MPRA

Risk aggregation, dependence structure and diversification benefit

Bürgi, Roland and Dacorogna, Michel M and Iles, Roger

15 August 2008

Risk Aggregation, Dependence Structure and Diversification Benefit

Roland Bürgi, Michel Dacorogna, Roger Iles[*]

August 15, 2008

Abstract

Insurance and reinsurance live and die from the diversification benefits or lack of it in their risk portfolio. The new solvency regulations allow companies to include them in their computation of risk-based capital (RBC). The question is how to really evaluate those benefits.

To compute the total risk of a portfolio, it is important to establish the rules for aggregating the various risks that compose it. This can only be done through modelling of their dependence. It is a well known fact among traders in financial markets that "diversification works the worst when one needs it the most". In other words, in times of crisis the dependence between risks increases. Experience has shown that very large loss events almost always affect multiple lines of business simultaneously. September 11, 2001, is an example of such an event: when the claims originated simultaneously from lines of business which are usually uncorrelated, such as property and life, at the same time that the assets of the company were depreciated due to the crisis on the stock markets.

In this paper, we explore various methods of modelling dependence and their influence on diversification benefits. We show that the latter strongly depend on the chosen method and that rank correlation grossly overestimates diversification. This has consequences on the RBC for the whole portfolio, which is smaller than it should be when correctly accounting for tail correlation. However, the problem remains to calibrate the dependence for extreme events, which are rare by definition. We analyze and propose possible ways to get out of this dilemma and come up with reasonable estimates.

[*]SCOR Switzerland, General Guisan - Quai 26, 8022 Zurich, Switzerland.

1   Introduction

The purpose of insurance or reinsurance is to offer sharing of risks among individuals. It is thus in its nature to deal with a variety of risks on its balance sheet. Due to its core business, the risks originating from the insurance business should dominate the other risk factors. Nevertheless, building an internal model for such a company implies that one is able to aggregate the risks in a portfolio. There are a number of purposes for which an aggregate model of the insurance business is needed: risk management, capital management, business steering, and profitability analysis are just a few examples.

For building a business plan of the company only the expected value of the insurance loss matters. In this case, the dependence between the lines of business is not relevant since the relationship

$$\mathrm{E}\left[\sum_i X_i\right] = \sum_i \mathrm{E}[X_i] \tag{1}$$

holds for any vector $\mathbf{X} = (X_1, \ldots, X_n)$ of random variables, independent of the relation between them. However, as soon as other properties implying the probability distribution of the losses, such as a risk measure, are of importance, the dependence structure is of utmost relevance. It will drive significantly the risk of the overall portfolio.

There are a variety of dependencies which influence insurance claims. There is the whole class of the intuitive dependencies which are traditionally thought of when modelling single risk factors. For instance, the dependence of natural catastrophe events on the single property policies is a well studied discipline. Especially when analysing large insurance claims, it becomes obvious that such events rarely trigger one policy only, or even just a single line. Most large losses have a strong multi-line character and trigger several product categories simultaneously.

The collapse of the World Trade Center towers on September 11 2001 has shown in a dramatic way how insurance products of lines of business, which had thought to be independent until then, can be triggered at once. The scenario of both towers collapsing had been assumed to be impossible. 9/11 caused a human drama, many people lost their lives in and around the Twin Towers. Several surrounding buildings have also collapsed or suffered severe damages. Insurance-wise, several property policies were triggered. Next to the damage to the buildings, there was a large amount of business interruption claims, and several life policies were triggered, while cars parked under the towers were also destroyed and triggered motor policies.

In times of human tragedies, insurance companies show a fair amount of goodwill. This could, for instance, be seen when hurricane Katrina had destroyed New Orleans. The wind had caused the city levees to break, and the city was flooded. Whereas most of the property insurance policies covered wind, a vast number of policies did not cover flood. However, a substantial part of the insurance loss was technically a flood loss. Nevertheless, most insurance companies paid for the flood loss even when it was not covered. This shows that, in extreme cases, insurance policies can also be triggered for perils they do not even cover.

There are also less obvious but none the more harmless dependencies between lines

of business. The most significant is probably the dependence of the product definition and the legislation associated with the product. Due to the globalisation of the world economy, product definitions adapt rather quickly across markets, and the legislation of most countries is interlinked with the ones from other countries through an organisation such as the European Union. The claims amounts can be influenced considerably by the exclusions in the policy and by the court practice. This can be seen very prominently in the case of the US liability products; insurance claims were exploding in the late 1990s.

Whilst it would be easy to find other examples of large multi-line losses, it is hard to imagine why there should be a dependence between small losses. For instance, if a given year has no large natural catastrophe events, why should it be a good year for fire insurance? If not many severe car accidents happen, would we necessarily expect the amount of medical malpractice claims to be low? The world seems to be asymmetric; whereas large events can cumulate large claims, low claims in one line do not imply a good year for another line. Moreover, as the subprime crisis is once again reminding us, dependence also increases among financial markets when confidence decreases. Whilst confidence is lost quickly, it takes much longer for the industry to regain confidence. Asymmetry is also in this case very relevant. We will show in this article that the asymmetry of the dependence is a driver for the capital, and that it is important to take the asymmetry into account when modelling the dependence. We thus find it very topical for a book about stress testing to study how to model changes in dependence during stress situations.

The article is organized as follows: After a section dealing with the measurement of dependence and risk, we briefly introduce the concept of copula and various forms of them in Section 3. Section 4 deals with the set up of our model, the type of analysis we perform and the methodology. The results are presented and discussed in Section 5. Because estimation of dependence parameters is crucial to portfolio modelling, we present our approach to this problem in Section 6. Conclusions about this study are drawn in Section 7.

2 Measuring and Modelling Risks

2.1 Measuring and Modelling Dependence between Risks

There are a variety of methods to measure and model dependence. For interested readers, an excellent overview is provided in the work of the risklab of the ETH Zurich around Paul Embrechts [Embrechts et al., 2002, Embrechts et al., 2003]. Here we shall briefly mention few relevant points to our study. The most popular measure of dependence between two random variables X_i and X_j, whose standard deviations exist, is probably the Pearson correlation, also known as linear correlation

$$\rho_{ij} = \frac{\mathrm{E}[(X_i - \mu_i)(X_j - \mu_j)]}{\sigma_i \sigma_j}, \tag{2}$$

3

where μ_i and μ_j denote the mean and σ_i and σ_j the standard deviation of X_i and X_j respectively. The expectation value in the numerator is also called the *covariance* of X_i and X_j.

Spearman's rank correlation measures the correlation between the ranks of X_i and X_j rather than their nominal values:

$$RC_{ij} = 12 \cdot \mathrm{E}[(U_i - 0.5)(U_j - 0.5)], \tag{3}$$

where U_i is the relative rank, in other words the cumulative probability, of X_i. This is a slightly more sophisticated measure, since it is independent of the distribution of the individual random variables and the existence of their second moments. The separation between the marginal distributions of X_i and X_j and the correlation measure facilitates the modelling of the risk.

A third popular risk measure is the Kendall τ

$$\tau_{ij} = \mathrm{E}[\mathrm{sign}(X_i - \tilde{X}_i)(X_j - \tilde{X}_j)], \tag{4}$$

where $(\tilde{X}_i, \tilde{X}_j)$ is a second pair of random variables with the same distribution as (X_i, X_j). It can be shown that, just as Spearman's rank correlation, the Kendall τ does not depend on the marginal distributions.

All those measures assume underlying a linear dependence model and do not sensitive to inhomogenity of the dependence structure for low and high quantiles. We have just argued above that claims in case of stress are more dependent on each other. To model such a behaviour, we need to use more sophisticated concepts such as copulas. In this article, we will model dependence through *copulas*. As some of the readers may not be familiar with copula modelling, we will introduce the topic in Section 3.

2.2 Measuring Risk

Risk can be defined as the deviation of reality from the expectation. There are a number of risk measures used in the financial industry. One of the most popular risk measures in both banking and insurance is the value-at-risk (VaR), which is defined as the negative quantile at a certain probability α:

$$\mathrm{VaR}^{(\alpha)}(X) = -x^{(\alpha)}(X) = -\sup\{x | \mathrm{P}[X \leq x] \leq \alpha\}. \tag{5}$$

This definition corresponds to the question: which is the minimum loss in the $\alpha\%$ worst cases of the portfolio? The VaR has also been chosen as the central risk measure for the European Solvency II regulation.

However, there are a number of shortcomings to the VaR as a risk measure, above all that the VaR is not a *coherent* risk measure: It violates the axiom of subadditivity of a risk measure ρ,

$$\rho(X_1) + \rho(X_2) \leq \rho(X_1 + X_2), \tag{6}$$

for certain types of probability distributions [Artzner et al., 1999].

Translating this axiom into the language of business, this would mean that the diversification gain when running two risks in one company as opposed to two separate companies must be positive. This is an elementary property when calculating capital requirements of legal entities and holdings. Due to the lack of subadditivity, it is not possible to allocate the overall VaR to individual risk factors, which is a necessity in companies with an active performance measurement and capital management process.

For these reasons, many insurance companies prefer to use as a risk measure the expected shortfall $ES^{(\alpha)}(X)$ to the VaR. We follow the definition of the expected shortfall as given in [Acerbi and Tasche, 2002]:

$$ES^{(\alpha)}(X) = -\frac{1}{\alpha}\left(\mathrm{E}[X\mathbf{1}_{\{X \leq x^{(\alpha)}\}}] - x^{(\alpha)}(\mathrm{P}[X \leq x^{(\alpha)}] - \alpha)\right). \qquad (7)$$

The expected shortfall is sometimes also called *tail-value-at-risk*, tVaR. The expected shortfall is the answer to the question: what is the *expected* loss, given the loss is in the worst $\alpha\%$ cases of the portfolio? In contrast to the VaR, the expected shortfall is a coherent risk measure [Artzner et al., 1999]. Furthermore, capital allocation to the various risk factors can easily be done using the Euler principle: the overall shortfall is subdivided into the contribution of each risk factor to the shortfall scenarios, i. e. the worst $\alpha\%$ loss scenarios of the portfolio. This yields a fair and consistent method for allocating the capital, K_X, to the risk X within the portfolio Z [Tasche, 2002]:

$$K_X = -ES^{(\alpha)}(X \mid Z \leq F_Z^{-1}(\alpha)), \qquad (8)$$

where F_Z is the cumulative probability distribution of the entire portfolio Z [1].

Going back to the risk-based capital (RBC) for each random variable X, we define it as the risk above the expected:

$$RBC(X) = -ES^{(\alpha)}(X) - \mathrm{E}[X]. \qquad (9)$$

We can then define the diversification gain as the capital which can be saved when running the risks together in a portfolio of a single company rather than each risk factor in a separate company:

$$DG(X) \equiv 100\% - \frac{RBC\left(\sum_i X_i\right)}{\sum_i RBC(X_i)}. \qquad (10)$$

This is the main variable of interest for our study. We will analyse the effect of dependence of random variables on the diversification gain.

3 Copula Modelling

One way of representing dependence between random variables are copulas. There are a number of good introductory readings to copula modelling, see for instance

---

[1] Since the expected shortfall, the risk measure we will use,is negative, it has a negative sign in the capital definition.

[Nelsen, 1999, Embrechts et al., 2003, Genest and Favre, 2007]. The basic idea of a copula is to translate any multivariate distribution $F(x_1, x_2, \ldots, x_n)$ into its marginal distributions $F_1(x_1), F_2(x_2), \ldots, F_n(x_n)$ and its copula $C(F_1(x_1), F_2(x_2), \ldots, F_n(x_n))$ describing the dependence between the random variables $X_1, \ldots, X_n$. Sklar's theorem [Sklar, 1958] shows the equivalence of multivariate distributions and the copula approach.

