Journal of Banking and Finance 70 (2016) 1–22

Contents lists available at ScienceDirect

Journal of Banking and Finance

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

Extreme risk modeling: An EVT–pair-copulas approach for financial stress tests

Lyes Koliai*

Paris-Dauphine University (LEDa-SDFi), Place du Maréchal de Lattre de Tassigny, 75016 Paris, France

ARTICLE INFO

Article history:
Received 15 October 2014
Accepted 15 February 2016
Available online 8 March 2016

JEL classification:
C5
G11
G17

Keywords:
Financial time series
EVT
Pair-copulas
R-vine
Stress testing
Scenario

ABSTRACT

This paper presents a semi-parametric copula-GARCH risk model for financial return series with a stress testing perspective. The marginal distributions of the returns are specified using the Extreme Value Theory (EVT), putting a specific emphasis on extreme returns. The joint distribution is then built up using the pair-copulas theorem, based on the marginal distributions and the pair dependence structures. The model performance is assessed for three sets of assets, namely equity indices, exchange rates, and commodity prices. The empirical results support a better static and dynamic properties of the presented model compared to most common specifications used in practice. The proposed model and the alternative specifications are then carried out to perform stress testing exercises on hypothetical portfolios, where financial returns are considered as risk factors. The results show that the use of a wide range of risk models produce significantly different results, in terms of the corresponding stress scenario and in the corresponding impact on the portfolios. Hence, considering flexible and consistent specifications, as in the proposed model, allows ensuring a better credibility of the stress scenario and enhances the usefulness of the stress testing results.

1. Introduction

Until quite recently, predicting the impact of extreme events on financial portfolios and the design of related management actions have been led through a set of conventional tools, such as the Value at Risk (VaR). These tools are often based on simple hypotheses whose consistency has been seriously questioned after a recurrence of extreme events with losses exceeding all expectations. These failures are mainly due to the incoherence of some tools as risk measures and/or the unsuitability of some of their underlying hypotheses especially during turmoil periods (Alexander and Sheedy, 2008; Haldane, 2009). Such hypotheses concern the *risk model* assumed to assess the dynamics of risk factors and the *pricing model* used to assess their impact on the portfolio. In a stress test, which exactly focuses on the impact of extreme events on the portfolio, the relevance of both models is vital. The related specifications should then be chosen with caution.

This paper analyzes the relevance of the existing risk models for stress testing purposes to which we include a new specification that places a special emphasis on extreme events.[1] A risk model is made up of assumptions designed to define the statistical properties of the risk factors facing a variable of interest. For a financial portfolio, risk factors are usually given by the returns, the volatilities, and the dependence structure of the underlying financial assets. Technically speaking, the model specifies the marginal and the joint distribution functions of the risk factors. It is then fitted to some data and carried out to simulate a set of *scenarios* among which those supposed as *severe yet plausible* are used for stress testing purposes. Hence, the stress test results are strongly influenced by the prior choice of the risk model. The misleading results of stress tests conducted before and after the last financial crisis are often due to a misspecification of these models.[2] The considered scenarios have then turned to be harmless as

* Tel.: +33 613 322 416.
E-mail address: lyes.koliai@dauphine.fr

http://dx.doi.org/10.1016/j.jbankfin.2016.02.004

[1] The analysis of the risk model is of particular importance for both portfolio (or micro-) and systemic-based stress tests. The pricing model instead is more studied for the second category. It includes broad hypotheses related to endogenous phenomena such as risk transmission channels to the financial system, the impact of private and public response functions, contagion, second-round effects, feedback effects on the real economy, etc.

[2] Alongside with the use of incomplete pricing models and the unsuitability of the considered scenarios (Borio and Drehmann, 2009; Haldane, 2009; Breuer and Csiszár, 2013).

unexpected variations in risk factors and portfolio losses have been recorded.

Since the seminal works of Mandelbrot (1963a,b), the literature highlighted a set of stylized facts shared by most financial return series. For example, it is shown that the empirical distributions of financial returns are left-skewed and leptokurtic. The return series also exhibit serial dependences (weak autocorrelation, strong heteroscedasticity, volatility clusters) with asymmetry and leverage effects of past returns on present volatilities. In the multivariate case, extreme returns present strong, nonlinear, and dynamic dependences, especially for bear markets. To capture these stylized facts, various models have been proposed. However, due to practical issues (data, expertise, and time requirements, communication issues, etc.), only simple model specifications have been considered. The latter often lack of flexibility and omit one or more of the stylized facts, which explains the gaps between the expected and the observed losses.

The recurrent and increasing losses generated by these errors have motivated a crucial need for more reliable modeling frameworks. Recent advances in quantitative methods, such as econometrics and software programming, have favoured this kind of initiatives. In this paper, we propose a sequential risk model for financial return series that captures the individual and the joint stylized facts. We put a specific emphasis on extreme returns of most interest in stress tests. More specifically, we first specify the marginal distributions of returns within a semi-parametric approach, with an extreme value distribution in the tails and an empirical distribution in the interior. The joint distribution of the multivariate system is captured by an R-vine model based on the pair-copulas theorem associating the marginal distributions and bi-variate copulas. The model is evaluated with respect to common univariate and multivariate specifications used in practice. It is then carried out to perform univariate and multivariate stress tests in a dynamic framework.

The rest of the paper is presented as follows. Section 2 introduces financial stress tests and highlights the desired properties in the underlying risk model. Section 3 reviews the main approaches used in the literature to model financial returns. Section 4 presents the two stages of the sequential model. The data used for estimation, evaluation, and stress testing are presented in Section 5. Section 6 presents the estimation results for the marginal model, evaluates its performance with respect to alternative common specifications, and compares their properties in a stress testing exercise. The next section adopts a similar sequence for the multivariate model. Section 8 summarizes and concludes.

2. Financial stress tests

Financial portfolio stress tests consist in estimating the likely impact of *harmful yet plausible* events – or scenario – on a financial portfolio.[3] The scenario can be seen to as a possible realization of risk factors drawn by simulating the underlying risk model. The latter specifies the marginal and the joint distribution functions of the risk factors. For a financial portfolio, a scenario may then consist of variations in the returns, the volatilities or the correlations of financial assets. The impact of the scenario is assessed over a given horizon using an adapted pricing model. Financial stress tests have been designed to complement the conventional risk management framework mainly centered on the VaR. Compared to other pure statistical tools, they allow personifying all the events that make the scenario. By doing so, the portfolio exposures are explicitly identified and each potential outcome is associated to the generating scenario; hence guiding decision-making. Moreover, the scenario and the pricing model can be augmented to account for more realistic features, such as second-round and feedback effects, private investment strategies, public response functions, etc.

As for the VaR, stress tests need a risk model to draw the scenarios. Since, in most cases, stress scenarios consist of extreme variations of risk factors, the related risk model should also present strong in-sample and out-of-sample properties in the tails of the underlying marginal and joint distributions. This is a quantitative criterion required for all stress testing risk models. However, given the involvement of different parties in the exercise, this condition may be insufficient and the choice of the model trickier. Indeed, stakeholders' viewpoints are often influenced by internal and external considerations. Čihák (2007) and Haldane, 2009, among others, show that risk managers in a financial institution are generally pessimistic and often tend to overestimate the risks facing the different portfolios. They thus confer more plausibility to the most severe scenarios. That is, considering fat-tailed distributions. On the other hand, Management and the Board of Directors are more optimistic and deem implausible that kind of specifications. Any use in a stress test would therefore be unnecessary. The risk manager should then convince Management of the accuracy of the considered assumptions. This task is often critical because of the subjective character of certain assumptions and the practical implications resulting from the choice of the most severe scenarios. Having given his approval, Management shall be liable to take the necessary actions in response to the test results. Depending on the severity of the scenario, these actions may present significant opportunity costs for the institution, which may result in a negative impact on its profits. In such cases, only scenarios of moderate severity will be admitted. This aspect is one of the main criticisms levelled by researchers and market players against most tests carried out before and during the 2007–09 crisis (Haldane, 2009; Borio and Drehmann, 2009).

To deal with such an issue, the flexibility of the risk model is therefore of most importance in stress tests. Limiting subjective considerations in the underlying hypotheses eases the adoption of more relevant risk models. This also prevents from external criticisms which may put in question the credibility of the scenario and the utility of the results. A good trade-off between the performance of the model and its flexibility is therefore vital in a stress test risk model. This deeply influences our choice of the model specification we shall present in this paper.

The existing literature on financial stress tests can be split into four main topics. First papers, going back to the early 2000s, have been dedicated to a general presentation of the related conceptual aspects of stress tests at the time considered as a relatively new tool in financial risk management (Berkowitz, 2000; Blaschke et al., 2001; Čihák, 2007). A second body of the literature has focused on model-based scenarios for portfolio stress tests (Kupiec, 1998; Breuer and Krenn, 1999; Bee, 2001; Kim and Finger, 2000; Aragonés et al., 2001; Breuer et al., 2002; Alexander and Sheedy, 2008; McNeil and Smith, 2012; Breuer and Csiszár, 2013). Later on, due to the widespread repercussions of the 2007–2009 financial crisis, researches have been more concerned about systemic stress tests (Boss, 2008; Alessandri et al., 2009; Aikman et al., 2009; van den End, 2010; van den End, 2012; Engle et al., 2014; Acharya et al., 2014). A final wave of papers have made a first diagnostic of the realized exercises since the 2007–2009 financial crisis, by outlining their main limits and the remaining challenges

[3] To be considered as harmful a scenario should be designed to capture the most adverse events for the portfolio. The severity is often linked to the range of variations applied to the risk factors. The plausibility indicates the confidence level associated with the scenario. To be considered as such, the scenario should consist of likely variations in risk factors. When explicitly set, the plausibility is usually assessed by the probability of these variations. The analysis of the trade-off severity-plausibility and the selection of stress scenarios is beyond the scope of this paper. We refer the interested reader the works of Breuer et al. (2002), McNeil and Smith (2012) and Breuer and Csiszár (2013).

(Haldane, 2009; Borio and Drehmann, 2009; Hirtle et al., 2009; IMF, 2012; Greenlaw et al., 2012; Borio et al., 2012).

In the light of the recent wave of the literature, our paper relates to the second category and precisely to Alexander and Sheedy (2008) were the first to study the relevance of alternative risk models for stress testing purposes. They also presented a methodological framework to carry out dynamic stress scenarios in a univariate framework. The first objective of this paper is to extend this approach to the multivariate framework. The second and major objective is to introduce a stress testing risk model with a good performance-flexibility balance in order to enhance the adoption of better stress testing risk models by professionals. We present the model in Section 4 after a discussion of the existing literature related to financial returns model specifications. The relevance of these models for stress testing purposes is studied in Sections 6 and 7.

3. Financial extreme returns modeling

In the financial literature, it is commonly admitted that normal-type distributions are unsuited to model individual financial assets. There is still no agreement for the proposed alternatives whose performances vary with the assets and the data samples under review. This is particularly true when it comes to model extreme returns. Student *t* (Praetz, 1972; Blattberg and Gonedes, 1974) and skewed Student *t* distributions (Hansen, 1994; Fernández and Steel, 1998; Branco and Dey, 2001; Azzalini and Capitanio, 2003; Jones and Faddy, 2003) are part of the first alternatives and the most used in practice. However, both Student *t* distributions are *two-fat-tails* specifications, which reduces their ability to specify skewed distributions exhibited by some returns. This leads to an underestimation of the probability of extreme negative returns of particular interest in stress tests. To overcome this rigidity, Aas and Haff (2006) have proposed a skewed *t* distribution, with a polynomial (fat) form on one tail and an exponential (semi-fat) form on the other. But yet again, this specification have proven to be unsuited for less asymmetric distributions (Banachewicz and van der Vaart, 2008; Jondeau, 2010).

Generalized hyperbolic distributions family (Barndorff-Nielsen, 1977; Barndorff-Nielsen, 1978; Eberlein and Keller, 1995; Eberlein and von Hammerstein, 2002; Bibby and Sørensen, 2003) offers another flexible alternative to these specification issues by considering five parameters (location, scale, skewness, kurtosis, and tail shape). They also present interesting theoretical (closed under conditionality, marginalization and affine transformations, finite moments, etc.) and practical (fitting performance for several financial returns, generalizing most of the usual distributions) properties (Eberlein and von Hammerstein, 2002; Bibby and Sørensen, 2003). Moreover, the estimation issues that have accompanied their introduction have been gradually reduced thanks to progress in programming software (hardware allowing to efficiently estimate the most complex models, free packages using the last advances in econometrics, etc.). In some cases, however, identification issues may stem between the generalized distribution and its specific forms. This generates specification issues, especially for extreme returns.

Extreme Value Theory (EVT) offers a specific solution to these specification issues. This is done by focusing on the distribution tails which makes them a potential candidate for stress testing risk model. According to the Central Limit theorem (CLT), and under some regular conditions, random variables are proven to share common asymptotic behaviors irrespective of their underlying theoretical distributions. This flexibility is also a desired property for stress tests as discussed above. The EVT can be used to specify extreme returns beyond a high quantile or *extreme threshold*. Extreme value distributions such as the Generalized Pareto (GPD) or the Generalized Extreme Value (GEV) are commonly used for this purpose as they arise naturally in key central limit theorems in EVT (McNeil and Saladin, 1997). These are proven to perform better than most of the parametric distributions presented above (Longin, 2000; McNeil and Frey, 2000; McNeil et al., 2005; Zhao et al., 2011). Moreover, focusing on the specification of distribution tails prevents possible misspecifications of the whole distribution. The main challenge of this approach consists in the choice of the extreme threshold. A low quantile alters the asymptotic properties of the EVT while a too high quantile creates estimation bias due to insufficient observations. A trade-off accuracy-efficiency is then required. However, the presence of a large data sample may allow meeting both criteria. This is precisely why we consider the EVT approach to model the marginal distributions in a stress testing perspective. It may indeed combine both the accuracy and the flexibility criteria required in the related risk models.

As the EVT can only be applied to independent and identically distributed (iid) random variables. A prior filter for the return series is often necessary. Univariate ARMA-GARCH (Engle, 1982; Bollerslev, 1986) or stochastic volatility (SV) models are usually used in the literature to filter financial data and, by the way, also to capture the dynamic stylized facts of financial returns such as autocorrelation, heteroscedasticity, and volatility effects. The remaining distributional stylized facts (skewness and leptokurtosis) are captured by applying the EVT to the filtered residuals of these models.

A flexible specification of the joint distribution in the multivariate framework is given by copulas functions (Sklar, 1959). This consists in factorizing the joint distribution into marginal distributions and a multivariate distribution called *copula*.[4] The marginal distributions specify the individual returns while the copula captures the dependence structure of the multivariate system. Therefore, this approach allows separating the shapes of the dependence structure from those of the marginal distributions. It also extends the concept of linear correlation coefficients to that of nonlinear dependence structure function (Cherubini et al., 2004, see; McNeil et al., 2005; Jondeau et al., 2007; Patton, 2009).

While different specifications can be found in the literature for bivariate copulas (Joe, 1997; Nelsen, 2006), elliptical and Archimedean copulas are still the most used for the multivariate case (Dißmann et al., 2013). Unfortunately, these multivariate or *meta-copulas* are often unsuited to reproduce the empirical dependence structure of extreme returns in higher dimensions (high and asymmetric dependence), which is a serious handicap to run multivariate stress scenarios. Elliptical multivariate copulas are relatively complex in terms of parameterization, but inflexible when it comes to specify the dependence structure. Some Archimedean copulas allow specifying the asymmetry of the dependence structure. However, they are often based on few parameters (typically one or two) to specify all the pair dependences within the system (Joe, 1997; McNeil and Nešlehová, 2009; Brechmann and Czado, 2011). This is a major issue to the use of these specifications in stress testing purposes where the accuracy of modeling extreme dependences is of most importance. Some extensions have been presented by introducing parameter restrictions on the classical Archimedean copulas (Joe, 1993; Joe and Hu, 1996; Kotz and Nadarajah, 2004; Demarta and McNeil, 2005). But often, the provided extra flexibility is insufficient for oversized multivariate systems on one side, and generates estimation issues on the other (Kurowicka and Joe, 2011).

[4] An alternative approach consists in generalizing the analysis of univariate distributions to the multivariate case. Among the parameters of this distribution is the covariance matrix that gives the linear correlations in the multivariate system.

Joe (1996) proposed an interesting method to specify multivariate copulas while overcoming technical issues. This consists in factorizing the multivariate copula into conditional and unconditional bivariate copulas, the so-called *pair-copulas*. This technique, further explored by Bedford and Cooke (2001); Bedford and Cooke, 2002, offers more flexibility as it specifies separately each of the pair dependences in the multivariate system. It also simplifies the estimation procedure as it is based on bivariate distributions. Kurowicka and Cooke (2006) have presented a graphical representation, called *Regular-vine* or R-vine model, that summarizes all alternative factorization schemes for a multivariate or R-vine copula (Kurowicka and Joe, 2011, see also). Each scheme is given by a sequence of trees starting from the marginal distributions, ending by the R-vine copula, through the pair-copulas (see Fig. 1). Dißmann et al. (2013) provide a statistical inference for R-vine models.

The choice of the decomposition scheme is the major challenge of vine copulas, especially for oversized systems. Two particular schemes have been studied in the literature, namely the textitcanonical (C-) and the *drawable* (D-) vine models. Aas et al. (2009) proposed a sequential procedure allowing to perform simultaneously a selection/estimation of the decomposition scheme and the pair-copulas.[5] Brechmann and Czado (2011) applied this procedure to the general R-vine model, for a multivariate system of 52 financial variables including the Euro Stoxx 50 securities. However, for oversized systems, the procedure can be computationally intensive (Heinen and Valdesogo, 2009; Mendes et al., 2010; Czado, 2010; Joe et al., 2010; Kurowicka and Joe, 2011). To deal with, Brechmann et al. (2012) have proposed two parsimonious versions of the R-vine model, by restricting the pair-copulas to be used in the last trees of the model (using only independent and Gaussian copulas) or by simply truncating the last trees in the model. Their results show insignificant differences in terms of performance between the restricted and truncated versions on one side, and the original model on the other. These findings deserve to be confirmed by further studies.

The principle of pair-copulas is consistent with the flexibility required in stress testing risk models. Moreover, recent studies have shown that the extra-flexibility drawn by pair-copulas allows for a better specification of the dependence structure (Aas et al., 2009; Aas and Berg, 2009; Chollete et al., 2009; Fischer et al., 2009; Min and Czado, 2010; Min and Czado, 2011; Czado et al., 2011; Hua and Joe, 2011; Nikoloulopoulos et al., 2011). We then use the general R-vine model to specify the dependence structure of the multivariate system. Alongside with the marginal distributions we present a unified framework for financial returns within the *copula-GARCH approach* (Jondeau and Rockinger, 2006). To the best of our knowledge, this study is the first to associate the EVT with the R-vine model to carry out financial stress tests. This would allow more flexible modeling of extreme returns in the univariate and the multivariate cases, as we shall see in Sections 6 and 7. The suitability of this approach and the gains drawn by its use for stress testing exercises are particularly pointed out. Before doing so, the next section presents the theoretical framework of the two-stage model.

