.046

Note: The table shows the rejection rates of a two-sided test of equal performance of the two non-exchangeable Clayton copulas for 0.05 significance level. The true DGP is based on a Student-t copula with $\nu=5$ degrees of freedom and all pair-wise correlations set to $\rho=0.5$. The true marginals follow an AR(1)-GARCH(1,1) process with standard normal innovations and are estimated using SCOMDY model. The tests are based on the copula region $[0,r]^d$ with variable threshold r and dimension d and use either censored or conditional scores. The number of observations in the moving in-sample estimation window is $R=1000$ and the number of out-of-sample evaluations is $P=1000$ and $P=5000$. The reported results are based on 1000 replications.

3.1. Size

In order to assess the size properties of the tests, a case is required with two competing copulas that are both 'equally (in) correct'. We achieve this with the following setup. We consider DGPs with a Student-t copula with degrees of freedom $\nu = 5$, 10 and 20, and all pair-wise correlations set to $\rho = 0.5$. We test the null hypothesis of equal predictive accuracy of two non-exchangeable Clayton copulas for one of which the tilting parameter is set to $\theta_1 = 0.5$ and all $\theta_j = 1,\ j = 2, \dots, d$ and for the other $\theta_d = 0.5$ and all $\theta_j = 1,\ j = 1, \dots, d-1$. The parameter α is estimated for both non-exchangeable Clayton copulas, based on the moving in-sample window. Due to the exchangeability of the Student-t copula, the two competing copula specifications are equally distant from the true copula. We consider the lower tail region $[0, r]^d$ for $r = 0.10$, $0.15, 0.25, 0.30$. For smaller r the number of observations falling within the regions of interest was not sufficient, especially in higher dimensions. We report results based on two-sided tests, for censored as well as conditional scores.

Table 1 shows the observed rejections rates for all considered dimensions and degrees of freedom at 0.05 significance level while Fig. 1 shows the discrepancy between the observed rejection rate (or actual size) and the nominal size of the test

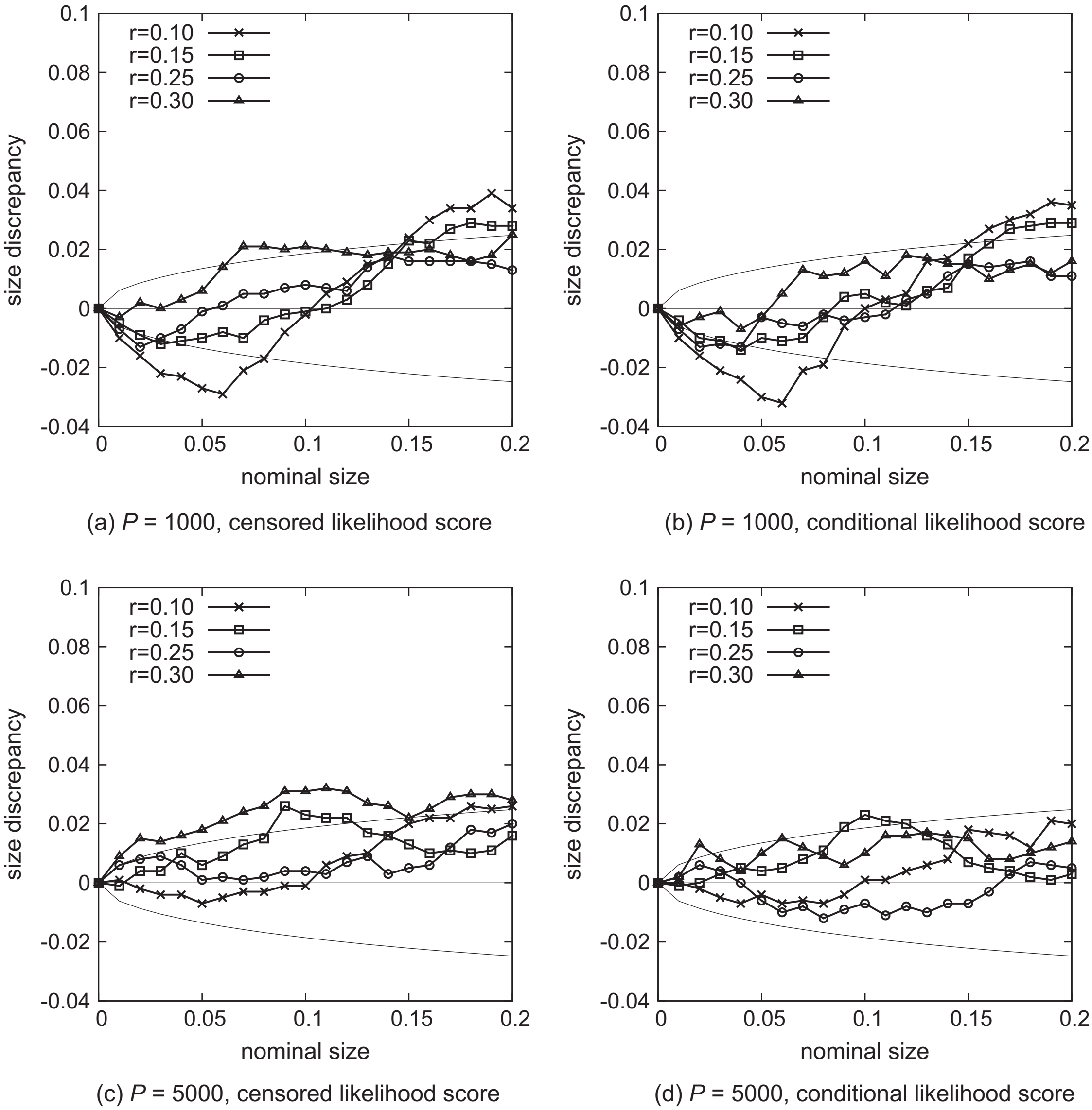

Fig. 1. Size discrepancy plots for the test of equal predictive accuracy for various levels of dependence in the DGP. The panels display the actual size – nominal size discrepancy of a two-sided asymptotic test of equal performance of the two non-exchangeable Clayton copulas. The true DGP is based on a Student-t copula with $\nu = 5$ degrees of freedom and all pair-wise correlations set to $\rho = 0.5$. The true marginals follow an AR(1)-GARCH(1,1) process with standard normal innovations and are estimated using SCOMDY model. The tests are based on the copula region $[0, r]^d$ ($d = 5$) and use either censored (left) or conditional scores (right). The number of observations in the moving in-sample estimation window is $R = 1000$ and the number of out-of-sample evaluations is $P = 1000$ and $P = 5000$. The reported results are based on $B = 1000$ replications. The thin lines indicate the 95% point-wise significance intervals (± 1.96 times the theoretical standard deviation $\sqrt{\alpha(1-\alpha)/B}$ of the estimated rejection rate for nominal size α under H_0).