For use of simplicity, we abbreviate $(F_1(x_1), F_2(x_2), \ldots, F_n(x_n))$ by $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ throughout this article. The copulas will generally just be defined for $\mathbf{u}$. In order to obtain the realisation vector $\mathbf{x} = (x_1, \ldots, x_n)$, the quantile functions of the marginal distributions will have to be applied on $\mathbf{u}$: $\mathbf{x} = (F_1^{-1}(u_1), F_2^{-1}(u_2), \ldots, F_n^{-1}(u_n))$.

An n-dimensional copula $C : [0,1]^n \rightarrow [0,1]$ is a function with the following properties

$$
\begin{align}
C(\mathbf{u}) &= 0 \qquad \text{whenever } \mathbf{u} \text{ has at least one component equal to } 0 & (11) \\
C(\mathbf{u}) &= u_i \qquad \text{whenever } \mathbf{u} = (1, \ldots, 1, u_i, 1, \ldots, 1) & (12) \\
V_C(B) &= \sum_{\mathbf{u} \text{ vertex of } B} \text{sgn}(\mathbf{u}) C(\mathbf{u}) \geq 0 \qquad \text{for any } B = [\mathbf{a}, \mathbf{b}]. & (13)
\end{align}
$$

In the last condition, $V_C(B)$ denotes the probability volume of B, which must not be negative. The term $\text{sgn}(\mathbf{u})$ is defined as

$$
\text{sgn}(\mathbf{u}) = \begin{cases} 1 & \text{if } u_k = a_k \text{ for an even number of } k, \\ -1 & \text{if } u_k = a_k \text{ for an odd number of } k. \end{cases} \tag{14}
$$

There exists a vast variety of copulas. However, there are only few families of them used in practice. The following subsections give an overview of the most popular families: the Archimedean and elliptical copulas.

3.1 Archimedean Copulas

There are a number of excellent papers on Archimedean copulas, see for instance [Savu and Trede, 2006] or [Armstrong, 2003]. Archimedean copulas are defined through their generating function ϕ:

$$
C(u_1, \ldots, u_n) = \phi^{-1}(\phi(u_1) + \cdots + \phi(u_n)). \tag{15}
$$

The function C is called an Archimedean copula if and only if

$$
(-1)^k \frac{\partial^k}{\partial u^k} \phi^{-1}(u) \geq 0 \qquad \text{for } k \in \mathbb{N}. \tag{16}
$$

Archimedean copulas are

- commutative, i.e. $C(u_1, u_2) = C(u_2, u_1)$ for all $u_1, u_2 \in [0,1]$,

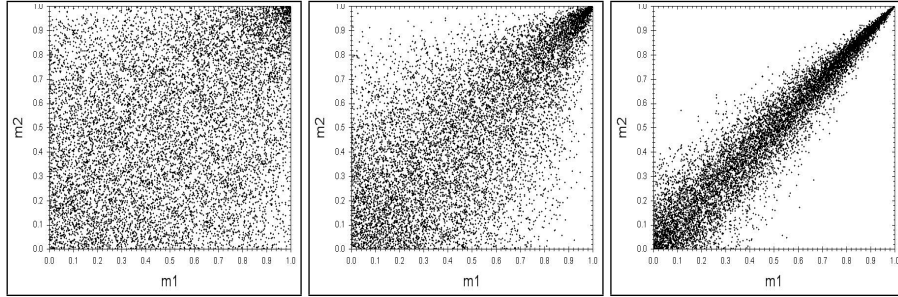- associative, i.e. $C(C(u_1, u_2), u_3) = C(u_1, C(u_2, u_3))$ for all $u_1, u_2, u_3 \in [0,1]$.

Figure 1: Rank scatter plots of the bivariate Clayton copula for $\theta = 0.5$, $\theta = 2$, $\theta = 10$.

3.1.1 The Clayton Copula

The Clayton copula, sometimes also called the Cook-Johnston copula, is defined by its generating function

$$\phi(t) = \frac{1}{\theta}\left(t^{-\theta} - 1\right).$$
(17)

In the bivariate case, this leads to the cumulative copula density

$$C(u,v) = \max\left(\left[u^{-\theta} + v^{-\theta} - 1\right]^{-1/\theta}, 0\right).$$
(18)

Scatter plots of the bivariate Clayton copula are shown for various values of θ in Figure 1. It can be seen that the higher the value of θ, the more the two random variables depend on each other. The limit $\theta \to 0$ corresponds to the uniform copula, i.e. the random variables are independent.

The Clayton copula is asymmetric. In its defined form, the Clayton copula acts on the lower tail of the distribution, whereas for upper tail the random variables are hardly dependent on each other. In insurance, however, the dependence should be modelled for the upper tails. This can easily be obtained by mirroring the copula by a transformation $(u,v) \to (1-u, 1-v)$. Throughout the following sections, we will consider the Clayton copula in its mirrored form.

3.1.2 The Gumbel Copula

The Gumbel copula is another representative of the family of Archimedean copulas. Its generating function is defined as

$$\phi(t) = (-\log t)^{\theta}.$$
(19)

For the bivariate case, this leads to the cumulative copula density of

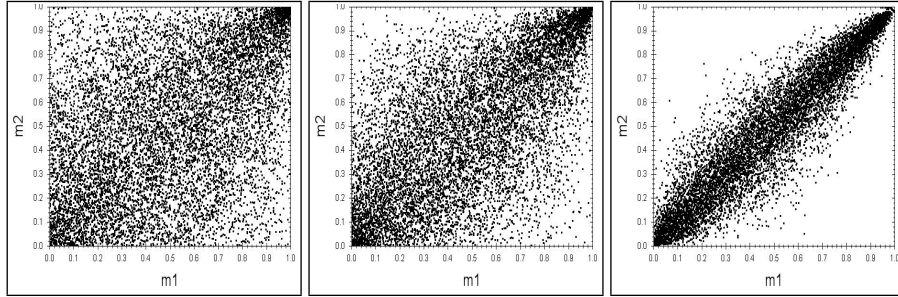$$C(u,v) = \exp\left(-\left[(-\log u)^{\theta} + (-\log v)^{\theta}\right]^{1/\theta}\right).$$
(20)

7

Figure 2: Scatter plots of the bivariate Gumbel copula for $\theta = 1.5$, $\theta = 2$, $\theta = 5$.

Scatter plots of the bivariate Gumbel copula are shown for various values of θ in Figure 2. As in the case of the Clayton copula, the dependence of the random variables grows with increasing value of θ. In contrast to the Clayton copula, the Gumbel dependence acts on both upper and lower tail. Nevertheless, the Gumbel copula is asymmetric: in its definition, the dependence of high quantiles is stronger than the one of low quantiles.

3.2 Elliptical Copulas

Elliptical copulas are based on multivariate elliptical distributions, which are described for instance in [Fang and Zhang, 1990]. Prominent members of the family of elliptical copulas are the Gauss copula and the Student's T copula. They will be described in the following subsections.

3.2.1 The Gauss Copula

The Gauss copula is the copula representation of a rank correlation process. It is based on the multivariate normal distribution. Assuming $\mathbf{Z} = (Z_1, \ldots, Z_n)$ to be a vector of multivariate normally distributed random variables, the joint probability density function is defined as

$$f(\mathbf{z}) = \frac{1}{(2\pi)^n |\Sigma|} \exp\left(-\frac{1}{2} \mathbf{z}' \Sigma^{-1} \mathbf{z}\right), \tag{21}$$

where $\mathbf{z} = (z_1, \ldots, z_n)$ is a realisation of $\mathbf{Z}$. The correlation matrix Σ is defined as

$$\Sigma = \begin{pmatrix} 1 & \rho_{12} & \ldots & \rho_{1n} \\ \rho_{21} & 1 & \ldots & \rho_{2n} \\ \vdots & \vdots & & \vdots \\ \rho_{n1} & \rho_{n2} & \ldots & 1 \end{pmatrix}, \tag{22}$$

where ρ_{ij} denotes the correlation between the random variables Z_i and Z_j. The correlation matrix is symmetric, i. e. $\rho_{ij} = \rho_{ji}$.
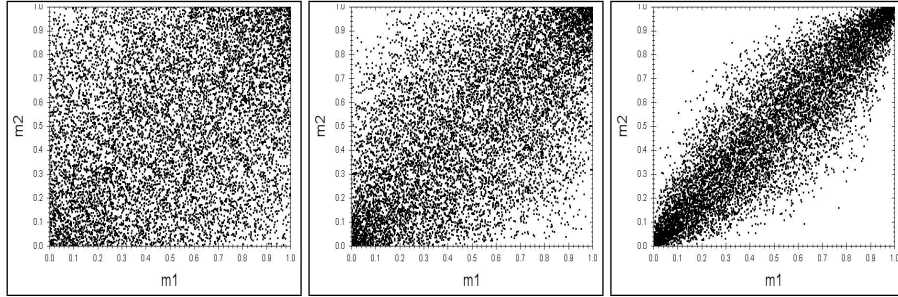
Figure 3: Rank scatter plot of the Gauss copula for $\rho = 0.3$, $\rho = 0.6$, and $\rho = 0.9$.

In order to obtain the cumulative probability of the marginal distributions, i. e. the vector $\mathbf{u}$ as described above, the cumulative density function of the univariate standard normal distribution has to be calculated:

$$u_i = \Phi(z_i) = \frac{1}{2\pi} \int_{-\infty}^{z_i} dt \exp\left(-\frac{t^2}{2}\right). \tag{23}$$

Scatter plots of the bivariate Gauss copula are shown for various values of the correlation parameter ρ in Figure 3. The symmetry and linearity of the Gauss copula can also be seen reflected in the scatter plots: the same dependence is homogeneous for low and high quantiles, i. e. the rank correlation within any ellipsoid around the diagonal of the scatter plot is the same. The Gauss copula has become popular in stochastic simulation, mainly due to the simplicity of calibrating it by measuring the rank correlation. Furthermore, the algorithms of simulating multivariate random variables using a Gauss copula are well known. Nevertheless, as we will see in this article, the symmetry and the linearity of the Gauss copula are strong restrictions for modelling the dependence; assuming that in reality we observe an asymmetric dependence of insurance losses, as we argue for in the introduction, the risk diversification gain is substantially overestimated.

3.2.2 The Student's T Copula

As the Gauss copula is based on the multivariate normal distribution, the Student's T copula is based on the multivariate Student's T distribution. A detailed discussion of the Student's T copula can be found in [Demarta and McNeil, 2005].