In the next section we present the two stages of the model. In Sections 6 and 7 we assess whether its extra-flexibility is accompanied by good in-sample and out-of-sample performances. We then study the implications of their use in stress testing exercises.

4. The model

This section presents the two stages of the sequential model used for the analysis of financial returns with a specific emphasis on extreme values. We first present the model for the marginal distributions then the dependence structure of the multivariate system, using the R-vine model based on the marginal distributions and adapted pair-copulas.

4.1. The marginal model

The application of the EVT requires a prior filter for the return series. ARMA-EGARCH models with a – temporary – normally distributed errors are used to capture the conditional mean and variance dynamics in each series.[6] The process $\{r_{i,t}\}_{t=1}^{T}$ of return $i = 1, \dots, N$ is defined by an $ARMA(p, q) - EGARCH(r, s)$ model, if

$$r_{i,t} = \mu_i + \sum_{j=1}^{p} \phi_{ij} r_{i,t-j} + \sum_{j=1}^{q} \theta_{ij} \varepsilon_{i,t-j} + \varepsilon_{i,t} \quad \varepsilon_{i,t} = z_{i,t}\sqrt{\sigma_{i,t}^2}$$

$$\ln(\sigma_{i,t}^2) = \omega_i + \sum_{j=1}^{s} \left(\alpha_{ij} z_{i,t-j} + \gamma_{ij}\left(|z_{i,t-j}| - \mathbb{E}|z_{i,t-j}|\right)\right) + \sum_{j=1}^{r} \beta_{ij} \ln(\sigma_{i,t-j}^2) z_{i,t} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1) \quad (1)$$

where σ_i is the conditional variance, μ_i is the long-run mean for return i;ϕ_{ij} and θ_{ij} capture, respectively, the *AR* and *MA* dynamics of the conditional mean; ω_i is the weighted long-run variance, β_{ij} is the GARCH dynamics of the conditional variance, α_{ij} and γ_{ij} capture, respectively, the asymmetry and the leverage effects of past returns on volatility, $\mathcal{N}(\cdot)$ is the standardized normal distribution of the filtered residuals $z_{i,t}$, and T is the size of the returns series.

The Generalized Pareto Distribution (GPD) is next used to accurately specify the extreme values of the standardized residuals $\hat{z}_{i,t}$ from the ARMA-EGARCH models. To do so, we use the Peaks over Threshold (POT) method according to which, beyond of a sufficiently high threshold, the distribution of excess returns (i.e. return minus threshold) follows a GPD. There exists different techniques to select the *optimal* extreme thresholds in an EVT-POT approach (Scarrott and MacDonald, 2012). In this paper, we use the 10% (left tail) and 90% (right tail) quantiles as extreme thresholds.[7] By doing so, we model the tails of the marginal distributions which holds for extreme negative and positive returns falling beyond the extreme thresholds.

To complete the specification of the whole marginal distribution, we model the returns falling between the extreme thresholds using the empirical cumulative distribution function (ecdf). This is performed by applying a linear interpolation to the corresponding returns. An alternative solution is to use a kernel density function with a predefined smoothing or *bandwidth* parameter. However, if the choice of the kernel density has a minor impact on the results, the choice of the smoothing parameter is proven to be crucial and can significantly impact the results.

The marginal distribution of each innovation is finally constructed through a semi-parametric approach, associating the GPD in the tails and the empirical distribution in the interior. The semi-

[5] For some data, the decomposition scheme is chosen among all alternative schemes of a C-vine or a D-vine model. The pair-copulas are selected from a predefined set of bivariate copulas.

[6] The ARMA model (Box and Jenkins, 1976) captures the autoregressive and the moving average components of the conditional mean in the return series. The GARCH model (Bollerslev, 1986) captures the conditional variance. ARCH-GARCH models are of comparable performance with stochastic volatility's most frequently used in continuous time modeling with high frequency data. The considered Exponential-GARCH (EGARCH) specification (Nelson, 1991) belongs to the *asymmetric volatility models*. It allows capturing the nonlinear dynamics in the conditional variance, based on a combination of ARCH and GARCH effects. ARCH effects are further split into an asymmetry (or sign) effect and a leverage (or amplitude) effect of past returns on actual volatilities. This provides an extra-flexibility of the EGARCH specification which presents comparable properties to other asymmetric versions of GARCH models.

[7] Results for the 5% and 95% quantiles are also provided for validation purposes.

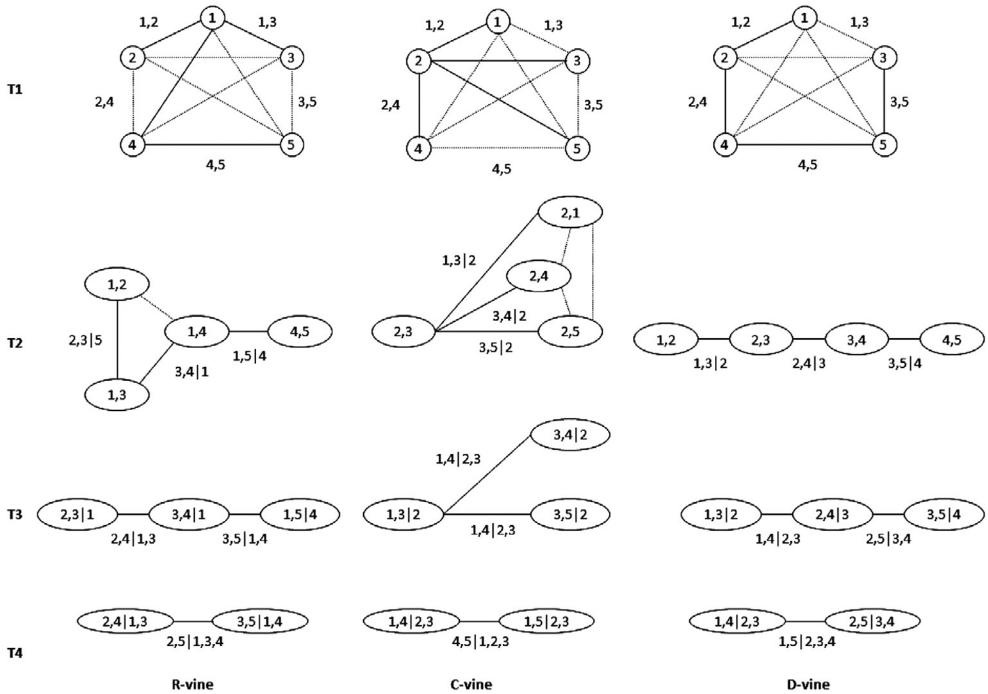

Fig. 1. An illustrative representation of vine copulas structures. This figure shows the structure of the R- (left panel), C- (middle panel), and D-vine (right panel) copulas for a five-dimensional system ($n = 5$). Each structure consists of $n-1$ spanning trees. The nodes in the first tree present the marginal distributions. Each edge corresponds to a bivariate copula density. In the following trees, the nodes become conditional bivariate densities while the edges are conditional bivariate (or pair-) copulas. The edges of tree $i = 1, \ldots, n-2$ determine the nodes of tree $i+1$, and so on up to setting the joint density in the last tree. The difference between the R-vine and the remaining two special cases consists in the structure of the trees, i.e. the adopted decomposition scheme for the joint density. In a D-vine, no node is allowed to connect to more than two edges, while, in a C-vine, each tree has a unique *key* node connected to all remaining nodes. In the R-vine instead, no particular restriction is considered, which confers to this specification a higher flexibility. Given these specifications, the selection procedure consists in retaining, at each stage, the spanning tree that maximizes the sum of absolute dependences.

parametric marginal distribution of return i is thus given as follows

$$F_i(\hat{z}_i) = \begin{cases} \frac{T_{u_i^L}}{T}\left(1+\xi_i^L \frac{u_i^L-\hat{z}_i}{\beta_i^L}\right) & \text{if} : \hat{z}_i < u_i^L \\ \varphi(\hat{z}_i) & \text{if} : u_i^U < \hat{z}_i < u_i^L, \ \ i = 1, \ldots, N \\ 1-\frac{T_{u_i^U}}{T}\left(1+\xi_i^U \frac{\hat{z}_i-u_i^U}{\beta_i^U}\right)^{-\frac{1}{\xi_i^U}} & \text{if} : \hat{z}_i > u_i^U \end{cases} \tag{2}$$

where u_i^L (resp. u_i^U) is the lower (resp. upper) threshold, $T_{u_i^L}$ (resp. $T_{u_i^U}$) is the number of exceedances on the left-side (resp. right-side) of the distribution, $\varphi(\cdot)$ is the empirical distribution function, ξ_i^L and β_i^L (resp. ξ_i^U and β_i^U) are, respectively, the shape and the scale parameters of the GPD on the lower (resp. upper) tail, and T is the size of the returns series.

4.2. The dependence structure model

We model the dependence structure of the multivariate system of returns using the R-vine model (Bedford and Cooke, 2001; Bedford and Cooke, 2002; Kurowicka and Cooke, 2006; Kurowicka and Joe, 2011; Dißmann et al., 2013). This supposes a factorization of the joint density of the multivariate system into marginal distributions on one side and an R-vine copula on the other. The latter specifies the conditional and the unconditional pair dependences as well as the joint dependence structures.

The joint density function $f(r_1, \ldots, r_N)$ of a multivariate system composed by N returns $r_1, \ldots, r_N$ can be factorized into successive conditional densities, as follows

$$f(r_1, \ldots, r_N) = f(r_1) \cdot f(r_2|r_1) \cdot f(r_3|r_1, r_2) \cdots f(r_N|r_1, \cdots, r_{N-1}) \tag{3}$$

Each but the first factor in the right-side product may, in turn, be decomposed into a bivariate density function or pair-copula and a conditional marginal density by the formula

$$f(r|v) = c_{r,v_j|v_{-j}}\big(F(r|v_{-j}), F(v_j|v_{-j})\big) \cdot (r|v-j) \tag{4}$$

where v is a v-dimensional vector (with $v = 1, \ldots, N-1$), v_j is a component of vector v, and v_{-j} is the vector v but v_j. $c_{r,v_j|v_{-j}}$ is the density function of the bivariate copula, and $F(\cdot)$ is the marginal conditional distribution function, given by

$$F(r|v) = \frac{\partial C_{r,v_j|v_{-j}}\big(F(r|v_{-j}), F(v_j|v_{-j})\big)}{\partial F(v_j|v_{-j})} \tag{5}$$

where $C_{r,v_j|v_{-j}}$ is the distribution function of the bivariate copula.

Depending on how the joint density has been factorized in Eq. (3), different decomposition schemes can be obtained. Kurowicka and Cooke (2006) have proposed a graphical structure, named Regular-vine or R-vine, representing the alternative schemes for a given multivariate density (see Fig. 1). The R-vine model is a sequence of trees $T_1, \ldots, T_{N-1}$. Each tree is given by nodes N_i, connected by edges E_i (for $i = 1, \ldots, N-1$). The nodes of the first tree correspond to the unconditional marginal densities. The related edges represent the unconditional bivariate copulas. For $i > 1$, the edges of a given tree become the nodes of the following one (i.e. $N_i = E_{i-1}$). Bedford and Cooke (2001) and Kurowicka and Cooke (2006) have shown that an edge $e \in E_i$ can be identified by only two nodes, namely conditioned nodes $j(e), k(e)$ and a conditioning set $D(e)$, such as $e = j(e), k(e)|D(e)$.

By associating to each edge $e = j(e), k(e)|D(e)$, the corresponding pair-copula $c_{e=j(e),k(e)|D(e)}$, Kurowicka and Cooke (2006) have shown that the density of the multivariate copula (or the R-vine

Table 1
Statistical features of the considered bivariate copulas. This table summarizes the key statistic features of the bivariate copulas we have used as candidates to model the pair dependence and tail dependence structures in the return series. Two elliptical copulas (Gaussian, Student t) and eight Archimedean copulas are considered. For the latter, and except for the Frank copula, we have also included the corresponding 90°, 180° (or survival), and 270° rotated versions. A total number of 31 specifications are fitted to the data. Those presenting the lowest AIC criterion are included in the R-vine trees at each stage of the sequential selection/estimation procedure.

Copula	Dependence		Tail dependence		Asymetry	Parameter space	
	Negative	Positive	Lower	Upper		1st parameter	2nd parameter
Gaussian	√	√				$(-1, 1)$	
Student t	√	√	√	√		$(-1, 1)$	$(2, \infty)$
Clayton		√	√		√	$(0, \infty)$	
Gumbel		√		√	√	$[1, \infty)$	
Frank	√	√				$\mathbb{R} - \{0\}$	
Joe		√		√	√	$(1, \infty)$	
BB1 (Clayton–Gumbel)		√	√	√	√	$(0, \infty)$	$[1, \infty)$
BB6 (Joe–Gumbel)		√		√	√	$[1, \infty)$	$[1, \infty)$
BB7 (Joe–Clayton)		√	√	√	√	$[1, \infty)$	$(0, \infty)$
BB8 (Joe–Frank)	√	√		√		$[1, \infty)$	$(0, 1]$

copula) can be written as follows

$$c(F_1(r_1), \ldots, F_N(r_N)) = \prod_{i=1}^{N-1} \prod_{e \in E_i} c_{e=j(e),k(e)|D(e)}\big(F_1(r_{j(e)}|r_{D(e)}), \ldots, F_N(r_{k(e)}|r_{D(e)})\big) \quad (6)$$

where $r_{D(e)}$ is a given subset of $D(e)$.

To estimate this model we consider the selection/estimation procedure proposed by Aas et al. (2009). This allows defining simultaneously the decomposition scheme and the adapted pair-copulas of the R-vine model. The method is based on a representation of the highest pair dependences on the first trees of the model. This allows for an explicit and a precise modeling of these dependences which are the most relevant to analyse. Each spanning tree is formed by a set of nodes connected by edges measuring the bivariate dependence of each pair of nodes. Taking into account the set of nodes given by the previous tree, the selection procedure consists in retaining, at each stage, the tree that maximizes the sum of absolute dependences. That is

$$\max \sum_{e=\{i,j\}} |\delta_{ij}|$$

where δ_{ij} is a measure of dependence and $1 \leq i < j \leq N$.

At each stage, after selecting the corresponding tree, the pair-copulas are introduced to estimate the conditional bivariate densities. Among a set of predefined bivariate copulas, those minimizing the AIC information criterion are selected. These copula functions form the nodes of the next step, according to which the corresponding spanning tree is selected, and so on. This iterative procedure (spanning tree – pair-copulas) ensures a better flexibility in the specification of the R-vine model and facilitates the related estimation. In this paper, we define the spanning tree using the Kendall's τ coefficient as a measure of dependence. We have then tested the suitability of thirty-one elliptical and Archimedean bivariate copulas exhibiting heterogeneous features in terms of dependence, extreme dependence structures, and asymmetry. Table 1 summarizes these specifications and their statistical properties (Brechmann and Czado, 2011). The estimations are performed using the `VineCopula` package in R, by Brechmann and Schepsmeier (2013). Introducing more flexibility through non-parametric copulas is explored in a forthcoming paper.

5. Data

To assess the performance of the presented model and its relevance for stress testing purpose, we consider three sets of daily time series: two sets of financial series and a set of nonfinancial series. The first class includes five equity indices, namely the CAC40 (France), the DAX30 (Germany), the FTSE100 (UK), the SP500 (USA), and the SPTSX (Canada). To account for the differences in the closure times between the European and the North American markets, we consider the closing prices for the latter two and the opening prices for the former three indices.[8] The series are collected daily from 2/9/1988 to 6/4/2015. The second set includes the nominal exchange rates of the six most traded currencies in the Forex spot market (BIS, 2013). The value of six currencies with respect to the U.S. dollar (USD) are considered: the euro (USD/EUR), the Japanese yen (USD/JPY), the British pound (USD/GBP), the Australian dollar (USD/AUD), the Canadian dollar (USD/CAD), and the Swiss franc (USD/CHF). We consider intermediate rates collected daily at 16:00 GMT on the London Forex market. The series cover the period 1/4/1999–6/4/2015. The third set includes the spot prices of five commodities, representing five distinct categories: the Brent (energy), gold (precious metal), copper (base metal), wheat (agricultural), and corn (agriculture-bioenergy). For each series, we consider daily closing prices - or intermediate prices for commodities traded OTC (gold and copper). The sample covers the period 7/5/1993–6/4/2015. All data are collected from Datastream. For all datasets, we have considered the maximum amplitude, given the data availability and synchronization issues related to the multivariate analysis. Last-observed-carried-forward (LOCF) method is considered for missing data. Table 2 summarizes the considered datasets.

We split each series into two sub-periods: the first one is used for the estimation procedure; the second, for forecasting and out-of-sample assessments. The first sub-period ends on 8/4/2011 for all datasets. The second sub-period consists of the last 1000 observations for each series.

To ensure the stationarity of our series, prices are transformed into log-differences. That is, considering stock returns (1st dataset), changes in exchange rates (2nd dataset), and changes in commodity prices (3rd dataset). To simplify the presentation, we use the term *return* regardless of the nature of the datasets. Table 3 presents the results of three stationarity tests. We have reported the statistics of the augmented Dickey–Fuller (ADF), the Phillips-Perron (PP), and the Zivot–Andrews (ZA) tests. The latter assumes, for each series, a presence of one possible structural break over the entire sample period. For all series, and regardless of the test, the null hypothesis of the presence of a unit root is rejected with a confidence level of 99%. The ZA test detects a structural break for all series. For stock returns, one can note the proximity of the dates on which these breaks have been detected. These occur in two particular periods shared by the series: early September 2000 (CAC40 and SPTSX) and March 2009 (FTSE100 and

[8] We are grateful to an anonymous referee for drawing our attention to perform the adjustment in this order.

Table 2
Data description. IPE: International Petroleum Exchange. LBM: London Bullion Market. LME: London Metal Exchange. CBOT: Chicago Board of Trade.

