MODELING DEPENDENCE BETWEEN LOSS TRIANGLES WITH HIERARCHICAL ARCHIMEDEAN COPULAS

BY

ANAS ABDALLAH, JEAN-PHILIPPE BOUCHER AND HÉLÈNE COSSETTE

ABSTRACT

One of the most critical problems in property/casualty insurance is to determine an appropriate reserve for incurred but unpaid losses. These provisions generally comprise most of the liabilities of a non-life insurance company. The global provisions are often determined under an assumption of independence between the lines of business. Recently, Shi and Frees (2011) proposed to put dependence between lines of business with a copula that captures dependence between two cells of two different runoff triangles. In this paper, we propose to generalize this model in two steps. First, by using an idea proposed by Barnett and Zehnwirth (1998), we will suppose a dependence between all the observations that belong to the same calendar year (CY) for each line of business. Thereafter, we will then suppose another dependence structure that links the CYs of different lines of business. This model is done by using hierarchical Archimedean copulas. We show that the model provides more flexibility than existing models, and offers a better, more realistic and more intuitive interpretation of the dependence between the lines of business. For illustration, the model is applied to a dataset from a major US property-casualty insurer, where a bootstrap method is proposed to estimate the distribution of the reserve.

KEYWORDS

Runoff triangles, copula, hierarchical Archimedean copula, maximum likelihood estimation, bootstrap.

1. INTRODUCTION

Reserves are a major component of the financial statements of a financial institution. With the advent of the new regulatory standards (e.g. Solvency II in Europe and the upcoming ORSA[1] guidelines in North America), insurance companies must better understand and quantify the risks associated with their

Astin Bulletin 45(3), 577-599. doi: 10.1017/asb.2015.6

activities as a whole, not just by risk classes. Thus, it is now necessary for an insurance company to not only assess a reserve for each line of business but also to better estimate the total reserves for all its insurance products. This involves taking into account dependence between lines of business. In this context, insurance companies must be particularly able to estimate the amount of provisions for the entire portfolio. For this purpose, different reserving approaches allowing dependence between lines of business must be investigated. We will focus on the parametric approach.

Parametric reserving methods have often involved copulas to model the dependence between lines of business. For example, Brehm (2002) uses a Gaussian copula to model the joint distribution of unpaid losses, while De Jong (2012) models dependence between lines of business with a Gaussian copula correlation matrix. Shi *et al.* (2012) and Wüthrich *et al.* (2013) have also used multivariate Gaussian copula to accommodate the correlation due to accounting years within and across runoff triangles. Bootstrapping is another popular parametric approach used to forecast the predictive distribution of unpaid losses for correlated lines of business. Kirschner *et al.* (2008) use a synchronized bootstrap and Taylor and McGuire (2007) extend this result to a generalized linear model context.

In this paper, we propose to use a parametric approach with multivariate Archimedean copulas and hierarchical Archimedean copulas. In the same vein as Frees and Shi's model, and following an idea proposed by Barnett and Zehnwirth (1998), we propose a model that allows a dependence relation between all the observations that belong to the same CY for each line of business using multivariate Archimedean copulas. We use another dependence structure that links the losses of CYs of different lines of business. We show that this complex dependence structure can be constructed using hierarchial Archimedean copulas. For illustration, the model is applied to a dataset from a major US property-casualty insurer, where we conclude that the proposed model can be considered as an interesting alternative of the model proposed by Shi and Frees (2011).

In Section 2, we review the modeling of runoff triangles, where notations are set and copulas briefly introduced. In Section 3, the model of Shi and Frees (2011) is implemented (again on their dataset from a major US property-casualty insurer), but with a different parametrization. The CY and hierarchical dependences are explained and applied to this data in Section 4. In Section 5, we use a parametric bootstrap to obtain the predictive distribution of unpaid losses and we propose a new method to simulate hierarchical copulas. Section 6 concludes the paper.

2. Preliminary

2.1. Modeling and reserves

Let us consider an insurance portfolio with ℓ lines of business ($\ell \in \{1, \ldots, L\}$). We define by $X_{i,j}^{(\ell)}$, the incremental payments of the i^{th} accident year

($i \in \{1, \ldots, I\}$), and the jth development period ($j \in \{1, \ldots, J\}$). To take into account the volume of each line of business, we will work with standardized data which we denote by $Y_{i,j}^{(\ell)} = X_{ij}^{(\ell)}/\omega_i^{(\ell)}$, where $\omega_i^{(\ell)}$ represents the exposure variable in the ith accident year for the ℓth line of business. The exposure variable can be the number of policies, the number of open claims, or the earned premiums. The latter option is the one chosen in this paper.