Let $\mathbf{Y} = (Y_1, \ldots, Y_n)$ be a random variable with a multivariate Student's T distribution with ν degrees of freedom. The probability density function is given by

$$f(\mathbf{y}) = \frac{\Gamma\left(\frac{\nu+n}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{(\pi\nu)^n|\Sigma|}} \left(1 + \frac{\mathbf{y}'\Sigma^{-1}\mathbf{y}}{\nu}\right)^{-\frac{\nu+n}{2}}. \tag{24}$$

The matrix Σ is the correlation matrix as defined also for the Gauss copula in Equation (22). The similarity between the Gauss copula and the Student's T copula becomes
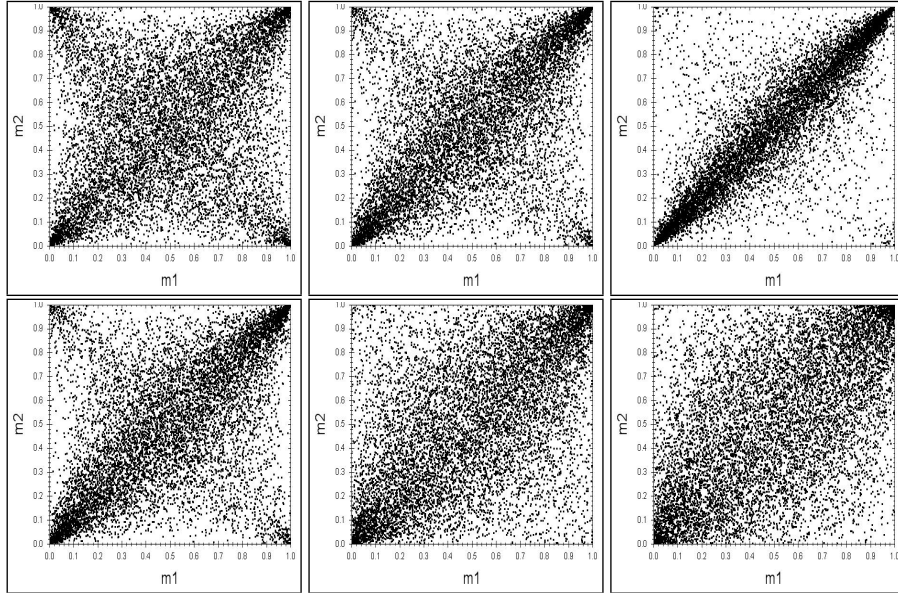
Figure 4: Rank scatter plots for Student's T copula with constant $\nu = 1$ and increasing $\rho = 0.3$, 0.6 and 0.9 shown in the top row; and for constant $\rho=0.6$ and increasing $\nu = 1$, 3 and 9 in the lower row

even more apparent when looking at the representation

$$\mathbf{Y} = \sqrt{W}\mathbf{Z}, \tag{25}$$

where $\mathbf{Z}$ is a random variable with a multivariate normal distribution as described in Equation (21) and W is independent from $\mathbf{Z}$ and is inverse gamma distributed $W \sim \text{Ig}(\nu/2, \nu/2)$. The degree of freedom parameter ν can assume any positive real value. However, it is more difficult to find fast numerical algorithms for non-integer values of ν than for integer values. Thus, many modellers prefer to work with integer values of ν.

Similarly to the case of the Gauss copula, the cumulative probability of the marginal distributions $\mathbf{u}$ can be obtained by calculating the cumulative distribution function of the univariate Student's T distribution

$$u_i = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi\nu}} \int_{-\infty}^{y_i} dt \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}. \tag{26}$$

Scatter plots of the bivariate Student's T copula are shown for various values of the correlation parameter ρ and the degree of freedom parameter ν are shown in Figure 4. As in the case of the Gauss copula, the scatter plot of the Student's T copula is symmetric: the dependence of high quantiles is the same as the one of low quantiles. The additional parameter ν allows for modelling a dependence orthogonal to the main dependence as given by ρ. For low values of ν and ρ, the scatter plot resembles an X-shape.

4 Setup of the Analysis and Methodology

4.1 Model Setup

In this article, we are interested in a variety of questions related to the estimation of diversification benefits when combining multiple risk factors. Foremost, we want to investigate the impact of the functional form of a copula. We will start our investigation by defining a benchmark sample that we will use as the reference to compare results.

Let us consider a standard property reinsurer whose products cover two perils: fire and windstorm. Our model reinsurer sells products in Germany and France. We will first have to think about how to set up a dependence structure for this business.

The windstorm peril underlies a strong dependence between Germany and France, as those are neighbouring countries, and the likelihood is big that a windstorm could pass over both countries in the same event. In an advanced company, the natural catastrophe peril windstorm would normally be modelled in an event set based way, which would include its dependence. However, for the sake of our investigation, we will assume that our model reinsurer has modelled the windstorm peril for both countries separately in the form of lognormal distributions and wants to model the windstorm dependence through a copula.

The fire peril underlies also a strong dependence between both countries. This may seem surprising at first glance: Why should a fire break out in Germany when there is a fire in France or vice versa? However, significant dependence between reinsurance products do not arise through the underlying peril itself but are caused by changes in legislation and insurance practice. Since both, France and Germany, are EU countries, their legislation will move somewhat in parallel. This implies a dependence between the fire losses.

Between the aggregate fire and windstorm perils, we want to model another dependence. The origin of this dependence is in both the reasons mentioned above: Whenever a windstorm event happens, it is likely to cause a number of fires in the corresponding country. Furthermore, both perils usually underlie a similar product definition and legislation and are therefore dependent. In conclusion, we model the dependence in the form of a hierarchical tree as shown in Figure 5: we subdivide the property portfolio of our model reinsurer into the perils fire and windstorm first and then into the peril regions Germany and France. This hierarchical structure is a natural way of expressing the dependence through our understanding of the business we are analysing. Moreover, there is nothing in this approach that prevents us to use an event loss set based model for windstorm to replace the lognormal model we will use throughout this article.

Since the main focus of this article is on modelling dependence rather than modelling the perils themselves, we will simplify the problem to maximise the effect of our investigation. The biggest effect of dependence can be seen when aggregating equal risks. Therefore, we choose equal marginal distributions for all our risk factors. A canonical type of aggregate loss model used in insurance is the lognormal distribution. Its
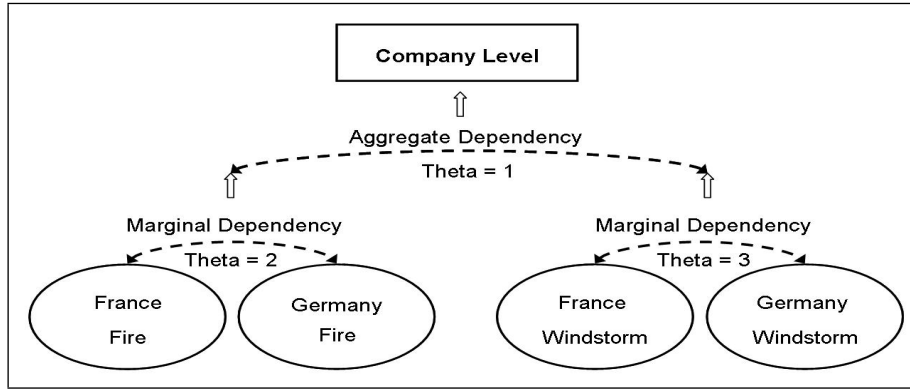
Figure 5: Model hierarchy used as a reference for comparing models.

probability density function (PDF) is defined as

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\log x - \mu)^2}{2\sigma^2}\right).$$ (27)

We choose $\mu = 10$ and $\sigma = 1$ for each of the marginal distributions modelling our risk factors. For our purposes, the choice of μ is hardly relevant; it is just a translational shift of the lognormal distribution and will not change our considerations on the diversification gain. The parameter σ influences the width and also the skewness of the distribution. The value of $\sigma = 1$ chosen for our study corresponds to a usual skewness of an aggregate insurance loss distribution.

As described in the introduction, there is strong evidence that the dependence of insurance claims are asymmetric: Whereas it is less likely that low claim sizes in one line lead to low claim sizes in the other, it is very likely that large claims in one line trigger large claims in other lines. Therefore, we will choose the most asymmetric and the most severe one of the four copulas we investigate for our reference model to produce our benchmark sample: the Clayton copula. We choose $\theta = 3$ as the dependence between the windstorm regions, $\theta = 2$ as the dependence between the fire regions, and $\theta = 1$ as the dependence between fire and windstorm. These parameters are higher than the ones observed in reality; the strong dependence of the reference model is chosen to strengthen the effects studied in this article and to increase the understanding. We deliberately choose an extreme case to better study the effect of dependence on the diversification gain. We want to see how the various methods fair in reproducing a rather difficult case of dependent variables.

It is important to note that the dependence on top of the tree refers to the dependence between the aggregate loss of the two underlying baskets. This is in contradiction to other definitions which can be found in the literature. For instance, [Savu and Trede, 2006] define the dependence of a node between the cumulative copula densities of the underlying nodes, which is not equivalent to the aggregate loss. The main advantage of modelling dependence between aggregate losses is that it is more intuitive to think

12

about aggregate losses than it is to think about copula densities.

In the following sections, we will investigate the behaviour of this reference model structure and how it can be modelled under *misspecification*, i. e. using the 'wrong' functional form of the copula or using different hierarchies.

4.2 Methodology

The difficulty about our model setup is that many computational algorithms have to be combined to a consistent set of methods. The essential features needed are the generation of hierarchical random numbers and fitting methods for copulas. Both topics are well documented in the literature.

Algorithms to generate random variables from multivariate Archimedean copulas can be found for instance in [Whelan, 2004]. Gauss copulas are best sampled from using the methodology described by [Wang, 1998]. Student's T copulas can be generated out of the Gauss copulas by extending Wang's algorithm with the method described by [Demarta and McNeil, 2005].

In our study, we do not only want to generate samples from multivariate copulas, but from a hierarchical copula tree. The literature on this topic becomes sparser. Most authors follow the definition and approach as defined in [Savu and Trede, 2006]. However, as mentioned above, our definition of a copula tree is slightly different: we define dependence between aggregate losses rather than between cumulative copula densities as in [Savu and Trede, 2006]. We have therefore developed a reordering schema to generate hierarchical copulas, the description of which would lead beyond the scope of this article. It will be described in detail in a succeeding article [Bürgi and Müller, 2008].

In the study, we define a benchmark set of observations with small theoretical diversification benefits. We then fit those observations with various models of dependence. In the following paragraphs, we describe which methods we use to fit the various dependence model we presented in Section 3.

The fitting of Archimedean copulas was done with maximum likelihood estimation, see for instance [Savu and Trede, 2008]. Given N vectors of observations $\mathbf{U_1}, \ldots, \mathbf{U_N}$, the log-likelihood function of θ is defined as

$$\log L(\theta) = \sum_{j=1}^{N} \log c(\mathbf{U_j}), \tag{28}$$

where $c(\mathbf{U_j})$ is the copula density of $\mathbf{U_j}$. The estimator of θ is then given as the maximum of the log-likelihood function

$$\hat{\theta} = \max_{\theta} \log L(\theta). \tag{29}$$

The correlation matrix of the Gauss copula was determined by measuring Spearman's rank correlation as given in Equation (3) and deriving the correlation coefficient ρ_{ij} as

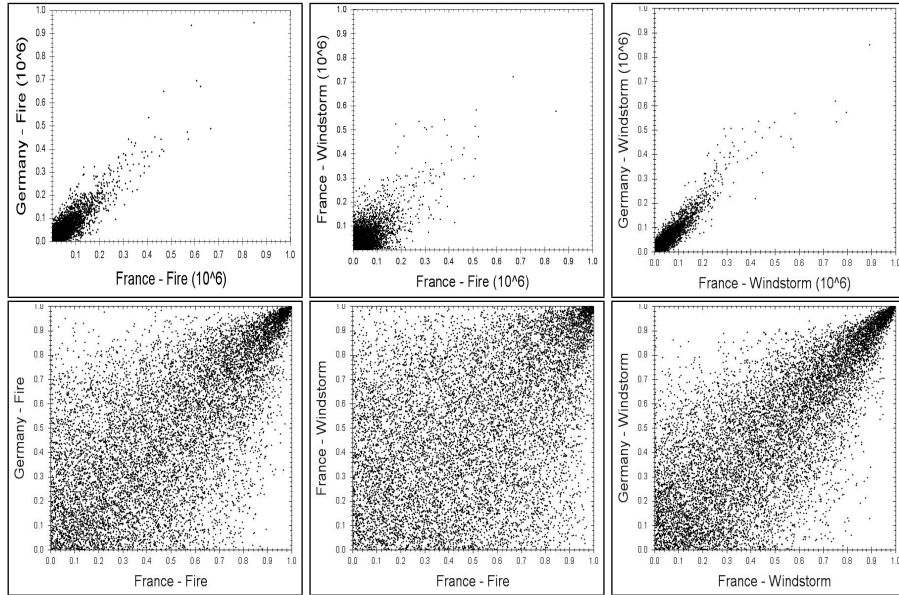$$\rho_{ij} = 2\sin\left(\frac{\pi}{6} RC_{ij}\right). \tag{30}$$

Figure 6: Value scatter (top) and rank scatter (bottom) of the reference perils.