Dataset	Series	Denomination	Mnemonic	Market (currency)	Period	# of obs.
Equity index	CAC40	Fance CAC 40	FRCAC40	Paris (EUR)	7/9/1987–6/4/2015	6980
	DAX30	DAX 30 Performance	DAXINDX	Frankfurt (EUR)	7/9/1987–6/4/2015	6980
	FTSE100	FTSE 100	FTSE100	London (GBP)	7/9/1987–6/4/2015	6980
	SP500	S&P 500 Composite	S&PCOMP	New York (USD)	7/9/1987–6/4/2015	6980
	SPTSX	S&P/TSX Composite	TTOCOMP	Toronto (CAD)	7/9/1987–6/4/2015	6980
Exchange rate	USD/EUR	Euro to USD (WMR&DS)	EUDOLLR	London (EUR)	1/4/1999–6/4/2015	4284
	USD/JPY	Japanese Yen to USD (WMR)	JAPAYE$	London (JPY)	1/4/1999–6/4/2015	4284
	USD/GBP	UK Pound to USD (WMR)	UKDOLLR	London (GBP)	1/4/1999–6/4/2015	4284
	USD/CAD	Canadian dollar to USD (WMR)	CNDOLL$	London (CAD)	1/4/1999–6/4/2015	4284
	USD/AUD	Australian dollar to USD (WMR&DS)	AUSTDO$	London (AUD)	1/4/1999–6/4/2015	4284
	USD/CHF	Swiss franc to USD (WMR)	SWISSF$	London (CHF)	1/4/1999–6/4/2015	4284
Commodity	Brent	Crude Oil-Brent Cur. Month FOB	OILBREN	London–IPE (USD/Barrel)	7/5/1993–6/4/2015	5719
	Gold	Gold Bullion LBM	GOLDBLN	London–LBM (USD/Troy Ounce)	7/5/1993–6/4/2015	5719
	Copper	LME-Copper Grade A Cash	LCPCASH	London–LME (USD/Metric Ton)	7/5/1993–6/4/2015	5719
	Wheat	Wheat No. 2 Soft Red	WHEATSF	Chicago–CBOT (Cents/Bushel)	7/5/1993–6/4/2015	5719
	Corn	Corn No. 2 Yellow	CORNUS2	Chicago–CBOT (Cents/Bushel)	7/5/1993–6/4/2015	5719

Table 3
Stationarity tests. This table reports the statistics of the unit root tests for the daily return series of stock indices, exchange rates, and commodity prices. Three tests are considered: the augmented Dickey–Fuller (ADF), the Phillips–Perron (PP) test, and the Zivot–Andrews (ZA) test accounting for one possible structural break in the series. The ordering and the date of the break are shown in the last two columns of the table. For the three tests, the symbols ** denotes a rejection of the null hypothesis of unit root at 1% confidence level.

	ZA					
	Obs.	ADF	PP	Statistic	Break point	Date
CAC40	6979	-18.95**	−7220.7**	−92.05**	3132	09/04/2000
DAX30	6979	−18.93**	−6639.4**	−82.80**	3003	03/06/2000
FTSE100	6979	−20.07**	−6899.6**	−88.71**	5348	03/02/2009
SP500	6979	−19.09**	−6873.6**	−88.86**	5352	03/06/2009
SPSTX	6979	−18.65**	−6583.8**	−81.67**	3131	09/01/2000
USD/EUR	4283	−15.15**	−4228.3**	−64.38**	472	10/25/2000
USD/JPY	4283	−16.36**	−4262.6**	−66.97**	3573	09/13/2012
USD/GBP	4283	−15.01**	−3971.4**	−63.07**	2308	11/08/2007
USD/AUD	4283	−15.50**	−4193.0**	−66.09**	709	09/21/2001
USD/CAD	4283	−16.40**	−4089.9**	−65.28**	2307	11/07/2007
USD/CHF	4283	−16.49**	−4089.0**	−63.90**	3287	08/09/2011
Brent	5718	−16.27**	−5754.6**	−75.83**	1418	12/10/1998
Gold	5718	−18.51**	−5745.4**	−75.76**	4740	09/05/2011
Copper	5718	−15.95**	−5903.5**	−78.43**	2177	12/05/2001
Wheat	5718	−18.75**	−5927.5**	−81.12**	12	07/21/1993
Corn	5718	−16.69**	−5826.1**	−75.06**	4991	08/21/2012

SP500) corresponding the early 2000s and fall 2009 market shock and financial crisis, respectively. For exchange rates and commodities, no evident link between the structural breaks can be noted. This may augur a weak dependence among the return series of these two datasets (see Appendix A).

6. Lessons from the marginal model

6.1. Empirical results

The considered lags for the ARMA-EGARCH models are those maximizing the AIC Akaike (1973) among all combinations up to maximum lags $p = q = r = s = 3$. Table 4 reports the estimation results and the related diagnostic tests. For all return series, most parameters are significant at 99% confidence level. Overall, the parameter estimates confirm the descriptive statistics presented in Appendix A, which justifies the considered ARMA-EGARCH models. All unconditional means are quasi-null. The heteroscedasticity is confirmed by the presence of ARCH and GARCH effects in all series. As shown in Fig. 2, the autocorrelation of returns is present for lower lags. This is given by significant AR and MA parameters. The specifications with the highest lags are even the best ones. Even though better specifications can still be found by relaxing the considered constraint of maximum lags in the ARMA processes, we have maintained our choice for two main reasons: a concern of parsimony, and because the considered specifications are proven to be sufficient to filter most of the return series which is our main concern here. This latter aspect is shown by the last three columns of the table. Overall, the Ljung-Box (LB) test rejects the null hypotheses of both autocorrelation and heteroscedasticity (autocorrelation of squared returns) in the standardized residuals. The ARCH test (Engle, 1983) rejects the null hypothesis of the presence of ARCH effects. These results are valid with confidence levels ranging from 10% to 100% and for lag orders from 1 to 50. Due to space limitation, however, only results for 20 lags are reported.

The estimation results of the marginal semi-parametric distributions are reported in Table 5. The values of the extreme negative (u^L) and positive (u^U) thresholds as well as those of the lower-tail (β^L) and the upper-tail (β^U) scale parameters of the GPD are quite comparable between the three datasets. However, the shape parameter varies substantially. GPD distributions of stock returns have a positive shape parameter on the lower tail (ξ^L) (i.e. a Pareto-type GPD). This value is always higher than the corresponding value in the upper tail (ξ^U), which is negative in three cases out of five (i.e. a GPD close to exponential or Pareto type II distribution). This result confirms the leptokurtic and skewed shape of all marginal distributions of stock returns. This is confirmed by the values of the scale parameter which has, for all series, a higher value on the lower tail, meaning a wider range of extreme negative returns. In three cases out of six, the exchange rate returns

Table 4
Estimation and diagnostic tests of the ARMA-EGARCH models. This table presents the estimation results of the ARMA-EGARCH models for the daily return series of stock indices, exchange rates, and commodity prices. The table reports: the estimates of the ARMA (columns 4–8) and EGARCH (columns 9–15) models. The corresponding robust standard errors are given in brackets. LLH is the estimation log-likelihood. LB(p) and ARCH(p) are the p-values of the Ljung-Box test for autocorrelation and the Engle (1983) test for ARCH effects in the models' residuals, respectively, for p lags.

	Obs.	Model	$\hat{\mu}$	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$	LLH	LB(20)		ARCH(20)
																$\hat{z}$	$\hat{z}^2$	$\hat{z}$
CAC40	5979	ARMA(3,3)-EGARCH(3,2)[a]	0.00	0.07	0.49	−0.13	0.49	−0.25	−0.09		0.16		0.97		17711.25	0.61	0.22	0.23
			(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.02)		(0.01)		(0.00)					
DAX30	5979	ARMA(3,2)-EGARCH(1,2)[b]	0.00	0.04	−0.99	−0.02	0.98	−0.17	−0.08		0.14		1.00	−0.12	18174.52	0.69	0.98	0.98
			(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.02)	(0.01)		(0.02)		(0.00)	(0.00)				
FTSE100	5979	ARMA(2,3)-EGARCH(3,2)[c]	0.00	−0.01	−0.99	−0.05	0.99	−0.15	−0.09		0.15		0.88	0.11	19198.56	0.64	0.35	0.49
			(0.00)	(0.04)	(0.00)	(0.01)	(0.00)	(0.09)	(0.01)		(0.02)		(0.00)	(0.01)				
SP500	5979	ARMA(2,3)-EGARCH(2,1)[d]	0.00	0.51	−0.84	−0.53	0.85	−0.14	−0.21	0.14	−0.08	0.19	0.98		19669.01	0.36	1.00	0.99
			(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)	(0.02)	(0.01)	(0.02)	(0.00)					
SPTSX	5979	ARMA(3,3)-EGARCH(2,1)[e]	0.00	−0.17	−0.99	0.27	1.01	−0.09	−0.12	0.07	0.08	0.05	1.00	−0.09	20621.41	0.52	0.89	0.92
			(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.03)	(0.02)	(0.02)	(0.03)	(0.03)	(0.00)	(0.01)				
USD/EUR	3283	ARMA(3,3)-EGARCH(1,2)[f]	−0.00	−0.01	−0.00			−0.03	0.04	−0.03	−0.11	0.17	0.99	0.24	12095.58	0.85	0.61	0.62
			(0.00)	(0.02)	(0.02)			(0.00)	(0.03)	(0.03)	(0.01)	(0.01)	(0.00)	(0.00)				
USD/JPY	3283	ARMA(2,3)-EGARCH(1,3)	−0.00					−0.21	−0.04		0.10		0.89	0.09	11901.89	0.42	0.72	0.70
			(0.00)					(0.05)	(0.01)		(0.00)		(0.01)	(0.02)				
USD/GBP	3283	ARMA(3,3)-EGARCH(1,3)[g]	−0.00	−0.46	0.41	0.47	−0.41	−0.08	0.02		0.09		1.00	0.00	12471.16	0.63	0.92	0.93
			(0.00)	(0.01)	(0.01)	(0.00)	(0.00)	(0.04)	(0.01)		(0.02)		(0.00)	(0.03)				
USD/AUD	3283	ARMA(3,3)-EGARCH(3,1)[h]	−0.00	−0.16	−0.44	0.18	0.42	−0.10	0.18	−0.05	0.01	0.05	0.99		11541.36	0.90	0.01	0.01
			(0.00)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.03)	(0.04)	(0.03)	(0.03)	(0.00)					
USD/CAD	3283	ARMA(3,2)-EGARCH(3,1)[i]	−0.00					−0.02	0.11	−0.11	−0.00	0.04	0.99	0.64	12670.55	0.93	0.02	0.03
			(0.00)					(0.00)	(0.04)	(0.04)	(0.00)	(0.00)	(0.00)	(0.00)				
USD/CHF	3283	ARMA(2,3)-EGARCH(3,3)	-0.00					−0.06	−0.03	0.02	−0.05	0.12	0.99	−0.01	11857.53	0.68	0.38	0.31
			(0.00)					(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.00)	(0.00)				
Brent	4718	ARMA(3,2)-EGARCH(1,3)[j]	0.00	−0.08	−0.67	0.08	0.66	−0.39	−0.03		0.12		0.72	0.27	11543.15	0.59	0.67	0.72
			(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.04)	(0.01)		(0.02)		(0.00)	(0.00)				
Gold	4718	ARMA(2,3)-EGARCH(3,2)[k]	0.00	−0.16	−0.20	0.16	0.20	−0.08	0.05		0.11		0.99		15895.58	0.12	0.00	0.00
			(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)		(0.01)		(0.00)					
Copper	4718	ARMA(3,3)-EGARCH(2,2)[l]	0.00	1.12	−0.97	−1.12	0.96	−0.10	80.07	0.04	0.05	0.13	0.99		13155.76	0.54	0.58	0.62
			(0.00)	(0.01)	(0.00)	(0.02)	(0.00)	(0.00)	(0.03)	(0.04)	(0.04)	(0.07)	(0.00)					
Wheat	4718	ARMA(2,3)-EGARCH(1,1)[m]	0.00	0.37	0.45	−0.41	−0.44	−0.05	0.05	−0.08	0.18	0.02	0.88	0.11	11413.4	0.39	0.87	0.92
			(0.00)	(0.02)	(0.02)	(0.01)	(0.01)	(0.03)	(0.04)	(0.03)	(0.04)	(0.06)	(0.00)	(0.00)				
Corn	4718	ARMA(3,3)-EGARCH(2,2)[n]	0.00	−0.73	−0.90	0.74	0.90	−0.10	−0.10	0.09	0.13	−0.00	0.99	−0.01	12634.53	0.16	0.66	0.67
			(0.00)	(0.02)	(0.01)	(0.01)	(0.01)	(0.08)	(0.03)	(0.03)	(0.04)	(0.14)	(0.01)	(0.01)				

[a] $\hat{\theta}_3 = 0.02(0.02)$.
[b] $\hat{\theta}_3 = 0.03(0.01)$, $\hat{\beta}_3 = 0.09(0.00)$.
[c] $\hat{\alpha}_3 = -0.04(0.00)$, $\hat{\theta}_3 = -0.03(0.01)$.
[d] $\hat{\phi}_3 = 0.66(0.01)$, $\hat{\theta}_3 = -0.70(0.01)$.
[e] $\hat{\theta}_3 = 0.11(0.00)$, $\hat{\beta}_3 = 0.08(0.01)$.
[f] $\hat{\beta}_3 = -0.25(0.00)$.
[g] $\hat{\phi}_3 = 0.96(0.01)$, $\hat{\theta}_3 = -0.97(0.00)$, $\hat{\beta}_3 = -0.01(0.04)$.
[h] $\hat{\phi}_3 = 0.25(0.01)$, $\hat{\theta}_3 = -0.28(0.01)$, $\hat{\alpha}_3 = -0.11(0.03)$, $\hat{\gamma}_3 = 0.07(0.04)$.
[i] $\hat{\beta}_3 = -0.64(0.00)$.
[j] $\hat{\phi}_3 = 0.38(0.01)$, $\hat{\theta}_3 = -0.39(0.01)$.
[k] $\hat{\phi}_3 = -0.92(0.01)$, $\hat{\theta}_3 = 0.93(0.00)$.
[l] $\hat{\theta}_3 = 0.00(0.00)$, $\hat{\alpha}_3 = 0.03(0.03)$, $\hat{\gamma}_3 = -0.06(0.04)$.
[m] $\hat{\theta}_3 = 0.01(0.01)$, $\hat{\alpha}_3 = 0.02(0.03)$, $\hat{\gamma}_3 = -0.07(0.05)$.
[n] $\hat{\theta}_3 = 0.02(0.02)$.

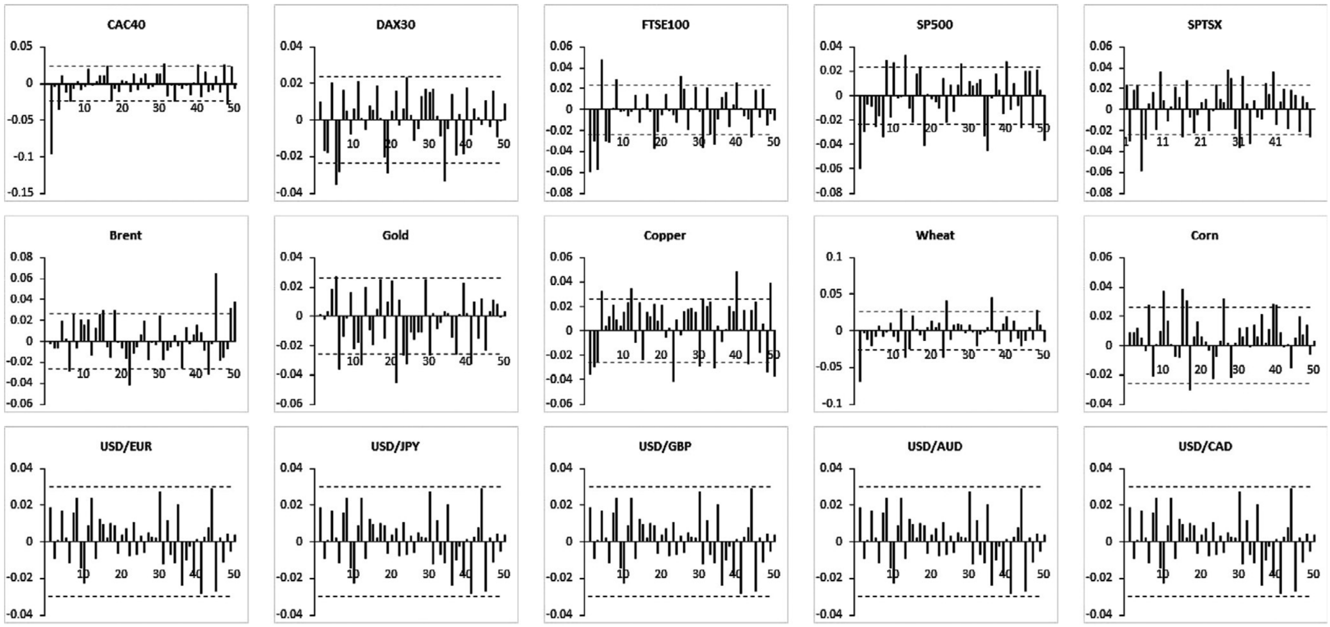

Fig. 2. Autocorrelation functions for the return series. This figure shows the autocorrelograms of the return series for lags 1–50. Horizontal dashed lines are the 5% and 95% confidence intervals. For the aestheticism of the graph, the sixth exchange rate series has been dropped.

present a negative shape parameter in both tails, even though, in most cases, this value is slightly greater in the lower tail. For the marginal distributions in this dataset, one can conclude to a non-fat lower tail and/or an unstable skewness. This remark also holds for commodities.

6.2. GoF tests

The suitability of the estimated semi-parametric marginal distributions to extreme returns is assessed through a battery of GoF tests. For each series, the test consists in: (i) obtaining a probability integral transformation (PIT) of the filtered residuals using the corresponding marginal distribution, and (ii) comparing the PIT distribution to a uniform distribution. The more the PIT distribution is close to the uniform distribution the more the marginal distribution used in the transformation is deemed close to the *true* distribution of the filtered residuals. Columns 11–13 of Table 5 report the confidence levels of three versions of the Kolmogorov–Smirnov (KS) GoF test, considering three alternative hypotheses: a PIT distribution which is different from (KS), above (KS+), or bellow (KS-) the uniform distribution. As we have done earlier for normality tests, we also considered the Anderson–Darling (AD) test, which is a version of the KS test with an overweighting of the distribution tails. The results do not allow rejecting the null hypothesis with confidence levels ranging from 74% to 100%, regardless of the test and the return series. This is even better when considering 5% and 95% as extreme thresholds in the GPD – with confidence levels from 85% to 100%. We can therefore conclude to a good accuracy of the considered semi-parametric specification to model the marginal distributions of financial returns.