Table 2
Power of a one-sided test of equal performance of the Gaussian and Student-*t* copulas with correctly specified Student-*t* copula.

Simulation parameters		$r=0.10$ cens	$r=0.10$ cond	$r=0.15$ cens	$r=0.15$ cond	$r=0.25$ cens	$r=0.25$ cond	$r=0.30$ cens	$r=0.30$ cond
$P=1000$									
$d=2$	$\nu=5$	0.269	0.352	0.369	0.404	0.478	0.496	0.501	0.506
	$\nu=10$	0.189	0.245	0.217	0.248	0.252	0.275	0.257	0.261
	$\nu=20$	0.188	0.237	0.210	0.240	0.171	0.180	0.163	0.170
$d=5$	$\nu=5$	0.238	0.413	0.412	0.553	0.702	0.718	0.784	0.786
	$\nu=10$	0.096	0.247	0.244	0.357	0.443	0.469	0.511	0.518
	$\nu=20$	0.067	0.229	0.204	0.308	0.312	0.346	0.324	0.337
$d=10$	$\nu=5$	0.105	0.205	0.335	0.480	0.748	0.775	0.857	0.873
	$\nu=10$	0.047	0.099	0.184	0.327	0.524	0.571	0.604	0.627
	$\nu=20$	0.030	0.062	0.157	0.258	0.365	0.410	0.383	0.412
$P=5000$									
$d=2$	$\nu=5$	0.779	0.786	0.866	0.867	0.958	0.958	0.976	0.976
	$\nu=10$	0.477	0.541	0.576	0.597	0.638	0.649	0.691	0.688
	$\nu=20$	0.405	0.454	0.452	0.478	0.421	0.426	0.374	0.377
$d=5$	$\nu=5$	0.885	0.932	0.985	0.987	1.000	1.000	1.000	1.000
	$\nu=10$	0.557	0.741	0.806	0.863	0.955	0.962	0.980	0.982
	$\nu=20$	0.368	0.566	0.633	0.708	0.767	0.785	0.795	0.803
$d=10$	$\nu=5$	0.820	0.942	0.962	0.983	1.000	1.000	1.000	1.000
	$\nu=10$	0.544	0.816	0.860	0.933	0.982	0.987	0.994	0.994
	$\nu=20$	0.315	0.618	0.762	0.882	0.927	0.940	0.923	0.933

Note: The table shows the observed rejection rates of a one-sided test of equal performance of the Gaussian and Student-*t* copulas, against the alternative hypothesis that the correctly specified Student-*t* copula has a higher average score. The tests are based on the left tail copula region $[0,r]^d$ and use either censored or conditional scores. The Student-*t* copula characterising the DGP has varying degrees of freedom parameter ν and varying dimensions d. The pair-wise correlation coefficients ρ_{ij}, $i \neq j$, are all set to 0.5. The true marginals follow an AR(1)-GARCH(1,1) process with standard normal innovations and are estimated using SCOMDY model. The test of equal predictive accuracy compares a Student-*t* copula (with both parameters ρ and ν estimated, rather than known) against a Gaussian copula with the parameter ρ also being estimated. The nominal size is 0.05, the number of observations in the moving in-sample estimation window is $R=1000$ and the number of out-of-sample evaluations is $P=1000$ and $P=5000$. The reported results are based on 1000 replications.

for $d=5$ and $v=5$. The test seems to be rather conservative for $d=2$ and gets closer to the actual size for higher dimensions. The deviations are mainly caused by relatively poor finite sample performance of the HAC variance estimator, which attempts to capture dependence between scores.[1] The tests based on censored and conditional likelihood scores exhibit similar properties.

We also considered other DGPs and regions to verify the size properties of the test. We do not report detailed results here for brevity; they can be briefly summarised as follows. The unifying theme is that the size distortion increases with the strength of dependence in the considered DGP and reduces for higher dimensions. We also considered different treatments of the marginal distributions in addition to the baseline case of semi-parametric marginals, namely, parametric marginals with known parameters and parametric marginals with estimated parameters. In some cases the semi-parametric models showed somewhat larger size distortions than the parametric models and models with fully known marginals.

3.2. Power

We evaluate the power of the tests of equal predictive accuracy by performing a simulation experiment where one of the competing copula specifications corresponds with that of the true DGP, while the distance of the alternative, incorrect copula specification to the DGP varies depending on a certain parameter of the DGP. Specifically, the DGP is a Student-*t* copula of varying dimension $d=2,5,10$ with all pair-wise correlation coefficients $\rho=0.5$. The number of degrees of freedom ν is varied over the interval $[5,20]$. We compare the predictive accuracy of the correct Student-*t* copula specification (with both parameters ρ and ν being estimated) against an incorrect Gaussian copula specification in the left tail region $[0,r]^d$ for $r=0.10, 0.15, 0.25, 0.30$. Hence, we focus on the question whether the proposed tests can distinguish between copulas with and without tail dependence. Note, however, that the Student-*t* copula approaches the Gaussian copula as ν increases, and the tail dependence disappears. Intuitively, the higher the value of ν in the DGP, the more difficult it becomes to distinguish between these two copula specifications.

