

Dependence Modelling in Insurance via Copulas with Skewed Generalised Hyperbolic Marginals

Vitali Alexeev^a, Katja Ignatieva^{b,*}, Thusitha Liyanage^{c,**}

^a*Finance Discipline Group, UTS Business School, University of Technology Sydney, Sydney, NSW 2007, Australia*

^b*School of Risk and Actuarial Studies, Business School, UNSW Australia, Sydney, NSW 2052, Australia*

^c*Portfolio and Market Risk Management Department, Commonwealth Bank of Australia, Sydney, NSW 2000, Australia*

Abstract

This paper investigates dependence among insurance claims arising from different lines of business (LoBs). Using bivariate and multivariate portfolios of losses from different LoBs, we analyse the ability of various copulas in conjunction with skewed generalised hyperbolic (GH) marginals to capture the dependence structure between individual insurance risks forming an aggregate risk of the loss portfolio. The general form skewed GH distribution is shown to provide the best fit to univariate loss data. When modelling dependency between LoBs using one-parameter and mixture copula models, we favour models that are capable of generating upper tail dependence, that is, when several LoBs have a strong tendency to exhibit extreme losses simultaneously. We compare the selected models in their ability to quantify risks of multivariate portfolios. By performing an extensive investigation of the in- and out-of-sample Value-at-Risk (VaR) forecasts by analysing VaR exceptions (i.e., observations of realised portfolio value that are greater than the estimated VaR), we demonstrate that the selected models allow to reliably quantify portfolio risk. Our results provide valuable insights with regards to the nature of dependence and fulfils one of the primary objectives of the general insurance providers aiming at assessing total risk of an aggregate portfolio of losses when LoBs are correlated.

Keywords: Dependence modelling, copula, insurance losses, skewed generalised hyperbolic distribution

JEL Classification: G22, C46, C15

*Corresponding author

**The opinions expressed in this article are the author's own and do not reflect the view of the Commonwealth Bank of Australia.

Email addresses: Vitali.Alexeev@uts.edu.au (Vitali Alexeev), k.ignatieva@unsw.edu.au (Katja Ignatieva), thusitha.liyanage@cba.com.au (Thusitha Liyanage)

1. Introduction

One of the main objectives of general insurance providers is to determine a risk capital needed to cover a series of simultaneous claims arising from losses in an unfavourable worst case scenario. When the total loss of the portfolio is generated by the sum of dependent losses, it depends not only on the individual losses, but also on the relationship among the individual risk factors. The need for a reliable dependence modelling framework is captured in Figure 1, which shows the relationship among fire losses arising from profit, building and content lines of business (LOBs) in the left panel and between content and profit losses in the right panel. The data is obtained from Copenhagen reinsurance for the period from 1980 to 1990.¹ Strong positive relationship is evident between all individual risks, which is also confirmed when computing Pearson correlation r , Spearman's ρ and Kendall's τ (refer to Table 1).² Thus, the insurer should have a realistic view not only on the distribution of the marginal losses, but also on the dependence between losses arising from different LOBs for risk management purposes as together they determine adequate capital requirements. If the allocated capital requirement is underestimated, an insurance company may become insolvent, while an unnecessary high capital level introduces unnecessary costs.

The risk of insurance losses has been examined by several authors, which include [McNeil et al. \(2004\)](#), [Jorion \(2007\)](#), [Dowd and Blake \(2006\)](#) and [Alemayehu et al. \(2013\)](#) among many others. The impact of the dependency between marginal risks on the sum of these dependent risks has been assessed by [Denuit et al. \(1999\)](#), [Kaas et al. \(2000\)](#), [Denuit et al. \(2001\)](#), [Cossette et al. \(2002\)](#), [Bolancé et al. \(2008\)](#) and [Bolancé et al. \(2014\)](#). The findings reported in these works confirm that when determining allocation of risk capital to each LoB, it is essential to have a reliable dependence modelling framework that can capture the dependence structure between individual insurance risks forming an aggregate portfolio risk.

Several research papers have examined the connection between insurance risks in the context of risk quantification (see, for example, [Lee and Lin, 2012](#); [Salzmann and Wüthrich, 2012](#); [Merz et al., 2013](#); [Alm, 2016](#)), but very few have utilised copulas to capture the possibly non-linear dependence structure between insurance losses (see, e.g., [De Jong, 2012](#);

¹We use data from [McNeil \(1997\)](#) that represents log-transformation of fire losses (reported in units of one million Danish Krone) that result from building, contents and profit claims, for details refer to Section 4.1 on data description.

²All pairwise correlation coefficients reported in the table are statistically significant: Testing pairwise the null hypothesis of zero correlation at 5% significance level, we strongly reject the null with p-values ranging from zero (for the pair contents-profit and Pearson correlation) to 6.77×10^{-5} (for the pair contents-building and Kendall coefficient).

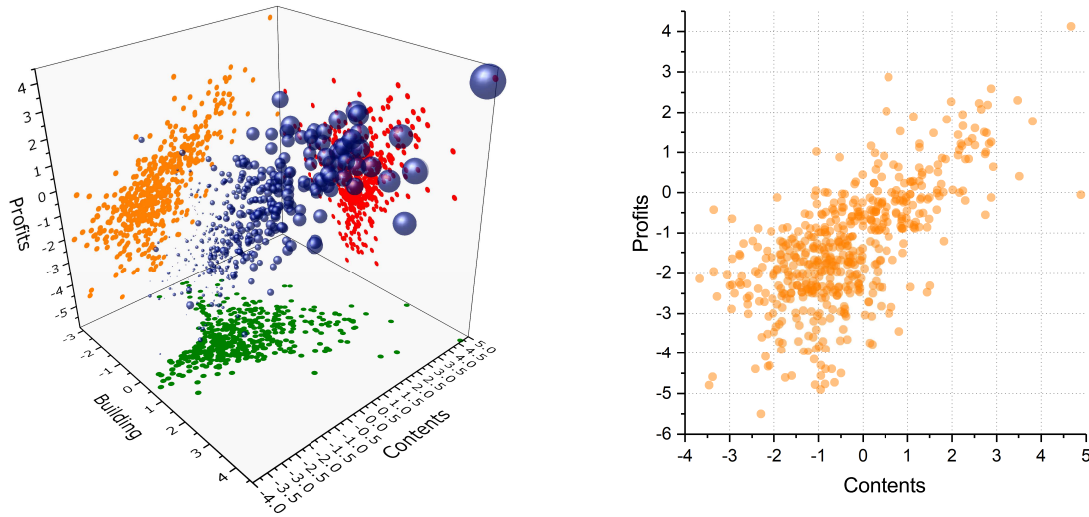


Figure 1: Profit, building and content losses data (in log-scale) are plotted in the **left panel** with XY , YZ , and XZ projections, representing our three-dimensional portfolio example. The size of the data points represents the total loss from all three LoBs. The right panel represents our two-dimensional portfolio example with a scatter of content and profit losses (in log-scale). The Danish fire data are from Copenhagen reinsurance for the period from 1980 to 1990.

Diers et al., 2012; Bolancé et al., 2014; Albrecher et al., 2017). Moreover, far less effort has been directed towards answering what marginal distributions ought to be used when modelling individual insurance losses. In Solvency II³ the standard formula for solvency capital requirement (SCR) suggests log-normal distribution for losses (i.e., normal distribution for logarithmic losses) in conjunction with parameter calibration to get an SCR corresponding to the 99.5% VaR of the insurer’s one-year loss distribution (refer to EIOPA, 2014). However, as demonstrated in the literature, insurance losses exhibit extreme tails and are severely right skewed (see, e.g., McNeil (1997), Embrechts et al. (2002), Bernardi et al. (2012), Eling (2012) and Miljkovic and Grün (2016)). Furthermore, even after applying a logarithmic transformation, the data is still positively right-skewed and cannot be reliably captured using normal distribution. This has been recognised by several authors (see, e.g., Lane (2000) and Ver-nic (2006)) and reconfirmed in this paper. As a consequence, skewed distributions have emerged in the insurance literature. Bolancé et al. (2008) and Eling (2012) are the pioneering papers that consider skewed normal and skewed Student-t distributions, respectively, to fit the insurance claims data and conduct risk analysis for non-life insurance portfolios.

To address the need of a reliable model for the marginals and a reliable non-linear dependence structure between losses, this paper goes beyond considering skewed elliptical

³The Solvency II is a Directive in European Union law that codifies and harmonises the EU insurance regulation. Primarily this concerns the amount of capital that EU insurance companies must hold to reduce the risk of insolvency.

Table 1: Pearson correlation coefficient r , Spearman's ρ and Kendall's τ computed for the profit, building and content losses are based on the Danish fire data obtained from Copenhagen reinsurance for the period from 1980 to 1990.

	Building	Contents	Profits
Panel A: Pearson correlation coefficient r			
Building	1.000000	0.282251	0.306523
Contents	0.282251	1.000000	0.667119
Profits	0.306523	0.667119	1.000000
Panel B: Spearman's ρ			
Building	1.0000000	0.1855800	0.2925118
Contents	0.1855800	1.0000000	0.6437903
Profits	0.2925118	0.6437903	1.0000000
Panel C: Kendall's τ			
Building	1.0000000	0.1173472	0.2009090
Contents	0.1173472	1.0000000	0.4621741
Profits	0.2009090	0.4621741	1.0000000

distributions (such as skewed Normal and skewed Student-t), and investigates a broader class of *skewed generalised hyperbolic* (GH) distributions. The skewed GH distributions are used as marginal distributions in conjunction with various copula models in context of the multivariate risk analysis for non-life insurance portfolios. Thus, the objectives of our paper can be formulated as follows.

The *first objective* of this paper is to investigate which marginal distributions from the class of skewed GH distributions capture reliably insurance losses that are severely right-skewed and exhibit excess kurtosis. As documented in [Nelsen \(2006\)](#), the choice of the appropriate marginal distributions may be crucial in modelling the dependence behaviour between variables. To model distributions of the marginals we consider the family of skewed GH distributions, which incorporates skewed Student-t, variance gamma (VG), normal inverse gaussian (NIG) and hyperbolic (HYP) distributions as special cases. We aim to identify the best performing distribution from the skewed GH family as a marginal distribution for univariate losses.

Our *second objective* is to find an appropriate multivariate model to capture the dependence structure that is essential to reliable estimation of the total risk of portfolio of losses. In order to model the dependency between insurance risks we resort to copulas, which are multivariate distribution functions allowing separation of the marginal distributions (i.e. distributions of losses from profit, losses from building and losses from content) from the multivariate dependence structure, see [Nelsen \(2006\)](#), [Embrechts \(2009\)](#), [Genest et al. \(2009\)](#). We then fit several copula models to the two-dimensional portfolio constructed of content and profit losses and a three-dimensional portfolio constructed of building, content and profit

losses. The considered models include one-parameter copulas: Clayton, survival Clayton, Gaussian, Gumbel, survival Gumbel, Student-t; as well as mixture copula models such as Clayton - survival Clayton mixture, Clayton - Gumbel mixture, survival Clayton - survival Gumbel mixture, Gumbel - survival Gumbel mixture. We aim to identify the best performing copula model for dependence modelling between individual losses in an aggregate portfolio.