To confirm this result, we have compared the performance of this specification to that of four alternative semi-parametric and five parametric distributions commonly used in practice. The KS test is used for this purpose. As semi-parametric specifications, we have considered marginal distributions with a GPD in the tails and kernel density in the bulk of the distribution. Normal, Epanechnikov, biweight (quartic), and triweight kernels are considered. The following parametric distributions are included: normal, Student *t*, skewed Student *t* (Fernández and Steel, 1998), generalized hyperbolic, and generalized hyperbolic skewed *t* (Aas and Haff, 2006). Columns 14–22 show a clear advantage of the semi-parametric approach over all alternative specifications, regardless of the return series. The normal and the generalized hyperbolic skewed Student *t* distributions are adequate anywhere. The performance of the remaining parametric distributions vary depending on the dataset. They are relatively well adapted to exchange rates whose empirical distributions are closer to normality. However, their performance drops significantly for stock and commodity returns whose empirical distributions have one or two fait tails. The performance of the alternative kernel-based semi-parametric specifications are rather similar. As for the parametric distributions, their performance is decreasing when the empirical distribution exhibits fat tails.

Finally, we carry out a backtesting procedure to check the dynamic properties of the semi-parametric distribution. Table 6 compares the results for four marginal specifications: the empirical distribution, the semi-parametric distribution, and the Normal and the Student *t* parametric specifications. VaR-95% and VaR-99% are considered, with corresponding theoretical number of exceedances 50 and 10, respectively. For each model, the table reports the number of exceedances (columns 3–6) and the corresponding *p*-values (columns 7–10) from the Weibull duration-based test of independence by Christoffersen and Pelletier, 2004. A long mono-asset portfolio is considered for each return. The models are re-estimated weekly using a rolling window of length equal to that of the first sub-period of each dataset. Again, the results show the suitability of the semi-parametric distributions presented in Section 4.1 to carry out stress testing scenarios given their flexibility and accuracy to model extreme returns. In this case, the empirical distribution seems to be too conservative, while the use of parametric distributions could be misleading with a number of exceedances clearly above the theoretical values.

6.3. Model-based univariate stress scenarios

We run three model-based univariate stress scenarios and measure their impact on three hypothetical portfolios. An equally weighted long portfolio is considered for each of our three

Table 5

Estimation and GoF tests for the marginal distributions. This table reports the parameter estimates of the marginal distributions, including skewness and kurtosis parameters for the filtered ARMA-EGARCH residuals (columns 3–4), values of extreme thresholds for quantiles 10–90% and 5–95% in square brackets (columns 5–6), and the shape and scale parameters of the GPD in the lower (columns 7 and 9) and upper (columns 8 and 10) tails. The adequacy of the estimated distributions is evaluated through goodness of fit (GoF) tests applied to the probability integration transformation (PIT) of the filtered residuals. The *p*-values of three versions of the Kolmogorov–Smirnov (KS) test are considered, with as alternative hypotheses: a PIT distribution different to (KS), greater than (KS+) or less than (KS−) the uniform distribution. Moreover, the Anderson–Darling (AD) uniformity test is used to check the accuracy of the specification for extreme returns (i.e. in the distribution tails). The performance of the presented semi-parametric specification is compared to that of four semi-parametric kernel-based distributions (normal, Epanechnikov, biweight, and triweight) and five parametric distributions commonly used in practice (normal (Nor), Student *t* (St), asymmetric Student *t* (SSt) by Fernández and Steel (1998), generalized hyperbolic distribution (GH), and generalized hyperbolic asymmetric *t* (GSt) by Aas and Haff (2006)).

	Obs.	Sk.	Ku.	u^L	u^U	ξ^L	ξ^U	β^L	β^U	Semi-par. distribution				KS for alternative distributions								
										KS	KS+	KS-	AD	Nor	Epa	Bi	Tri	Nor	St	SSt	GH	GSt
CAC40	5979	−0.49	5.47	−1.22	1.20	0.07	−0.07	0.60	0.50	1.00	0.90	0.91	0.99	0.15	0.12	0.13	0.14	0.00	0.04	0.11	0.27	0.23
										[1.00]	[0.88]	[0.95]	[0.99]									
DAX30	5979	−0.63	7.59	−1.26	1.18	0.07	−0.02	0.58	0.48	1.00	0.84	0.74	0.99	0.15	0.12	0.13	0.14	0.00	0.00	0.00	0.01	0.01
										[1.00]	[0.85]	[0.93]	[0.99]									
FTSE100	5979	−0.14	4.83	−1.23	1.22	0.02	0.06	0.62	0.47	1.00	0.77	0.75	1.00	0.29	0.23	0.25	0.26	0.00	0.02	0.17	0.04	0.14
										[1.00]	[0.85]	[0.94]	[0.99]									
SP500	5979	−0.46	5.78	−1.22	1.23	0.06	−0.08	0.62	0.54	1.00	0.84	0.91	1.00	0.16	0.14	0.15	0.15	0.00	0.00	0.02	0.30	0.36
										[1.00]	[0.96]	[0.97]	[1.00]									
SPTSX	5979	−0.52	5.73	−1.24	1.18	0.07	−0.08	0.60	0.53	1.00	0.78	0.90	0.98	0.31	0.26	0.28	0.28	0.00	0.05	0.01	0.05	0.08
										[1.00]	[0.94]	[0.96]	[0.97]									
USD/EUR	3283	−0.13	3.82	−1.26	1.25	−0.01	−0.16	0.56	0.61	1.00	0.92	0.88	0.99	0.56	0.54	0.55	0.55	0.00	0.25	0.42	0.56	0.00
										[1.00]	[0.90]	[0.90]	[0.99]									
USD/JPY	3283	−0.29	4.98	−1.22	1.19	0.08	0.01	0.58	0.53	1.00	0.92	0.93	1.00	0.62	0.62	0.62	0.62	0.00	0.79	0.73	0.79	0.00
										[1.00]	[0.97]	[0.94]	[0.99]									
USD/GBP	3283	−0.03	3.76	−1.23	1.30	0.01	−0.13	0.52	0.59	1.00	0.90	0.94	1.00	0.61	0.61	0.61	0.61	0.02	0.34	0.43	0.58	0.00
										[1.00]	[0.97]	[0.98]	[1.00]									
USD/AUD	3283	0.36	4.41	−1.22	1.24	−0.05	0.03	0.49	0.61	1.00	0.83	0.94	1.00	0.72	0.72	0.73	0.72	0.00	0.57	0.82	0.92	0.00
										[1.00]	[0.94]	[0.95]	[1.00]									
USD/CAD	3283	0.01	3.49	−1.28	1.24	−0.05	−0.08	0.52	0.58	1.00	0.95	0.84	1.00	0.70	0.69	0.70	0.70	0.18	0.82	0.83	0.83	0.54
										[1.00]	[0.97]	[0.93]	[1.00]									
USD/CHF	3283	−0.25	4.04	−1.28	1.24	−0.02	−0.06	0.61	0.51	1.00	0.86	0.93	1.00	0.74	0.75	0.75	0.74	0.00	0.21	0.62	0.88	0.00
										[1.00]	[0.96]	[0.95]	[0.99]									
Brent	4718	−0.16	4.40	−1.24	1.21	−0.02	0.07	0.63	0.48	.00	0.89	0.94	0.98	0.09	0.08	0.09	0.09	0.00	0.05	0.06	0.09	0.00
										[1.00]	[0.95]	[0.97]	[0.98]									
Gold	4718	−0.14	7.38	−1.17	1.14	0.02	0.11	0.65	0.58	1.00	0.96	0.92	0.99	0.02	0.01	0.02	0.02	0.00	0.00	0.00	0.00	0.00
										[1.00]	[0.92]	[0.92]	[0.98]									
Copper	4718	−0.33	6.03	−1.17	1.20	0.15	−0.02	0.55	0.57	1.00	0.92	0.94	0.99	0.37	0.31	0.33	0.33	0.00	0.43	0.49	0.42	0.40
										[1.00]	[0.97]	[0.96]	[0.99]									
Wheat	4718	0.19	5.85	−1.14	1.19	0.04	0.07	0.60	0.58	1.00	0.95	0.87	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
										[1.00]	[0.93]	[0.93]	[0.98]									
Corn	4718	−0.19	5.30	−1.19	1.24	0.09	0.05	0.57	0.52	1.00	0.93	0.94	0.97	0.01	0.01	0.01	0.01	0.00	0.00	0.02	0.00	0.00
										[1.00]	[0.98]	[0.96]	[0.98]									

Table 6
Assessment of the dynamic properties of the marginal distributions (baktesting). This table compares the results of the backtesting procedure for four marginal specifications, namely: (i) the empirical distribution, (ii) the semi-parametric distribution, and (iii) two parametric specifications (Normal and Student *t*). For each model and for two VaR's confidence levels (column 2), the table reports the number of exceedances (columns 3–6) and the corresponding *p*-values (columns 7–10) from the Weibull duration-based test of independence by Christoffersen and Pelletier (2004). VaR-95% and VaR-99% are considered, with corresponding theoretical number of exceedances 50 and 10, respectively. The models are re-estimated weekly using a rolling window of 1000 observations. A long mono-asset portfolio is considered for each return.

	VaR (%)	Exceedances				Weibull statistic			
		Empirical	S-parametric	Normal	Student *t*	Empirical	S-parametric	Normal	Student *t*
CAC40	95	35	55	69	66	0.447	0.165	0.592	0.461
	99	8	11	14	13	0.340	0.218	0.431	0.409
DAX30	95	25	53	52	54	0.698	0.123	0.100	0.164
	99	7	10	16	11	0.409	0.148	0.515	0.218
FTSE100	95	35	56	81	60	0.447	0.199	0.859	0.357
	99	6	12	16	18	0.431	0.340	0.515	0.688
SP500	95	32	55	88	62	0.555	0.165	0.966	0.412
	99	7	12	14	15	0.409	0.340	0.431	0.451
SPTSX	95	42	51	71	60	0.245	0.091	0.617	0.357
	99	8	10	14	13	0.340	0.148	0.431	0.409
USD/EUR	95	26	50	72	55	0.666	0.065	0.636	0.165
	99	9	11	16	11	0.218	0.218	0.515	0.218
USD/JPY	95	35	55	70	57	0.447	0.165	0.600	0.240
	99	8	13	18	16	0.340	0.409	0.688	0.515
USD/GBP	95	33	51	70	60	0.480	0.091	0.600	0.357
	99	6	10	15	12	0.431	0.148	0.451	0.340
USD/AUD	95	31	53	82	73	0.592	0.123	0.218	0.637
	99	8	11	17	14	0.340	0.218	0.578	0.431
USD/CAD	95	37	54	69	66	0.435	0.164	0.592	0.461
	99	8	12	17	17	0.340	0.340	0.578	0.578
USD/CHF	95	33	52	64	57	0.480	0.100	0.441	0.240
	99	5	10	12	13	0.451	0.148	0.340	0.409
Brent	95	33	52	73	68	0.480	0.100	0.637	0.555
	99	7	10	16	12	0.409	0.148	0.515	0.340
Gold	95	34	54	70	66	0.461	0.164	0.600	0.461
	99	8	12	17	13	0.340	0.340	0.578	0.409
Copper	95	32	50	68	53	0.555	0.065	0.555	0.123
	99	5	9	14	11	0.451	0.218	0.431	0.218
Wheat	95	31	51	64	66	0.592	0.091	0.441	0.461
	99	6	11	15	17	0.431	0.218	0.451	0.578
Corn	95	35	53	78	59	0.447	0.123	0.812	0.246
	99	7	13	20	16	0.409	0.409	0.959	0.515

Table 7
The impact of stress scenarios in terms of portfolio' return loss (static analysis). This table shows the return losses (in percentage) for three linear portfolios formed by long equally weighted equity indices, exchange rates, and commodity prices, respectively. These losses are computed for three shocks of probabilities 0.01%, 0.02%, and 0.05%, as given by the marginal distributions of the underlying returns. Four alternative marginal distributions are used to define the size of the shock corresponding to each return series and probability.

	Marginal distribution				
	Probability (%)	Empirical	Semi-parametric	Normal	Student *t*
Equity indices	0.01	−11.56	−7.01	−3.75	−5.81
	0.02	−8.09	−6.13	−3.55	−5.19
	0.05	−5.33	−5.20	−3.27	−4.46
Exchange rates	0.01	−5.82	−5.23	−3.55	−5.41
	0.02	−5.04	−4.75	−3.32	−4.74
	0.05	−4.44	−4.18	−3.20	−4.22
Commodity prices	0.01	−15.33	−6.22	−3.68	−7.11
	0.02	−14.41	−5.55	−3.46	−6.17
	0.05	−10.82	−4.89	−3.27	−5.12

datasets. The portfolio's return is given as a linear function of the individual returns, assumed as risk factors. The specification of the initial shock on the risk factors and the related probability is based on the marginal distributions of the return series. The evolution of the shock over an arbitrary stress horizon is given, for each return series, by the marginal model presented in Section 4.1. The result of the test is assessed in terms of losses in the portfolio and in terms of the regulatory capital required to cover these losses. However, to avoid repetitions with the multivariate framework to be studied in the next section, we focus here on static stress scenarios.

Three scenarios are set. In each scenario, the initial shock is given by a joint extreme variation in all returns of a same dataset. We have considered negative variations given by quantiles $0.0005, 0.0002$, and 0.0001 in the marginal distributions.[9] These quantiles are, respectively, equivalent to 0.05%, 0.02%, and 0.01% percentiles. This allows considering extreme risks not captured by the conventional VaR where the coverage level is usually set at 5% or 1%. For each portfolio, Table 7 reports the resulting returns. To show the importance of the risk model in the stress result, we have also reported the results obtained from alternative marginal models: the empirical distribution, the normal and the Student *t* distributions, and the semi-parametric distribution of Eq. (2). The results show a significant difference in the extent of the loss for the same quantile. The most severe changes are generated by the empirical distribution, followed by the semi-parametric

[9] For simplicity reasons, the shocks are obtained from the marginal distributions rather than from the joint distribution.

Table 8
Sequential estimation of the R-vine model. This table reports the estimation results of the dependence structures in the daily return series of stock indices, exchange rates, and commodity prices as specified by three R-vine models. The table reports: the tree number in the graphical model (column 1), the names of the conditional (columns 3–4) and conditioning (column 5) series, the name of the corresponding pair-copula (column 6), the corresponding parameter estimates (columns 7–8) and robust standard errors (in brackets), as well as the empirical and theoretical rank coefficients (columns 9–10) measuring the unconditional dependence on the first tree and the conditional dependence on the following trees. LLH denotes the log-likelihood of the sequential estimation of the pair-copulas (in light) and that of the multivariate system (in bold).

Tree	Obs.	Series 1	Series 2	Conditioning series	Pair-copula	1st parameter	2nd parameter	Empirical τ	Theoretical τ	LLH
1	5979	FTSE100	DAX30		Student t	0.58(0.01)	8.39(1.01)	0.39	0.39	1223.72
	5979	DAX30	CAC40		Student t	0.73(0.01)	4.68(0.35)	0.53	0.52	2359.38
	5979	CAC40	SP500		S-BB1	0.20(0.02)	1.54(0.02)	0.42	0.41	1433.05
	5979	SPTSX	SP500		S-BB1	0.23(0.03)	1.60(0.02)	0.44	0.44	1668.94
2	5979	FTSE100	CAC40	DAX30	Student t	0.22(0.01)	12.66(1.01)	0.14	0.14	176.67
	5979	DAX30	SP500	CAC40	Student t	0.19(0.01)	12.09(1.99)	0.13	0.12	133.38
	5979	CAC40	SPTSX	SP500	Student t	0.09(0.01)	20.70(5.09)	0.05	0.06	34.63
3	5979	FTSE100	SP500	CAC40, DAX30	Student t	0.20(0.01)	13.08(2.36)	0.13	0.13	14.19
	5979	DAX30	SPTSX	SP500, CAC40	Student t	0.06(0.01)	27.53(9.27)	0.04	0.04	13.40
4	5979	FTSE100	SPTSX	SP500, CAC40, DAX30	Student t	0.12(0.01)	29.99(−)	0.08	0.08	48.52
										7236.33
1	3283	USD/GBP	USD/EUR		S-BB1	0.22(0.03)	1.71(0.03)	0.49	0.47	1064.73
	3283	USD/CAD	USD/AUD		Student t	0.53(0.01)	8.38(1.38)	0.36	0.36	552.66
	3283	USD/AUD	USD/EUR		Student t	0.56(0.01)	6.62(0.88)	0.39	0.38	641.26
	3283	USD/EUR	USD/CHF		Student t	0.90(0.00)	2.23(0.15)	0.72	0.72	2858.13
	3283	USD/CHF	USD/JPY		Student t	0.43(0.02)	5.97(0.73)	0.28	0.28	357.89
2	3283	USD/GBP	USD/AUD	USD/EUR	Student t	0.18(0.02)	18.93(6.18)	0.12	0.12	62.15
	3283	USD/CAD	USD/EUR	USD/AUD	Student t	0.13(0.02)	11.59(2.51)	0.09	0.08	38.32
	3283	USD/AUD	USD/CHF	USD/EUR	Student t	−0.12(0.02)	15.43(3.93)	−0.07	−0.07	34.09
	3283	USD/EUR	USD/JPY	USD/CHF	Clayton-270	−0.14(0.02)		−0.07	−0.07	34.50
3	3283	USD/GBP	USD/CHF	USD/AUD, USD/EUR	Student t	0.04(0.02)	13.07(6.18)	0.04	0.03	13.22
	3283	USD/CAD	USD/CHF	USD/EUR, USD/AUD	Gaussian	−0.08(0.02)		−0.05	−0.05	9.89
	3283	USD/AUD	USD/JPY	USD/CHF, USD/EUR	Student t	0.05(0.02)	7.19(1.02)	0.04	0.03	33.27
4	3283	USD/GBP	USD/CAD	USD/CHF, USD/AUD, USD/EUR	Frank	0.19(0.11)		0.02	0.02	1.64
	3283	USD/CAD	USD/JPY	USD/CHF, USD/EUR, USD/AUD	BB8−90	−1.11(0.08)	−0.92(0.10)	−0.04	−0.04	6.82
5	3283	USD/GBP	USD/JPY	USD/CAD, USD/CHF, USD/AUD, USD/EUR	Student t	0.02(0.02)	19.55(6.16)	0.02	0.01	6.04
										5714.62
1	4718	Brent	Gold		Student t	0.16(0.01)	20.13(6.70)	0.10	0.10	62.62
	4718	Gold	Copper		S-Gumbel	1.15(0.01)		0.14	0.13	120.15
	4718	Copper	Corn		S-Gumbel	1.07(0.01)		0.07	0.06	30.81
	4718	Corn	Wheat		Student t	0.50(0.01)	9.11(1.32)	0.34	0.34	698.14
2	4718	Brent	Copper	Gold	Student t	0.12(0.01)	18.48(5.48)	0.08	0.08	43.22
	4718	Gold	Corn	Copper	S-Clayton	0.07(0.02)		0.04	0.03	9.13
	4718	Copper	Wheat	Corn	BB8	1.19(0.12)	0.77(0.17)	0.04	0.04	9.96
3	4718	Brent	Corn	Copper, Gold	Gaussian	0.06(0.01)		0.04	0.04	9.99
	4718	Gold	Wheat	Corn, Copper	Frank	0.32(0.09)		0.03	0.04	6.39
4	4718	Brent	Wheat	Corn, Copper, Gold,	Student t	0.04(0.02)	21.67(7.19)	0.02	0.02	8.30
										997.71

distribution, then the Student t and the normal distributions. Yet, as pointed out earlier, one of the main challenges in stress testing exercises is to choose among all alternative specifications, the distribution that determines the plausibility of the scenario - and thus the credibility of the stress test. However, in Table 5, we have seen the poor performance of most parametric specifications to capture the distribution of the return series. The use of these distributions in a stress testing exercise may then generate an overestimation in the probability of the scenario, i.e. considering relatively benign scenarios as harmful or, in the same spirit, discarding plausible most severe scenarios from the analysis. Stress tests based on this kind of distributions may then suffer a lack of efficiency, reflected by the misleading results observed in practice. On the other side, considering distributions that underestimate the probability of extreme scenarios can be costly. They may indeed create a false alarmism and require expensive hedging actions which in the future could turn to be unnecessary.