The full results are shown in Table 2 and are visualised for the censored likelihood-based scores in Fig. 2 in the form of power plots, showing the observed rejection rates (for a nominal size of 0.05) as a function of the degrees of freedom parameter, ν, in the DGP. The results displayed are for the null hypothesis that the Gaussian and Student-*t* copulas perform equally well, against the one-sided alternative hypothesis that the correctly specified Student-*t* copula has a higher average

[1] We have tried various implementations of the HAC estimator available in the literature (see den Haan Wouter and Levin Andrew, 1996 for a review), but no implementation was fully satisfactory for the considered DGPs and sample sizes.

score. Since the true DGP uses the Student-t copula, we might expect the Student-t copula to perform better. Note, however, that as the number of degrees of freedom ν in the copula describing the DGP becomes large, the Gaussian copula might outperform the Student-t copula. This is a consequence of the fact that the Gaussian copula is very close to the Student-t copula for large values of ν, but requires one parameter less to be estimated. Indeed, the rejection rates become smaller than the nominal size for very large values of ν. Consequently, Fig. 2 shows that the test has higher power for smaller values of ν. For higher dimensions we typically observe increased power, at least in the situations when a sufficient number of observations falls in the considered tail region. We also observe that in most of the cases, for higher r, that is, a larger tail region the power is higher. However, in the situation when both copulas are close to each other, $\nu = 20$, the tests based on lower regions may provide a slightly higher power, especially in the lower dimensions when the number of observations falling within these regions is sufficiently high. Hence, in practice, the selection of the appropriate region will depend on the specific DGP and considered dimension. It is important to verify that a sufficient number of observations falls within a considered region. Naturally, the test based on the larger number of out-of-sample evaluations, P, shows higher rejection rates.

We also considered other DGPs to verify the power properties and generally for higher levels of dependence our test showed higher power. In summary, although the suggested tests of predictive accuracy in the selected region of support exhibit moderate discrepancy from the nominal size, they have satisfactory statistical power.

4. Empirical application

We examine the empirical usefulness of our predictive accuracy test with an application to modeling yield changes in major world economies. Specifically, we consider daily changes of yields on the 10 year benchmark government bonds of the

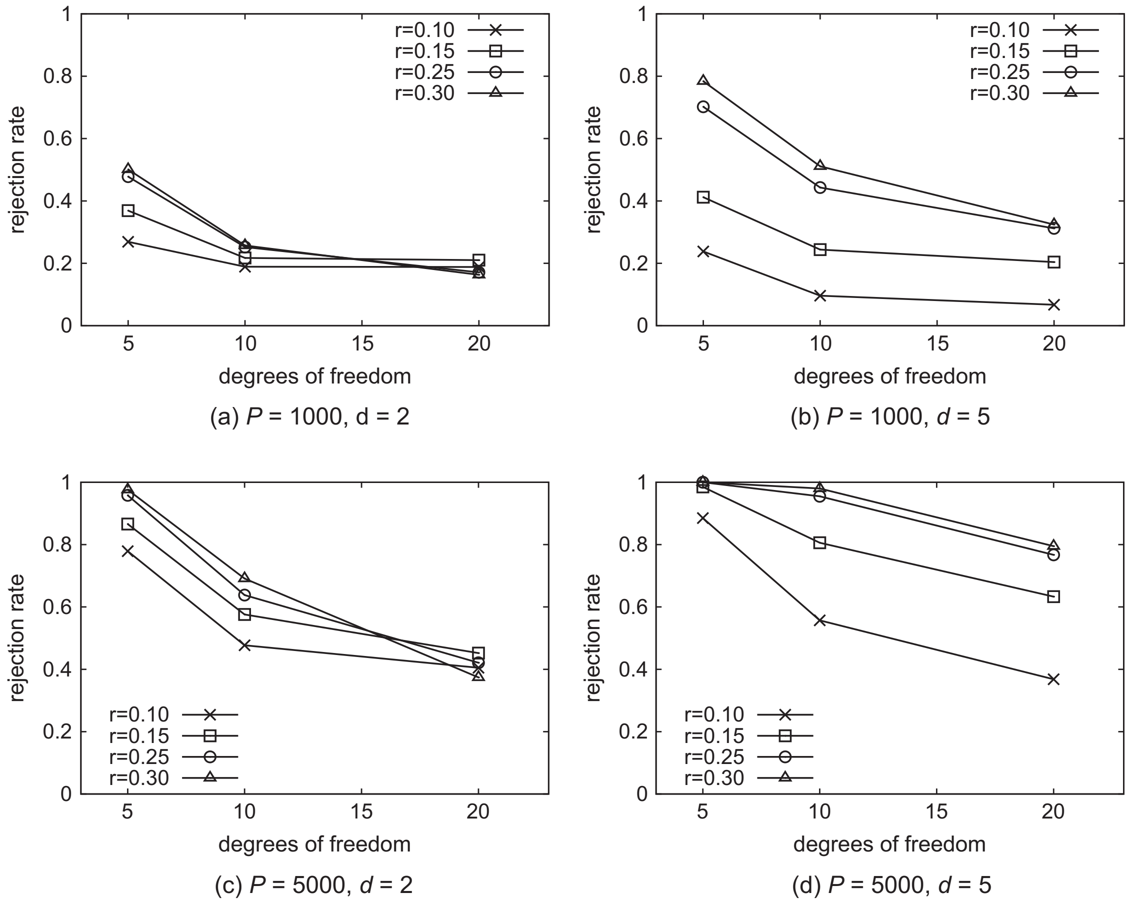

Fig. 2. Power of the test of equal predictive accuracy for various levels of tail dependence. The panels show the observed rejection rates (on the vertical axis) of a one-sided test of equal performance of the Gaussian and Student-t copulas, against the alternative hypothesis that the correctly specified Student-t copula has a higher average score. The tests are based on the left tail copula region $[0, r]^d$ and use censored-likelihood based scores. The horizontal axis displays the degrees of freedom parameter of the Student-t copula characterising the DGP. The pair-wise correlation coefficients ρ_{ij}, $i \neq j$, are all set to 0.5. The true marginals follow an AR(1)-GARCH(1,1) process with standard normal innovations and are estimated using SCOMDY model. The test of equal predictive accuracy compares a Student-t copula (with both parameters ρ and ν estimated, rather than known) against a Gaussian copula with the parameter ρ also being estimated. The nominal size is 0.05, the number of observations in the moving in-sample estimation window is $R=1000$ and the number of out-of-sample evaluations is $P=1000$ and $P=5000$. The reported results are based on 1000 replications.