In addition to providing an extensive analysis of the copula models with skewed GH marginals, our *final goal* is to test the performance of the applied copula models with selected marginals in the in- and out-of-sample risk management application where we estimate the Value-at-Risk (VaR) and the expected shortfall (ES) for bivariate and multivariate portfolios of losses arising from different LoBs. Using Monte Carlo simulations, we perform an extensive analysis by investigating the number of exceptions (i.e., observations of realised portfolio value that are greater than the estimated VaR), the size-adjusted loss function, and test for unconditional coverage and independence of VaR exceptions. Our results are of importance to market participants wishing to select the best performing model for computing total risk of an aggregate portfolio of losses, assuming that business lines are correlated.

Our findings highlight important insights into the structure of dependence between insurance losses arising from different LoBs as well as their univariate behaviour, and provide guidance to insurance companies on the risks inherited in the portfolios of such losses. When assessing the quality of fit to the univariate losses data arising from building, contents and profit lines, we demonstrate that all special cases from the family of skewed GH distributions are capable to reliably capture a heavy tailed and right skewed insurance data. We find that the most general form GH distribution outperforms all other distributions, which is due to its flexibility achieved through the number of estimated parameters. Regarding the nature of dependence, we find that those models that are capable of generating upper tail dependence (i.e. when LoBs have a strong tendency to jointly exhibit extreme losses) provide the best multivariate fit to the data. These models include Clayton and survival Clayton mixture, survival Clayton and survival Gumbel mixture, as well as one-parameter models survival Clayton copula and Gumbel copula. Finally, when comparing these copula models in their ability to quantify risks of the two- or three-dimensional portfolios we find that for both, in- and out-of-sample analysis, the one-parameter Gumbel copula in conjunction with skewed generalised hyperbolic marginals leads to the best performance across all considered models, resulting in the most reliable VaR and ES forecasts. Our results are important for insurance companies wishing to reliably quantify portfolio risks, assuming correlation between insurance business lines within this portfolio.

The remainder of the paper is organised as follows. Section 2 provides an overview of

multivariate copulas. Section 3 introduces the family of skewed generalised hyperbolic distributions that will be used as marginal distributions. Empirical analysis, which includes results on the parametric estimation of marginal distributions and the copula parameters is presented in Section 4. Section 5 deals with quantification of risks of multivariate portfolios using Value-at-Risk and expected shortfall risk measure, and provides comprehensive analysis of VaR exception. Section 6 concludes the paper.

2. Copula methodology

In this section we provide an overview of parametric copula families that will be used in our empirical analysis. Copulas are multivariate distributions which allow to connect d one-dimensional uniform-(0,1) marginals to a joint cumulative distribution. Pursuant to *Sklar's theorem* (see [Joe, 1997](#)), if F is a d -dimensional distribution function with marginals $F_1 \dots, F_d$, then there exists a copula C with

$$F(x_1, \dots, x_d) = C\{F_1(x_1), \dots, F_d(x_d)\} \quad (2.1)$$

for every $x_1, \dots, x_d \in \overline{\mathbb{R}}$. Alternatively, if we let $u_1 = F_1(x_1), \dots, u_d = F_d(x_d)$, then Eq. (2.1) can be written as

$$C(u_1, \dots, u_d) = F(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)). \quad (2.2)$$

Copulas provide a very flexible way to capture the dependence structure between random variables. They allow to model non-linear dependencies that the linear correlation coefficient fails to capture, see e.g. [Embrechts et al. \(2001\)](#); [Dias \(2004\)](#).

Additionally, in our empirical analysis, we resort to various survival copulas. A *survival copula* C^* corresponding to a copula C is defined as:

$$\bar{F}(x_1, \dots, x_d) = C^*\{\bar{F}_1(x_1), \dots, \bar{F}_d(x_d)\}, \quad (2.3)$$

where $\bar{F}(x_1, \dots, x_d) = P(X_1 > x_1, \dots, X_d > x_d)$. For the bivariate case, C^* is given by

$$C^*(u_1, u_2) = u_1 + u_2 - 1 + C(1 - u_1, 1 - u_2), \quad (2.4)$$

see [Nelsen \(2006\)](#).

When modelling a non-linear dependence, one may be particularly interested in modelling the occurrence of joint extreme events that can be captured by the coefficient of tail dependence. For insurance markets this scenario refers to the situation when extreme losses

occur simultaneously in several LoBs (in our case, when extreme building, contents and profits losses occur simultaneously).

The tail dependence measure describes the amount of comovements in the tails of the distributions. For a pair of uniform variables (U_1, U_2) on the unit square $[0, 1]^2$, we can define the upper and the lower tail dependence coefficients as

$$\lambda_u = \lim_{u \rightarrow 1^-} P(U_1 > u | U_2 > u) = \lim_{u \rightarrow 1^-} \frac{1 - 2u + C(u, u)}{1 - u} \quad (2.5)$$

and

$$\lambda_l = \lim_{u \rightarrow 0^+} P(U_1 \leq u | U_2 \leq u) = \lim_{u \rightarrow 0^+} \frac{C(u, u)}{u}, \quad (2.6)$$

respectively.

If $\lambda_u \in (0, 1]$, then U_1 and U_2 are said to be asymptotically dependent in the upper tail, and if $\lambda_u = 0$, then U_1 and U_2 are said to be asymptotically independent in the upper tail. Similarly, if $\lambda_l \in (0, 1]$ or $\lambda_l = 0$, then U_1 and U_2 are said to be asymptotically dependent, or independent, respectively, in the lower tail. The tail dependence measures for mixture copula models are detailed in [Hu \(2006\)](#).

We focus our attention on two popular copula families: the *elliptical family of copulas* and the *Archimedean family of copulas*. *Elliptical copulas* have dependence structures generated by elliptical distributions, including, for example, the normal and Student-t distributions (see [Lindskog et al., 2003](#)). The Gaussian copula generates the dependence structure given by the multivariate normal distribution, allowing the choice of any marginal distribution, for example, normal or Student-t. For the case of normal marginals, that is, if $X_j \sim N(0, 1)$ and $X = (X_1, \dots, X_d)^\top \sim N_d(0, \Psi)$, where Ψ denotes a correlation matrix, an explicit expression for the Gaussian copula is given by

$$C_{\Psi}^{Ga}(u_1, \dots, u_d) = F_X\{\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_d)\}. \quad (2.7)$$

Thus, combining normal marginals via a Gaussian copula leads to the multivariate normal distribution. The Student-t copula generates the dependence structure from the multivariate Student-t distribution. If $X = (X_1, \dots, X_d)^\top \sim t_d(\nu, \mu, \Sigma)$ has a multivariate Student-t distribution with ν degrees of freedom (d.f.), mean vector μ and positive-definite dispersion or scatter matrix Σ , the Student-t copula is given by

$$C_{\nu, \Psi}^t(u_1, \dots, u_d) = t_{\nu, \Psi}\{t_{\nu}^{-1}(u_1), \dots, t_{\nu}^{-1}(u_d)\}, \quad (2.8)$$

where t_v^{-1} is the quantile function from the univariate Student-t distribution, and Ψ is the correlation matrix associated with Σ .⁴ The Student-t copula generates symmetric tail dependence with the tail dependence coefficients defined by

$$\lambda_u = \lambda_l = 2 \left(-t_{v+1} \sqrt{(v+1)(1-\rho)/(1+\rho)} \right), \quad (2.9)$$

where t_v denotes the Student-t distribution function, v is the number of d.f., and ρ is the correlation coefficient. Further details in modelling dependence using elliptical distributions can be found in [Hult and Lindskog \(2002\)](#), [Fang et al. \(2002\)](#) and [Frahm et al. \(2003\)](#). Applications in risk management are considered, for instance, in [Breymann et al. \(2003\)](#) and [Dias and Embrechts \(2010\)](#).

Furthermore, we apply the Gumbel and Clayton copulas from the family of *Archimedean copulas*. The *Clayton copula* with the dependence parameter $\theta \in (0, \infty)$ is defined by

$$C_\theta(u_1, \dots, u_d) = \left\{ \left(\sum_{j=1}^d u_j^{-\theta} \right) - d + 1 \right\}^{-1/\theta}. \quad (2.10)$$

As $\theta \rightarrow \infty$, the dependence becomes maximal whilst the independence in the tail is achieved when $\theta \rightarrow 0$. The Clayton copula can mimic lower tail dependence with the tail dependence coefficient of $\lambda_l = 2^{-1/\theta}$ but no upper tail dependence, that is, $\lambda_u = 0$.

The *Gumbel copula* with dependence parameter $\theta \in [1, \infty)$ is given by

$$C_\theta(u_1, \dots, u_d) = \exp \left[- \left\{ \sum_{j=1}^d (-\log u_j)^\theta \right\}^{1/\theta} \right]. \quad (2.11)$$

For $\theta > 1$ this copula generates an upper tail dependence with the tail dependence coefficient of $\lambda_u = 2 - 2^{1/\theta}$ but no dependence in the lower tail, that is, $\lambda_l = 0$. For $\theta = 1$ it reduces to the product copula (i.e., exhibiting independence in the tail): $C_\theta(u_1, \dots, u_d) = \prod_{j=1}^d u_j$. Maximal dependence is achieved when $\theta \rightarrow \infty$. Applications of Archimedean copulas to financial markets and insurance have been considered in [Dias \(2004\)](#), [Wu et al. \(2007\)](#) and [Chavez-Demoulin and Embrechts \(2010\)](#).

Apart from the Archimedean and elliptical families of copulas discussed above we con-

⁴Since copula functions remain invariant under any series of strictly increasing transformations of X (e.g., standardisation of the marginal distributions), the copula of a $t_d(\nu, \mu, \Sigma)$ distribution is identical to that of a $t_d(\nu, 0, \Psi)$. For details, refer to [Nelsen \(2006\)](#).

sider some *mixture models* of Archimedean copulas as introduced in Joe (1993). Mixture copulas usually take the form of a convex combination of two or more copulas. Denoting C^A and C^B copulas with dependence parameters θ_1 and θ_2 , respectively, the mixture model takes the following form:

$$C_X(u_1, \dots, u_d, \boldsymbol{\theta}) = \theta_3 C_X^A(u_1, \dots, u_d, \theta_1) + (1 - \theta_3) C_X^B(u_1, \dots, u_d, \theta_2). \quad (2.12)$$

Empirical applications of mixture copulas in financial and insurance markets can be found in Dias (2004) and Hu (2006). In our empirical analysis we will consider four mixture models of copulas, namely the Clayton & survival Clayton, Clayton & Gumbel, survival Clayton & survival Gumbel and the Gumbel & survival Gumbel mixture copulas.

Note that the literature suggests different approaches to the estimation of copulas (e.g., see Joe, 1997; Cherubini et al., 2004). In our empirical analysis, we estimate copula parameters using the so-called *inference for marginals* (IFM) method, which is a sequential two-step maximum likelihood method. The method recommends estimating marginal parameters in the first step, and then, substitute those into the copula to obtain the pseudo log-likelihood function, which is then maximised with respect to the copula dependence parameter θ . For details on the IFM estimation refer to Appendix (Section 7). An overview of different estimation techniques, including IFM, can be found in McLeish and Small (1988) and Joe (1997).

3. Specification of marginal distributions: Generalised hyperbolic family

Barndorff-Nielsen and Stelzer (1997) has introduced the family of generalised hyperbolic (GH) distributions with its general form discussed in Barndorff-Nielsen and Stelzer (2005) and McNeil et al. (2015). GH distributions provide a sufficiently rich class of distributions, which can be represented as a normal mean-variance mixture where the mixture variable follows the generalised inverse gaussian (GIG) distribution. A multivariate representation of skewed GH distributions as well as its applications in insurance risk management can be found in Ignatieva and Landsman (2017). For the purpose of our empirical analysis, we will only require the univariate skewed GH distributions to model marginals (insurance resulting from individual days on building, contents and profits). The univariate skewed GH distributions, its special and limiting cases are presented below.