This procedure is described in detail in [Wang, 1998].

The correlation matrix of the Student's T copula was estimated by estimating the Kendall τ_{ij} between the random variables X_i and X_j as given in Equation (4) and calculating the correlation

$$\rho_{ij} = \sin\left(\frac{\pi}{2}\tau_{ij}\right). \tag{31}$$

The degree of freedom parameter ν is estimated using maximum likelihood estimation. This method is described in more detail in [Demarta and McNeil, 2005].

5   Results and Discussion

To start the discussion on risk aggregation, we will analyse the reference model structure as described in Section 4.2. Value scatter plots of the risk factors of our reference model are shown in the top row of Figure 6. In the bottom row, the rank scatter plots, i. e. the plot of the cumulative probabilities of the marginal distributions, are shown. The left column is dedicated to the dependence between fire Germany and fire France, the right column to the dependence between windstorm Germany and windstorm France. In the middle column, the scatter plot of fire France with windstorm France can be seen. Note that the dependence on the latter perils is imposed through the top-level dependence in the reference tree, acting on the aggregate fire and windstorm losses. Since it is difficult to see structures on the value scatter plots, we will

Table 1: Parameters of the Gumbel (first row), the Gauss (second row), and the Student's T (third row) hierarchies fitted to the Clayton reference model for fire Germany / fire France (left), fire / windstorm (middle), and windstorm France / windstorm Germany (right).

	Fire	Fire/Windstorm	Windstorm
Gumbel	$\hat{\theta} = 2.07$	$\hat{\theta} = 1.54$	$\hat{\theta} = 2.61$
Gauss	$\hat{\rho} = 0.7$	$\hat{\rho} = 0.5$	$\hat{\rho} = 0.8$
Student's T	$\hat{\rho} = 0.71; \hat{v} = 6$	$\hat{\rho} = 0.51; \hat{v} = 9$	$\hat{\rho} = 0.81; \hat{v} = 4$

only use the rank scatter plots for the further discussions and we will call them simply scatter plots dropping the 'rank'.

5.1 Modelling under Misspecification

First of all, we will study the impact of the functional form for modelling dependence on the risk. For this purpose, we generate 100'000 samples from the reference structure, our benchmark sample, which we will use as the realisations, i. e. the 'observations'. The other copulas described in Section 3 are fitted to those realisations, one for each node of the dependence tree.

The fitted parameters can be seen in Table 1. Since the reference hierarchy is calibrated with a strong Clayton copula, also the fitted copula types show high values of their parameters. As a first analysis, let us look at the scatter plots in Figure 7. The Gumbel copula models the Clayton reference quite well. Due to its asymmetry, the fit of the upper tail looks optically similar. For low quantiles, the Gumbel copula models a weak dependence, whereas the Clayton implies almost independence in the lower tail. The Gauss and the Student's hierarchies, however, deviate quite significantly from the Clayton reference. Both these copulas are symmetric, and therefore, the fits of the lower tail (striving to uncorrelated) and of the upper tail (striving to high correlation) are competing with each other. Looking at the upper tail, which is of particular importance to us, we notice that the tail of the Clayton is much more pointed than the ones of the elliptic copulas. Due to its additional degree of freedom parameter, the Student's T copulas model more points in the upper left and lower right corners of the scatter plot than the Gauss copulas.

As a measure of the goodness of fit, let us analyse the quantile / quantile plots. They are shown in Figure 8. The picture we have found in the scatter plots is confirmed: The Gumbel fits the Clayton reference the best, whereas the Gauss and the Student's T deviate more pronouncedly from the Clayton reference. In the Gumbel plot, we notice the stronger dependence up to around the 0.2 quantile, which gets corrected in the upper range. In the case of the Gauss and the Student's T copula, the dependence in the lower left part of the quantile / quantile plot is much stronger, and therefore needs
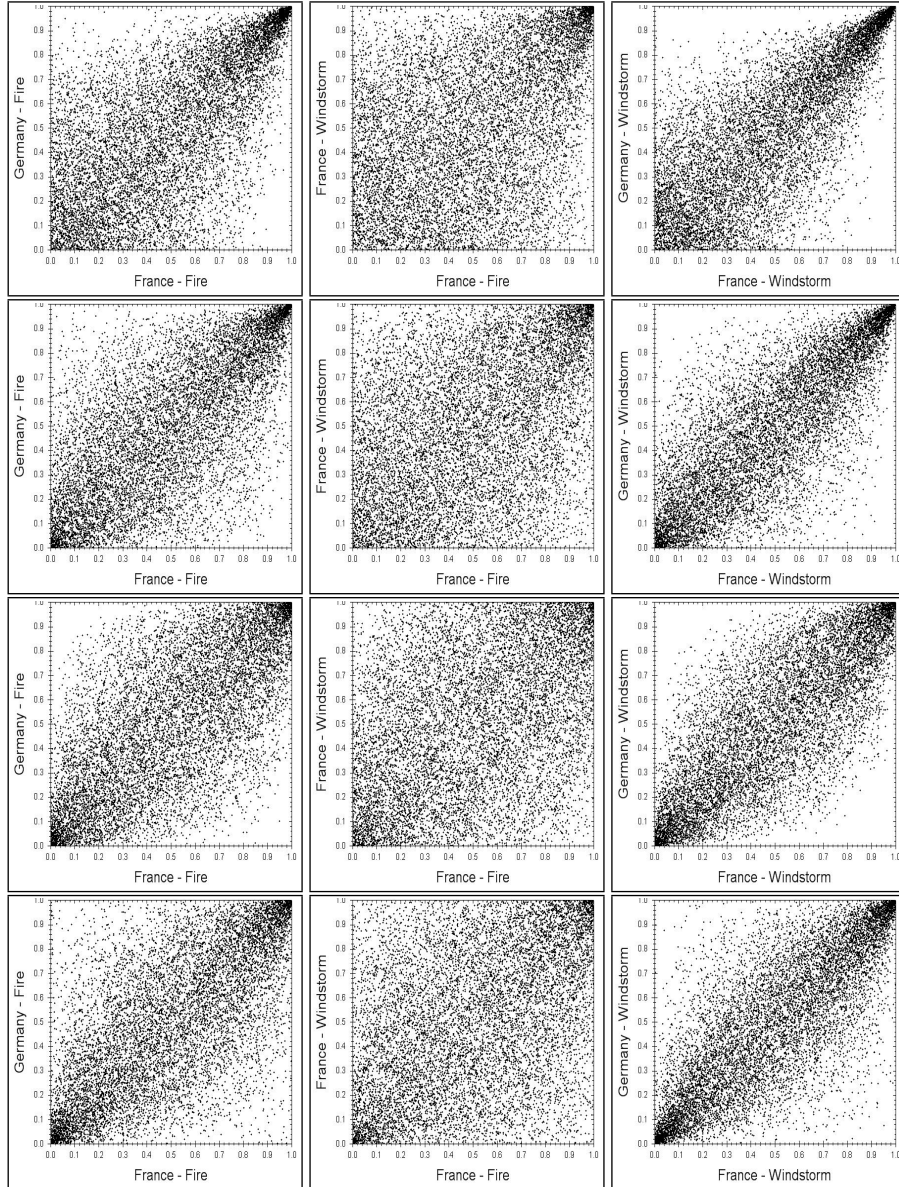
Figure 7: The hierarchical Clayton reference model (first row), and fitted Gumbel (second row), Gauss (third row), and Student's T (fourth row) hierarchies. In the columns, the plots for fire Germany / fire France (left), fire France / windstorm France (middle), windstorm Germany / windstorm France (right) are shown.
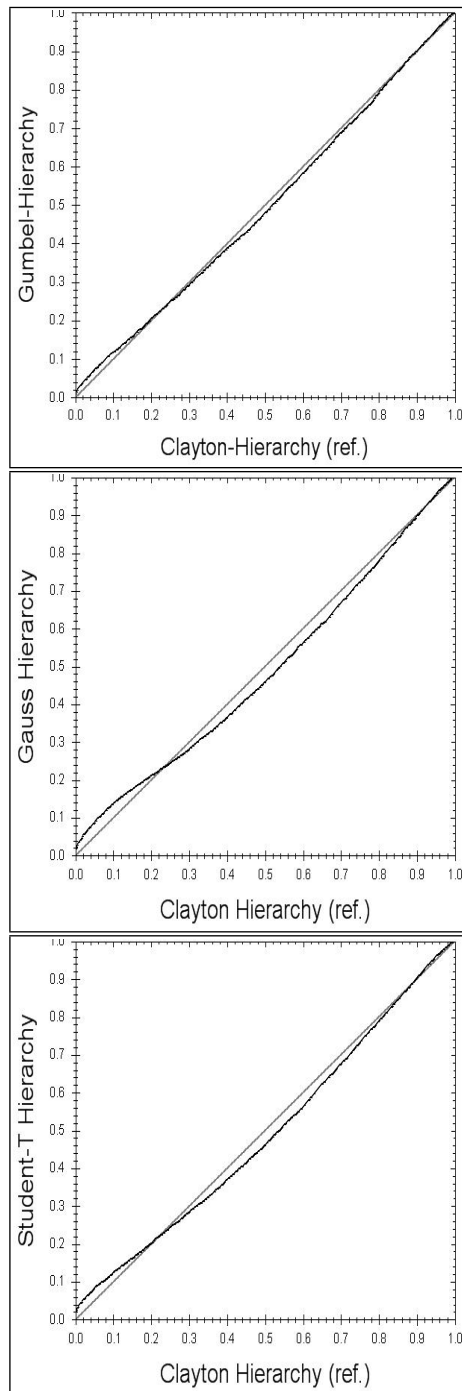
Figure 8: Quantile / quantile plots for the Gumbel (top-left), Gauss (top-right), and Student's T (bottom-middle) hierarchies versus the Clayton reference hierarchy.

17

Table 2: Key statistics for the Clayton reference and the fitted hierarchies. All figures, except for the diversification gain, are in thousands.

in [thousands]	Clayton	Gumbel	Student's T	Gauss
Mean	145	145	145	145
Std. Dev.	173	165	150	146
99.5% VaR	1'070	1'021	909	877
99% Shortfall	1'248	1'195	1'045	990
RBC	1'103	1'050	845	900
RH Div. Gain	**8.2%**	**11.7%**	**25.0%**	**30.3%**
LH Div. Gain	9.6%	5.7%	3.1%	3.9%

stronger correction in the upper part. It can be seen that the diagonal is approached at a higher value than in the case of the Gumbel. There are a number of more sophisticated goodness-of-fit methods used for copula fitting, as can be found for instance in [Savu and Trede, 2008] or [Berg, 2007]. However, as we are mostly interested in the diversification gain, we will focus on analysing the statistical properties of the fitted models.

As the distribution to measure the statistical properties and the diversification gain, we take the aggregate distribution as resulting from 250'000 simulations of the corresponding dependency trees. The key statistics of the Clayton reference and the fitted hierarchies can be found in Table 2. The mean of the aggregate distribution is independent of the dependence as is shown in Equation (1). However, in the standard deviation it can be seen that the asymmetric Archimedean copulas show significantly more volatility than the symmetric elliptical ones.