G7 countries: United States, Canada, Japan, United Kingdom, Germany, France, and Italy. The data on the yields are obtained from Datastream. We base our analysis on the daily data over the period from March 6, 1991, until May 21, 2014, comprising 6055 observations.[2]

Recalling that yields are inversely related to prices of fixed income securities, notice that yield increases correspond to drops in bond prices, while yield decreases correspond to growth in bond prices. Since multiple fixed income instruments depend crucially on their domestic yield curves, the yield changes influence trades involving positions in multiple sectors of the fixed income market. Understanding and managing the market risks of fixed income portfolios is of significant interest to pension funds, insurance companies, and hedge funds. From the macroeconomic perspective, yields on government debt represent an important measure of the interest rates in the economy. With the post-crisis increase in the role of the monetary policy (Quantitative Easing), the yield curve has become a major indicator of the economic policy and the stimulus applied to the economy. The dependence between yield changes in the G7 economies reflects the contemporary views held by the central banks of these countries.

We employ a GARCH framework to model the marginal characteristics of the daily yield changes. For the conditional mean and the conditional variance of the change of yield j we use an AR(5)-GARCH(1,1) specification, given by

$$Y_{j,t} = c_j + \sum_{l=1}^{5} \phi_{j,l} Y_{j,t-l} + \sqrt{h_{j,t}}\varepsilon_{j,t} \tag{16}$$

$$h_{j,t} = \omega_j + \alpha_j \left(Y_{j,t-1} - c_j - \sum_{l=1}^{5} \phi_{j,l} Y_{j,t-1-l} \right)^2 + \beta_j h_{j,t-1}, \tag{17}$$

where $\omega_j > 0$, $\beta_j \geq 0$, $\alpha_j > 0$ and $\alpha_j + \beta_j < 1$. These lag lengths and GARCH orders were sufficient to remove AR and GARCH effects in the marginals. No attempts were made to optimise lag lengths of the model.

The joint distribution of the standardised innovations $\varepsilon_{j,t}$ combines empirical univariate marginal distributions F_j with a parametric copula C. We consider a substantial number of different copula specifications. In particular, we consider the Gaussian (Ga) and Student-t (St-t) elliptic copulas and the classic Clayton (Cl), Clayton survival (Cl-s), and a mixture of Clayton and Clayton survival (Cl/Cl-s) copulas. Moreover, we extend the standard set of copulas by also analyzing the predictive performance of mixtures of elliptic and Clayton family copulas: Gaussian and Clayton (Ga/Cl), Gaussian and Clayton survival (Ga/Cl-s), Student-t and Clayton (St-t/Cl), Student-t and Clayton survival (St-t/Cl-s). In the application we have considered a number of copulas as well as some of their mixtures. It might be possible to obtain better results by considering combinations of more than two copulas, but this is left for future research.

We compare the one-step ahead density forecasting performance of the different copula specifications using a rolling window scheme. The length of the rolling estimation window is set to $R=1000$ observations, such that $P=5055$ observations (from January 5, 1995, until May 21, 2014) are left for out-of-sample forecast evaluation. For comparing the accuracy of the resulting copula-based density forecasts we use the Diebold–Mariano–West type test based on the conditional likelihood in (11) and the censored likelihood in (12). As both scoring rules give qualitatively similar results, to save space we only report results of the tests based on the censored likelihood.[3]

We focus on three specific regions of the copula support. The first region, labeled D, corresponds to a simultaneous fall of all the yields, and is defined as

$$D = \{(u_1, \dots, u_d) \in [0,1]^d | u_j < r \text{ for all } j = 1, \dots, d\},$$

where $d=7$ and the threshold $r \in \{0.15, 0.20, 0.25, 0.30\}$. Note that we only consider regions with identical thresholds for all yield changes. This admittedly is an arbitrary choice, but it is made in order to limit the number of regions under consideration. Below, we present detailed results for $r=0.25$ only, but consider the alternative values in the verification of the robustness of overall results. The second region, denoted U, is the mirror image of D in the sense that it represents a simultaneous increase of all the considered yields

$$U = \{(u_1, \dots, u_d) \in [0,1]^d | u_j > 1 - r \text{ for all } j = 1, \dots, d\},$$

where again the threshold $r \in \{0.15, 0.20, 0.25, 0.30\}$. The third region concerns the central part of the copula support. This region M is defined as

$$M = \{(u_1, \dots, u_d) \in [0,1]^d | r < u_j < 1 - r \text{ for all } j = 1, \dots, d\},$$

where we use the same values of r as for the regions D and U. Region M corresponds to ‘regular’ economic conditions.

As an additional diagnostic tool, we use the model confidence set (MCS) concept of Hansen et al. (2011) to identify the collection of models which includes the best copula specification with a certain level of confidence. Starting with the full set of models, at each iteration we test the null hypothesis that all considered models have equal predictive ability on average across the out-of-sample period. If the null hypothesis is rejected, the worst performing model is omitted, and equal predictive ability is tested again for the remaining models. This procedure is repeated until the null hypothesis can no longer

[2] The choice of the particular range of dates reflects data availability for yields on Italian government debt and the date the sample was obtained.

[3] Detailed results based on the conditional likelihood are available upon request.

be rejected, and the collection of models that remains at this point is defined to be the MCS. In our implementation of the MCS procedure, we continue to exclude the worst performing model at each iteration and repeat the algorithm until only one model remains in the MCS. In this way we obtain a complete ranking of the competing models, together with corresponding p-values. We report the MCS p-values at every iteration. The critical values of the MCS are based on the stationary bootstrap of Politis and Romano (1994) with a probability of sampling the consecutive observation equal to 0.90. The ranking of the different copula specifications and the MCS p-values are robust with respect to the choice of this probability. Due to space considerations, the MCS results are reported in the tables, but not discussed in detail.

We first conduct the test of forecasting accuracy based on the entire support, that is, without distinguishing its particular regions. Table 3 reports the values of the pairwise $Q_{R,P}$ test statistic based on the log-scoring rule $\mathcal{S}^l(\boldsymbol{y}_{t+1})$ as given in (13). Obviously, the matrices in the two panels are antisymmetric, that is $Q_{R,P}(i,j) = -Q_{R,P}(j,i)$ for copula specifications i and j. We nevertheless report the full matrices as this allows for an easy assessment of the relative performance of the various copulas. Given that in each panel the (i,j)th entry is based on the score difference $d^l_{t+1} = \mathcal{S}^l_{j,t+1}(\boldsymbol{y}_{t+1}) - \mathcal{S}^l_{i,t+1}(\boldsymbol{y}_{t+1})$, positive values of the test statistic indicate that the copula in column j achieves a higher average score than the one in row i. Hence, the more positive values in a given column, the higher the ranking of the corresponding copula specification.