The distribution of a random variable $X \sim GH(\lambda, \chi, \psi, \mu, \sigma, \gamma)$ is characterised by its

probability density function (pdf)

$$f_X(x) = \frac{\psi^\lambda (\psi + \gamma^2 \sigma^{-2})^{\frac{1}{2} - \lambda} (\sqrt{\chi \psi})^{-\lambda}}{\sqrt{2\pi\sigma} K_\lambda(\sqrt{\chi \psi})} \times \frac{K_{\lambda - \frac{1}{2}}\left(\sqrt{(\chi + Q)(\psi + \gamma^2 \sigma^{-2})}\right)}{\left(\sqrt{(\chi + Q)(\psi + \gamma^2 \sigma^{-2})}\right)^{\frac{1}{2} - \lambda}} e^{(x - \mu)\gamma\sigma^{-2}} \quad (3.1)$$

where parameters $\chi \geq 0$, $\psi \geq 0$ and the quadratic form corresponds to $Q = (x - \mu)^2 \sigma^{-2}$. Function $K_\lambda(\cdot)$ denotes a modified Bessel function of the third kind with index λ (refer to [Abramowitz and Stegun, 1972](#) for more details). In Eq. (3.1), μ is interpreted as a location parameter, σ as a dispersion parameter and γ as a skewness parameter. If γ is set to zero, the distribution is symmetric around μ . We notice that a linear transformation of GH random variable is again a GH random variable, see [McNeil et al. \(2015\)](#). Special and limiting cases of GH distributions are discussed below.

3.1. Special and limiting cases of GH distributions

The GH family contains several special and limiting cases. The limiting cases include (skewed) Student-t distribution obtained by setting $\lambda < 0$, $\psi = 0$, $\chi = \nu - 2$ and $\nu = -2\lambda$, see [Praetz \(1972\)](#); and Variance Gamma (VG) distribution obtained when $\lambda > 0$, $\chi = 0$ and $\psi = 2\lambda$, see [Madan and Seneta \(1990\)](#). In what follows the densities of these limiting cases are presented. The other two cases, a hyperbolic (HYP) and a normal inverse gaussian (NIG) distributions are special cases that can be obtained from Eq. (3.1) by simply setting the parameter λ to 1 or -0.5, respectively. For details refer to [Eberlein and Keller \(1995\)](#) and [Barndorff-Nielsen and Stelzer \(2005\)](#).

For the non-symmetric *Student-t distribution* ($\gamma \neq 0$) the density results from Eq. (3.1) when setting $\psi = 0$, $\chi = \nu - 2$ and $\nu = -2\lambda$ and assuming $\lambda < 0$ ⁵:

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{(\nu - 2)^{\frac{\nu}{2}} (\boldsymbol{\gamma}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\gamma})^{\frac{\nu+d}{2}}}{(2\pi)^{\frac{d}{2}} |\boldsymbol{\Sigma}|^{\frac{1}{2}} \Gamma(\frac{\nu}{2}) 2^{\frac{\nu}{2} - 1}} \times \frac{K_{\frac{\nu+d}{2}}(\sqrt{(\nu - 2 + Q(\mathbf{x})) \boldsymbol{\gamma}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\gamma}})}{(\sqrt{(\nu - 2 + Q(\mathbf{x})) \boldsymbol{\gamma}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\gamma}})^{\frac{\nu+d}{2}}} e^{(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\gamma}}. \quad (3.2)$$

For the case when $\gamma = 0$ an asymptotic expansion of $K_\lambda(x)$ reduces the density in Eq. (3.2) to a classical form of non-skewed Student-t distribution, refer to [Ignatieva and Landsman \(2015\)](#) for further details. The *Variance Gamma (VG)* density can be obtained by setting $\chi = 0$, $\psi = 2\lambda$ and assuming $\lambda > 0$ in Eq. (3.1):

⁵For simplicity of notation we omit parameters in the notion of $f(x, \cdot)$.

$$f_{\mathbf{x}}(\mathbf{x}) = \frac{2\lambda^\lambda (2\lambda + \boldsymbol{\gamma}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\gamma})^{\frac{d}{2} - \lambda}}{(2\pi)^{\frac{d}{2}} |\boldsymbol{\Sigma}|^{\frac{1}{2}} \Gamma(\lambda)} \times \frac{K_{\lambda - \frac{d}{2}}(\sqrt{Q(\mathbf{x})(2\lambda + \boldsymbol{\gamma}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\gamma})})}{(\sqrt{Q(\mathbf{x})(2\lambda + \boldsymbol{\gamma}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\gamma})})^{\frac{d}{2} - \lambda}} e^{(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\gamma}}. \quad (3.3)$$

4. Empirical analysis

In this section we demonstrate how the methodology outlined above can be applied for modelling dependence between insurance losses using copulas with (skewed) generalised hyperbolic marginals. First, we describe the data used in our empirical analysis, then proceed with the estimation of the marginals and estimation of the copula functions.

4.1. Data

We use *Danish fire data* collected and used in [McNeil \(1997\)](#) to estimate the tails of loss severity distributions using extreme value theory. The data are obtained from Copenhagen reinsurance and comprise a set of fire losses that result from building (1990 non-zero observations), contents (1679 non-zero observations) and profit (616 non-zero observations) claims for the period from 1980 to 1990. In order to model dependency, we have selected only those observations for which all three entries (building, contents and profit losses) are non-zero. This results in a total of 517 observations. The data are in units of one million Danish Krone and represent large losses, where an aggregate loss (i.e. the sum of building, profit and fire losses) typically exceeds one million Danish Krone. Similarly to [McNeil \(1997\)](#) we use the log-transformation.⁶ Figure 2 shows the empirical cdf of the total log-loss data evaluated at each of the data points. The empirical cdf for a sample of size n is defined as $F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{X_i \leq x\}}$, where $\mathbf{1}_{\{\cdot\}}$ is an indicator function taking the value of one if $X_i \leq x$ and zero otherwise. The vertical lines represent the 95%, 99% and 99.5% quantiles of the empirical distribution. The x -axis is shown in log-scale and includes losses occurring beyond these extreme, which are chosen to show extreme behaviour in the tail. While these extreme observations will be used in estimating in- and out-of-sample Value-at-Risk and Expected Shortfall of an aggregate portfolio in Section 5, in our next section, Section 4.2, we discuss the fit of the univariate GH family of distributions to the data on building, contents, profits and total losses.

⁶In the following, where we say "losses" or "claims" we refer to logarithmic losses.

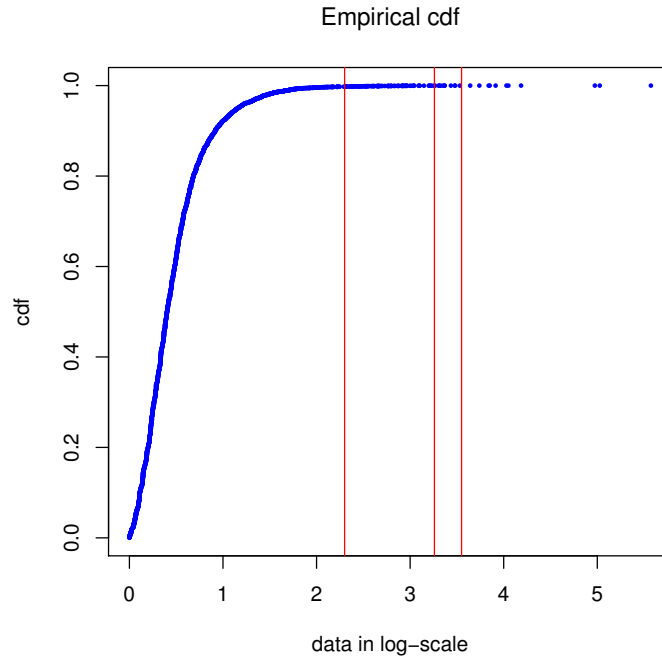


Figure 2: Empirical cdf for the (total) Danish fire losses (x-axis is in log-scale). Vertical lines represent the 95%, 99% and 99.5% quantiles of the empirical distribution, respectively.

4.2. Estimation of the marginals

In order to fit skewed GH distributions to the univariate data, i.e. to estimate model parameters for individual data on building, contents or profit losses, we use the maximum likelihood technique. For different special cases of GH family we fix those parameter values that correspond to the Student-t, VG, NIG and HYP distributions, respectively, and maximise the log-likelihood function with respect to the remaining parameters.⁷ We fit univariate GH distribution and its special cases to the logarithmic losses resulting from buildings, contents and profit claims, in order to identify the best performing distribution from the GH family. For comparison, we also fit normal distribution to the logarithmic losses, which can serve as a benchmark used in Solvency II to calculate the SCR as the 99.5% Value-at-Risk (VaR) of the insurer's one year loss distribution (refer to [EIOPA, 2014](#)). Table 2 reports estimated parameters for buildings, contents and profit losses in Panels A, B and C respectively. The results for the general form GH distribution are reported in the second column. Additionally, columns three to six contain estimated parameters for each of the considered special cases (Student-t, VG, NIG and HYP), and column seven reports the results

⁷We use R package on generalised hyperbolic distributions (**ghyp**) by David Luethi and Wolfgang Breyman from the Institute of Data Analysis and Process Design at Zurich University of Applied Sciences.

for the normal distribution.⁸ One observes that the skewness parameter γ for the general form GH distribution as well as its special cases is positive (if transformed from log-scale to a standard scale), indicating that the loss data is right skewed. The average loss resulting from building line in case of the GH distribution corresponds to $\exp(0.2255) = 1.2529$, which is above one million Danish Krone. The average contents and profit losses are substantively smaller, and correspond to $\exp(-1.7050) = 0.1817$ and $\exp(-1.9031) = 0.1491$ (i.e. 181 thousand and 149 thousand), respectively. The number of degrees of freedom ν in case of the Student-t distribution can be computed as $\nu = -2\lambda$.

In order to assess the quality of fit of each of the distributions we apply three goodness-of-fit tests proposed in [Stephens \(1974\)](#) and [D'Agostino and Stephens \(1986\)](#). We test the null hypothesis that the sample comes from a distribution $F(x)$, which is one of the distributions from the family of skewed GH distributions, that is, $H_0 : \hat{F}_n(x) = F(x)$, where $\hat{F}_n(x)$ represents the empirical cumulative density function (cdf) and $F(x)$ is the theoretical cdf of GH, Student-t, VG, NIG or HYP distribution. The required test statistics calculated from the z-values $z_i = F_n(x_i)$, $i = 1, \dots, n$ where $\hat{F}_n(x)$ contains estimated parameters, have the following form:⁹

1. The Anderson-Darling statistic A^2 is obtained as

$$A^2 = -n - 1/n \sum_{i=1}^n (2i - 1)(\log(z_i) + \log(1 - z_{n+1-i})). \quad (4.1)$$

2. The Kolmogorov statistic D is obtained as $D = \max(D^+, D^-)$ where $D^+ = \max_{1 \leq i \leq n} [(i/n) - z_i]$ and $D^- = \max_{1 \leq i \leq n} [z_i - (i - 1)/n]$. The modified form statistics which is proposed in [Stephens \(1974\)](#) together with the critical values corresponds to $D(\sqrt{n} + 0.12 + 0.11/\sqrt{n})$.
3. The Carmér-von Mises statistic W^2 is given by

$$W^2 = 1/12n + \sum_{i=1}^n (z_i - (2i - 1)/2n)^2. \quad (4.2)$$

The modified form statistic reported in [Stephens \(1974\)](#) is given by $(W^2 - 0.4/n + 0.6/n^2)(1 + 1/n)$.