A regression model with two independent explanatory variables, accident year and development period, will be used. Assume that $\alpha_i^{(\ell)}$ ($i \in \{1, 2, \ldots, I\}$) and $\beta_j^{(\ell)}$ ($j \in \{1, 2, \ldots, J\}$) characterize respectively the accident year effect and the development period effect. In such a context, a systematic component for the ℓth line of business can be written as:

$$\eta_{ij}^{(\ell)} = \zeta^{(\ell)} + \alpha_i^{(\ell)} + \beta_j^{(\ell)},\ \ell = 1, \ldots, L,$$

where $\zeta^{(\ell)}$ is the intercept, $I = J = n$, and for parameter identification, the constraint $\alpha_1^{(\ell)} = \beta_1^{(\ell)} = 0$ is supposed.

In our empirical illustration, we work with the runoff triangles of cumulative paid losses exhibited in Tables 1 and 2 of Shi and Frees (2011). They correspond to paid losses of Schedule P of the National Association of Insurance Commissioners (NAIC) database. These are data of 1997 for personal auto and commercial auto lines of business, and each triangle contains losses for accident years 1988–1997 and at most ten development years.

Shi and Frees (2011) show that a lognormal and a gamma distribution provide a good fit for the Personal Auto and the Commercial Auto line data respectively. To demonstrate the reasonable model fits for the two triangles, the authors exhibit the qq-plots of marginals for personal and commercial auto lines. We work with their conclusion and then continue with the same continuous distributions for each line of business. More specifically, we consider the form $\mu_{ij}^{(1)} = \eta_{ij}^{(1)}$ for a lognormal distribution with location (log-scale) parameter $\mu_{ij}^{(1)}$ and shape parameter σ. However, for the gamma distribution, we change the parametrization and we do not use the canonical inverse link $\mu_{ij}^{(2)} = \frac{1}{\eta_{ij}^{(2)}\phi}$ with location (scale) parameter $\mu_{ij}^{(2)}$ and shape parameter ϕ. Such a parametrization can lead to undesirable negative values for the lower right part of the runoff triangle, especially when one uses the bootstrap technique. To assure positive means of all the cells of the runoff triangle, we use the exponential link $\mu_{ij}^{(2)} = \frac{\exp(\eta_{ij}^{(2)})}{\phi}$, which is always positive, even for the prediction values of the runoff triangle.

With both parametrizations, the estimated total reserve is $\sum_{\ell=1}^{2} \sum_{i=2}^{n} \sum_{j=n-i+2}^{n} \widehat{y}_{ij}^{(\ell)} \omega_i^{(\ell)}$, where $\widehat{y}_{ij}^{(\ell)}$ is the projected unpaid loss ratio, and $\omega_i^{(\ell)}$ represents the net premiums earned in the corresponding accident year i. For the lognormal distribution, we have $\widehat{y}_{ij}^{(1)} = \exp^{\widehat{\mu}_{ij}^{(1)} + (\widehat{\gamma}^{(1)})^2/2}$, and for the gamma distribution, $\widehat{y}_{ij}^{(2)} = \widehat{\mu}_{ij}^{(2)}\widehat{\gamma}^{(2)}$, where $\widehat{\mu}_{ij}^{(\ell)}$ and $\widehat{\gamma}^{(\ell)}$ are respectively the scale (location) and the shape parameters. Also, $\widehat{\gamma}^{(1)} = \widehat{\sigma}$ and $\widehat{\gamma}^{(2)} = \widehat{\phi}$.

2.2. Copulas

Copulas are a useful and flexible tool to model a dependence relation between runoff triangles of different lines of business. They allow a separate interpretation of the relationship (linear and non-linear) between linked random variables and their marginals. See Joe (1997) for further details. We briefly recall below definitions and results that will be used later.

A multivariate copula $C(u_1, u_2, \ldots, u_n)$ is an application from $[0, 1]^n$ to $[0, 1]$, that has the same properties as a joint cumulative distribution. In other words, a copula is a function that links a multidimensional distribution to its one-dimensional margins. Let F be a n-dimensional cumulative joint function with margins $F^{(1)}, F^{(2)}, \ldots, F^{(n)}$. Then, if the margins are all continuous, the joint distribution of n random variables $(Y^{(1)}, Y^{(2)}, \ldots, Y^{(n)})$, can be represented by a unique copula function:

$$F(y^{(1)}, y^{(2)}, \ldots, y^{(n)}) = C(F^{(1)}, F^{(2)}, \ldots, F^{(n)}; \theta),$$

where $F^{(i)}$, with $i \in \{1, 2, \ldots, n\}$, are the respective distribution functions of $Y^{(i)}$, and θ is the dependence parameter, also called the association parameter.