From Table 3, we observe that the mixture of the Student-t and Clayton copulas (St-t/Cl) performs best when the full support is taken into account. Based on the pairwise $Q_{R,P}$ test statistic, the null of equal predictive accuracy is rejected at better than the 0.01 significance level for all competing copula specifications, with the exception of the mixture of the Student-t and Clayton survival copulas (St-t/Cl-s) (against which the test statistic has a p-value of 0.108). Thus, adding either lower or upper tail dependence structure of the Clayton family copulas improves upon the simple Student-t copula.

Next we proceed to analyze the source of this result by focusing on sub-regions. Table 4 shows results of the pairwise $Q_{R,P}$ test statistic based on the censored likelihood scoring rule for regions D, M, and U with $r = 0.25$. These demonstrate that the mixture of the Student-t and Clayton copulas (St-t/Cl) is strongly preferred for the D region, while the mixture of the Student-t and Clayton survival copulas (St-t/Cl-s) performs marginally better in the M and U regions. The numbers of observations in each of the sub-regions across the out-of-sample period are 90 for D, 280 for M and 72 for U, respectively. Next, we describe and discuss these results in more detail.

The dependence structure in region D strongly favours the mixture of the Student-t and Clayton copulas (see Panel A in Table 4). Notably, this specification outperforms the Student-t copula at the 0.05 significance level. This signifies a departure from the purely elliptical copula specification. Moreover, the p-value of the test of the Student-t and Clayton mixture (St-t/Cl) against the Student-t and Clayton survival (St-t/Cl-s) mixture is 0.016.

Region M of the copula support corresponds to modest changes in the yields. The Student-t copula and its mixtures with the Clayton family copulas decisively outperform (at significance levels of less than 0.01) all other competing specifications in this region of support, see Panel B in Table 4. For region U, associated with a common increase in the yields, the mixture of the Student-t and Clayton survival copulas (St-t/Cl-s) emerges as the marginally better option, see Panel C in Table 4. However, the Student-t copula mixture specifications (St-t/Cl and St-t/Cl-s) are statistically indistinguishable in the M and U regions. This supports the conjecture that the value added by mixing in the Clayton copula with the Student-t copula stems from its performance in region D. Thus, the addition of the Clayton copula allows one to better capture the dependence

Table 3
Daily yield changes: pair-wise tests of equal predictive accuracy of copulas for full copula support, based on the test developed by Diks et al. (2010).

Copula	St-t	Ga	Cl	Cl-s	Cl/Cl-s	Ga/Cl	Ga/Cl-s	St-t/Cl	St-t/Cl-s
St-t		−5.75	−24.99	−26.01	−22.40	−3.30	−3.56	3.30	3.21
Ga	5.75		−14.36	−14.92	−11.11	3.73	3.67	5.56	5.54
Cl	24.99	14.36		−0.67	16.91	30.36	30.07	25.94	25.92
Cl-s	26.01	14.92	0.67		18.36	31.93	31.89	27.06	27.04
Cl/Cl-s	22.40	11.11	−16.91	−18.36		28.12	27.92	23.52	23.50
Ga/Cl	3.30	−3.73	−30.36	−31.93	−28.12		−1.86	5.46	5.27
Ga/Cl-s	3.56	−3.67	−30.07	−31.89	−27.92	1.86		5.83	5.63
St-t/Cl	−3.30	−5.56	−25.94	−27.06	−23.52	−5.46	−5.83		−1.24
St-t/Cl-s	−3.21	−5.54	−25.92	−27.04	−23.50	−5.27	−5.63	1.24	
MCS order	7	4	2	1	3	6	5		8
MCS p-value	0.004	0.00	0.00	0.00	0.00	0.00	0.00		0.191

Note: Values of the Diks et al. (2010) test statistic. The test statistic is based on one-step ahead density forecasts for daily yield changes during the period from January 5, 1995, until May 21, 2014. The length of the rolling estimation window set equal to $R = 1000$ observations. Consequently, the number of forecasts is $P = 5055$. In each panel the (i,j)th entry is based on the score differences such that positive values of the test statistic indicate that the model in column j achieves a higher average score than the model in row i. Acronyms used for referring to copula specifications: Ga – Gaussian; St-t – Student-t; Cl – Clayton; Cl-s – Clayton survival; Cl/Cl-s – Clayton–Clayton survival mixture; Gaussian and Clayton mixture (Ga/Cl), Gaussian and Clayton survival mixture (Ga/Cl-s), Student-t and Clayton mixture (St-t/Cl), Student-t and Clayton survival mixture (St-t/Cl-s). MCS order is the iteration, at which the model is omitted from the MCS, while the MCS p-val is the corresponding p-value.

Table 4
Daily yield changes: pair-wise tests of equal predictive accuracy of copulas for selected regions of support, based on the censored likelihood scoring rule.