⁸We compared our results with those obtained using pseudo data for estimating copula parameters. Our estimation procedure leads to the parameter estimates that are close to those specified in the simulations. Details are available from the authors upon request.

⁹Note, that observations are sorted in ascending order: $x_1 \leq x_2 \leq \dots \leq x_n$.

In addition to the above tests for the performance of each model fitted, we use the Akaike information criterion (AIC), introduced by Akaike (1974), to judge the quality of fit for each of the distributions:

$$AIC = -2\hat{l} + 2k. \quad (4.3)$$

In Eq. (4.3), \hat{l} denotes the maximised value of the log-likelihood and k is the number of parameters for the distribution fitted. Since AIC is defined as minus twice the log-likelihood plus the penalty term which accounts for the effective number of estimated parameters, smaller values of AIC indicate a better fit.

Table 2: Parameter estimates (and fixed parameter values) from the univariate fit of GH family of distributions and normal distribution to the building, contents and profits losses. For the GH family, parameters are computed using maximum likelihood estimation and refer to $X \sim GH(\lambda, \chi, \psi, \mu, \sigma, \gamma)$ with the pdf characterised by Eq. (3.1). In case of normal distribution, μ and σ denote mean and standard deviation, respectively.

Param.	GH	Student- <i>t</i>	VG	NIG	HYP	Normal
Panel A: Building						
λ	0.695969	-1.679011	1.088627	-0.500000	1.000000	-
χ	0.096606	1.358022	0.000000	0.654990	0.015613	-
ψ	1.702319	0.000000	2.177253	0.654990	2.062127	-
μ	0.225508	0.274102	0.238307	0.258622	0.215621	0.264184
σ	0.919386	0.991112	0.915322	0.929960	0.917694	0.916806
γ	0.038933	-0.010576	0.025916	0.005662	0.048738	-
Panel B: Contents						
λ	5.036248	-24.43728	5.046387	-0.500000	1.000000	-
χ	0.009183	46.87456	0.000000	8.437966	1.594113	-
ψ	10.08395	0.000000	10.09277	8.437966	4.191727	-
μ	-1.705063	-5.035216	-1.703853	-2.511307	-1.166411	-0.351062
σ	1.252766	4.684146	1.252835	1.174368	1.389138	1.387794
γ	1.353993	4.684146	1.352802	2.160381	0.824639	-
Panel C: Profits						
λ	6.132114	-10.33926	6.300678	-0.500000	1.000000	-
χ	0.133929	18.67852	0.000000	6.994209	4.716896	-
ψ	12.42330	0.000000	12.601360	6.994209	7.503495	-
μ	-1.903191	-2.073762	-1.906130	-1.953652	-1.857375	-1.354839
σ	1.441250	1.434851	1.441199	1.439395	1.457313	1.455903
γ	0.548302	0.719008	0.551495	0.599140	0.503120	-

The results for the test statistics and the AIC are reported in Table 3 for the univariate data on building, contents and profit losses in Panels A, B and C, respectively. The 5% and 1% critical values reported in the last two columns are obtained from Stephens (1974). We observe that normal distribution provides the poorest fit to the data; it is rejected at 1% significance level by all three statistics in case of building losses, and at 5% significance level by the Carmér-von Mises statistic in case of contents losses. Although normal distribution is not rejected in case of profit losses, it results in significantly larger values of all three test statistics and thus, the poorest fit to the data. We conclude that although normal distribution is extensively used in practice to describe distribution of log-losses, it fails to appropriately

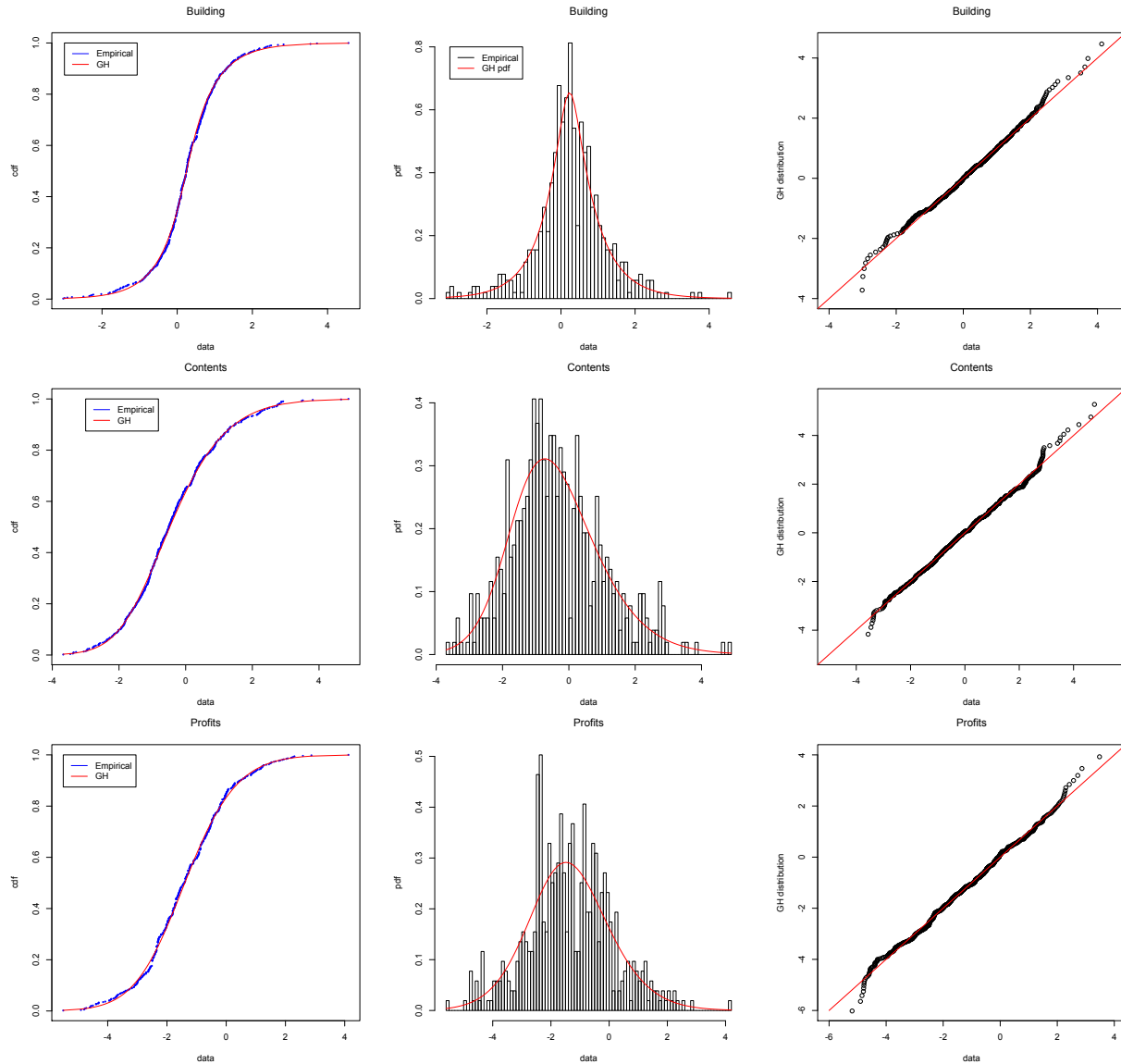


Figure 3: Left panels: Empirical vs. theoretical cdf; middle panels: histogram representing the empirical distribution (black) vs. GH pdf (red); right panels: QQ plots. Top, middle and low panels correspond to the building, contents and profit losses, respectively.

capture underlying characteristics of the data. At the same time, none of the theoretical GH distributions under assessment could be rejected for building, contents and profit losses, indicating a very good fit of the data with the distributions we selected. The smallest values of the test statistics are typically achieved for the general form GH distribution (second column). Figure 3 shows in the (i) left panels: empirical (blue) vs. theoretical GH (red) cdf, (ii) middle panels: histogram of losses (black) and the best performing fitted pdf of GH distribution (red); (iii) right panels: the quantile-to-quantile (QQ) which contrasts empirical quantiles against quantiles from the best performing GH distribution. One observes an excellent fit of the GH distribution to the building, contents and profit claims data, whereby

Table 3: Distributional tests based on Anderson-Darling, Kolmogorov and Carmér-von Mises statistics (D and W^2 are the modified form statistics), with critical values for the Danish fire losses data. * and ** indicate rejections of the hypothesised distribution at the 5% and 1% significance levels, respectively. Akaike information criterion (AIC) is also reported. None of the skewed GH distributions could be rejected while the two smallest values of the test statistics (indicating the best fit) are shown in **bold** numerals. General form GH distribution provides the best fit to the data.

Stat.	GH	Student- t	VG	NIG	HYP	Normal	Crit.5%	Crit.1%
Panel A: Building								
A^2	0.364672	0.437590	0.368478	0.349796	0.394148	5.165004**	2.492	3.857
D	0.685501	0.751948	0.683768	0.653612	0.736907	1.891970**	1.358	1.628
W^2	0.141017	0.158344	0.142800	0.143113	0.143985	0.977103**	0.461	0.743
AIC	=1318.231	1323.964	1318.349	1319.461	1318.286	1381.371	—	—
Panel B: Contents								
A^2	0.272998	0.392018	0.273097	0.332737	0.400620	2.374551	2.492	3.857
D	0.508442	0.573274	0.507646	0.545062	0.561523	1.341274	1.358	1.628
W^2	0.124142	0.146454	0.124151	0.135361	0.142049	0.502741*	0.461	0.743
AIC	1792.133	1792.528	1792.133	1792.262	1796.088	1810.040	—	—
Panel C: Profit								
A^2	0.471268	0.515215	0.470944	0.490424	0.490329	0.867412	2.492	3.857
D	0.755952	0.826130	0.755254	0.789400	0.808248	0.932532	1.358	1.628
W^2	0.152217	0.155600	0.152144	0.153686	0.156071	0.212967	0.461	0.743
AIC	1859.318	1859.850	1859.307	1859.609	1859.653	1869.580	—	—

even the tails of the losses are captured fairly well. The result obtained using the values of the AIC statistics are consistent with those reported for the A^2 , D and W^2 test statistics, indicating that the GH distribution results in the best fit to the data in the majority of the cases, with an exception of the VG for contents and building data, where it leads to a comparable fit. Overall, since all data are fit best by means of the GH distribution, we will use this distribution for all marginals entering the copula function, which will be discussed in the next section.

4.3. Copula analysis

In this section we fit various parametric copulas, assuming that marginals follow GH distribution (general form). As emphasised above, we use the IFM method which suggests that once marginals have been estimated, they are plugged-into the copula function, and the pseudo-likelihood function is maximised with respect to the dependence parameter. We consider various Archimedean and elliptical copula, which include Clayton, survival Clayton, Gaussian, Gumbel, survival Gumbel and Student- t ; as well as mixture copula models, which include Clayton & survival Clayton, Clayton & Gumbel, survival Clayton & survival Gumbel and Gumbel & survival Gumbel copulas. Copula parameters estimates with standard errors, together with the coefficients of lower and upper tail dependence, the Akaike Information Criterion (AIC) for model performance, and model ranking are reported in Table 4. We notice that in case of the mixture models, θ_1 and θ_2 are the dependence parameters for

the first and second terms of the mixture, respectively, while θ_3 is the parameter which gives the proportion of the first term in the mixture. For Student-t copula θ_1 is the dependence parameter and θ_2 is the number of degrees of freedom. The standard errors are reported in parentheses.