The accuracy of the semi-parametric specification presented in Section 7.1 for extreme returns allows overcoming these issues. Moreover, its flexibility allows preventing subjective considerations when it comes to the choice of a specific marginal distribution which may turn to be unsuited. These two properties makes this specification more suited for stress testing exercises, at least for the static scenario framework studied in this subsection. In the next section, we study whether the flexibility added by the R-vine model is also accompanied with good performance and draw the relevant implications for dynamic and multivariate stress tests.

7. The R-vine model

7.1. Empirical results

The results of the specification/estimation procedure of the R-vine model are reported in Table 8. The three bottom trees of Fig. 3 provide a graphical overview of the results. R-vine models are estimated for each of the three considered datasets. The results confirm the heterogeneity of the dependence structures, depending on the dataset and, within the same dataset, between the return series. Whatever the considered multivariate system is, the first tree

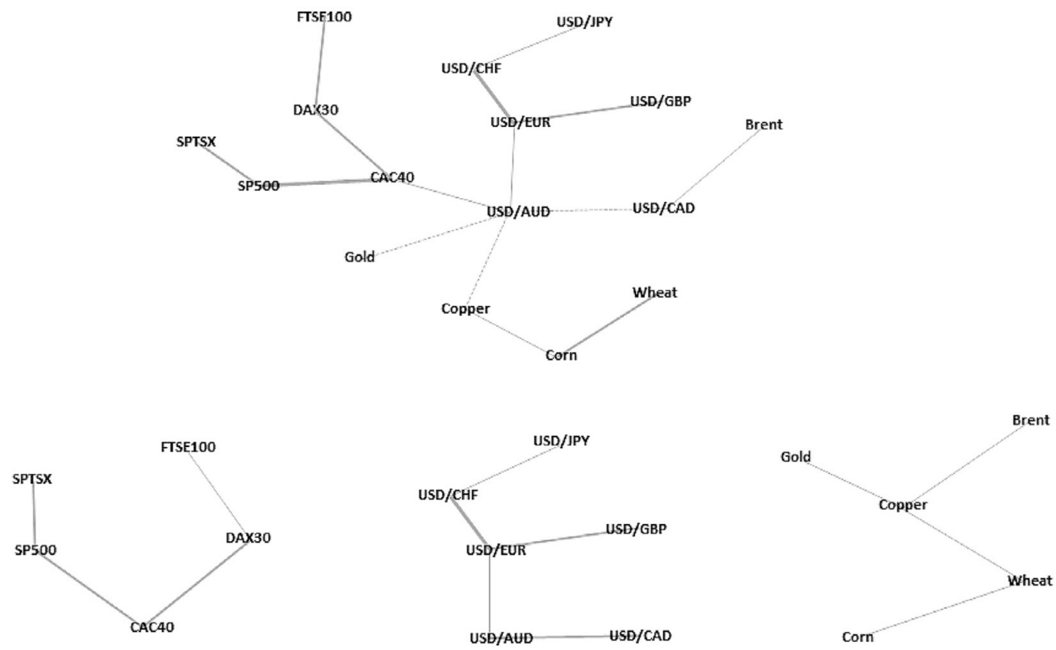

Fig. 3. Graphical representation of the R-vine estimation results. This figure gives a graphical view of the dependence structure in the first tree of four R-vine models estimated for: equity indices, exchange rates, commodity prices, and a mixed model of all datasets. For two given nodes, the width of the corresponding edge is proportional the degree of dependence between the underlying returns.

in the model – capturing the unconditional dependences between the return series – shows that extreme returns exhibit a strong dependence (Student *t*, survival Gumbel, and survival BB1 copulas). The dependence is often symmetric, except for three pairs of equities (S-BB1 copulas). These results fairly confirm those of the non-parametric tests (see Appendix A).

The lower trees of the R-vine model capture the conditional dependences. Overall, these trees are consisting of symmetric copulas. Among all candidates, the Student *t* copula seem to better fit the extreme conditional dependence structures – in the last tree – for the three datasets. However, except exchange rates at a lower degree, the joint copula's degree of freedom (the second parameter of the Student *t* copula) seem to be too high. This suggests a fairly weak dependence, given the lower levels of the corresponding Kendall's τ coefficients. Hence, the returns seem to be asymptotically weakly correlated in turmoil periods (stress and/or euphoria). This result is part of the academic debate around correlations in financial markets. The empirical evidence of a correlation increase during turmoil periods is not always obvious, as it has been shown since the Asian crisis.[10]

7.2. Validation tests

To check the robustness of our results and the performance of the R-vine model to capture the extreme dependence structure of the return series and hence for stress testing, we have first simulated a sample of 1 million bootstrap multivariate realizations for the estimated R-vine models. For each sample, the ECs have been calculated. These coefficients are then compared to the empirical ECs calculated for the filtered residuals of the ARMA-EGARCH models, and to the ECs calculated for a 1 million multivariate sample each from two models: a Gaussian meta-copula and a meta-copula with individual Student *t* copulas, both fitted to the standardized residuals. We report graphical results for the stock return series in Fig. 4, for quantiles ranging from 1% to 10% (negative returns) and from 90% to 99% (positive returns).

Three main results can be drawn. First, the estimates obtained by the R-vine and the Student *t* models are quite similar, and are clearly distinguished from those of the Gaussian model. This result is due to the fact that, for the data under review, the R-vine model is essentially composed of Student *t* pair-copulas (see Table 8). Second, both R-vine and the Student *t* models returned increasing ECs trend as the negative quantile moves away to most negative values. The corresponding ECs of the R-vine are always larger than those returned by the Student *t* copula's. Overall, the R-vine model seems to be more suited for extreme negative quantiles. Only in three cases out of ten this performance does drop, essentially because empirical ECs fall drastically for quantiles below 0.04. The lack of data points for these quantiles may indeed make the empirical ECs shift below the simulated trend. The fact remains, however, that we cannot conclude at this stage to an absolute accuracy of the R-vine model over the Student *t* model. Finally, the partial performance of the R-vine model also holds for extreme positive quantiles. In this case, the R-vine model performs well up to say quantile 0.96. Beyond, the Gaussian copula seems to dominate the other two models. This is due to the same reasons as for extreme negative quantiles.

To overcome the possible bias generated by the lack of extreme values in the data, we run an in-sample analysis to compare the performances of the R-vine, the Student *t* and the Gaussian models, as well as the C-vine and the D-vine pair-copulas models (see Brechmann and Czado (2011) for a similar analysis). To capitalize on the extra-flexibility of the vine copulas, we also report the results for a 16-dimension model (hereafter "mixed system") including the five equity indices, the five commodity prices, and the six exchange rate returns.[11] Results are displayed in Table 9.

[10] See, e.g., Chiang et al. (2007) and the references therein.

[11] The graphical output of the estimated R-vine model for the mixed system is shown in the upper panel of Fig. 3. The data sample used for the estimation runs from 1/4/1999 to 8/4/2011.

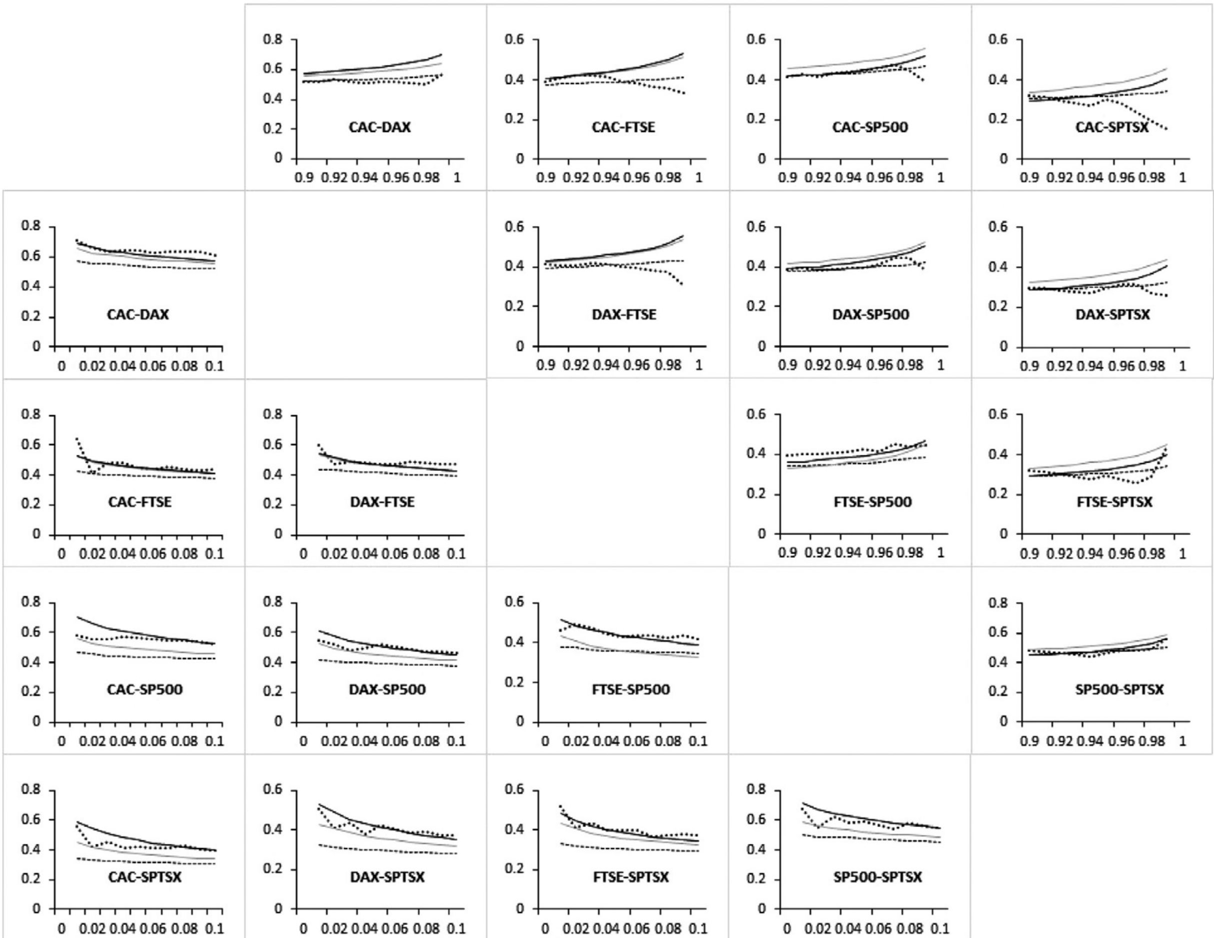

Fig. 4. Out-of-sample performance of multivariate risk models (equity indices). This figure shows empirical pair exceedance correlation (ECs) for the equity returns (dotted line) as well as the estimated ECs from a 1 million multivariate sample each obtained by fitting the R-vine (solid line), the Student *t* copula (gray line), and the Gaussian copula (dashed line) to the filtered residuals from the ARMA-EGARCH models. The Pearson's ρ (y axis) is given for the extreme negative (bellow the diagonal) and positive quantiles (over the diagonal).

Based on the estimation log-likelihood and common information criteria, the results show a dominance of the R-vine specification in 15 cases out of 16. It is followed by the other two vine and the Student *t* models whose performances vary depending on the dataset. The Gaussian copulas are clearly outperformed. To confirm these results, we carried out the Vuong (1989) and the Clarke (2007) tests designed for non-nested models. They test the null hypothesis that the performance of an alternative model is indistinguishable or superior to a baseline model. Considering the R-vine as the baseline model, we have reported the test statistics of both tests. The latter is followed by an asterisk when the null hypothesis cannot be rejected at 95% confidence level. To account for the difference in the number of parameters between the models, we also reported the corrected AIC-based and BIC-based adjusted versions. Except for commodity prices, the Clarke test rejects the null hypothesis for all alternative models. That is, the R-vine outperforms the remaining models, with and without a correction by the number of parameters. The Vuong test moderates this optimism, for the three datasets, as it only confirms the R-vine's dominance over the Gaussian model. We then perform a second round of assessment by increasing the size of the multivariate system. Even though the considered mixed system is not excessively oversized, both tests confirm the dominance of the R-vine model over all alternative specifications. This confirms the gains drawn from the flexibility of the R-vine model which becomes a serious advantage when the dimension increases as, in the same time, the alternative specifications become more restrictive. Indeed the more the system includes variables the more the flexibility drawn from the R-vine model is apparent. This is why with systems of say 5 variables, models such as C-vine, D-vine, or Student *t* could compete with the R-vine model. This is no longer the case with the mixed system of 16 variables as the competing specifications become too restrictive. To test whether the dominance of the R-vine model is independent of the used marginal distributions, we run the same exercise for the mixed system by assuming normal (light-tails) marginal distributions for the return series. Again the R-vine model clearly dominates the alternative specifications with an even higher margin with respect to the lower-sized systems.

Finally, to assess the accuracy of the R-vine model to perform dynamic stress tests, we have compared its out-of-sample performance with the that of DCC-GARCH model (Engle, 2002) commonly used in practice. The comparison is based on usual backtesting procedures for six multivariate specifications, including: three versions of the DCC model (DCC-Normal, and DCC-Student *t* with 4 and 7 degrees of freedom), a meta-copula with individual Student *t* copulas, and two pair-copulas models (R-vine, C-Vine). An equally weighted long portfolio is considered for each dataset. The models are re-estimated weekly using a rolling window of observations. The procedure is detailed in Appendix B for the R-vine model. Table 10 reports the number of exceedances (column 4) and the

Table 9
Assessment of the static properties of the R-vine model. This table compares the static (in sample) properties of the R-vine model to those of a C-vine and a D-vine pair-copulas models, a Gaussian meta-copula model, and a meta-copula model with individual Student t bivariate copulas. The comparison is based on the estimation log-likelihood (LLH), the number of parameters, and the AIC and BIC information criteria. For non-nested models, the Vuong (1989) and Clarke (2007) tests are also performed, with and without correction by the number of parameters – according to AIC and BIC criteria. Both tests assume, under the null hypothesis, that the considered model has indistinguishable or higher performance with respect to the R-vine. These statistics are marked with an asterisk when the null hypothesis cannot be rejected at 95% confidence level.

	Model	Obs.	LLH	pars.	AIC	BIC	Clarke			Vuong		
							wo/cor.	AIC	BIC	wo/cor.	AIC	BIC
Equity indices	R-vine	5979	7236.33	20	−14432.67	−14298.75						
	C-vine	5979	7227.93	20	−14415.86	−14281.94	2569	2569	2569	1.16*	1.16*	1.16*
	D-vine	5979	7233.06	19	−14428.11	−14300.89	2868	2859	2830	0.75*	0.82*	0.92*
	Student t	5979	7214.07	20	−14388.14	−14254.21	2893	2893	2893	1.89*	1.89*	1.89*
	Gaussian	5979	6822.96	10	−13624.96	−13558.00	3436	3410	3350	10.32	10.07	9.23
Exchange rates	R-vine	3283	5714.62	27	−11375.24	−11210.64						
	C-vine	3283	5694.78	27	−11335.55	−11170.95	1583	1583	1583	2.09	2.09	2.09
	D-vine	3283	5535.21	30	−11011.21	−10828.32	1853	1858	1875	7.25	7.37	7.74
	Student t	3283	5709.20	27	−11364.40	−11199.80	1576	1576	1576	0.60*	0.60*	0.60*
	Gaussian	3283	5073.74	15	−10117.47	−10026.03	2104	2097	070	11.54	11.33	11.67
Commodity prices	R-vine	4718	997.71	15	−1965.42	−1868.53						
	C-vine	4718	996.05	15	−1962.10	−1865.22	2302*	2302*	2302*	0.27*	0.27*	0.27*
	D-vine	4718	998.02	16	−1964.03	−1860.69	2387*	2394*	2427	−0.04*	0.10*	0.57*
	Student t	4718	992.69	20	−1953.37	−1850.37	2408*	2417*	2463	0.65*	0.77*	1.19*
	Gaussian	4718	931.38	10	−1842.75	1778.16	2683	2670	2610	4.53	4.19	3.09
Mixed system (M.S.)	Rvine	3283	12899.39	160	−25478.79	−24503.34						
	C-vine	3283	12740.08	169	−25142.15	−24111.84	1764	1770	1800	5.46	5.77	6.71
	D-vine	3283	12722.89	162	−25121.79	−24134.15	1754	1757	1767	5.21	5.27	5.44
	Student t	3283	12831.07	166	−25330.14	24318.12	1732	1742	1783	3.54	3.85	4.80
	Gaussian	3283	11868.88	120	−23497.76	−22766.18	2104	2075	1998	14.06	13.94	12.23
M.S. (normal marginals)	Rvine	3283	9349.77	139	−18417.54	−17557.93						
	C-vine	3283	9300.91	145	−18311.83	−17427.83	1699	1703	1733	2.21	2.40	2.95
	D-vine	3283	9116.13	123	−17944.25	−17066.35	1944	1948	1961	7.41	7.50	7.79
	Student t	3283	9246.40	143	−18206.80	−17335.00	1878	1883	1896	4.17	4.25	4.49
	Gaussian	3283	8513.65	120	−16787.60	−16055.72	2115	2100	2061	12.19	11.88	10.95

p-values of two duration-based backtesting procedures (columns 5–7): the Weibull test of independence by Christoffersen and Pelletier (2004) and the GMM test of conditional coverage by Candelon et al. (2011) with 2 and 5 orders. We use VaR-95% and VaR-99% confidence levels, with corresponding theoretical number of exceedances 50 and 10, respectively. The results confirm the poor performance of the DCC-Normal model, regardless of the confidence level. The DCC-Student t and the Student t models are fairly conservative above the 99% confidence level. Their use for risk management can then generate substantial opportunity costs related to the overestimated hedging costs of potential losses. Overall, copula-based models seem more efficient than DCC models. However, the meta-copula model is less performant than the pair-copulas'. This is mainly due to the relatively lack of flexibility of the former, assuming a common dependence structure for all pairs of returns. This becomes a serious handicap in a dynamic framework, as we have seen for oversized systems in the previous subsection. Moreover, the R-vine model slightly dominates the C-vine model, which is relatively less flexible.