The left handed (LH) diversification gain is shown as a measure for the effect of the dependence in the lower tail of the aggregate distribution. The LH diversification gain of 9.6% of the Clayton reference may appear surprisingly low, since the risk factors are nearly independent in the lower tail. This is due to the asymmetry of the lognormal marginal distributions: The bulk of the risk of the lognormal distribution is in the upper tail. Not surprisingly, the Gumbel (5.7%) underestimates the LH diversification gain; this corresponds to the observation we have made for the scatter plots that the Gumbel models a slight dependence in the lower tail. The symmetric elliptical copulas underestimate the LH diversification gain considerably: The Gauss (3.9%) by 60% and the Student's T (3.1%) by almost 70%. However, the LH diversification gain is only of limited importance to us. Let us now focus on the risk in the upper tail.

Considering the 99.5% value at risk (VaR), the differences in shapes of the copula models become even more apparent. Whereas the VaR of the Gumbel (1'021'000) deviates only by a few percent from the Clayton reference (1'070'000), the Gauss (877'000)

underestimates the 99.5% VaR by more than 20%. The Student's T (909'000) fits the Clayton reference slightly better than the Gauss due to its additional parameter ν. However, the deviation is still more than 10%.

The picture is similar and even accentuated when considering the 99% shortfall: The Gumbel (1'195'000) approximates the Clayton reference (1'248'000) the best, the Student's T (1'045'000) underestimates the shortfall by more than 10%, and the Gauss (990'000) shows the lowest shortfall, deviating from the reference by more than 20%. The deviations of the risk-based capital (RBC), being defined as the difference between the shortfall and the mean in Equation (9), worsen the picture even more dramatically: by subtracting the same constant of all the shortfalls, the relative deviations become higher.

The diversification gain is defined as the gain in the capital when combining all four risk factors in one company as opposed to running four companies with the risk factors individually, see Equation (10). In this projection of our capital figures, a highly relevant projection for all insurance companies, the deviation becomes the worst: The Clayton reference shows a diversification gain of 8.2%. The Gumbel (11.7%) overestimates the diversification gain by over 40%, whereas the Student's T (25%) overestimates it by over 200%, and the Gauss (30.3%) overestimates it by 270%. Using the elliptical copulas to model the dependencies of the risk model would lead to a gross underestimation of the risk of this company.

To illustrate the consequences such an underestimation of the risk based capital can have, assume that our model reinsurance company has implemented an active capital management. If the company calculated its capital using the wrong functional form of the copula, it would pay back too much capital to its shareholders, up to 20% of its total risk-based capital in the case of the Gauss. If reality follows the highly asymmetric Clayton copula, the model company will need to increase its capital again when the losses need to be paid. Since this capital increase is not foreseen, the company might have to pay a high price to raise funds under financial distress.

5.2 The Effect of the Hierarchy

An important aspect when modelling dependence of insurance risks is to decide which underlying risks can be modelled separately. The advanced (re)insurance companies have integrated systems which can provide information about a high granularity of risk factors. However, there is currently still a number of companies where risk factors are classified in a higher level of aggregation. Most risk models will also contain a model basket labelled 'miscellaneous' in which all the business is allocated that is too small to be separately modelled. The 'true' dependence of these miscellaneous risk factors is then often approximated by one dependence over the entire basket. Moreover, any internal model will contain many risk factors, of the order of a few hundreds. Trying to model the dependence of each factor with each other would require the modelling of tens of thousands of parameters. This would be impracticable in terms of calibration, and the interpretation of most of these parameters would make little economical sense. A hierarchical tree reduces the number of parameters to estimate drastically. For all

these reasons, the influence of the dependence structure is of high relevance. This is what we study in this subsection.
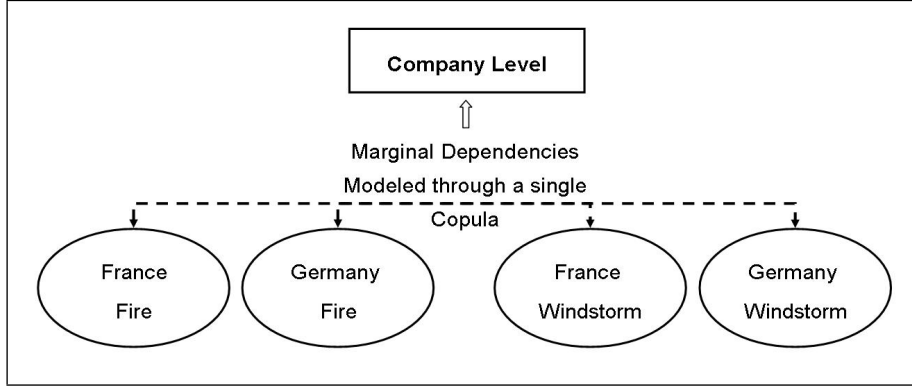

Figure 9: Portfolio modelled with a flat dependence.

We will now investigate how our model portfolio behaves if the dependence is modelled in one flat basket as shown in Figure 9 rather than the hierarchical tree as considered in the previous section. First of all, we fit a single flat Clayton copula to the hierarchical reference tree. We obtain $\hat{\theta} = 1.2$. We repeat the same process for a flat Gumbel and obtain $\hat{\theta} = 1.55$. For the Gauss and Student's T, we allow individual correlations for each pair of risk factors, which are estimated as

$$
\hat{\Sigma}_G = \begin{pmatrix} 1.0 & 0.7 & 0.45 & 0.46 \\ 0.7 & 1.0 & 0.45 & 0.46 \\ 0.45 & 0.45 & 1.0 & 0.8 \\ 0.46 & 0.46 & 0.8 & 1.0 \end{pmatrix}; \quad \hat{\Sigma}_S = \begin{pmatrix} 1.0 & 0.71 & 0.45 & 0.46 \\ 0.71 & 1.0 & 0.45 & 0.46 \\ 0.45 & 0.45 & 1.0 & 0.81 \\ 0.46 & 0.46 & 0.81 & 1.0 \end{pmatrix}. \quad (32)
$$

The degree of freedom parameter of the Student's T is estimated as $\hat{\nu} = 10$.

For the elliptical copulas, the main difference between the hierarchy and the flat structure consists in the correlation between the fire and the windstorm risk factors. The reason for this is that the original Clayton reference tree models a dependence between the sum of the fire and windstorm perils. The flat structure, however, assumes a correlation between the underlying risk factors, and therefore, these correlations are underestimated. Note also that the degree of freedom parameter of the Student's T is higher than in any of the three previous nodes of the hierarchical tree. Thus, the Student's T copula can compensate slightly for the underestimated correlations.

Looking at the scatter plots in Figure 10, we see that the flat Clayton and Gumbel have the same scatter plot for all pairs of risk factors, since we have only allowed one parameter in the model. The elliptical copulas show a similar scatter plot as shown in Figure 7 for their hierarchical variants; the differences in their parameters are rather small compared to the hierarchical case.

In Figure 11, we show the quantile / quantile plots for the flat fitted baskets versus the hierarchical Clayton reference tree. It is interesting to note that, despite the flat Clayton
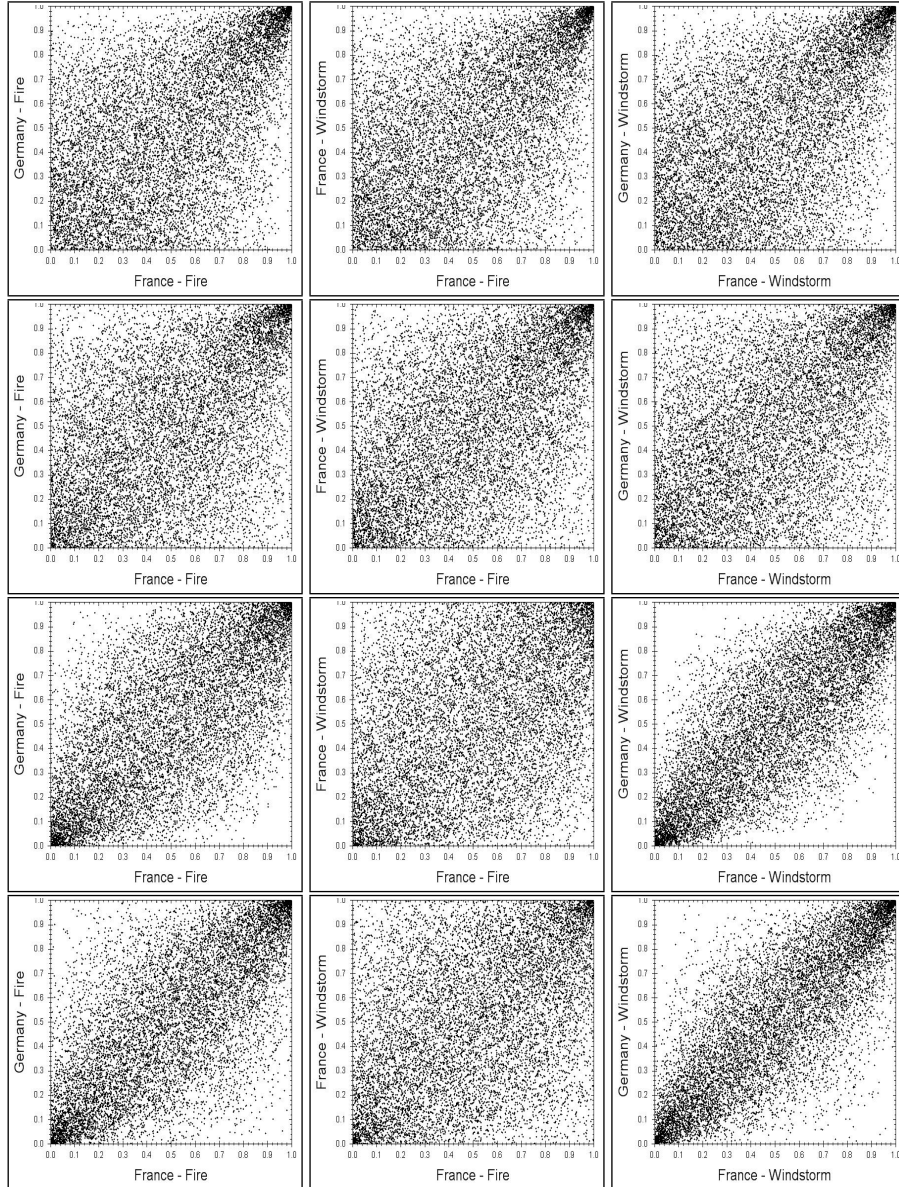
Figure 10: The fitted flat Clayton (first row), flat Gumbel (second row), flat Gauss (third row), and flat Student's T (fourth row) hierarchies. The columns show the plots for fire Germany / fire France (left), fire France / windstorm France (middle), windstorm Germany / windstorm France (right).
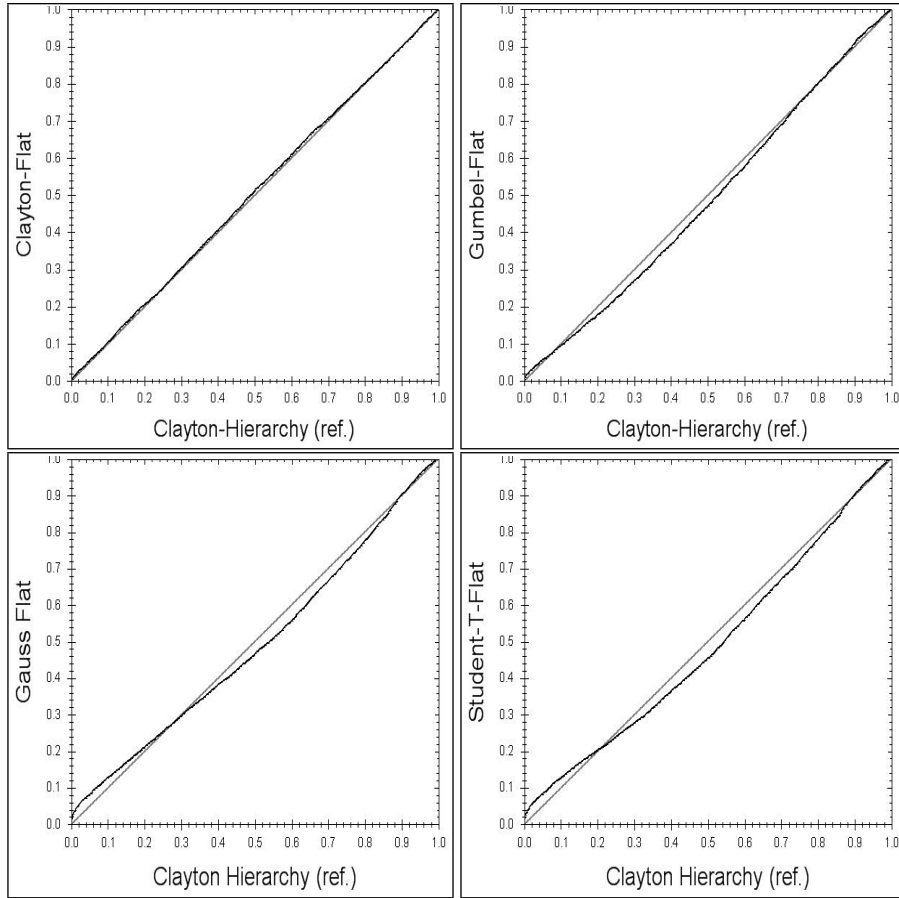
Figure 11: Quantile / quantile plots for the flat Clayton (top left), Gumbel (top right), Gauss (bottom left), and Student's T (bottom right) versus the Clayton reference hierarchy.