Table 4: Estimated copula dependence parameters with standard errors, coefficients of lower and upper tail dependence, AIC and model ranking for the two-dimensional portfolio consisting of content and profit losses, and a three-dimensional portfolio consisting of building, content and profit losses. In the case of Student-t copula θ_2 is the number of degrees of freedom. In case of the mixture models, θ_1 and θ_2 are the dependence parameters for the first and second terms of the mixture, respectively, while θ_3 is the parameter which gives the proportion of the first term in the mixture.

Copula model	$\hat{\theta}_1$ (s.e.)	$\hat{\theta}_2$ (s.e.)	$\hat{\theta}_3$ (s.e.)	λ_l	λ_u	AIC	Rank
Panel A: Portfolio of Content and Profit Losses							
Clayton	0.7126(0.0662)	-	-	0.3780	0.0000	-112.2661	(10)
Surv. Clayton	1.5723(0.1083)	-	-	0.0000	0.6435	-342.8681	(3)
Gaussian	0.6405(0.0218)	-	-	0.0000	0.0000	-271.0229	(8)
Gumbel	1.8604(0.0777)	-	-	0.0000	0.5485	-329.8440	(4)
Surv. Gumbel	1.6056(0.0535)	-	-	0.4601	0.0000	-183.1137	(9)
Student-t	0.6447(0.0233)	14.988(10.927)	-	0.0816	0.0816	-271.2900	(7)
Clayton and surv. Clayton	3.8904(2.0044)	1.6788(0.1295)	0.0775(0.0350)	0.0648	0.6104	-349.3858	(1)
Clayton and Gumbel	6.3754(7.6484)	1.8588(0.0674)	0.0094(0.0323)	0.0084	0.5429	-325.9332	(5)
Surv. Clayton and surv. Gumbel	1.7499(0.1531)	1.8811(0.4398)	0.8924(0.0525)	0.0596	0.6005	-347.9786	(2)
Gumbel and surv. Gumbel	1.8635(0.0724)	1.8550(2.6991)	0.9900(0.0592)	0.0054	0.5439	-325.5352	(6)
Panel B: Portfolio of Building, Content and Profit Losses							
Clayton	0.2128(0.0299)	-	-	0.0385	0.0000	-34.20380	(10)
Surv. Clayton	0.8377(0.0774)	-	-	0.0000	0.4371	-373.8009	(1)
Gaussian	0.3877(0.0334)	-	-	0.0000	0.0000	-208.4459	(8)
Gumbel	1.3124(0.0426)	-	-	0.0000	0.3042	-281.0174	(4)
Surv. Gumbel	1.2825(0.0269)	-	-	0.2831	0.0000	-108.9461	(9)
Student-t	0.3908(0.0261)	15.423(6.9814)	-	0.0161	0.0161	-212.3526	(7)
Clayton and surv. Clayton	0.0001(0.0001)	0.8377(0.0544)	0.0001(0.0007)	0.0000	0.4371	-369.7963	(2)
Clayton and Gumbel	0.0001(0.0001)	1.3124(0.0284)	0.0001(0.0008)	0.0000	0.3042	-277.0168	(5)
Surv. Clayton and surv. Gumbel	0.8467(0.0559)	1.0001(0.0001)	0.9900(0.0165)	0.0000	0.4366	-369.3233	(3)
Gumbel and surv. Gumbel	1.3169(0.0294)	1.0001(0.0001)	0.9900(0.0216)	0.0000	0.3042	-276.9506	(6)

We observe large values for the copula dependence parameters for the two-dimensional portfolio (Panel A of Table 4) consisting of content and profit losses, indicating strong dependence between content and profit losses. We notice that in the cases of mixture models, the mixture parameter θ_3 suggests allocating higher weight to either Gumbel or survival Clayton part in the mixture, suggesting the importance of modelling upper tail in the joint distribution of content and profit losses. This is also confirmed by examining tail dependence coefficients: Copulas that are able to generate upper tail dependence exhibit high values for the upper tail dependence coefficient λ_u . Copula dependence parameters θ_1 and θ_2 are typically lower for the three-dimensional portfolio consisting of building, content and profit losses (Panel B). However, we still observe that the entire weight in the mixture copula models is assigned to the Gumbel or survival Clayton parts, and coefficients of upper tail dependence are high for those copulas that are capable of generating upper tail dependence.

In order to quantify the performance of the considered copula models, we compute the

Akaike Information Criterion (AIC) introduced in [Akaike \(1974\)](#):

$$AIC = -2\ln(\hat{L}) + 2q, \quad (4.4)$$

where \hat{L} is the maximum value of the likelihood function, and q is the number of parameters for the distribution fitted. From the definition of the AIC it is obvious that smaller values of AIC indicate better fit. The last two columns of [Table 4](#) report the AIC and relative ranking of the models.

For the two-dimensional portfolio we observe that the best two models ranked based on AIC are Clayton & survival Clayton and survival Clayton & survival Gumbel mixtures, followed by the one-parameter survival Clayton and Gumbel copulas. Notice that both survival Clayton and Gumbel copulas are capable of generating upper tail dependence and appear in one-parameter copulas as well as in the mixture models. When these two copulas form parts of the mixture model, the survival Clayton and Gumbel parts are weighted with the highest weight. Note that the best performing mixture models allow for asymmetric tail dependence and the estimated models provide clearly higher estimates for the upper tail of λ_u of above 0.6, while the dependence in the lower tail (captured by λ_l) is much smaller, corresponding to around 0.06 for both mixtures. This confirms our findings that extreme losses in the upper tail of the distribution often occur jointly in both LoBs. On the other hand, there seems to be nearly no dependence across LoBs for losses in the left tail of the distribution. It is also worth noting that both elliptical copulas (Gaussian and Student-t) result in poor fit to the data. While it is not surprising that Gaussian copula cannot reliably capture the dependence structure due to its inability to control tail dependence, Student-t copula generates symmetric dependence in the upper and the lower tail, with $\lambda = \lambda_u = \lambda_l = 0.08$, while the number of degrees of freedom ($\nu = \theta_2$) is close to 15. Note that the degrees of freedom parameter of the Student-t copula has a significant impact on the fatness of the tails of the distribution. Therefore, with a decrease in ν , the tail dependence coefficient λ increases. On the other hand, when the degrees of freedom parameter increases, the multivariate distribution becomes more normal with less tail dependence, which is what we observe in our results. This explains why all results (dependence and tail dependence parameters and model ranking) reported for the Student-t copula are consistent with those for the Gaussian copula. The results for the three-dimensional portfolio reported in [Panel B of Table 4](#) are generally in line with those for the two-dimensional case, with the only difference in the order of ranking for the four best models: the survival Clayton copula is ranked first, followed by the two mixture models and the Gumbel copula. We conclude that the best fitting copulas for the insurance losses are those capable of capturing upper

tail dependence in the joint distribution of portfolio losses. Our results about the nature of dependence, and particularly, the strong tendency of several LoBs to jointly exhibit extreme losses are important for insurance companies that are interested in the dependence between losses for the purpose of risk management. We detail some of its applications in our next section.

5. Risk management applications

The dependence structure between insurance losses is of a great importance for risk management decisions of insurance providers who are interested in modelling extreme losses occurring simultaneously in several LoB. Risks from extreme losses can be quantified using appropriate risk measures. In this section we apply the techniques discussed earlier and estimate the in-sample and prediction of the out-of sample Value-at-Risk (VaR) and expected shortfall (ES) for the two- and three-dimensional portfolio examples constituting (a) content and building losses and (b) building, content and profit losses.

5.1. Backtesting procedure

In order to assess how different models compare with each other in terms of their risk management performance, we explore a set of complimentary measures. The first measure compares the percentage of exception observations (i.e., observations of realised portfolio values that are greater than the estimated $VaR_{1-\alpha}$) to the specified quantile level α . To formally define exception observation, we use a binary loss function proposed by [Lopez \(1998\)](#):

$$L = \begin{cases} 1, & \text{if } V_{real} > VaR_{1-\alpha} \\ 0, & \text{if } V_{real} \leq VaR_{1-\alpha}, \end{cases} \quad (5.1)$$

where V_{real} represents the realised portfolio values, and $VaR_{1-\alpha}$ is the estimated VaR at level $1 - \alpha$ with $\alpha = \{0.10, 0.05, 0.025, 0.01, 0.005\}$. To express exception in percentages, we divide the total number of exception observations by the total number of observations. For the model that provides a reliable risk quantification, one would expect the percentage of exceptions under the estimated $VaR_{1-\alpha}$ to be exactly $100 \times \alpha\%$.

For our second measure we use size-adjusted loss function, which, contrary to the percentage of exception observations, takes the magnitude of the loss in the tail into account

(refer to [Blanco and Ihle, 1999](#)):

$$L = \begin{cases} (V_{real} - VaR_{1-\alpha}) / VaR_{1-\alpha}, & \text{if } V_{real} > VaR_{1-\alpha} \\ 0, & \text{if } V_{real} \leq VaR_{1-\alpha}. \end{cases} \quad (5.2)$$

In contrast to Eq. (5.1), in Eq. (5.2) each VaR exception is given a weight proportional to the realised loss relative to the estimated VaR. Note that due to the VaR being in the denominator of the loss function, Eq. (5.2) is only defined when the VaR is non-zero. Given our applications to the portfolios of insurance losses and the estimated models this condition is satisfied for all of our VaR estimates. We refer to the obtained measure as Blanco score.

In addition to calculating the percentage of exceptions and the Blanco score, we conduct statistical tests to evaluate the quality of the VaR forecasts by (i) comparing the nominal number of exceptions of the models to the true number of exceptions in a statistical test, and, (ii) testing for independence or clustering of the observed VaR exceptions. Both tests are documented in [Christoffersen \(1998\)](#). The first test (referred to in [Christoffersen \(1998\)](#) as the likelihood test of unconditional coverage) examines whether the percentage of exception observations is significantly higher than the expected percentage of exceptions given by the VaR quantile levels, α . The null hypothesis is that the model provides the same number of exceptions as expected, and its rejection indicates that the VaR estimate at a specified level $1 - \alpha$ is significantly miss-specified. The second test (referred to in [Christoffersen \(1998\)](#) as the test on independence) assesses whether there are patterns or clustering in the exceptions values, with the null hypothesis that VaR exceptions should be independent.¹⁰ Thus, the probability that the next observation is an exception should be independent of whether any previous observation was an exception or not. A model might pass the test for unconditional coverage, but may still be inadequate because it produces clusters of exceptions that should not arise when the model is appropriately specified.

5.2. In-sample performance

In the following we analyse in-sample VaR exceptions for a two-dimensional portfolio constructed from content and profit losses and a three-dimensional portfolio based on building, content and profit losses. We use the following four models identified in Section 4.3 as the best performing: two mixtures, namely, Clayton & survival Clayton and survival Clay-

¹⁰Both likelihood ratio test statistics, LR_{unc} for the number of exceptions/unconditional coverage and LR_{ind} for the independence of VaR exceptions, follow $\chi^2(1)$ distribution with critical values corresponding to 2.706, 3.841 and 6.635 at 10%, 5% and 1% significance level, respectively.