The confirmed performance and the flexibility of the presented model framework justifies its use for extreme risk management in a dynamic and multivariate stress testing framework. Their suitability over competing specifications also rises for multidimensional multivariate systems.

7.3. An application to financial stress tests

The model specification impacts the stress testing results through the size of the corresponding initial shock (see Section 7.1). For dynamic scenarios, this specification also impacts the results through the after-shock recorded over the test horizon. To illustrate this aspect, we have considered the initial shocks estimated by the semi-parametric marginal distributions as explained above. We have then analyzed the evolution of returns for three equally weighted long portfolios (one for each datset) over a horizon of $H = 20$ daily sub-periods. We assume, during this horizon, that the portfolio cannot be liquidated nor hedged. The initial shock occurs at date $h = 1$ corresponding to date $T + 1$, where T is the size of the data sample. The post-shock is assessed for the next $H - 1$ periods. For a given period, the R-vine model captures the shock transmission across the different returns of the portfolio, given their extreme dependence structure. Moreover, the ARMA-EGARCH models allow measuring the shock transmission for the same return series across the horizon sub-periods. The combined effect of both models allows estimating the portfolio's return at each sub-period (see Appendix B for a formal presentation of the procedure). For each initial shock, we have compared the resulting impact on the portfolio's return as given by three alternative multivariate models: the R-vine, the Student t meta-copula, and a DCC-Student t with four degrees of freedom. The results are shown in Table 11.

They show a significant difference across models. By using the Student t meta-copula model, with high extreme dependences, most of the initial shock is transmitted across the multivariate system. This generates significant losses on the portfolios as the horizon goes on. However, this loss is fairly overestimated as actually not all returns do present a Student t-type pair dependence. Moreover, the use of a DCC-Student t model can be confusing, as it does not allow distinguishing between the form of the dependence structure and that of the underlying marginal distributions

Table 10
Assessment of the dynamic properties of the R-vine model (baktesting). This table compares the results of the backtesting procedure for different multivariate specifications including: (i) three versions of the DCC model (DCC-Normal, and DCC-Student t with 4 and 7 degrees of freedom), representing the direct approach, (ii) a meta-copula model with individual Student t copulas, for the classical indirect approach, and (iii) two pair-copulas models (R-vine, C-Vine), for the modern indirect approach. For each model and for two VaR's confidence levels (column 3), the table reports the number of exceedances (column 4). VaR-95% and VaR-99% are considered, with corresponding theoretical number of exceedances 50 and 10, respectively. Moreover, we report the p-values of two duration-based becktesting procedures (columns 5–7): the Weibull test of independence by Christoffersen and Pelletier (2004) and the GMM test of conditional coverage by Candelon et al. (2011) with 2 and 5 orders. The models are re-estimated weekly using a rolling window of 1000 observations. An equally weighted long portfolio is considered for each dataset.

	Model	Var (%)	Exceedances	Weibull	GMM (2 orders)	GMM (5 orders)
Equity indices	R-vine	95	53	0.123	0.073	0.101
		99	12	0.340	0.226	0.252
	C-vine	95	56	0.199	0.086	0.181
		99	14	0.431	0.262	0.286
	Student t meta-copula	95	59	0.246	0.142	0.194
		99	16	0.515	0.251	0.402
	DCC-Normal	95	66	0.461	0.247	0.381
		99	21	0.971	0.562	0.801
	DCC-Student$t$ (df = 4)	95	39	0.400	0.304	0.292
		99	6	0.431	0.308	0.344
	DCC-Student$t$ (df = 7)	95	45	0.165	0.080	0.130
		99	8	0.340	0.248	0.289
Exchange rates	R-vine	95	54	0.164	0.098	0.146
		99	13	0.409	0.228	0.364
	C-vine	95	57	0.240	0.144	0.176
		99	14	0.431	0.236	0.348
	Student t meta-copula	95	58	0.245	0.136	0.172
		99	17	0.578	0.372	0.411
	DCC-Normal	95	68	0.555	0.355	0.444
		99	20	0.959	0.613	0.885
	DCC-Student$t$ (df = 4)	95	42	0.245	0.147	0.181
		99	5	0.451	0.222	0.374
	DCC-Student$t$ (df = 7)	95	45	0.165	0.099	0.129
		99	6	0.431	0.281	0.383
Commodity prices	R-vine	95	52	0.100	0.059	0.072
		99	11	0.218	0.134	0.175
	R-vine	95	53	0.123	0.074	0.104
		99	12	0.340	0.185	0.302
	Student t meta-copula	95	60	0.357	0.206	0.330
		99	15	0.451	0.233	0.364
	DCC-Normal	95	72	0.636	0.387	0.433
		99	24	0.977	0.484	0.724
	DCC-Student$t$ (df = 4)	95	42	0.245	0.123	0.175
		99	6	0.431	0.222	0.426
	DCC-Student$t$ (df = 7)	95	45	0.165	0.065	0.124
		99	7	0.409	0.262	0.352

Table 11
Impact of the scenarios in terms of portfolios' return. This table shows the return losses (in percentage) for three linear portfolios formed by long equally weighted equity indices, exchange rates, and commodity prices, respectively. These losses are computed according to three multivariate models, 3 and 10 days after two initial shocks of probabilities 0.02% and 0.05%.

	Multivariate model				
	Probability (%)	Horizon (days)	R-vine	DCC-Student t	Student t meta copula
Equity indices	0.02	3	−8.96	−8.22	−9.31
	0.02	10	−12.21	−10.32	−12.32
	0.05	3	−8.44	−7.87	−8.88
	0.05	10	−9.45	−8.69	−9.82
Exchange rates	0.02	3	−8.91	−8.33	−10.18
	0.02	10	−12.04	−10.28	−12.32
	0.05	3	−8.99	−7.06	−9.55
	0.05	10	−10.42	−9.96	−10.30
Commodity prices	0.02	3	−8.53	−7.91	−8.80
	0.02	10	−11.89	−10.36	−11.88
	0.05	3	−7.86	−6.71	−8.23
	0.05	10	−10.24	−9.91	−10.59

(Garcia and Tsafack, 2011). Instead, the use of pair-copulas models allows avoiding these two drawbacks, being based on an explicit and a customized modeling of all pair dependences that form the joint dependence structure. Here again, capitalizing on the extra-flexibility and the performance of R-vine models can bring more credibility to stress tests, leading to a better use of their results and to minimizing management risks.

Given the impact of the shock and the after-shock on our three equally weighted portfolios, we have measured the required regulatory capital to cover the generated losses. Depending on the used method (so-called *internal risk model*), the Basel Committee defines the required amount of regulatory capital (RC) to hedge market risk losses revealed by the VaR (BCBS, 2005). This amount is given by the following formula

$$RC(t) = \max\left(VaR_{10-day|t}(0.01), k\frac{\Sigma_{i=1}^{60} VaR_{10-day|t-i+1}(0.01)}{60}\right) + c$$

where $VaR_{10-day|j}(0.01)$ is the VaR for a 10-day holding period with 99% confidence level, estimated given the information available information at date j. The multiplier k, with $3 \leq k \leq 4$, refers to the *stress factor*, defined by the regulator depending on the internal risk model, and c captures the idiosyncratic risks but market risk.

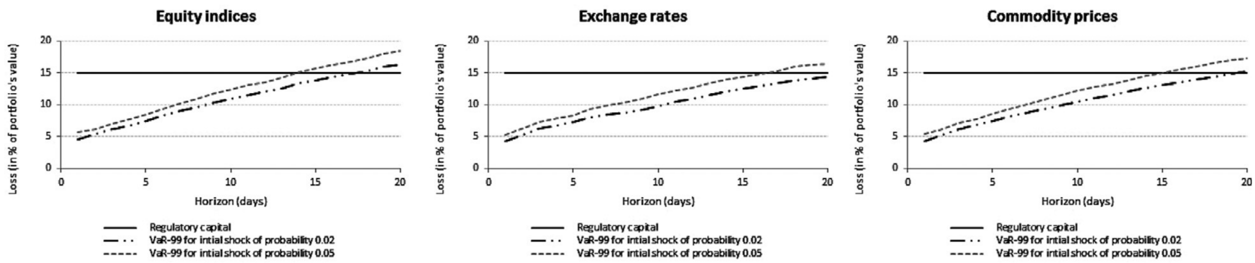

Fig. 5. Impact of the stress scenarios in terms of regulatory capital. This figure shows the impact of two stress scenarios of probabilities 0.02% and 0.05% for three long equally weighted portfolios (one for each dataset) over a horizon period of 20 trading days. For each period, the 99% VaR corresponding to the loss implied by the scenario is compared to the regulatory capital required to cover extreme losses.

The results are reported in Fig. 5, for $k = 3$ and $c = 0$. The required capital is indicated by a horizontal line. Whatever the considered shock is, the estimated losses increase with the test horizon. Overall, this occurs in comparable proportions for the three portfolios. Yet, the impact is relatively harmless for exchange rates and commodities, given the low (extreme-) dependence of this category of returns. This imply a low transmission of the initial shocks across the returns of these datasets. For equities, the losses remain below the regulatory capital up to say 14 (resp. 17) days after the initial shock of probability 0.02% (resp. 0.05%). For the remaining two portfolios, this occurs around day 17 and 20, respectively. In theory, the portfolios may be hedged or liquidated before these dates, to avoid any increase in the regulatory capital. In practice, however, the ability to perform these actions depends on several factors (e.g. market liquidity, portfolio size, strategic decisions, etc.). To account for these factors, regulators often recommend a 10-day non-transferability assumption.

8. Conclusion

In this paper, we have presented a flexible and robust model specification to model the extreme values and dependence structure of financial return series. The major objective of this paper is to present an adapted specification for model-based stress tests that meets the accuracy and flexibility criteria required in such exercises. Our aim is to overcome the recurrent misleading results drawn by common used specifications due to their simplicity and/or restrictive hypotheses. Often, the latter have led to false alarmism that exaggerates potential losses and increases hedging costs or, in the opposite situation, to an illusion of safety. The presented model is based on flexible specifications of the conditional marginal distributions and the dependence structure of the return series, based on a copula-GARCH model framework. The latter combines the EVT, for the univariate processes, and an R-vine model based on pair-copulas, for the multivariate system. The model is experimented for three datasets, including equity indices, bilateral exchange rates, and commodity prices. The results show fairly good performances in static and dynamic situations. The performance of this model also increases for extreme returns and dependence structure as well as for oversized portfolios.

In a second time, we presented a univariate and multivariate framework to run model-based static and dynamic scenarios. We have compared stress testing results from the R-vine model and alternative model specifications. The results show a significant difference of losses stemming from different risk models. The choice of a more powerful and flexible specification ensures the credibility of the stress scenario and the usefulness of the related results. We have shown, in this respect, that the R-vine model has interesting properties.

Two main areas of future research could be considered as an extension of this work. First, introducing dynamic and nonparametric specifications for the dependence structure, distinguishing between normal and stressful periods in financial markets (Giacomini et al., 2009; Manner and Reznikova, 2012; Stöber and Czado, 2012). The second possible extension consists in introducing market variables such as volatility and liquidity as risk factors to allow for more realism in the stress testing scenarios and results.

Acknowledgments

The author gratefully acknowledges conversations with Prof. Sanvi-Avouyi-Dovi and the members of his Ph.D dissertation committee. I thank the Editor, the Associate Editor and two anonymous reviewers from Journal of Banking and Finance for their constructive comments and suggestions that helped to considerably improve the manuscript. A special dedication goes to Profs. Eike Christian Brechmann and Alexios Ghalanos for their technical support. I am also grateful to comments from the audiences of the 11th INFINITI Conference on International Finance, the 62nd AFSE Annual Meeting, the GDRE-Leo Workshop on Financial Risk, and the XXXe Journées de Microéconomie Appliquée.

Appendix A. Statistical evidence of stylized facts

A.1. Univariate analysis

We have analyzed the descriptive statistics of each return series. For each series, Table 12 reports the historical extreme values, the first empirical moments, and the results of normality, autocorrelation, and heteroscedasticity tests. All series have a quasi-null conditional mean. Their variances vary depending on the dataset. Hence, commodities have the highest standard deviation (with a maximum of 0.024 for wheat), followed by equities (between 0.011 and 0.015).The two datasets have comparable extreme values with the wheat return as the most dispersed. Exchange rates are less volatile (between 0.06 and 0.08) and have lower extremes compared to the previous two sets.

In general, the series exhibit a negative skewness (except for USD/GBP, USD/AUD, and USD/CHF) and a kurtosis excess (except for USD/EUR). This confirms the stylized facts that the distributions of financial returns are leptokurtic and left-skewed. This is particularly true for negative returns. To confirm this result, we have reported the results of two statistical normality tests, namely the Jarque–Bera (JB) and the Anderson–Darling (AD) tests. The former is based on the values of the 3rd and 4th empirical moments of the return distributions. The second, considers the distance between the empirical and the normal distributions, with an overweighting of extreme quantiles. Both tests reject the null hypothesis of normality for all series, with a confidence level of 99%. Moreover, the test statistics show heterogeneous features between the empirical distributions on one side, and the normal distribution on the other. The distributions of stock returns are thus the most disconnected from the normal distribution, especially in the tails.

Table 12
Descriptive statistics. This table presents the descriptive statistics (columns 3–8) and the results of normality (columns 9–10), autocorrelation (columns 11–14), and heteroscedasticity (columns 15–17) tests for the daily return series of stock indices, exchange rates, and commodity prices. For each series, the table reports: (i) the empirical extrema and the first four moments, (ii) the statistics of the Jarque–Bera (JB) and the Anderson–Darling (AD) normality tests, and (iii) the statistics of the Ljung–Box LB(p) and the Diebold D(p) autocorrelation/heteroscedasticity tests for p lags. The symbols * and ** denote a rejection of the null hypothesis of normality, no autocorrelation, or no heteroscedasticity (depending on the test) at the 5% and 1% confidence levels, respectively.

	Obs.	Min	Max	Mean	Std. dev.	Sk.	Ku.	JB	AD	LB(25)r	LB(50)r	D(25)r	D(50)r	LB(25)r²	LB(50)r²	LB(25)|r|
CAC40	6979	−0.133	0.119	0.000	0.014	−0.430	6.969	14382.96**	77.37**	96.14**	137.03**	33.22	60.26	3376.4**	4138.7**	6275.6**
DAX30	6979	−0.103	0.101	0.000	0.014	−0.370	5.553	9133.03**	92.44**	45.24**	74.06*	19.41	35.24	5867.6**	7990.3**	9180.3**
FTSE100	6979	−0.093	0.104	0.000	0.011	−0.004	9.954	14106.72**	82.98**	115.01**	169.33**	40.20*	70.12**	5667.3**	7378.4**	6752.8**
SP500	6979	−0.094	0.110	0.000	0.011	−0.269	9.371	25698.56**	128.76**	94.14**	164.08**	25.59	52.59	9972.9**	13887.0**	10268.0**
SPTSX	6979	−0.098	0.094	0.000	0.010	−0.760	11.83	41511.33**	146.26**	85.25**	162.89**	18.94	42.81	11234.0**	16730.0**	12778.0**
USD/EUR	4283	−0.046	0.038	0.000	0.006	−0.172	2.509	1151.90**	17.62**	16.46	39.13	12.24	29.06	832.96**	1592.5**	920.29**
USD/JPY	4283	−0.046	0.037	0.000	0.006	−0.237	3.628	2402.78**	22.65**	29.99	71.4*	23.10	57.40	468.03**	642.57**	710.22**
USD/GBP	4283	−0.045	0.039	0.000	0.006	0.042	4.349	3395.91**	19.38**	67.34**	100.71**	35.73*	53.67	2613.2**	4603.8**	2407.4**
USD/AUD	4283	−0.067	0.088	0.000	0.008	0.841	12.51	28467.24**	53.18**	56.93**	110.29**	14.19	38.89	4753.5**	5732.7**	4828.6**
USD/CAD	4283	−0.050	0.043	0.000	0.006	−0.069	5.740	5919.53**	31.80**	51.73**	91.67**	22.38	42.01	3397.8**	5330.3**	3787.4**
USD/CHF	4283	−0.114	0.085	0.000	0.007	−0.886	24.49	108148.73**	29.90**	35.66*	61.92	23.43	45.67	233.42**	240.34**	527.88**
Brent	5718	−0.136	0.135	0.000	0.021	−0.083	3.228	2500.47**	38.62**	46.80**	110.53**	30.61	75.19*	1583.8**	2619.0**	1985.1**
Gold	5718	−0.102	0.074	0.000	0.010	−0.321	7.393	13170.05**	106.88**	63.12**	94.59**	32.12	49.90	1569.4**	2562.4**	3729.3**
Copper	5718	−0.105	0.117	0.000	0.016	−0.209	5.122	6317.02**	67.03**	61.02**	139.9**	26.57	73.28*	4631.5**	6195.0**	5255.9**
Wheat	5718	−0.226	0.226	0.000	0.024	−0.155	8.398	16891.93**	71.40**	71.86**	104.0**	34.76*	57.38	974.3**	1257.8**	2209.0**
Corn	5718	−0.121	0.109	0.000	0.018	−0.179	3.437	2858.53**	54.85**	43.13*	72.71*	28.60	48.96	1446.5**	2321.3**	1954.5**

The introduced flexibility introduced by the marginal model allows considering the specificity of each distribution of returns which varies on the dataset and even between the returns of the same dataset.