Copula	St-*t*	Ga	Cl	Cl-s	Cl/Cl-s	Ga/Cl	Ga/Cl-s	St-*t*/Cl	St-*t*/Cl-s
Panel A: region D									
St-*t*		−3.42	−5.61	−6.10	−5.27	−2.77	−3.43	1.80	0.91
Ga	3.42		−5.45	−6.10	−4.58	0.84	−1.47	3.49	3.46
Cl	5.61	5.45		−4.44	2.16	5.99	5.49	5.61	5.60
Cl-s	6.10	6.10	4.44		4.66	6.24	6.13	6.10	6.10
Cl/Cl-s	5.27	4.58	−2.16	−4.66		5.01	4.59	5.27	5.26
Ga/Cl	2.77	−0.84	−5.99	−6.24	−5.01		−1.08	2.82	2.79
Ga/Cl-s	3.43	1.47	−5.49	−6.13	−4.59	1.08		3.50	3.46
St-*t*/Cl	−1.80	−3.49	−5.61	−6.10	−5.27	−2.82	−3.50		−2.14
St-*t*/Cl-s	−0.91	−3.46	−5.60	−6.10	−5.26	−2.79	−3.46	2.14	
MCS order	7	5	2	1	3	6	4		8
MCS *p*-val	0.125	0.002	0.00	0.00	0.00	0.009	0.001		0.125
Panel B: region M									
St-*t*		−5.98	−8.67	−8.71	−8.23	−6.01	−6.01	0.52	0.68
Ga	5.98		−9.60	−9.64	−9.21	−5.23	−5.67	6.08	6.06
Cl	8.67	9.60		−0.31	9.21	9.63	9.63	8.72	8.71
Cl-s	8.71	9.64	0.31		9.25	9.67	9.67	8.76	8.75
Cl/Cl-s	8.23	9.21	−9.21	−9.25		9.24	9.23	8.29	8.28
Ga/Cl	6.01	5.23	−9.63	−9.67	−9.24		0.22	6.10	6.08
Ga/Cl-s	6.01	5.67	−9.63	−9.67	−9.23	−0.22		6.11	6.09
St-*t*/Cl	−0.52	−6.08	−8.72	−8.76	−8.29	−6.10	−6.11		0.17
St-*t*/Cl-s	−0.68	−6.06	−8.71	−8.75	−8.28	−6.08	−6.09	−0.17	
MCS order	7	6	2	1	3	5	4	8	
MCS *p*-val	0.737	0.00	0.00	0.00	0.00	0.00	0.00	0.879	
Panel C: region U									
St-*t*		−0.84	−5.02	−5.62	−5.64	−0.86	−0.87	1.96	2.10
Ga	0.84		−5.13	−6.10	−5.47	−0.96	−1.33	0.97	0.97
Cl	5.02	5.13		3.57	3.44	5.15	5.15	5.02	5.02
Cl-s	5.62	6.10	−3.57		0.04	6.15	6.14	5.63	5.63
Cl/Cl-s	5.64	5.47	−3.44	−0.04		5.49	5.49	5.63	5.63
Ga/Cl	0.86	0.96	−5.15	−6.15	−5.49		−3.05	0.99	1.00
Ga/Cl-s	0.87	1.33	−5.15	−6.14	−5.49	3.05		1.00	1.00
St-*t*/Cl	−1.96	−0.97	−5.02	−5.63	−5.63	−0.99	−1.00		0.43
St-*t*/Cl-s	−2.10	−0.97	−5.02	−5.63	−5.63	−1.00	−1.00	−0.43	
MCS order	7	6	3	1	2	5	4	8	
MCS *p*-val	0.349	0.349	0.00	0.00	0.00	0.349	0.349	0.661	

Note: Values of the Diebold–Mariano–West type test statistic $t_{R,P}$ defined in (1) based on the censored likelihood score (12) for the regions D, M and U with the threshold $r=0.25$. The test statistic is based on one-step ahead density forecasts for daily yield changes during the period from January 5, 1995, until May 21, 2014. The length of the rolling estimation window set equal to $R=1000$ observations. Consequently, the number of forecasts is $P=5055$. In each panel the (i,j)th entry is based on the score difference $d^{csl}_{t+1} = \mathcal{S}^{csl}_{j,t+1}(\boldsymbol{y}_{t+1}) - \mathcal{S}^{csl}_{i,t+1}(\boldsymbol{y}_{t+1})$ such that positive values of the test statistic indicate that the model in column j achieves a higher average score than the model in row i. Acronyms used for referring to copula specifications: Ga – Gaussian; St-*t* – Student-*t*; Cl – Clayton; Cl-s – Clayton survival; Cl/Cl-s – Clayton–Clayton survival mixture; Gaussian and Clayton mixture (Ga/Cl), Gaussian and Clayton survival mixture (Ga/Cl-s), Student-*t* and Clayton mixture (St-*t*/Cl), Student-*t* and Clayton survival mixture (St-*t*/Cl-s). MCS order is the iteration, at which the model is omitted from the MCS, while the 'MCS *p*-val' is the corresponding *p*-value.

structure between falling yields. In this respect, it is interesting to note that a common adoption of QE-like monetary policies by the central banks of the G7 countries contributed to the dependence structure between negative yield changes. Finally, the obtained results are sufficiently robust with respect to the choice of the cutoff value defining the regions. In particular, our conclusions about the chief sources of the outstanding forecasting performance of the Student-*t* and Clayton mixture copula (St-*t*/Cl) mostly hold.

We conclude this section with a brief description of the dynamics of the copula parameters for the case of the Student-*t* and Clayton mixture specification. The richness of the problem is illustrated in Fig. 3 with three groups of correlations between yield changes: the high correlation group (typified by the US–Canada and Germany–France pairs), the medium correlation group (exemplified by the US and Canada paired with European countries), and the low correlation group (composed of Japan paired with other G7 countries). This heterogeneity explains the inadequacy of a single parameter Clayton family copulas (and the three-parameter Cl/Cl-s mixture) for the task of capturing the dependence structure between yield changes. In the run-up to and during the 2007 financial crisis and the great recession, we observe increasing dependence between yield changes. As shown in Fig. 4, this is reflected in the higher probability of yield changes visiting the D region of support. Recall that the superior performance in this region, corresponding to jointly falling yields, is what ensured the leading role of the Student-*t* and Clayton mixture copula. Fig. 5 contains the dynamics of the mixture parameter of the Student-*t* and Clayton copulas. Note that the Clayton copula comes into play during the period of strong dependence immediately before and during the latest economic recession.