Table 5: In-sample results for the percentage of exceptions, Blanco scores and likelihood ratio test statistics, LR_{unc} for the number of exceptions/unconditional coverage and LR_{ind} for the independence of VaR exceptions for the portfolio of content and profit losses (Panel A) and the portfolio of building, content and profit losses (Panel B). *, ** and *** indicate rejections of an appropriate model specification at the 10%, 5% and 1% significance levels, respectively. The results for the best model with respect to each of the considered criteria are emphasised using **bold** numerals.

Copula	Measure	0.9	0.95	0.975	0.99	0.995
Panel A: Portfolio of Content and Profit Losses						
Surv. Clayton	Exceptions	0.1121	0.0754	0.0541	0.0212	0.0096
	Blanco	77.7236	27.1024	10.0420	2.8072	1.6165
	LR_{unc}	1.7678	9.0071***	16.9570***	5.0171**	7.3026***
	LR_{ind}	0.1805	0.8632	5.0754**	1.6395	0.2206
Gumbel	Exceptions	0.1063	0.0657	0.0290	0.0058	0.0019
	Blanco	56.1553	13.7605	2.9531	0.6161	0.3372
	LR_{unc}	0.2297	2.4713	0.0695	1.0836	1.2754
	LR_{ind}	0.3728	3.2067*	1.3373	0.0234	0.0000
Clayton and surv. Clayton	Exceptions	0.1063	0.0715	0.0541	0.0154	0.0077
	Blanco	57.1157	20.9469	8.3607	2.5572	1.3025
	LR_{unc}	0.2297	4.4920**	13.5957***	1.3407	1.7784
	LR_{ind}	0.3728	2.1296	3.4787*	0.2206	0.0783
Surv. Clayton and surv. Gumbel	Exceptions	0.1121	0.0754	0.0541	0.0174	0.0096
	Blanco	64.7827	22.9837	10.1405	2.5620	1.2839
	LR_{unc}	0.5862	6.1328**	16.9570***	2.3470	1.7784
	LR_{ind}	0.1653	1.5482	5.0754**	0.2843	0.0783
Panel B: Portfolio of Building, Content and Profit Losses						
Surv. Clayton	Exceptions	0.1199	0.0870	0.0676	0.0406	0.0232
	Blanco	71.2910	31.3779	16.1705	5.9246	2.9786
	LR_{unc}	2.1576	12.3482***	26.5640***	27.7040***	18.1877***
	LR_{ind}	0.0488	0.8451	2.4603	2.3089	0.6609
Gumbel	Exceptions	0.1179	0.0735	0.0406	0.0096	0.0019
	Blanco	68.4416	21.6251	6.0874	1.1050	0.5468
	LR_{unc}	2.1576	6.1328**	4.3651**	0.0057	1.2754
	LR_{ind}	0.1041	1.5482	1.3678	0.0783	0.0000
Clayton and surv. Clayton	Exceptions	0.1276	0.0986	0.0754	0.0580	0.0290
	Blanco	89.5201	41.3032	23.3093	8.7417	4.7746
	LR_{unc}	4.0780**	20.3209***	37.7083***	57.0640***	43.4985***
	LR_{ind}	0.0066	0.9472	1.2957	5.0754	1.9473
Surv. Clayton and surv. Gumbel	Exceptions	0.1199	0.0812	0.0638	0.0406	0.0232
	Blanco	61.7935	27.2201	13.6933	4.9977	3.5147
	LR_{unc}	1.4153	9.0071***	22.5250***	27.7040***	28.2218***
	LR_{ind}	0.0108	0.8632	3.6247*	1.3678	0.6609

ton & survival Gumbel mixtures, as well as one-parameter models, namely, survival Clayton and Gumbel copulas. Table 5 reports the results for the proportion of exceptions, the Blanco score, the likelihood ratio test statistics LR_{unc} for the number of exceptions/unconditional coverage and LR_{ind} for the independence of VaR exceptions for the portfolio of content and profit losses (Panel A) and the portfolio of building, content and profit losses (Panel B). To compute VaR we generate 100,000 loss trajectories from each of the considered copula models, assuming general form GH marginals. The results for the best model with respect to each of the considered criteria are indicated in bold numerals.

Let us first examine the models with respect to the empirically observed fraction of VaR exceptions and the size-adjusted loss function (Blanco score). Overall, we notice that all models, in both, the two- and three-dimensional examples, tend to underestimate the VaR, resulting in a larger number of exceptions than expected. This result can be explained by a large number of extreme losses occurring in the tail of the loss distribution. Our results in Table 5 Panel A for the two-dimensional portfolio indicate that Gumbel copula provides the best fit in terms of the fraction of the observed VaR exceptions at 90%, 95% and 97.5% level, while the Clayton & survival Clayton mixture performs best for the higher quantiles of 99% and 99.5%. Gumbel copula also leads to the smallest Blanco scores, and the tests of independence and unconditional coverage are generally not rejected (with an exception of independence of $VaR_{0.95}$ exceptions at 10% significance level). Survival Clayton model is the worst performing model leading to the largest underestimation of the VaR, the largest Blanco scores, and a large number of rejections for both tests. Overall, we favour Gumbel across all considered models for the two-dimensional portfolio case consisting of content and profit losses. Our selection is reconfirmed when we consider a three-dimensional portfolio (Panel B): Gumbel copula leads to the closest percentage of VaR exceptions to the target values, the smallest Blanco score and the smallest value of the LR_{unc} statistic (with an exception of $VaR_{0.95}$). It also leads to the best results for LR_{inc} in case of $VaR_{0.99}$ and $VaR_{0.995}$. Furthermore, we document the smallest number of rejections of the test results for the Gumbel copula: LR_{unc} rejects unconditional coverage for $VaR_{0.95}$ and $VaR_{0.975}$ at 5% significance level (although it is not rejected at 1%). Nevertheless, under the Gumbel copula, the considered test statistics have the smallest values across all considered models. We conclude that Gumbel copula in conjunction with GH marginals leads to the most reliable in-sample VaR estimates.

In addition to analysing VaR exceptions, we show in Figure 4 a bivariate sample with 100,000 observations generated from Gumbel copula (left) and corresponding K-plot (right). We set the dependence parameter to be 1.6056 (refer to Table 4). The generated sample clearly shows the ability of the Gumbel copula to generate upper tail dependence. The K-plot in the right panel can be seen as the bivariate copula equivalent to QQ-plots. If the points of a K-plot lie approximately on the diagonal then two variables are approximately independent. In our case we observe a clear deviation from the diagonal line, which points towards dependence. Since the points of the K-plot are located above the diagonal line, the dependence is positive. Finally, in Figure 5, for a range of quantile levels from 0.8 to 1.00, we plot (i) in the top panel: 2-dimensional VaR (left) and expected shortfall (right); (ii) in the middle panel: 3-dimensional VaR (left) and expected shortfall (right); (iii) in the

bottom panel: VaR and ES comparison for the 2-dimensional case (left) and 3-dimensional case (right) using Gumbel copula. From the top and middle panel we observe that Gumbel copula (red line) leads to the largest values for the VaR and ES in the tail of the distribution, which is consistent with the smallest misspecification and selection of Gumbel copula as a preferred model in Table 5. The survival Clayton copula (black line) and the mixture survival Clayton - survival Gumbel model (green line) result in almost identical VaR and ES estimates, which is, again, consistent with the results in Table 5 (same number of exceptions for both models). Since VaR is defined as the quantile of the loss distribution, while the ES represents the expected loss which will occur in an unfavourable scenario (thus, providing a more reliable risk quantification), ES is always larger than the VaR, which is evident from the comparison of VaR estimates (left panel) and ES estimates (right panel). This discrepancy is particularly evident when plotting VaR and ES for the best performing Gumbel copula model in the bottom panel.

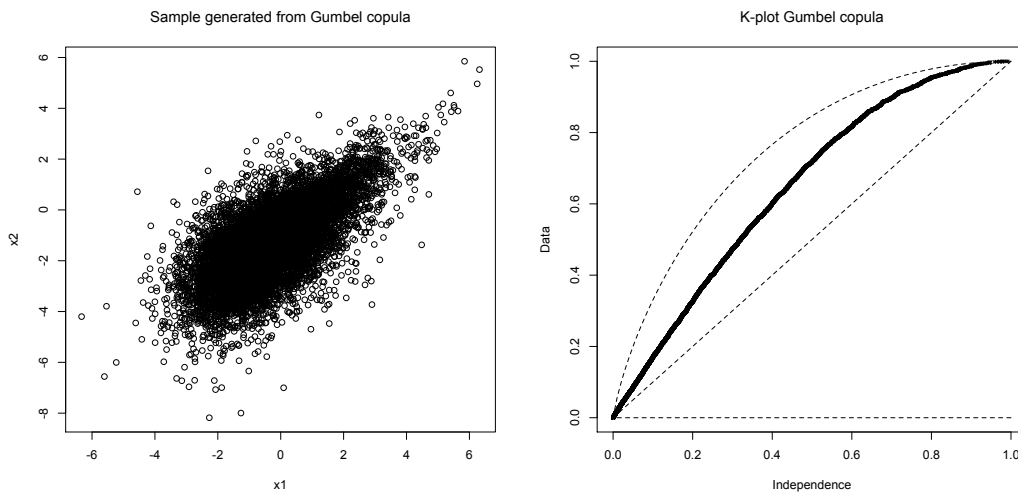


Figure 4: Bivariate sample with 100,000 observations generated from Gumbel copula (left) and corresponding K-plot (right).

5.3. Out-of-sample performance

In this section we assess VaR exceptions for the two- and three-dimensional portfolios in out-of-sample. We split our sample into in-sample set (350 observations) and out-of-sample set (the remaining 167 observations).¹¹ We use in-sample observations and re-estimate pa-

¹¹Since we only have 517 observations, we inevitably face a tradeoff by splitting the data sample into in-sample and out-of-sample periods. On the one hand, having a large number of in-sample data points is required for reliable estimation of the marginals and the copula. On the other hand, a larger number the out-of sample data points is desired to establish exception observations at higher quantiles. We have only considered quantiles up to 99% to ensure adequate estimation of exceptions at this extreme quantile.