The presence of autocorrelation and heteroscedasticity in the return series.[12] we carried out the Ljung–Box (LB) and the Diebold (D) tests. The latter also allows separating autocorrelation and heteroscedasticity features. The statistics of the tests are reported for 25 and 50 lags in Table 12 (see Fig. 2 for the autocorrelation function of lags 1–50). Except for USD/EUR and USD/JPY, the LB test rejects the null hypothesis of no autocorrelation with confidence levels ranging from 95% to 99% for 25 and 50 lags (except for USD/CHF). Yet, the D statistic shows that only the FTSE100, USD/GBP and wheat (respectively, FTSE100, Brent, and copper) are actually autocorrelated for 25 (respectively, 50) lags. This suggests that the series are rather heteroscedastic than autocorrelated. To check this aspect, we have reported the LB statistics for nonlinear transformations of the returns, namely squared and absolute values. The LB test rejects the null hypothesis of no autocorrelation for the transformed returns (i.e. of no heteroscedasticity in the returns) with a confidence level of 99%. Yet, the presence of heteroscedasticity is known to impact both the tails of the return distributions and their dependence structure (Jondeau, 2010). The ARMA-EGARCH models allow isolating potential autocorrelation and heteroscedasticity effects.

A.2. Analysis of the dependence structure

Regarding the joint dynamics of the series, Table 13 analyzes their empirical dependence structures. Within each dataset, and for each pair of returns, we have reported the Pearson's ρ linear correlation coefficient and the Kendall's τ nonlinear rank coefficient, as well as the extreme correlation (EC) $\rho^{\pm}$ and tail dependence coefficient (TDC) $\tau^{\pm}$.[13] The latter measure the correlation and the rank correlation of extreme positive and negative returns beyond extreme quantiles. Positive (resp. negative) extreme quantiles of 10%, 5%, and 1% (resp. 90%, 95%, and 99%) are considered.

The results show that, overall, equities and exchange rates are more correlated than commodities. Equity indices of the same geographic area are more correlated even though the SP500 is rather highly linked to European indices. European currencies are the most correlated in the system (between 0.66 and 0.79). However, except with the yen, the Australian dollar is positively correlated to all currencies (between 0.39 and 0.62). The USD/CAD series may then occupy a central position in the first tree of the R-vine copula. The JPY is weakly – and even negatively – correlated to the rest of currencies. This explains inconsistencies between the values of ρ and τ for JPY/AUD and JPY/CAD. The analysis of the dependence through the rank coefficients confirms all the previous results.

Table 13 also provides a first indication for the dependence structure of extreme returns. The results show that, overall, the magnitude of the extreme coefficient follows that of the coefficient calculated on the whole distribution. Thus, highly correlated returns in normal times also exhibit high correlations in more volatile periods (stress or euphoria). Instead, returns with a (very) low correlation coefficient show a low or negative extreme correlation. This is particularly the case of commodity prices. This result is still valid regardless of the considered extreme quantile. Instead, the asymmetry of the extreme dependence structure is less obvious. Indeed, a comparison of the ECs and TDCs for positive and negative returns of the same quantile leads to different results, depending on the considered quantile and the dependence measure.

[12] Autocorrelation and heteroscedasticity are also referred to as linear (or first-order) and nonlinear (second-order) serial dependences, respectively.

[13] See Jondeau (2010) for a formal presentation and analysis of these measures.

Table 13
Nonparametric EC, TDC, and test for asymmetry of extreme correlations. This table reports, for each pair of series: (i) the linear correlation (ρ) and rank (τ) coefficients, (ii) the extreme correlation ($\rho^{\pm}$) and tail dependence ($\tau^{\pm}$) coefficients for 10–90%, 5–95%, and 1–99% quantiles, and (iii) the corresponding statistics and p-values of the Hong et al., 2007 test (J_ρ) for the symmetry of extreme correlations. A p-value close to zero indicates a rejection of the null hypothesis of symmetric extreme correlations.

Series 1	Series 2	Obs.	ρ	τ	$\rho^-(1\%)$	$\rho^+(1\%)$	$J_\rho(1\%)$	p-value	$\rho^-(5\%)$	$\rho^+(5\%)$	$J_\rho(5\%)$	p-value	$\rho^-(10\%)$	$\rho^+(10\%)$	$J_\rho(10\%)$	p-value	$\tau^-(1\%)$	$\tau^+(1\%)$
CAC40	DAX30	6979	0.77	0.56	0.78	0.54	1.65	0.20	0.76	0.47	2.22	0.14	0.76	0.55	0.94	0.33	0.54	0.53
CAC40	FTSE100	6979	0.54	0.37	0.66	0.18	0.73	0.39	0.55	0.29	2.37	0.12	0.48	0.37	4.20	0.04	0.33	0.27
CAC40	SP500	6979	0.65	0.43	0.71	0.21	1.72	0.19	0.70	0.42	3.23	0.07	0.69	0.53	6.54	0.01	0.47	0.42
CAC40	SPTSX	6979	0.49	0.30	0.29	0.14	4.03	0.04	0.63	0.17	5.90	0.02	0.57	0.29	2.45	0.12	0.33	0.27
DAX30	FTSE10	6979	0.62	0.41	0.50	0.72	0.00	0.95	0.53	0.53	0.00	0.98	0.54	0.54	0.55	0.56	0.46	0.34
DAX30	SP500	6979	0.61	0.39	0.54	0.65	1.26	0.26	0.61	0.36	3.04	0.08	0.60	0.47	0.11	0.74	0.47	0.39
DAX30	SPTSX	6979	0.49	0.30	0.23	0.52	1.89	0.17	0.44	0.15	3.33	0.07	0.45	0.26	10.96	0.00	0.31	0.24
FTSE100	SP500	6979	0.51	0.33	0.33	0.53	0.11	0.74	0.57	0.50	0.16	0.69	0.47	0.52	0.50	0.48	0.36	0.33
FTSE100	SPTSX	6979	0.49	0.30	0.15	0.06	0.32	0.57	0.46	0.38	0.20	0.66	0.50	0.41	0.08	0.77	0.34	0.26
SP500	SPTSX	6979	0.70	0.46	0.67	0.70	0.04	0.85	0.65	0.66	0.01	0.94	0.47	0.51	0.01	0.91	0.49	0.48
USD/EUR	USD/JPY	4283	0.25	0.19	0.51	0.31	0.29	0.59	0.82	0.19	1.11	0.29	0.80	−0.41	7.01	0.00	0.14	0.22
USD/EUR	USD/GBP	4283	0.66	0.47	0.69	0.15	0.31	0.58	0.68	0.39	0.68	0.41	0.51	0.31	1.50	0.22	0.40	0.29
USD/EUR	USD/AUD	4283	0.55	0.38	0.54	−0.05	1.53	0.22	0.59	0.16	2.05	0.15	0.58	0.28	6.83	0.00	0.33	0.26
USD/EUR	USD/CAD	4283	0.45	0.27	0.19	−0.05	1.18	0.28	0.59	0.30	1.43	0.23	0.56	0.32	0.90	0.34	0.28	0.21
USD/EUR	USD/CHF	4283	0.79	0.71	0.85	0.86	0.17	0.68	0.77	0.69	0.07	0.79	0.79	0.71	0.00	0.98	0.58	0.48
USD/JPY	USD/GBP	4283	0.15	0.14	–	–	0.77	0.38	0.69	−0.17	1.86	0.17	0.43	0.08	0.00	1.00	0.08	0.19
USD/JPY	USD/AUD	4283	0.03	0.12	1.00	0.99	0.14	0.71	0.78	0.50	0.16	0.69	0.57	0.39	0.00	1.00	0.06	0.15
USD/JPY	USD/CAD	4283	0.01	0.04	1.00	–	1.14	0.29	0.69	0.02	1.37	0.24	0.57	0.08	0.00	1.00	0.04	0.06
USD/JPY	USD/CHF	4283	0.35	0.26	0.25	0.40	0.00	0.98	0.56	0.25	0.74	0.39	0.36	0.36	0.14	0.71	0.15	0.18
USD/GBP	USD/AUD	4283	0.54	0.34	0.83	0.10	0.51	0.48	0.67	0.28	1.23	0.27	0.55	0.36	4.45	0.03	0.24	0.35
USD/GBP	USD/CAD	4283	0.44	0.25	0.50	0.23	0.27	0.60	0.68	0.37	0.92	0.34	0.62	0.48	0.84	0.36	0.19	0.25
USD/GBP	USD/CHF	4283	0.52	0.41	0.67	0.96	0.18	0.67	0.53	0.22	0.55	0.46	0.41	0.29	0.20	0.65	0.24	0.18
USD/AUD	USD/CAD	4283	0.62	0.39	0.87	0.56	0.43	0.51	0.80	0.47	0.76	0.38	0.71	0.52	0.64	0.42	0.35	0.40
USD/AUD	USD/CHF	4283	0.39	0.30	−0.02	−0.19	3.19	0.07	0.37	0.09	1.95	0.16	0.49	0.17	0.24	0.63	0.16	0.18
USD/CAD	USD/CHF	4283	0.31	0.21	−0.53	0.28	0.39	0.53	0.31	0.20	0.32	0.57	0.38	0.26	5.19	0.02	0.19	0.18
Brent	Gold	5718	0.18	0.11	1.00	−0.08	4.72	0.03	0.34	0.19	0.25	0.61	−0.01	0.35	0.00	1.00	0.04	0.11
Brent	Copper	5718	0.23	0.12	0.10	0.23	0.11	0.74	0.34	0.45	0.28	0.60	0.35	0.30	0.15	0.70	0.15	0.12
Brent	Wheat	5718	0.10	0.05	−0.34	−0.03	0.72	0.40	−0.08	0.37	2.26	0.13	0.09	0.26	0.00	1.00	0.07	0.06
Brent	Corn	5718	0.13	0.06	0.77	0.99	0.00	0.99	0.34	0.34	0.00	0.99	0.32	0.32	0.38	0.54	0.11	0.08
Gold	Copper	5718	0.27	0.16	−0.34	−0.03	0.56	0.45	0.12	0.09	0.04	0.85	0.27	0.18	1.05	0.31	0.19	0.16
Gold	Wheat	5718	0.09	0.05	−0.03	−1.00	0.69	0.41	0.15	0.04	0.20	0.66	0.15	0.03	0.00	1.00	0.13	0.05
Gold	Corn	5718	0.10	0.06	−0.38	−1.00	0.93	0.33	0.04	0.04	0.00	0.98	0.25	0.13	0.00	1.00	0.10	0.08
Copper	Wheat	5718	0.13	0.06	0.27	0.24	0.02	0.88	0.30	0.25	0.03	0.87	0.13	0.06	0.39	0.53	0.13	0.08
Copper	Corn	5718	0.16	0.07	0.35	0.35	0.00	0.99	0.39	0.55	0.15	0.70	0.16	0.07	1.06	0.30	0.16	0.08
Wheat	Corn	5718	0.48	0.35	0.41	0.28	0.58	0.44	0.53	0.36	0.45	0.50	0.48	0.35	7.43	0.00	0.26	0.27

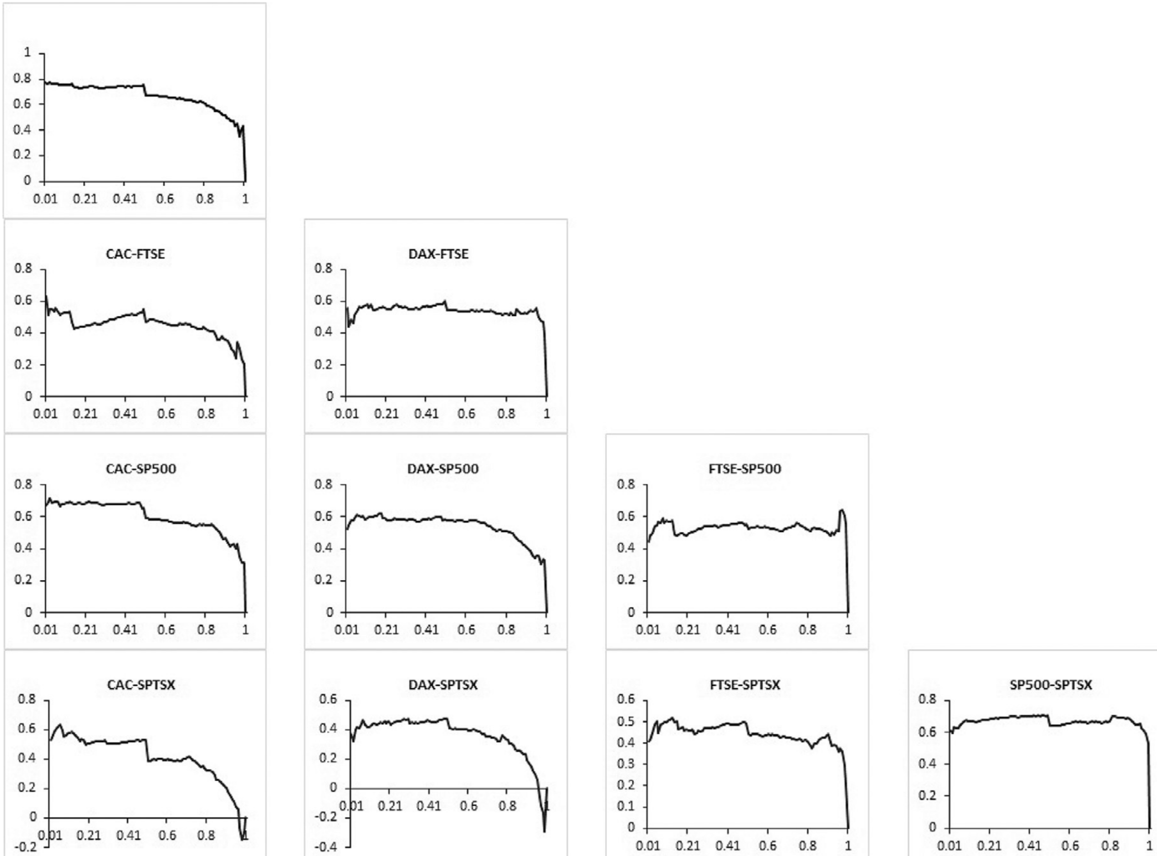

Fig. 6. Empirical exceedance correlations (equity indices). This figure presents the pair exceedance correlations (EC) for the equity returns. The Pearson' ρ (y axis) is given for empirical quantiles from 1% to 99%.

Except for the Brent-gold and the Brent-wheat pairs, the correlation of negative returns below the 10% quantile are thus greater than or equal to that of positive returns beyond the 90% quantile. However, such a result is no longer valid when reducing the size of extreme quantiles to 1% and 99%, respectively, by considering the most extreme returns. In this case, negative returns are more correlated than positive returns for some pairs and less correlated for the remaining pairs, even though the TDC in the lower tail is often larger than in the upper. This ordering also varies with the considered dependence measure. The limited number of observations for the most extreme returns may justify this results' instability. Fig. 6 shows the profile of the empirical extreme correlation for equity indices, for quantiles ranging from 0.01 to 0.99. Visually, the asymmetry in extreme dependence is yet apparent in most cases.

To get a more precise idea on the extent of the (a-) symmetry in extreme returns correlation, we have carried out the nonparametric statistical test proposed by Hong et al. (2007). It allows comparing all positive and negative ECs, respectively, beyond an extreme quantile, with a null hypothesis of symmetric extreme dependence. We have reported the test statistics and the corresponding p-values for 10%–90%, 5%–95%, and 1%–99% quantiles. Whatever the threshold is, the test does not allow rejecting, in most cases, the null hypothesis with a confidence level of 95%. This may seem at odds with previous works (Longin and Solnik, 1995; Longin and Solnik, 2001; Ang and Bekaert, 2002a; Ang and Bekaert, 2002b; Ang and Chen, 2002; Patton, 2004; Patton, 2006a; Patton, 2006b). However, as shown by the ECs and in Fig. 6 this asymmetry does exist. Yet, it is often too weak to be significantly captured by the considered statistical test. The variability of extreme correlations according to the considered quantile may also explain this result. As the extreme dependence is of particular interest for stress testing models, assuming that all returns share the same extreme dependence function may be misleading. In this paper, this issue is addressed by introducing more flexibility in the specification of pair dependences of the return series, using bivariate copulas with different dependence and extreme dependence shapes.

Appendix B. The R-vine model for risk management

Let p a portfolio of N assets and w the related weighting vector, with $w' = [w_1 \dots w_i \dots w_N]$.[14] For each asset, we collect T observations of the underlying returns. We split this sample into two sub-periods: the first T–M observations are used to estimate the models; the remaining M observations to perform out-of-sample analysis. For the latter, we carry out one-period ahead estimations of the portfolio's VaR using a rolling window of T-M observations. The models are re-estimated weekly.[15] Hereafter is a technical description of the procedure.[16]

B.1. Value-at-risk assessment

Let τ the last period in the rolling data sample used to estimate the models. The one-period ahead forecast of the portfolio's VaR is obtained as follows:

[14] In this paper, we have considered $w_i \in \mathbb{R}^+$, $\precapprox\beth = \precapprox\beth$ for $i \neq j$, and $\sum_{i=1}^{N} w_i = 1$.

[15] In this paper, M is set to 1000 observations. In the backtesting analysis, we compare the real data spanning from T–M + 1 and T to the simulated data for the same period.

[16] See also Brechmann and Czado, 2011 for a similar approach.

1. For each return series $i = 1, \ldots, N$.
   i. Fit weekly the ARMA-EGARCH model to the T-M last observed data, as in Eq. (1).
   ii. Get the corresponding standardized residuals $\hat{z}_{i,t}$, with $t = \tau - (\mathrm{T} - \mathrm{M} - 1), \ldots, \tau$.
   iii. Use this series to get a one-step ahead conditional variance $\hat{\sigma}_{i,\tau+1}$, as in Eq. (1).
   iv. Fit the semi-parametric marginal distribution of Eq. (2) to the standardized residuals.
   v. Use the estimates of the marginal distribution to get the PIT series.
2. Use the PIT series of all returns to estimate the R-vine model.
3. To get the VaR's confidence intervals, simulate R random vectors from the R-vine model.[17]
4. For each $r = 1, \ldots, R$.
   i. Use the marginal distributions to get a one-step ahead vector of residuals $(\hat{z}_{i,\tau+1})_{i=1,\ldots,N}$.[18]
   ii. Use the one-step ahead vectors of residuals and conditional variances to get the one-step ahead vector of returns $(\hat{r}_{i,\tau+1})_{i=1,\ldots,N}$, as in Eq. (1).
   iii. Get the portfolio's one-step ahead return using the weighting vector, as follows

$$\hat{r}_{p,\tau+1} = \sum_{i=1}^{N} w_i \cdot \hat{r}_{i,\tau+1}$$

5. Get the portfolio's return distribution using the one-step ahead returns for $r = 1, \ldots, R$.
6. The VaR(α) is the α-quantile of the one-step ahead portfolio's return distribution.
7. The simulated VaR(α) is compared to real returns in the back-testing analysis.