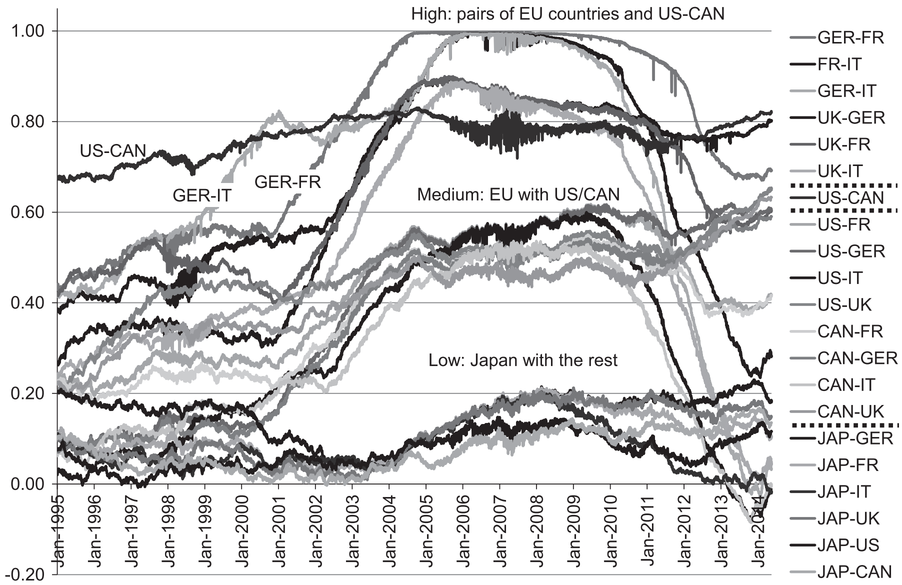

Fig. 3. Rolling window estimates of the correlation matrix of the Student-*t* component of the Student-*t* and Clayton mixture copula. The size of the rolling estimation window is $R = 1000$ observations. The legend labels are listed in descending order of the correlations on January 1, 2005, and partitioned into groups based on the strength of correlations. The correlations are reported for one day ahead of the end of estimation window, which ranges from January 5, 1995, until May 21, 2014.

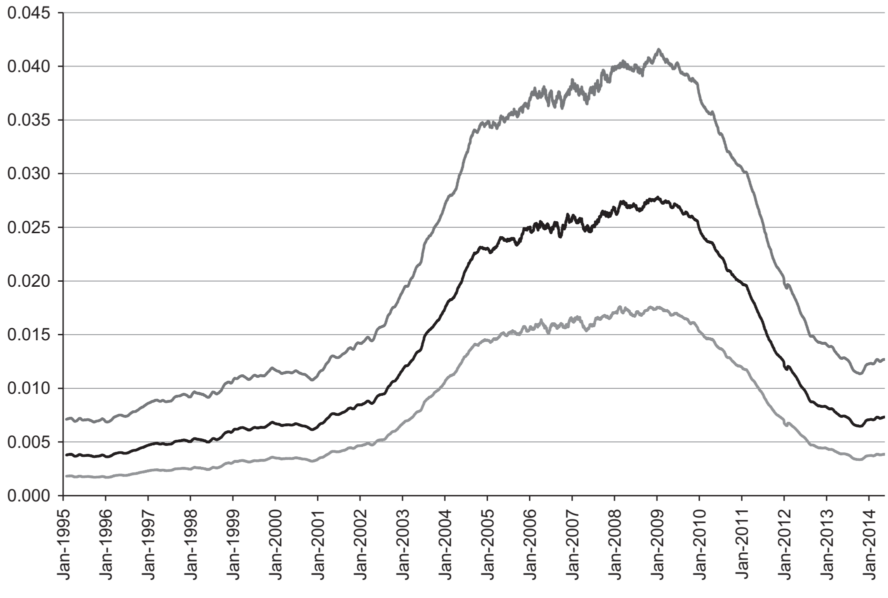

Fig. 4. Rolling window estimates of the one-day-ahead probability of the yield differences jointly visiting region *D* of the support. The probabilities are calculated in accordance with the Student-*t* and Clayton mixture copula specification and the SCOMDY model with a rolling estimation window of $R = 1000$ observations. The probabilities are smoothed with a 20-day moving average. The dark gray line is the probability of yield changes visiting region *D* defined by the 0.30 cutoff; the black line is the probability of yield changes visiting region *D* defined by the 0.25 cutoff; the light gray line is the probability of yield changes visiting region *D* defined by the 0.20 cutoff. The considered time period is from January 5, 1995, until May 21, 2014.

5. Summary and conclusions

Many practical applications involving joint density forecasts of multivariate asset returns focus on a particular part of the domain of support. Given that the dependence structure may vary, for example, with the sign and magnitude of changes, it becomes imperative to identify the best forecast method for the targeted part of the distribution. Copula modeling allows for a straightforward construction of a flexible multivariate distribution via their decomposition into the dependence structure, represented by a copula function, and marginal distributions. Thus, copulas represent an indispensable

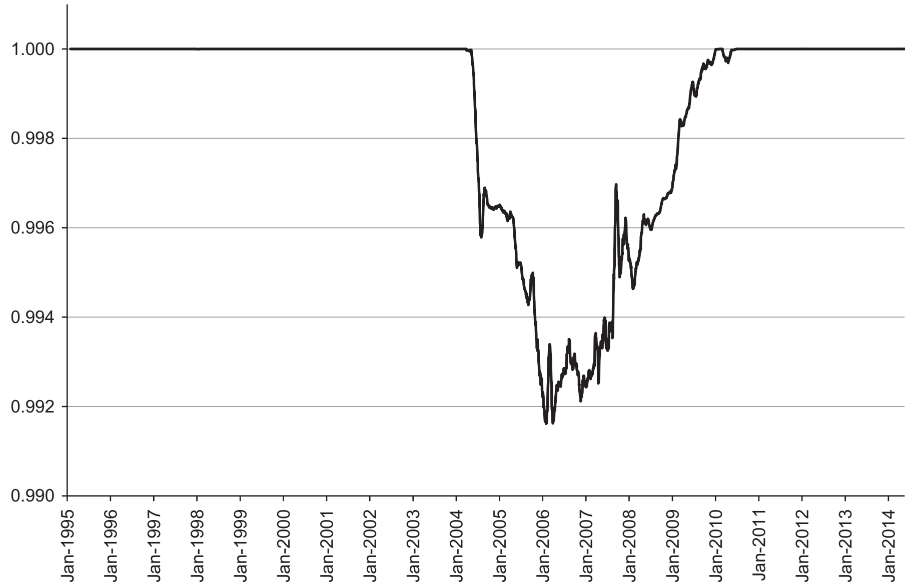

Fig. 5. Rolling window estimates of the mixture parameter, which take the value of 1 for a pure Student-t copula and 0 for a pure Clayton copula. The size of the rolling estimation window is $R=1000$ observations. The parameter is smoothed using a 20-day moving average. The parameter is reported for one day ahead of the end of estimation window, which ranges from January 5, 1995, until May 21, 2014.

instrument in the arsenal of modern risk management. However, the multitude of their specifications poses the challenge of selecting the best copula for forecasting risk under different types of market dynamics. This paper develops tests that take on this challenge and provide a reliable solution as indicated by a Monte Carlo study of their power and size.