Table 3: Key statistics for the Clayton reference and the fitted flat baskets. All figures, except for the diversification gain, are in thousands.

in [thousands]	Clayton ref.	Clayton	Gumbel	Student's T	Gauss
Mean	145	145	145	145	145
Std. Dev.	173	173	163	151	144
99.5% VaR	1'070	1'078	1'029	902	870
99% Shortfall	1'248	1'251	1'189	1'043	978
RBC	1'103	1'106	1'044	898	832
RH Div. Gain	**8.2%**	**7.8%**	**12.7%**	**25.4%**	**30.3%**
LH Div. Gain	9.6%	9.14%	7.2%	3.2%	4.0%

having only one parameter, and despite the model difference of fitting the dependence to all underlying risk factors rather then the sum of fire and windstorm, and despite the obvious differences in the scatter plots, the flat Clayton model fits the hierarchical reference tree almost perfectly. The quantile / quantile plots of the other fits resemble the ones of their corresponding hierarchical variants shown in Figure 10, with slightly amplified deviations due to the poorer modelling.

The key statistics of the Clayton reference and its fitted flat structures, sampled from 250'000 simulations, are shown in Table 3. The strong similarity observed in the quantile / quantile plot can also be seen in the statistical figures: The flat Clayton (7.8%) models the Clayton reference extremely accurately; it even underestimates the reference slightly as the only model. All other flat copulas yield slightly worse but very similar results as seen for their hierarchical counterparts shown in Table 2. All deviations between the flat and the hierarchical variants are below 10%.

In conclusion it can be said that the error of modelling flat structures rather than hierarchical structures does not have a strong effect on the diversification gain, even when three parameters are replaced by one as in the case of the Clayton and the Gumbel. This shows that the choice of the functional form of the copula is the key driver of the diversification gain. The asymmetry of the Archimedean copulas is the essential property which allows for accurate representation of the asymmetric reference data. Even though the flat Gauss and the flat Student's T have more parameters than the flat Clayton and the flat Gumbel, they overestimate the diversification gain considerably.

5.3   Extending the Reference Model

In section 5.2 we investigated the effect of misspecifying the hierarchical tree of dependent risks with a flat model structure. The results indicated that, in the case of a two-level hierarchy with four risk factors, the diversification benefit can be well mod-

elled by a flat structure. In this section we extend the investigation to see under which circumstances the modelling of the hierarchical structure as opposed to a flat model matters.

First of all, we extend the model from four to eight risk factors. This increases the variety of dependencies which can be modelled. We remain with the choice of equal risk factors, all lognormal distributed with $\mu = 10$ and $\sigma = 1$, as we are mostly interested in the dependence. As the reference hierarchy, we choose a three-level tree as shown in Figure 12.

The eight risk factors are grouped in four pairs, each with a separate dependence. The four aggregate distributions are grouped in two pairs with a dependence, and their two aggregate distributions underlie the top-level dependence.

We stick to the Clayton copula as the functional form of the copula of the reference model to produce a benchmark sample. In order to find out where the misspecified flat model becomes inaccurate, we investigate six different reference scenarios defined in table 4 by various combinations of the Clayton θ. The goal is to see if some combinations would have stronger effects than others.

Scenario 1 corresponds to a straight-forward extension of our original four risk-factor model hierarchy, assigning $\theta = 1, 2, 3$ to the nodes with θ decreasing towards the top level. In Scenario 2, we leave the very strong dependencies at the bottom level of the tree but relax the parameters considerably for higher aggregations. In Scenario 3 we choose generally lower θ values, which corresponds to a moderately dependent risk model in insurance. Scenario 4 describes a highly asymmetric situation: one half of the risk factors diversify much better than the other half of the risk factors; however, there is a strong top-level dependency. In Scenario 5, we try a 'sandwich' hierarchy: high θ for the bottom level, low θ for the middle level, and high θ for the top level. Scenario 6 is the reverse of Scenario 5: Low dependency at the bottom, strong dependency in the middle, and low dependency on top.

As for the functional form of the flat, misspecified structure, we will reduce our investigation to the Clayton and the Gauss copulas. For the four risk factors, we have seen

Table 4: Clayton θ parameters for the six reference scenarios.

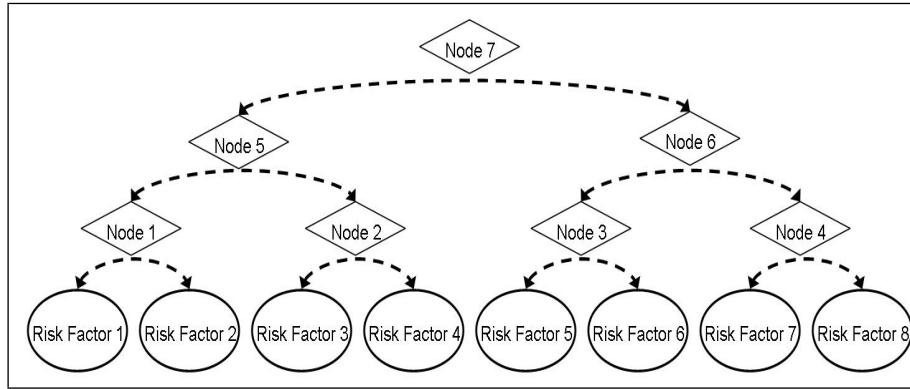| Scenarios	Node 1	Node 2	Node 3	Node 4	Node 5	Node 6	Node 7
Scenario 1	$\theta = 2.0$	$\theta = 3.0$	$\theta = 3.0$	$\theta = 3.0$	$\theta = 1.0$	$\theta = 2.0$	$\theta = 1.0$
Scenario 2	$\theta = 2.0$	$\theta = 3.0$	$\theta = 3.0$	$\theta = 4.0$	$\theta = 0.75$	$\theta = 0.5$	$\theta = 0.3$
Scenario 3	$\theta = 0.5$	$\theta = 0.4$	$\theta = 0.5$	$\theta = 0.3$	$\theta = 0.3$	$\theta = 0.2$	$\theta = 0.1$
Scenario 4	$\theta = 0.1$	$\theta = 0.2$	$\theta = 2.0$	$\theta = 3.0$	$\theta = 0.5$	$\theta = 3.0$	$\theta = 3.0$
Scenario 5	$\theta = 1.5$	$\theta = 2.0$	$\theta = 3.0$	$\theta = 4.0$	$\theta = 0.2$	$\theta = 0.3$	$\theta = 2.0$
Scenario 6	$\theta = 0.1$	$\theta = 0.2$	$\theta = 0.1$	$\theta = 0.3$	$\theta = 2.5$	$\theta = 3.0$	$\theta = 0.1$

Figure 12: The hierarchical reference model setup. The eight risk factors are grouped in pairs over three hierarchy levels. Each node can be modelled with a different copula.

that the flat Clayton structure matched the hierarchical reference almost perfectly in terms of its aggregate distribution. We will investigate to which degree this similarity holds for the six scenarios described in this subsection. The Gauss copula is chosen as the one fitting the Clayton reference the worst. As for the other two copulas described in this article, the Gumbel and the Student's T, the results would be similar as the ones of the four marginal cases: they are in between the Clayton and the Gauss, the Gumbel closer to the Clayton, and the Student's T closer to the Gauss.

The diversification gain of the hierarchical Clayton reference model, as well as the flat Clayton and flat Gauss models, are shown for all six scenarios in Table 5. It is apparent that all the diversification gains of the flat Clayton structures are in the same range as the ones for the hierarchical reference models¡indexreference model. Most deviations are of the order of 10%. The flat structure overestimates the diversification gain in two scenarios and underestimates it in four scenarios. Nevertheless, the difference

Table 5: Diversification Gain results of the Clayton Hierarchy reference model and the fitted flat Clayton and flat Gauss models.

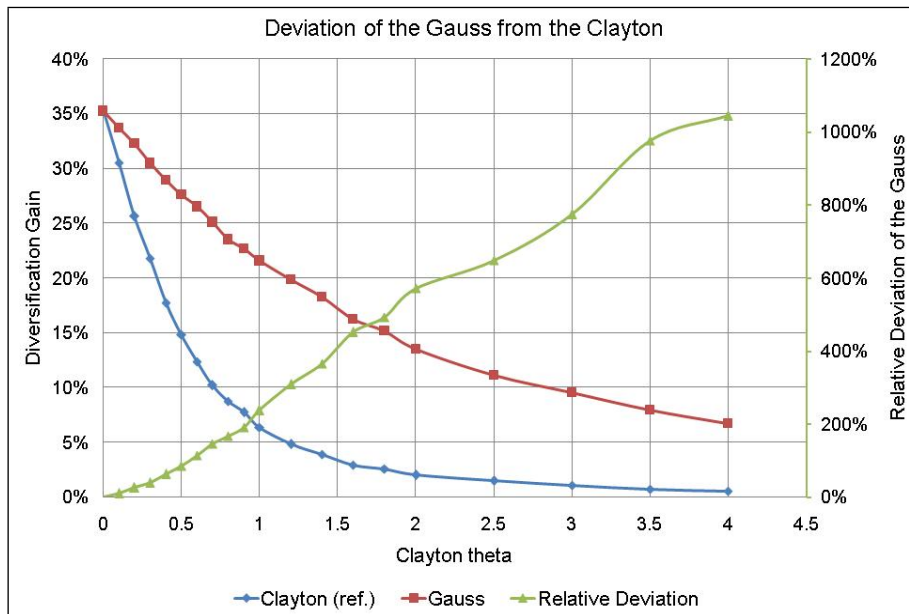|	Reference	Clayton-Flat	*Deviation*	Gauss-Flat	*Deviation*
Scenario 1:	12.1%	14.2%	*17.4%*	42.4%	*250.0%*
Scenario 2:	32.0%	29.3%	*-8.4%*	54.3%	*69.7%*
Scenario 3:	55.6%	58.8%	*5.8%*	68.7%	*23.6%*
Scenario 4:	21.5%	19.9%	*-7.4%*	44.9%	*108.8%*
Scenario 5:	26.7%	21.2%	*-20.6%*	50.5%	*89.1%*
Scenario 6:	50.1%	45.3%	*-9.6%*	63.4%	*26.5%*

Figure 13: The absolute and relative deviation between the diversification benefit of the Clayton and the Gauss for increasing dependency between two risk factors.

is more pronounced than in the simpler structure above indicating that if the structure becomes more complex the flat scenario will deviate more and more from the structure. Scenario 1, the one close to the one we have investigated for the case of four risk factors, shows one of the biggest deviations of the flat Clayton versus the reference structure: 17.4%. This is mainly due to the overall strong Clayton parameters chosen. In Scenario 3, the one with moderately dependent risk factors, the deviation between the hierarchical structure and its flat variant is only 5.8%.