B.2. Stress testing

To run stress testing exercises, a slight modification is to be introduced at step 4i in the first period of the procedure above. Among the R simulated vectors of residuals, we keep only a few of them for stress testing purposes. In Section 7.3, three (3) particular vectors are considered. The first one is formed by the residuals corresponding to the 0.0001 quantiles of each of the N marginal distributions. The remaining vectors are drawn from quantiles 0.0002 and 0.0005.[19]

The procedure terminates at the end of step 4 for the first period, while steps 1–6 are repeated for the following periods up to the stress horizon.[20] The obtained simulations of the VaR are compared to the regulatory capital to draw Fig. 5.

References

Aas, K., Berg, D., 2009. Models for construction of multivariate dependence: a comparison study. The European Journal of Finance 5 (7–8), 639–659.
Aas, K., Haff, I.H., 2006. The generalized hyperbolic skew student's t-distribution. Journal of Financial Econometrics 4 (2), 275–309.
Aas, K., Czado, C., Frigessi, A., Bakken, H., 2009. Pair-copula constructions of multiple dependence. Insurance: Mathematics and Economics 44 (2), 182–198.

[17] In this paper, R is set to 1 million multivariate draws.

[18] The simulated random vector obtained using the R-vine model is in the unit hypercube $[0, 1]^N$. This step simply consists in transforming this vector into real data, using the inverse of the marginal distributions.

[19] Note that the probabilities of the multivariate scenarios are different from these individual quantiles, as the former are drawn from the R-vine joint distribution. This aspect has been omitted for more simplicity, given that the objective of the paper is to analyse the impacts of the scenarios rather than their probabilities.

[20] Unless a new shock is introduced in between. In such a case, we proceed the same as in the first period.

Acharya, A., Engle, R., Pierret, D., 2014. Testing macroprudential stress tests: the risk of regulatory risk weights. Journal of Monetary Economics 65 (2014), 36–53.
Aikman, D., Alessandri, P., Eklund, B., Gai, P., Kapadia, S., Martin, E., Mora, N., Sterne, G., Willison, M., 2009. Funding Liquidity Risk in a Quantitative Model of Systemic Stability, Bank of England, Working Paper No. 372.
Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. In: Petrov, B.N., Csaki, F. (Eds.), Second International Symposium on Information Theory. Academiai Kiado, Budapest, pp. 267–281.
Alessandri, P., Gai, P., Kapadia, S., Mora, N., Puhr, C., 2009. Towards a framework for quantifying systemic stability. International Journal of Central Banking 5 (3), 47–81.
Alexander, C., Sheedy, E., 2008. Developing a stress testing framework based on market risk models. Journal of Banking and Finance 32 (10), 2220–2236.
Ang, A., Bekaert, G., 2002. Regime switches in interest rates. Journal of Business and Economic Statistics 20 (2), 163–182.
Ang, A., Bekaert, G., 2002. International asset allocation with regime shifts. The Review of Financial Studies 15 (4), 1137–1187.
Ang, A., Chen, J., 2002. Asymmetric correlations of equity portfolios. Journal of Financial Economics 63 (3), 443–494.
Aragonés, J.R., Blanco, C., Dowd, K., 2001. Incorporating stress tests into market risk modeling. Derivative Quartely 7 (3), 44–49.
Azzalini, A., Capitanio, A., 2003. Distributions generated by pertubation of symmetry with emphasis on a multivariate skew t distribution. Journal of the Royal Statistical Society: Series B 65 (2), 367–389.
Banachewicz, K., van der Vaart, A., 2008. Tail dependence of skewed grouped t-distributions. Statistics and Probability Letters 78 (15), 2388–2399.
Barndorff-Nielsen, O.E., 1977. Exponentially decreasing distributions for the logarithm of particle size. Proceedings of the Royal Society of London: Series A 353 (1674), 401–419.
Barndorff-Nielsen, O.E., 1978. Hyperbolic distributions and distributions on hyperbolae. Scandinavian Journal of Statistics 5 (3), 151–157.
Basel Committee on Banking Supervision (BCBS), 2005. Amendment to the Capital Accord to Incorporate Market Risks, BIS, November.
Bedford, T., Cooke, R.M., 2001. Probability density decomposition for conditionally dependent random variables modeled by vines. Annals of Mathematics and Artificial Intelligence 32 (1–4), 245–268.
Bedford, T., Cooke, R.M., 2002. Vines – a new graphical model for dependent random variables. The Annals of Statistics 30 (4), 1031–1068.
Bee, M., 2001. Mixture Models for VaR and Stress Testing, ALEA Tech Reports, Università di Trento, No. 12.
Berkowitz, J., 2000. A coherent framework for stress-testing. Journal of Risk 2 (2), 5–15.
Bibby, B.M., Sørensen, M., 2003. Hyperbolic processes in finance. In: Rachev, S. (Ed.), Handbook of Heavy Tailed Distributions in Finance. Elsevier Science, 211–48.
Bank for International Settlements (BIS), 2013. Triennial Central Bank Survey of Foreign Exchange and Derivatives Market Activity in April 2013, Preliminary global results, Monetary and Economic Department, September.
Blaschke, W., Jones, M.T., Majnoni, G., Martinez Peria, S., 2001. Stress Testing of Financial Systems: An Overview of Issues, Methodologies, and FSAP Experiences, International Monetary Fund, Working Paper No. 01/88.
Blattberg, R.C., Gonedes, N.J., 1974. A comparison of the stable and student distributions as statistical models for stock prices. The Journal of Business 47 (2), 244–280.
Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31 (3), 307–327.
Borio, C., Drehmann, M., 2009. Towards an Operational Framework for Financial Stability: "Fuzzy" Measurement and its Consequences, Central Bank of Chile, Working Paper No. 544.
Borio, C., Drehmann, M., Tsatsaronis, K., 2012. Stress Testing Macro Stress Testing: Does it Live Up to Expectations? Bank for International Settlements, Working Paper No. 369.
Boss, M.G., 2008. Stresstests Für das Osterreichische FSAP-Update 2007: Methodik, Szenarien und Ergebnisse, Oesterreichische Nationalbank. OeNB Finanzmarktsstabilitätsbericht 15, 72–99.
Box, G.E.P., Jenkins, G.M., 1976. Time Series Analysis: Forecasting and Control. San Francisco Holden-Day (Revised Edition).
Branco, M.D., Dey, D.K., 2001. A general class of multivariate skew-elliptical distributions. Journal of Multivariate Analysis 79 (1), 99–113.
Brechmann, E.C., Czado, C., 2011. Risk Management with High-Dimensional Vine Copulas: An Analysis of the Euro Stoxx 50, Center for Mathematical Sciences. Technische Universität München, Working Paper.
Brechmann, E.C., Schepsmeier, U., 2013. Modeling dependence with C- and D-vine copulas: the R package CDVine. Journal of Statistical Software 52 (3), 1–27.
Brechmann, E.C., Czado, C., Aas, K., 2012. Truncated regular vines in high dimensions with application to financial data. Canadian Journal of Statistics 40 (1), 68–85.
Breuer, T., Csiszár, I., 2013. Systematic stress tests with entropic plausibility constraints. Journal of Banking and Finance 37 (5), 1552–1559.
Breuer, T., Krenn, G., 1999. Stress Testing: Guidelines on Market Risk. Oesterreichische Nationalbank vol. 5.
Breuer, T., Krenn, G., Pistovc˘ák, F., 2002. Stress tests, maximum loss, and value at risk. In: Britzelmaier, B., Geberl, S., Menichetti, M., Kaufmann, H.-R. (Eds.), Regulierung oder Deregulierung der Finanzmärkte: Strategieimplikationen für Finanzdienstleister. Physica Verlag (Springer), pp. 301–314.
Candelon, B., Colletaz, G., Hurlin, C., Tokpavi, S., 2011. Backtesting value-at-risk: a GMM duration-based test. Journal of Financial Econometrics 9 (2), 314–343.

Cherubini, U., Luciano, E., Vecchiato, W., 2004. Copula Methods in Finance. Wiley, Chichester.
Chiang, Th.C., Jeon, B.N., Li, H., 2007. Dynamic correlation analysis of financial contagion: evidence from asian markets. Journal of International Money and Finance 26 (2007), 1206–1228.
Chollete, L., Heinen, A., Valdesogo, A., 2009. Modeling international financial returns with a multivariate regime-switching copula. Journal of Financial Econometrics 7 (4), 437–480.
Christoffersen, P.F., Pelletier, D., 2004. Backtesting value-at-risk: a duration-based approach. Journal of Financial Econometrics 2 (1), 84–108.
Čihák, M., 2007. Introduction to Applied Stress Testing, Technical Report, International Monetary Fund, Working Paper No. 07/59.
Clarke, K.A., 2007. A simple distribution-free test for nonnested model selection. Political Analysis 15 (3), 347–363.
Czado, C., 2010. Pair-copula constructions of multivariate copulas. In: Jaworski, P., Durante, F., Härdle, W.K., Rychlik, T. (Eds.), Copula Theory and its Applications, Lecture Notes in Statistics, 198. Springer, pp. 93–109.
Czado, C., Schepsmeier, U., Min, A., 2011. Maximum likelihood estimation of mixed C-vines with application to exchange rates. Statistical Modelling 12 (3), 229–255.
Demarta, S., McNeil, A.J., 2005. The t copula and related copulas. International Statistical Review 73 (1), 111–129.
Dißmann, J., Brechmann, E.C., Czado, C., Kurowicka, D., 2013. Selecting and estimating regular vine copulae and application to financial returns. Computational Statistics and Data Analysis 59, 52–69.
Eberlein, E., Keller, U., 1995. Hyberbolic distributions in finance. Bernoulli 1 (3), 281–299.
Eberlein, E., von Hammerstein, E.A., 2002. Generalized hyperbolic and inverse gaussian distributions: limiting cases and approximation of processes. In: Dalang, R.C., Dozzi, M., Russo, F. (Eds.), Seminar on Stochastic Analysis, Random Fields and Applications IV, Progress in Probability 58. Birkhäuser Verlag, pp. 221–264.
Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. Econometrica 50 (4), 987–1007.
Engle, R.F., 1983. Estimates of the variance of U.S. inflation based on the ARCH model. Journal of Money Credit and Banking 15 (3), 286–301.
Engle, R.F., 2002. Dynamic conditional correlation – a simple class of multivariate GARCH models. Journal of Business and Economic Statistics 20 (3), 339–350.
Engle, R.F., Jondeau, E., Rockinger, M., 2014. Systemic Risk in Europe. Review of Finance, forthcoming.
Fernández, C., Steel, M.F.J., 1998. On bayesian modelling of fat tails and skewness. Journal of the American Statistical Association 93 (441), 359–371.
Fischer, M., Köck, C., Schlüter, S., Weigert, F., 2009. An empirical analysis of multivariate copula models. Quantitative Finance 9 (7), 839–854.
Garcia, R., Tsafack, G., 2011. Dependence structure and extreme comovements in international equity and bond markets. Journal of Banking and Finance 35 (8), 1954–1970.
Giacomini, E., Härdle, W., Spokoiny, V., 2009. Inhomogeneous dependence modeling with time-varying copulae. Journal of Business and Economic Statistics 27 (2), 224–234.
Greenlaw, D., Kashyap, A., Schoenholtz, K., Shin, H., 2012. Stressed Out: Macroprudential Principles for Stress Testing, Chicago Booth School of Business, Working Paper No. 71.
Haldane, A., 2009. Why Banks Failed the Stress Test, Speech to Marcus-Evans Conference on Stress Testing, London, February 9–10.
Hansen, B.E., 1994. Autoregressive conditional density estimation. International Economic Review 35 (3), 705–730.
Heinen, A., Valdesogo, A., 2009. Asymmetric CAPM Dependence for Large Dimensions: The Canonical Vine Autoregressive Model, Center for Operations Research and Econometrics, Université Catholique de Louvain, CORE Discussion Papers No. 2009069.
Hirtle, B., Schuermann, T., Stiroh, K., 2009. Macroprudential Supervision of Financial Institutions: Lessons from the SCAP. Federal Reserve Bank of New York Staff Report No. 409.
Hong, Y., Tu, J., Zhou, G., 2007. Asymmetries in stock returns: statistical tests and economic evaluation. The Review of Financial Studies 20 (5), 1547–1581.
Hua, L., Joe, H., 2011. Tail order and intermediate tail dependence of multivariate copulas. Journal of Multivariate Analysis 102 (10), 1454–1471.
International Monetary Fund (IMF), 2012. Macrofinancial Stress Testing, Principles and Practices, Working Paper, August 22.
Joe, H., 1993. Parametric families of multivariate distributions with given margins. Journal of Multivariate Analysis 46 (2), 262–282.
Joe, H., 1996. Families of m-variate distributions with given margins and m(m-1)/2 biariate dependence parameters. In: Räuschendorf, L., Schweitzer, B., Taylor, M., (Eds.), Distributions with Fixed Marginals and Related Topics, Lecture Notes-Monograph Series, vol. 28, pp. 120–41.
Joe, H., 1997. Multivariate Models and Dependence Concepts. Chapman and Hall, New York.
Joe, H., Hu, T., 1996. Multivariate distributions from mixtures of max-infinitely divisible distributions. Journal of Multivariate Analysis 57 (2), 240–265.
Joe, H., Li, H., Nikoloulopoulos, A.K., 2010. Tail dependence functions and vine copulas. Journal of Multivariate Analysis 101 (1), 252–270.
Jondeau, E., 2010. Asymmetry in Tail Dependence of Equity Portfolios, FINRISK, Working Paper No. 658.
Jondeau, E., Rockinger, M., 2006. The copula-GARCH model of conditional dependencies: an international stock market application. Journal of International Money and Finance 25 (5), 827–853.
Jondeau, E., Poon, S.-H., Rockinger, M., 2007. Financial Modeling under Non-Gaussian Distributions. Springer, London.
Jones, M.C., Faddy, M.J., 2003. A skew extension of the t-distribution, with applications. Journal of Royal Statistical Society: Series B 65 (1), 159–174.
Kim, J., Finger, C.C., 2000. A stress test to incorporate correlation breakdown. RiskMetrics Journal 1, 61–75.
Kotz, S., Nadarajah, S., 2004. Multivariate *t* Distributions and their Applications. Cambridge University Press, Cambridge.
Kupiec, P.H., 1998. Stress-testing in a value at risk framework. Journal of Derivatives 6 (1), 7–24.
Kurowicka, D., Cooke, R., 2006. Uncertainty Analysis with High Dimensional Dependence Modelling. Wiley, Chichester.
Kurowicka, D., Joe, H., 2011. Dependence Modeling: Vine Copula Handbook. World Scientific, London.
Longin, F., 2000. From VaR to stress testing: the extreme value approach. Journal of Banking and Finance 24, 1097–1130.
Longin, F., Solnik, B., 1995. Is the correlation in international equity eeturns constant: 1960–1990? Journal of International Money and Finance 14 (1), 3–26.
Longin, F., Solnik, B., 2001. Extreme correlation of international equity markets. The Journal of Finance 56 (2), 649–676.
Mandelbrot, B., 1963. New methods of statistical economics. The Journal of Political Economy 71 (5), 421–440.
Mandelbrot, B., 1963. The variation of certain speculative prices. Journal of Business 36 (4), 394–419.
Manner, H., Reznikova, O., 2012. A survey on time-varying copulas: specification, simulations and Estimation. Econometric Reviews 31 (6), 654–687.
McNeil, A.J., Frey, R., 2000. Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. Journal of Empirical Finance 7 (3–4), 271–300.
McNeil, A.J., Nešlehová, J., 2009. Multivariate archimedean copulas, d-monotone functions and l1-norm symmetric distributions. The Annals of Statistics 37 (5B), 3059–3097.
McNeil, A.J., Saladin, T., 1997. The peaks over thresholds method for estimating high quantiles of loss distributions. In: Proceedings of 28th International ASTIN Colloquium.
McNeil, A.J., Smith, A.D., 2012. Multivariate stress scenarios and solvency. Insurance: Mathematics and Economics 50 (3), 299–308.
McNeil, A.J., Frey, R., Embrechts, P., 2005. Quantitative Risk Management: Concepts, Techniques and Tools. Princeton University Press, Princeton, New Jersey.
Mendes, B.V.d.M., Semeraro, M.M., Leal, R.P.C., 2010. Pair-copulas modeling in finance. Financial Markets and Portfolio Management 24 (2), 193–213.
Min, A., Czado, C., 2010. Bayesian inference for multivariate copulas using pair-copula constructions. Journal of Financial Econometrics 8 (4), 511–546.
Min, A., Czado, C., 2011. Bayesian model selection for D-vine pair-copula constructions. Canadian Journal of Statistics 39 (2), 239–258.
Nelsen, R.B., 2006. An Introduction to Copulas. Springer, New York.
Nelson, D.B., 1991. Conditional heteroskedasticity in asset returns: a new approach. Econometrica 59 (2), 347–370.
Nikoloulopoulos, A.K., Joe, H., Li, H., 2011. Vine copulas with asymmetric tail dependence and applications to financial return data. Computational Statistics and Data Analysis 56 (11), 3659–3673.
Patton, A., 2004. On out-of-sample importance of skewness and asymmetric dependence for asset allocation. Journal of Financial Econometrics 2 (1), 130–168.
Patton, A., 2006. Estimation of multivariate models for time series of possibly different lengths? Journal of Applied Econometrics 21 (2), 147–173.
Patton, A., 2006. Modelling asymmetric exchange rate dependence. International Economic Review 47 (2), 527–556.
Patton, A., 2009. Copula-based models for financial time series. In: Andersen, T., Davies, R., Kreiss, J., Mikosch, T. (Eds.), Handbook of Financial Time Series. Springer, New York, pp. 767–785.
Praetz, P.D., 1972. The distribution of share price changes. The Journal of Business 45 (1), 49–55.
Scarrott, C., MacDonald, A., 2012. A review of extreme value threshold estimation and uncertainty quantification. Statistical Journal 10 (1), 33–60.
Sklar, A., 1959. Fonctions de Répartition à n Dimensions et leurs Marges. Publications de l'Institut de Statistique de l'Université de Paris.
Stöber, J., Czado, C., 2012. Detecting Regime Switches in the Dependence Structure of High Dimensional Financial Data, mimeo.
van den End, J.W., 2010. Liquidity stress-tester: a macro model for stress-testing banks' liquidity risk. CESifo Economic Studies 56 (1), 38–69.
van den End, J.W., 2012. Do basel III and unconventional monetary policy work? Applied Financial Economics 22 (15), 1233–1257.
Vuong, Q.H., 1989. Ratio tests for model selection and non-nested hypotheses. Econometrica 57 (2), 307–333.
Zhao, X., Scarrott, C.J., Oxley, L., 2011. GARCH dependence in extreme value models with bayesian inference. Mathematics and Computers in Simulation 81, 1430–1440.