In this paper we have developed Kullback–Leibler Information Criterion (KLIC) based tests of equal (out-of-sample) forecasting accuracy of different copula specifications in a selected region of the copula support. Proper scoring rules were obtained by making use of censored and conditional logarithmic scoring rules. To highlight the importance of the suggested tests, we have considered an application to the daily changes of yields on the 10 year benchmark government bonds of the G7 countries. We discovered the origins of the dominating forecasting performance of the Student-t and Clayton mixture specification across the full copula by considering the area of support corresponding to jointly falling yields. The negative tail dependence contributed by the Clayton copula when mixed with the t-copula is crucial to obtain high forecasting accuracy in periods of jointly falling yields.

Although throughout we have focused on one-step ahead forecasts, the methodology can easily be adjusted to the context of multi-step ahead forecasts. In the application we have compared a number of copulas as well as some of their mixtures. It might be possible to obtain better results by considering combinations of more than two copulas, but this is left for future research.

Acknowledgements

One of us (C.D.) acknowledges the support of the Netherlands Organisation for Scientific Research (NWO) under the Complexity Programme. V.P. acknowledges the support under the Australian Research Council's Discovery Projects funding scheme (Project no. DP120104014).

References

Amisano, G., Giacomini, R., 2007. Comparing density forecasts via weighted likelihood ratio tests. J. Bus. Econ. Stat. 25, 177–190.
Bao, Y., Lee, T.-H., Saltoğlu, B., 2004. A Test for Density Forecast Comparison with Applications to Risk Management. Working Paper 04-08, UC Riverside.
Bao, Y., Lee, T.-H., Saltoğlu, B., 2007. Comparing density forecast models. J. Forecast. 26, 203–225.
Berg, D., 2009. Copula goodness-of-fit testing: an overview and power comparison. Eur. J. Finance 15, 675–701.
Chen, X., Fan, Y., 2005. Pseudo-likelihood ratio tests for semiparametric multivariate copula model selection. Can. J. Stat. 33, 389–414.
Chen, X., Fan, Y., 2006. Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification. J. Econom. 135, 125–154.
Chib, S., Omori, Y., Asai, M., 2009. Multivariate stochastic volatility. In: Andersen, T.G., Davis, R.A., Kreiss, J.-P., Mikosch, T. (Eds.), Handbook of Financial Time Series, Springer-Verlag, Berlin, pp. 365–400.
den Haan Wouter, J., Levin Andrew, T., 1996. A Practitioner's Guide to Robust Covariance Matrix Estimation. NBER Technical Working Papers 0197. National Bureau of Economic Research Inc.
Diebold, F.X., Gunther, T.A., Tay, A.S., 1998. Evaluating density forecasts with applications to financial risk management. Int. Econ. Rev. 39, 863–883.

Diebold, F.X., Lopez, J.A., 1996. Forecast evaluation and combination. In: Maddala, G.S., Rao, C.R. (Eds.), Handbook of Statistics, vol. 14. , North-Holland, Amsterdam, pp. 241–268.
Diebold, F.X., Mariano, R.S., 1995. Comparing predictive accuracy. J. Bus. Econ. Stat. 13, 253–263.
Diks, C., Panchenko, V., van Dijk, D., 2010. Out-of-sample comparison of copula specifications in multivariate density forecasts. J. Econ. Dyn. Control 34, 1596–1609.
Diks, C., Panchenko, V., van Dijk, D., 2011. Likelihood-based scoring rules for comparing density forecasts in tails. J. Econom. 163, 215–230.
Genest, C., Gendron, M., Bourdeau-Brien, M., 2009. The advent of copulas in finance. Eur. J. Finance 15, 609–618.
Giacomini, R., White, H., 2006. Tests of conditional predictive ability. Econometrica 74, 1545–1578.
Gneiting, Tilmann, Raftery, A.E., 2007. Strictly proper scoring rules prediction and estimation. J. Am. Stat. Assoc. 102 (477), 359–378.
Granger, C.W.J., Pesaran, M.H., 2000. Economic and statistical measures of forecast accuracy. J. Forecast. 19, 537–560.
Hansen, P.R., Lunde, A., Nason, J.M., 2011. The model confidence set. Econometrica 79, 453–497.
Liebscher, E., 2008. Construction of asymmetric multivariate copulas. J. Multivar. Anal. 99 (10), 2234–2250.
Mitchell, J., Hall, S.G., 2005. Evaluating comparing and combining density forecasts using the KLIC with an application to the Bank of England and NIESR 'fan' charts of inflation. Oxf. Bull. Econ. Stat. 67, 995–1033.
Nelsen, R.B., 2006. An Introduction to Copulas, second ed. Springer Verlag, New York.
Patton, A.J., 2006. Modelling asymmetric exchange rate dependence. Int. Econ. Rev. 47, 527–556.
Patton, A.J., 2009. Copula-based models for financial time series. In: Andersen, T.G., Davis, R.A., Kreiss, J.-P., Mikosch, T. (Eds.), Handbook of Financial Time Series, Springer-Verlag, Berlin, pp. 767–786.
Patton, A.J., 2012. A review of copula models for economic time series. J. Multivar. Anal. 110, 4–18.
Patton, A.J., 2013. Copula methods for forecasting multivariate time series. In: Elliott, G., Timmermann, A. (Eds.), Handbook of Economic Forecasting, vol. 2. , Elsevier, Amsterdam, pp. 899–960.
Politis, D.N., Romano, J.P., 1994. The stationary bootstrap. J. Am. Stat. Assoc. 89, 1303–1313.
Rivers, D., Vuong, Q., 2002. Model selection tests for nonlinear dynamic models. Econom. J. 5 (1), 1–39.
Silvennoinen, A., Teräsvirta, T., 2009. Multivariate GARCH models. In: Andersen, T.G., Davis, R.A., Kreiss, J.-P., Mikosch, T. (Eds.), Handbook of Financial Time Series, Springer-Verlag, Berlin, pp. 201–229.
Sklar, A., 1959. Fonction de répartitions à n dimensions et leurs marges. Publications de l'Institut Statistique de l'Université de Paris 8, 229–231.
Vuong, Q., 1989. Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica 57, 307–333.
West, Kenneth D., 1996. Asymptotic inference about predictive ability. Econometrica 64, 1067–1084.