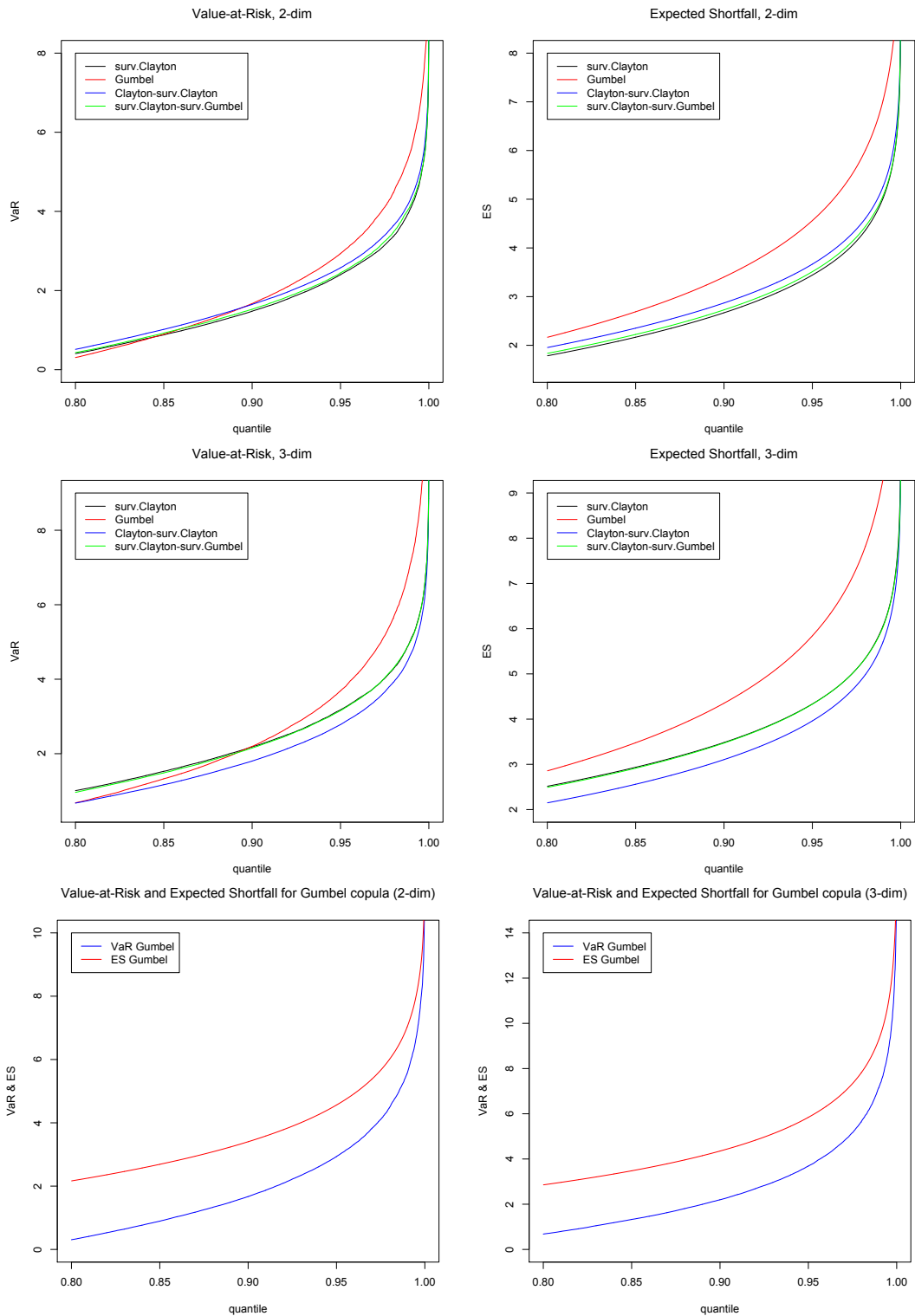


Figure 5: Top panel: 2-dimensional VaR (left) and expected shortfall (right); middle panel: 3-dimensional VaR (left) and expected shortfall (right); bottom panel: VaR and expected shortfall comparison for the 2-dimensional case (left) and 3-dimensional case (right) using Gumbel copula.

rameters for the marginals and the copula functions, and then generate 100,000 loss trajectories to estimate the out-of-sample VaR at different quantiles.

Table 6: Out-of-sample results for the percentage of exceptions, Blanco score and likelihood ratio test statistics, LR_{unc} for the number of exceptions/unconditional coverage and LR_{ind} for the independence of VaR exceptions for the portfolio of content and profit losses (Panel A) and the portfolio of building, content and profit losses (Panel B). **, * and *** indicate rejection of an appropriate model specification at the 10%, 5% and 1% significance levels, respectively. The results for the best model with respect to each of the considered criteria are emphasised using **bold** numerals.

Copula	Measure	0.9	0.95	0.975	0.99
Panel A: Portfolio of Content and Profit Losses					
Surv. Clayton	Exceptions	0.1078	0.0719	0.0539	0.0180
	Blanco	18.5663	6.4898	2.3447	0.3004
	LR_{unc}	0.1099	1.4881	4.3204**	0.8655
	LR_{ind}	2.5111	1.5557	3.6317	5.6388
Gumbel	Exceptions	0.1018	0.0599	0.0180	0.0000
	Blanco	14.6332	3.3542	0.3037	0.0000
	LR_{unc}	0.0060	0.3237	0.3754	3.3568*
	LR_{ind}	3.2180*	2.8046*	5.6388**	0.0000
Clayton and surv. Clayton	Exceptions	0.1086	0.0714	0.0457	0.0171
	Blanco	37.4213	12.8705	4.8828	1.8121
	LR_{unc}	0.2788	3.0042*	4.9683**	1.4860
	LR_{ind}	0.2971	0.9228	1.4441	0.1752
Surv. Clayton and surv. Gumbel	Exceptions	0.1078	0.0719	0.0539	0.0120
	Blanco	18.1883	6.2731	2.2908	0.2448
	LR_{unc}	0.1099	1.4881	4.3204**	0.0620
	LR_{ind}	2.5111	1.5557	3.6317*	0.0245
Panel B: Portfolio of Building, Content and Profit Losses					
Surv. Clayton	Exceptions	0.1198	0.0778	0.0539	0.0419
	Blanco	18.9085	7.7555	3.7356	1.2009
	LR_{unc}	0.6859	2.3475	4.3204**	9.5769***
	LR_{ind}	1.3918	1.0948	3.6317*	1.4361
Gumbel	Exceptions	0.1198	0.0659	0.0419	0.0000
	Blanco	18.8496	5.1637	1.1870	0.0000
	LR_{unc}	0.6859	0.8085	1.6344	3.3568*
	LR_{ind}	1.3918	2.1203	1.4361	0.0000
Clayton and surv. Clayton	Exceptions	0.1286	0.0971	0.0771	0.0571
	Blanco	65.0122	28.6300	14.9927	6.2098
	LR_{unc}	2.9392*	12.9955***	25.3402***	37.5173***
	LR_{ind}	0.6221	0.0194	0.4990	2.6698
Surv. Clayton and surv. Gumbel	Exceptions	0.1198	0.0838	0.0539	0.0419
	Blanco	19.1258	7.9832	3.7438	1.3229
	LR_{unc}	0.6859	3.3739*	4.3204**	9.5769***
	LR_{ind}	1.3918	2.8489*	3.6317*	1.4361

The results are summarised in Table 6. We first notice similarly to the in-sample back-testing procedure, all models (for two- or three-dimensional portfolio) tend to generally underestimate the VaR, resulting in a larger number of exceptions than expected. We observe that the out-of-sample results for the empirically observed fraction of VaR exceptions and the size-adjusted loss function (Blanco score) are consistent with the in-sample results: Gumbel copula results in the smallest number of exceptions (the closest percentage of VaR

exceptions to the target values) for the two- and three-dimensional portfolios, with an exception of the two-dimensional 99%-VaR, which is outperformed by the survival Clayton and survival Gumbel copula model. Furthermore, the Gumbel copula passes the test for unconditional coverage for the two- and three-dimensional case, resulting in the smallest LR_{unc} statistics for the majority of the cases (with an exception of the two-dimensional portfolio and 99% quantile). We notice that the null of independence of exceptions is not rejected by the Gumbel copula models for the three-dimensional portfolio; it is rejected at 10% significance (but not rejected at 5% significance) for the two-dimensional portfolio. Overall, we conclude that Gumbel copula leads to the best performance across all considered models, leading to the most reliable out-of-sample VaR forecasts.

6. Conclusion

In this paper we modelled the dependence between insurance claims arising from different LoBs using copula functions in conjunction with skewed generalised hyperbolic (GH) marginals. Using the loss data arising from building, contents and profit claims we considered a number of distributions from the class of skewed generalised hyperbolic distributions to investigate which distribution provides the best possible fit to losses arising from the individual LoB. We compared the fit of skewed GH distributions to normal distribution, which is an underlying distribution for log-losses used within Solvency II framework to determine solvency capital requirement for non-life insurance companies. We found that normal distribution provides the worst fit to the logarithmic losses data while none of the considered special cases (Student-t, variance gamma, normal inverse gaussian and hyperbolic) could be rejected by conventional distributional tests (Anderson-Darling, Kolmogorov and Carmér-von Mises). Among the cases considered, the general form generalised hyperbolic distribution provided the best fit to each of our data sets (individual losses from building, contents and profit claims). Conceivably, this is due to its flexibility achieved through the number of estimated parameters. We considered a number of copula models, including one-parameter and mixture copula models, and compared their ability to capture the dependence structure within the bivariate portfolio consisting of profit and content losses, as well as in the three-dimensional case constituting building, content and profit losses. Our results indicate that Clayton and survival Clayton mixture, survival Clayton and survival Gumbel mixture, as well as one-parameter models survival Clayton copula and Gumbel copula provide the best fit to the data based on the Akaike Information Criterion. All these models are capable of generating upper tail dependence in the joint distribution of portfolio losses. We assessed and compared the ability of selected copula models to quantify risk of portfolios of losses by

performing extensive analyses of in-sample estimates and out-of-sample Value-at-Risk forecasts, focusing on analysing the number of exceptions (i.e., observations of realised portfolio value that are greater than the estimated VaR), the size-adjusted loss function, and considering tests for unconditional coverage and independence of VaR exceptions. We document that for both, in- and out-of-sample analysis, the one-parameter Gumbel copula in conjunction with skewed generalised hyperbolic marginals leads to the best performance across all considered models, resulting in the most reliable VaR forecasts. We demonstrate that computing VaR and ES risk measures using multivariate copulas (capturing the dependence structure) in conjunction with skewed generalised hyperbolic marginals (capturing extreme tail and kurtosis in each of the loss distributions) allows to reliably quantify portfolio risk.

Our paper provides valuable insights with regards to the nature of dependence, and particularly, the strong tendency of several LoBs to jointly exhibit extreme losses. It fulfils one of the primary objectives of the general insurance providers, namely, assessing total risk of an aggregate portfolio of losses, assuming that LoBs are correlated. The proposed framework allows to determine a risk capital within a risk management framework that is needed to cover a series of simultaneous claims arising from losses in an unfavourable case scenario under the realistic assumption on the marginals and the dependence structure. It would be interesting to investigate other copula models, including pair-copula models based on vines, which allow to select pairs of assets to define vines prior to considering pair - copula decomposition, and analyse how the proposed structures affect VaR estimates, see e.g., [Bolancé and Padilla-Barreto \(2018\)](#). We leave this investigation of dependence between insurance losses using vine copulas for future research.

References

- Abramowitz, M. and I. A. Stegun (1972): *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, New York: Dover Publications.
- Akaike, H. (1974): "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, 19(6), 716–723.
- Albrecher, H., J. Beirlant, and J. L. Teugels (2017): *Reinsurance: Actuarial and Statistical Aspects*, Series: Statistics in practice, Hoboken, NJ: John Wiley & Sons, 1st edition.
- Aleman, R., C. Bolancé, and M. Guillén (2013): "A nonparametric approach to calculating Value-at-Risk," *Insurance: Mathematics and Economics*, 52, 255–262.
- Alm, J. (2016): "Signs of dependence and heavy tails in non-life insurance data," *Scandinavian Actuarial Journal*, 10, 859–875.