As observed in all previous studies, the Gauss always overestimates the diversification gain of the reference structure in all scenarios. In Scenario 1, the deviation of 250% is similar to the one observed in the case of four risk factors. In Scenario 3, the deviation of 23.6% is still considerable. Considering the rather strong deviations, it may also be interesting to note that we have allowed for fitting 28 correlation parameters in the case of the Gauss, as opposed to the single parameter in the case of the flat Clayton. Nevertheless, the diversification gains of the flat Gauss models deviate generally more than the ones of the flat Clayton models.

Concluding, the error of approximating a hierarchical Clayton reference model with eight risk factors by a flat Clayton structure is small in terms of the overall diversification gain. For moderately dependent risk factors, it is below 10%. Nevertheless, this conclusion is to be taken with care: in this study, we have exclusively focused on the aggregate distribution and the overall diversification gain. As soon as *capital allocation* or the diversification gain by risk factor is of interest, the detailed dependencies of the

hierarchical reference model would matter significantly. This can be easily seen when considering that the capital allocation in the case of a flat Clayton structure would be the overall capital divided by eight, whereas in the case of the hierarchical structure, the risk factors with the stronger dependencies would attract by far more capital. Moreover, we see a tendency of increasing difference when we increase the complexity of the structure.

In the above case we have examined the effect of complexity on diversification gain but it is difficult to get a feel for how the strength of the dependence between risk factors influences the ability of the models to cope. Generally we have modelled extreme dependence between risk factors to demonstrate the significance of misspecification, however, it is important to also understand the effect for weaker dependencies between risk factors as these better reflect reality. To investigate this effect we show in Figure 13 how the diversification gain modelled by the Gauss deviates from the Clayton reference as a function of increasing dependence between two risk factors. To help relate the mathematical world of copulas to real-world situations; two risk factors are typically considered to be strongly dependent when they are modelled with a tail-dependency of 50%, which corresponds to a Clayton $\theta = 1$. It is interesting to note that the largest absolute deviation of 15.2% of RBC between the Gauss and the Clayton occurs for $\theta = 1$. As the dependence increases further the absolute deviation between the Gauss and the Clayton declines again as diversification gain of the Clayton tends towards zero. For $\theta < 1$, a region that can be considered to represent tail-dependencies for most other real-world (re)insurance scenarios, there also exists a notable relative deviation that is already at 10% at $\theta = 0.1$. The relative deviation rapidly increase as the dependence between the two risk factors grows stronger, reaching 237% at $\theta = 1$ and 1044% at $\theta = 4$!

We have so far demonstrated the importance of selecting the correct functional form when selecting a copula for modelling and the effect of complexity on the ability to obtain a good fit. We have also tried to relate these results to the real-world. In the following two sections we now deal with the practical aspects of fitting; convergence and error; and give some insights into how to go about deriving the correct dependence between risk factors.

5.4   Fitting Copulas: How Many Points are Needed?

One of the big problems when calibrating internal models is the estimation of dependence parameters. The data at hand in insurance are often sparse and it is difficult to obtain relevant samples over more than a decade or two in the best cases. It is thus important to explore what are the size of samples that are needed to reasonably fit copulas. This is the last question we are going to address in this study: how many points are needed to fit the copulas? It is interesting to see if there is a systematic small sample bias for certain copula types and whether certain copulas are easier to fit than others.

We investigate all the hierarchical as well as the flat structures described for four risk factors in sections 5.1 and 5.2. The procedure is the following:

1. Simulate *N* observations from the reference tree,

2. Fit the corresponding structure to the *N* observations,

3. Re-sample the fitted scenario with 100'000 simulations,

4. Measure the diversification gain,

5. Repeat 1. – 4. twenty times.

6. Measure the mean diversification gain and the standard deviation from the twenty calculations.

The fitting convergence plot produced according to this procedure is shown in Figure 14 for all structures.

For each point, the diversification gain plus/minus one standard deviation is drawn (top). Since it is difficult to read the error bars, the fitting errors, i.e. the standard deviation of the diversification gain measurements, are plotted separately (bottom).

The main observation of the fitting convergence plot is a confirmation of our preceding analysis: There is little difference between the hierarchical and the flat structures. The choice of the functional form is by far more important than the number of observations the copulas are fitted to.

The small sample bias seen in the plots is not significant. Not surprisingly, the diversification gain is slightly overestimated for a small number of observations, as not sufficient extreme events are used for the fits. Nevertheless, the error bars are sufficiently large to include the value of the diversification gain as fitted for a high number of observations. When fitting the structures to 100 observations, the small sample bias can be hardly seen in the average estimates.

In order to have a better understanding of the fitting error, let us analyse the bottom chart of Figure 14. All errors decrease significantly when moving from 20 to 100 observations. There is still a considerable drop of the errors when moving from 100 to 500 observations, whereas the errors remain rather stable for higher numbers of observations. It is interesting to note that the hierarchical Archimedean copulas have significantly higher errors for a low number of observations than their flat variants. This can be explained by the flat structures considering all observation points for fitting one parameters, whereas the parameters of the hierarchical structures are fitted each on a subset of the observation. These differences, however, vanish when moving to 100 and to 500 observations.

The flat Archimedean copulas show in tendency a lower error for a low number of observations than the elliptical copulas. The reason for this is the following: Considering any bivariate copula, and dividing the scatter plot into four quadrants, the probability is point symmetric around the center point. This means that the probability mass in the lower left and in the upper right quadrants, as well as the lower right and the upper left quadrants, are equal. For a low number of points, a slight shift of the number of observations in the upper or lower tail can have a significant effect on the fit of a symmetric elliptical copulas: more points in the upper tail will lead to high correlation,
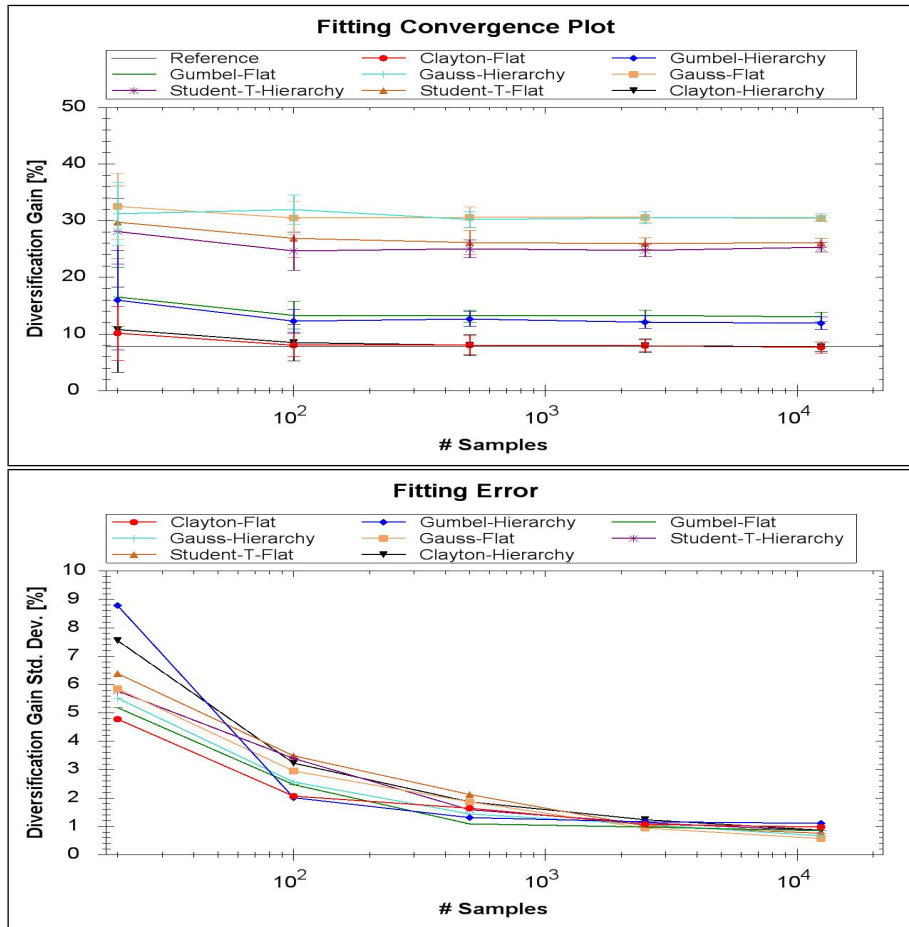
Figure 14: Fitting convergence plot (top) for the hierarchical and flat copula models. The mean diversification gain ± one standard deviation is shown as a function of the number of samples. The errors are shown separately (bottom).

more points in the lower tail to low correlation. The asymmetric Archimedean copulas can absorb a shift in the number of lower and upper tail points better, since they can account for the asymmetry.

In general, the convergence behaviour of the diversification gain when varying the number of observations is similar for all copulas investigated. The average estimate of the diversification gain when fitting to 20 observations is not significantly different from the average fitting to a high number of observation. The errors of the estimates of all structures converge quickly to a similar value.
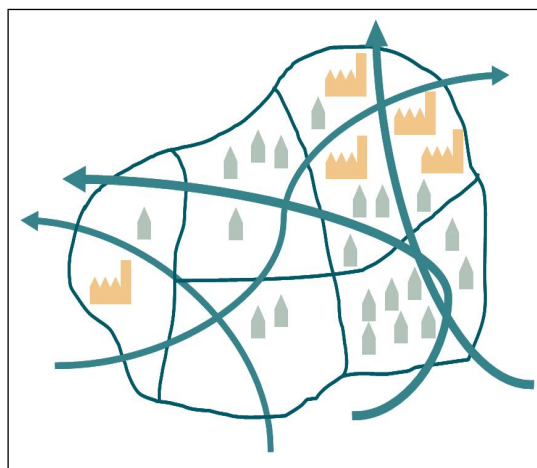
Figure 15: Example of an event set based wind model.

6   Estimating Dependence

When considering the hundreds of observations needed to fit the dependence of 4 risk factors, in reality this would mean collecting (aggregate) loss experience during hundreds of years. Even if we could wait to collect the observations needed, it would be questionable what the significance of the fitted model is; the definitions and behaviour of insurance products would change significantly over this period of time, and the assumption that these observations underlie the same model would not be valid. Therefore, alternative ways of modelling dependence have to be sought.