- Barndorff-Nielsen, O. and R. Stelzer (1997): "Exponentially decreasing distributions for the logarithm of particle size," *Proceedings of the Royal Society London A*, 353, 410–419.
- Barndorff-Nielsen, O. and R. Stelzer (2005): "Absolute moments of generalized hyperbolic distributions and approximate scaling of normal inverse gaussian Levy-processes," *Scandinavian Journal of Statistics*, 32(4), 617–637.
- Bernardi, M., A. Maruotti, and A. Petrella (2012): "Skew mixture models for loss distributions: A bayesian approach," *Insurance: Mathematics and Economics*, 51, 617–623.
- Blanco, C. and G. Ihle (1999): "How good is your VaR? using backtesting to assess system performance," *Financial Engineering News*, 11, 1–2.
- Bolancé, A. R., C. and A. E. Padilla-Barreto (2018): "Impact of d-vine structure on risk estimation," *Journal of Risk*, 20, 1–32.
- Bolancé, C., Z. Bahraoui, and M. Artís (2014): "Quantifying the risk using copulae with nonparametric marginals," *Insurance: Mathematics and Economics*, 58, 46–56.
- Bolancé, C., M. Guillén, E. Pelican, and R. Vernic (2008): "Skewed bivariate models and nonparametric estimation for CTE risk measure," *Insurance: Mathematics and Economics*, 43, 386–393.
- Breymann, W., A. Dias, and P. Embrechts (2003): "Dependence structures for multivariate high-frequency data in finance," *Quantitative Finance*, 3, 1–14.
- Chavez-Demoulin, V. and P. Embrechts (2010): *Copulas in Insurance*, John Wiley & Sons.
- Cherubini, U., E. Luciano, and W. Vecchiato (2004): *Copula Methods in Finance*, Wiley Finance Series.
- Christoffersen, P. (1998): "Evaluating interval forecasts," *International Economic Review*, 39, 841–862.
- Cossette, H., P. Gaillardetz, E. Marceau, and J. Rioux (2002): "On two dependence individual risk models," *Insurance: Mathematics and Economics*, 30, 153–166.
- D'Agostino, R. and M. Stephens (1986): *Goodness-of-fit Techniques, Statistics: a Series of Textbooks and Monographs*, volume 68, Marcel Dekker, 1st ed edition.
- De Jong, P. (2012): "Modeling dependence between loss triangles," *North American Actuarial Journal*, 16, 74–86.
- Denuit, M., J. Dhaene, and C. Ribas (2001): "Does positive dependence between individual risks increase stop-loss premiums?" *Insurance: Mathematics and Economics*, 28, 305–308.
- Denuit, M., C. Genest, and E. Marceau (1999): "Stochastic bounds on sums of dependent risks," *Insurance: Mathematics and Economics*, 25, 85–104.
- Dias, A. (2004): "Copula inference for finance and insurance," *Doctoral Thesis*.

- Dias, A. and P. Embrechts (2010): "Modelling exchange rate dependence at different time horizons," *Journal of International Money and Finance*, 8, 1687–1705.
- Diers, D., M. Eling, and S. D. Marek (2012): "Dependence modelling in non-life insurance using the Bernstein copula," *Insurance: Mathematics and Economics*, 50, 430–436.
- Dowd, K. and D. Blake (2006): "After VaR: the theory, estimation, and insurance applications of quantile-based risk measures," *Journal of Risk and Insurance*, 73, 193–229.
- Eberlein, E. and U. Keller (1995): "Hyperbolic distributions in finance," *Bernoulli*, 1, 131–140.
- EIOPA (2014): *The underlying assumptions in the standard formula for the Solvency Capital Requirement calculation*, the European Insurance and Occupational Pensions Authority, https://eiopa.europa.eu/Publications/Standards/EIOPA-14-322_Underlying_Assumptions.pdf.
- Eling, M. (2012): "Fitting insurance claims to skewed distributions: Are the skew-normal and skew-student good models?" *Insurance: Mathematics and Economics*, 51, 239–248.
- Embrechts, P. (2009): "Copulas: a personal view," *Journal of Risk and Insurance*, 76(3), 639–650.
- Embrechts, P., A. McNeil, and D. Straumann (2001): "Correlation and dependency in risk management: properties and pitfalls," in M. Dempster and H. Moffatt, eds., *Risk Management: Value at Risk and Beyond*, University Press, Cambridge.
- Embrechts, P., A. McNeil, and D. Straumann (2002): *Correlation and Dependency in Risk Management: Properties and Pitfalls*, In U. Press (ed.), *Risk Management: Value at Risk and Beyond*, M. Dempster and H. Moffatt, Cambridge.
- Fang, H., K. Fang, and S. Kotz (2002): "The meta-elliptical distributions with given marginals," *Journal of Multivariate Analysis*, 82(1), 1–16.
- Frahm, G., M. Junker, and A. Szimayer (2003): "Elliptical copulas: applicability and limitations," *Statistics and Probability Letters*, 63, 275–286.
- Genest, C., M. Gendron, and M. Bourdeau-Brien (2009): "The advent of copulas in finance," *The European Journal of Finance*, 15(7-8), 609–618.
- Hu, L. (2006): "Dependence patterns across financial markets: a mixed copula approach," *Applied Financial Economics*, 16, 717–729.
- Hult, H. and F. Lindskog (2002): "Multivariate extremes, aggregation and dependence in elliptical distributions," *Advances in Applied Probability*, 34(3).
- Ignatieva, K. and Z. Landsman (2015): "Estimating the tails of loss severity via conditional risk measures for the family of symmetric generalised hyperbolic distributions," *Insurance: Mathematics and Economics*, 65, 172–186.
- Ignatieva, K. and Z. Landsman (2017): "Conditional tail risk measures for skewed generalised hyperbolic family," *Working paper*.

- Joe, H. (1993): "Parametric families of multivariate distributions with given margins," *Journal of Multivariate Analysis*, 46(2), 262 – 282.
- Joe, H. (1997): *Multivariate Models and Dependence Concepts*, Chapman & Hall, London.
- Jorion, P. (2007): *Value-at-Risk. The New Benchmark for Managing Financial Risk*, The McGraw-Hill Companies.
- Kaas, R., J. Dhaene, and M. Goovaerts (2000): "Upper and lower bounds for sums of random variables," *Insurance: Mathematics and Economics*, 27, 151–168.
- Lane, M. (2000): "Pricing risk transfer transactions," *ASTIN Bulletin, The Journal of the International Actuarial Association*, 30(2), 259–293.
- Lee, S. C. K. and S. X. Lin (2012): "Modeling dependent risks with multivariate Erlang mixtures," *ASTIN Bulletin, The Journal of the International Actuarial Association*, 42, 153–180.
- Lindskog, F., A. McNeil, and U. Schmock (2003): "Kendall's tau for elliptical distributions," in *Credit Risk: Contributions to Economics*, Springer-Verlag, 149–156.
- Lopez, J. (1998): "Regulatory evaluation of Value-at-Risk models," *Economic Policy Review*, 4(3), 119–124.
- Madan, D. and E. Seneta (1990): "The variance gamma model for share market returns," *Journal of Business*, 63, 511–524.
- McLeish, D. L. and C. G. Small (1988): *The Theory and Applications of Statistical Inference Functions. Lecture Notes in Statistics*, volume 44, Springer-Verlag, New York.
- McNeil, A. (1997): "Estimating the tails of loss severity distributions using extreme value theory," *ASTIN Bulletin, The Journal of the International Actuarial Association*, 27(1), 117–137.
- McNeil, A., R. Frey, and P. Embrechts (2004): *Quantitative Risk Management: Concept, Techniques and Tools*, Springer, Princeton University Press.
- McNeil, A., R. Frey, and P. Embrechts (2015): *Quantitative Risk Management: Concepts, Techniques and Tools. Revised Edition*, Princeton University Press.
- Merz, M., M. Wüthrich, and E. Hashorva (2013): "Dependence modelling in multivariate claims run-off triangles," *Annals of Actuarial Science*, 7, 3–25.
- Miljkovic, T. and B. Grün (2016): "Modeling loss data using mixtures of distributions," *Insurance: Mathematics and Economics*, 70, 387–396.
- Nelsen, R. (2006): *An Introduction to Copulas, second edition*, Springer, Portland, OR, USA.
- Praetz, P. D. (1972): "The distribution of share price changes," *Journal of Business*, 45, 49–55.
- Salzmann, R. and M. V. Wüthrich (2012): "Modelling accounting year dependence in runoff triangles," *European Actuarial Journal*, 2, 227–242.

Stephens, M. (1974): “EDF statistics for goodness of fit and some comparisons,” *Journal of the American Statistical Association*, 69, 730–737.

Vernic, R. (2006): “Multivariate skew-normal distributions with applications in insurance,” *Insurance: Mathematics and Economics*, 38, 413–426.

Wu, F., A. Valdez, and M. Sherris (2007): “Simulating from exchangeable archimedean copulas,” *Communications in Statistics - Simulation and Computation*, 36(5).

7. Appendix: Copula estimation

Generally, the maximum likelihood technique is applied to estimate parametric copulas. From (2.1), the density of the random vector $X = (X_1, \dots, X_d)^\top$ is given by

$$f(x_1, \dots, x_d; \delta_1, \dots, \delta_d, \theta) = c\{F_{X_1}(x_1; \delta_1), \dots, F_{X_d}(x_d; \delta_d); \theta\} \prod_{j=1}^d f_j(x_j; \delta_j), \quad (7.1)$$

where

$$c(u_1, \dots, u_d) = \frac{\partial^d C(u_1, \dots, u_d)}{\partial u_1 \dots \partial u_d} \quad (7.2)$$

denotes a copula density. Denoting a vector of parameters $\alpha = (\delta_1, \dots, \delta_d, \theta)^\top \in \mathbb{R}^{d+1}$, the likelihood function is given by

$$L(\alpha; x_1, \dots, x_T) = \prod_{t=1}^T f(x_{1,t}, \dots, x_{d,t}; \delta_1, \dots, \delta_d, \theta). \quad (7.3)$$

Combining (7.1) and (7.3), the corresponding log-likelihood function is given by

$$\ell(\alpha; x_1, \dots, x_T) = \sum_{t=1}^T \ln [c\{F_{X_1}(x_{1,t}; \delta_1), \dots, F_{X_d}(x_{d,t}; \delta_d); \theta\}] + \sum_{t=1}^T \sum_{j=1}^d \ln [f_j(x_{j,t}; \delta_j)]. \quad (7.4)$$

To maximize this log-likelihood numerically, we perform the *inference for marginals* (IFM) method, which is a sequential two-step maximum likelihood method, see e.g. [McLeish and Small \(1988\)](#) and [Joe \(1997\)](#). Parameters from the marginals are estimated in the first step as

$$\hat{\delta}_j = \arg \max_{\delta} \ell_j(\delta_j), \quad (7.5)$$

where

$$\ell_j(\delta_j) = \sum_{t=1}^T \ln f_j [x_{j,t}; \delta_j] \quad (7.6)$$

denotes the log-likelihood function for each of the marginal distribution $j = 1, \dots, d$. Their estimates are then substituted into the copula to obtain the pseudo log-likelihood function

$$\ell(\theta, \hat{\delta}_1, \dots, \hat{\delta}_d) = \sum_{t=1}^T \ln [c\{F_{X_1}(x_{1,t}; \hat{\delta}_1), \dots, F_{X_d}(x_{d,t}; \hat{\delta}_d); \theta\}], \quad (7.7)$$

which is then maximised with respect to θ to obtain the estimator $\hat{\theta}$. The estimates $\hat{\alpha}_{IFM} = (\hat{\delta}_1, \dots, \hat{\delta}_d, \hat{\theta})^\top$ solve the first order condition

$$(\partial \ell_1 / \partial \delta_1, \dots, \partial \ell_d / \partial \delta_d, \partial \ell / \partial \theta) = 0. \quad (7.8)$$