An elegant means of modelling dependence is to use an explicit physical model whenever one is known. This is the case, for instance, for natural hazards. The contemporary models are event loss set based, i. e. the single wind, earthquake, or flood events are modelled together with a frequency of their occurrence. Given these events, the whole portfolio of risks can be evaluated yielding the aggregate loss by event. Figure 15 illustrates an example of an event set based wind model. Individual risk factors can be considered by describing the vulnerability of the buildings, their insured value, as well as their insurance conditions. By aggregating over all events, the dependence of the risks is implicitly reflected.

The example of the natural catastrophe perils is the most prominent and most advanced one in terms of modelling. However, many (re)insurance companies have started constructing similar individual risk or scenario models, for instance for pricing auto liability policies. There is still a significant potential to establish further individual risk and scenario based models for insurance products.

If no physical dependence model can be found, the dependence has to be described by means of statistical properties, e. g. by using a copula. As discussed in this article, the choice of the functional form is paramount even if the parameter can not be precisely

30

Table 6: Relationship between the tail dependence and the θ-parameter for Clayton and Gumbel copulas.

Tail dep.	1%	2%	5%	10%	20%	30%	40%	50%	60%
Clayton θ	0.15	0.18	0.23	0.30	0.43	0.58	0.76	1.00	1.36
Gumbel θ	1.01	1.02	1.04	1.08	1.18	1.31	1.48	1.71	2.06

estimated. The form of the hierarchical tree is fairly comfortable in its usage, as there is an intuitive meaning associated with each node.

Since the copula parameters often lack an intuitive interpretation, considerations can be made about the *tail dependence* of the lines aggregated in each node. The *upper* tail dependence λ_U and *lower* tail dependence λ_L of a bivariate distribution are defined as

$$\lambda_U \quad = \quad \lim_{q \to 1} P(X_2 > F_2^{-1}(q) | X_1 > F_1^{-1}(q)), \tag{33}$$

$$\lambda_L \quad = \quad \lim_{q \to 0} P(X_2 < F_2^{-1}(q) | X_1 < F_1^{-1}(q)). \tag{34}$$

Detailed discussion on tail dependence can be found in [Embrechts et al., 2003]. It can be shown that the tail dependence is a property of the bivariate copula only:

$$\lambda_U \quad = \quad \lim_{q \to 1} (1 - 2u + C(u, u))/(1 - u), \tag{35}$$

$$\lambda_L \quad = \quad \lim_{q \to 0} (2u - 1 + C(u, u))/u. \tag{36}$$

The idea behind the tail dependence is to ask the question with which probability a line of business has an extreme loss given that another line of business has an extreme loss. The answer to this question can be given by thinking about adverse scenarios in the portfolio. In this context, it might be especially interesting to investigate the causal relations between the individual risk factors. Interviews with the risk specialists regarding quantitative impact of adverse scenarios can help understanding extreme scenarios and their dependence. This is stress testing by analysing possible outcomes of extreme scenarios.

The relationship between the tail dependence and the copula parameters can be found by evaluating Equations (35) and (36) for the given copula. For the Gauss copula, it can be found that the tail dependence does not exist: the dependence implied by the rank correlation is not strong enough to yield a value $\lambda_U > 0$. For the Clayton and Gumbel copulas, however, the tail dependence does exist, and the parameters can be associated with it as follows:

$$\theta_{\text{Clayton}} \quad = \quad -\frac{\log 2}{\log \lambda}, \tag{37}$$

$$\theta_{\text{Gumbel}} \quad = \quad \frac{\log 2}{\log(2 - \lambda)}. \tag{38}$$

Table 6 shows a range of values of the tail dependence and the corresponding Clayton and Gumbel θ values.

As an example, consider a product which is highly dependent on the legislation and court practice, such as a professional liability product. Two legal entities within the European Union write substantial volumes of this product in two different countries. The specialists define an extreme scenario in which a change in legislation would cause an extreme loss in the one of the countries. Since all the EU countries will develop their legislation somewhat in parallel, the specialists estimate that – given the legislation is changed in one country – the legislation would be changed in the other country with a probability of 50%. If the risk modeller decides to use a Clayton copula to model the dependence, the choice of the parameter would be around θ = 1.

7 Conclusions

In this article, we have analysed the impact of the functional form of a copula and the hierarchical dependence structure on the risk-based capital and on the diversification gain of a portfolio of risks, which is characterized by a strong dependence in the extreme cases. This is a typical situation in insurance and as we have experienced it in financial markets.

The main conclusion to be drawn is that the functional form of the copula chosen for modelling has a significant impact on the diversification gain: Given an asymmetric Clayton reference model (diversification gain of 8.2%), the use of symmetric copulas grossly overestimate the diversification gain (diversification gain of over 30% in the case of the Gauss copula overestimating the 'true' diversification gain by 270%). Even though the fitted parameters of the symmetric Gauss and Student's T copulas are perceived as rather conservative, they do not manage to compensate for the symmetry effects and their lack of tail dependence in the shortfall. The Gumbel copula, which models a slight dependence in the lower tail, still yields a fairly accurate diversification gain, as it can deal better with the asymmetry of the reference model.

The impact of the hierarchy is less severe than the choice of the functional form. By reducing the hierarchical structure of the reference model to one flat node, the resulting diversification gain is slightly higher than in the case of the hierarchical variants in most cases. The only structure which slightly underestimates the diversification gain is the flat Clayton structure, which manages to model the aggregate distribution almost perfectly. In general, the asymmetric Archimedean copulas model the reference structure more accurately than the symmetric elliptical copulas even if they are reduced to a single parameter, whereas many parameters are allowed for the elliptical copulas (three and four in the case of four risk factors, 28 and 29 in the case of eight risk factors). The functional form of the copula is the decisive choice and by far more important than the dependence structure.

When varying the number of realisations based on which the copulas are fitted, we note that the small sample bias is not significant. On average, the level of the diversification gain is reached even when fitting to 20 observations, however with an error of 60% -

100% of the reference diversification gain. The error drops considerably to around 20% - 30% of the reference diversification gain when fitting to hundreds of observations, which provides a reasonable basis for fitting dependence of four risk factors.

The conclusions of these observations are to be taken with care; let us bear in mind that a book of business is often subdivided in significantly more than four risk factors. Furthermore, we have only looked at the convergence of the overall diversification gain. As soon as partial results are in the focus of interest, for instance the capital allocation by risk factor, the detailed dependence structure of the risk model matters. Also the convergence behaviour of the fits will be slower. Therefore, the number of observations needed to fit the dependence of an entire risk model would increase. Thus a hierarchical structure, which focuses on few dependencies, is well suited to calibrate the internal model.

In practice, it is hardly possible to fit copulas for insurance risks based on historical observations. Alternatives have to be sought for estimating dependence. For some of the insurance products, it is possible to create explicit physical models which implicitly contain their dependence structure. This is an elegant alternative to using copulas. However, if no physical models can be found, the modelling with copulas is a good option.

The tail dependence is an intuitive quantity relating probabilities as defined in extreme scenarios with copula parameters. The modelling of scenarios is often easier for risk specialists than the modelling of statistical properties. Since the Gauss copula, i. e. the rank correlation, does not have tail dependence, this approach cannot be used. However, when using Clayton or Gumbel copulas, easy relationships between the tail dependence and the copula θ can be found, and the extreme scenarios can be used to calibrate the copulas. So much for stress testing!

References

[Acerbi and Tasche, 2002] Acerbi, C. and Tasche, D. (2002). On the coherence of expected shorfall. *Journal of Banking & Finance*, 26(7):1487–1503.

[Armstrong, 2003] Armstrong, M. (2003). Copula catalogue, part 1: Bivariate Archimedean copulas. `http://www.cerna.ensmp.fr/Documents/MA-CopulaCatalogue.pdf`.

[Artzner et al., 1999] Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999). Coherent risk measures. *Mathematical Finance*, 9:203–228.

[Berg, 2007] Berg, D. (2007). Copula goodness-of-fit testing: an overview and power comparison. *Statistical Research Report, University of Oslo*.

[Bürgi and Müller, 2008] Bürgi, R. and Müller, U. (2008). Aggregating risks in hierarchical copula trees. *To be published*.

[Demarta and McNeil, 2005] Demarta, S. and McNeil, A. J. (2005). The T copula and related copulas. *International Statistical Review*, 73(1):111–129.

[Embrechts et al., 2003] Embrechts, P., Lindskog, F., and McNeil, A. (2003). *Handbook of Heavy Tailed Distributions in Finance*, chapter Modelling Dependence with Copulas and Applications to Risk Management. Elsevier.

[Embrechts et al., 2002] Embrechts, P., McNeil, A., and Straumann, D. (2002). *Risk Management: Value at Risk and Beyond*, chapter Correlation and Dependence in Risk Management: Properties and Pitfalls. Cambridge University Press.

[Fang and Zhang, 1990] Fang, K. T. and Zhang, Y. T. (1990). *Generalized Multivariate Analysis*. Springer.

[Genest and Favre, 2007] Genest, C. and Favre, A.-C. (2007). Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering*, pages 347–368.

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

[Savu and Trede, 2006] Savu, C. and Trede, M. (2006). Hierarchical Archimedean copulas. `http://www.uni-konstanz.de/micfinma/conference/Files/papers/Savu_Trede.pdf`.

[Savu and Trede, 2008] Savu, C. and Trede, M. (2008). Goodness-of-fit tests for parametric families of Archimedean copulas. *Quantitative Finance*, 8(2):109–116.

[Sklar, 1958] Sklar, A. (1958). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de L'Université de Paris*, 8:229–231.

[Tasche, 2002] Tasche, D. (2002). Expected shortfall and beyond. *Journal of Banking & Finance*, 26(7):1519–1533.

[Wang, 1998] Wang, S. S. (1998). Aggregation of correlated risk portfolios: Models and algorithms. *Proceedings of the Casualty Actuarial Society*, pages 848–939.

[Whelan, 2004] Whelan, N. (2004). Sampling from Archimedean copulas. *Quantitative Finance*, 4:339–352.

Index

Dr Roland Bürgi is a senior risk consultant in Group Financial and Risk Modelling at SCOR. His current role is as business architect of Asset and Liability Management Programme, a project portfolio with the scope of implementing an integrated quantitative risk management in SCOR Group. Before joining Converium, which later became SCOR Roland Bürgi has worked on the pricing methodology at Swiss Re, where he has led and successfully completed several large-scale pricing projects for property and casualty business. He holds a MSc in theoretical physics from the University of Berne and a PhD in physical chemistry from the ETH Zurich.

Dr Michel M. Dacorogna , member of the senior management of SCOR, is currently heading its Group Financial Analysis and Risk Modeling team. His main responsibilities are: develop the Asset and Liability Management models for the SCOR group, assist customers in determining the best reinsurance structure taking into account the full portfolio of asset and liabilities. Besides his main tasks, he actively participates in customers' contacts and in the design of complicated contracts implying risk transfer to financial markets. He also conducts research in the field of insurance and reinsurance: capital allocation to risk, pricing and optimizing reinsurance covers, optimal asset allocation of investments, evaluation of credit risk, risk management of investments, forecasting models of long term economic trends. The author of more than 60 publications in refereed scientific journals, he often presents his results in international conferences and specialized seminars. He received his Habilitation, PhD and MSc in Theoretical Physics from the University of Geneva in Switzerland.

Dr Roger Iles is a risk consultant for SCOR reinsurance. After attaining his PhD in Astrophysics at the University College London, Roger continued his work in research for a number of years before moving to Switzerland. Roger gained valuable experience and knowledge in modelling weather based services while based at Syngenta headquarters in Basel. Roger is now applying his physics and modelling expertise and commercial experience to the reinsurance business, in particular in asset and liability modelling.