In this paper, we choose to use the Archimedean family of copulas, given its several interesting properties. This family of copulas offers a wide choice of copulas for which many have a closed form expression in a multivariate setting. This last property will prove to be useful in what follows. Finally, Archimedean copulas can be constructed easily with a simple generator. Formally, we can define multivariate Archimedean copulas as

$$C(u_1, u_2, \ldots, u_n) = \phi^{-1}(\phi(u_1) + \cdots + \phi(u_n)), \tag{1}$$

where the function ϕ is called the generator of the copula. From (1), one can derive the expression for the multivariate density function of an Archimedean copula. According to McNeil and Nešlehová (2009), an Archimedean copula C admits a density c if and only if $\phi^{(n-1)}$ exists and is absolutely continuous on $(0, \infty)$. In such a case, c is given by

$$c(u_1, u_2, \ldots, u_n) = \phi^{(n)}(\phi^{-1}(u_1) + \cdots + \phi^{-1}(u_n)) \prod_{i=1}^{n} (\phi^{-1})'(u_i),$$

where functions $\phi^{(n)}$ and ϕ^{-1} correspond to the nth derivative of the generator function of the copula and the inverse generator respectively. Hofert *et al.* (2012) derive closed form expressions for the multivariate density function of a few Archimedean copulas, notably the Clayton, the Frank and the Gumbel copula used in this paper.

3. Pairwise dependence

Dividing a portfolio into homogeneous sub-portfolios and deriving the total reserve by summing the reserve for each segment implicitly assumes independence between risks. It is generally admitted that common social or economic factors may affect several lines of business simultaneously. Allowing a possible dependence relation between the runoff triangles of different lines of business of a portfolio provides a better representation of the portfolio's behavior as a whole and hence permits to take better advantage of diversification. It is also helpful to risk managers in determining the risk capital for an insurance portfolio.

Shi and Frees (2011) propose a model that incorporates a dependence structure between two runoff triangles in a pairwise manner. More precisely, the dependence between two lines of business is based on an identical association between cells of a given accident year and development period, coming from different lines of business. This means that two paid loss ratios $Y_{i,j}^{(1)}$ and $Y_{i,j}^{(2)}$ are correlated for a given couple (i, j). This form of dependence goes back to Braun (2004). Throughout the paper, we refer to Frees and Shi's model as the pairwise dependence model (PWD).

3.1. Modeling

The PWD model associates two elements of the same accident year and development period, $(Y_{i,j}^{(1)}, Y_{i,j}^{(2)})$ with a bivariate copula. Mathematically, and following Sklar's theorem, the joint distribution of normalized incremental payments $(Y_{i,j}^{(1)}, Y_{i,j}^{(2)})$ will be represented by the unique copula function:

$$F_{ij}(y_{ij}^{(1)}, y_{ij}^{(2)}) = \Pr(Y_{ij}^{(1)} \leq y_{ij}^{(1)}, Y_{ij}^{(2)} \leq y_{ij}^{(2)}) = C(F_{ij}^{(1)}, F_{ij}^{(2)}; \theta), \qquad (2)$$

where $C(., \theta)$ denotes the copula function with parameter θ, that captures the dependence between two runoff triangles. Also, this model has the flexibility of choosing a different cumulative density function for each line of business. The log-likelihood expression can be easily derived from equation (2):

$$L = \sum_{i=1}^{I} \sum_{j=1}^{I-i+1} \log(f_{ij}^{(1)}) + \log(f_{ij}^{(2)}) + \sum_{i=1}^{I} \sum_{j=1}^{I-i+1} \log \ c(F_{ij}^{(1)}, F_{ij}^{(2)}; \theta), \qquad (3)$$

where $c(.)$ denotes the probability density function corresponding to the copula distribution function $C(.)$, $f_{ij}^{(\ell)}$ denotes the density of marginal distribution $F_{ij}^{(\ell)}$, for $\ell = 1, 2$. These marginals are noted as:

$$F_{ij}^{(\ell)} = \text{Prob}(Y_{ij}^{(\ell)} \leq y_{ij}^{(\ell)}) = F^{(\ell)}(y_{ij}^{(\ell)}; \eta_{ij}^{(\ell)}, \gamma^{(\ell)}),$$

for $i = 1, \ldots, I$, $j = 1, \ldots J$ and $\ell = 1, \ldots, L$. Shi and Frees (2011) choose the Gaussian and the Frank copula to model dependence, as well as the product

https://doi.org/10.1017/asb.2015.6 Published online by Cambridge University Press

TABLE 1

FIT STATISTICS OF PWD MODEL.

Fit Statistics	Independence	Copula Frank		Copula Gaussian	
Dependence Parameter	.	−2.7978	(1.0243)	−0.3655	(0.1190)
Log-Likelihood	346.6	350.3		350.5	
AIC	−613.2	−618.5		−618.9	
BIC	−505.2	−507.9		−508.3	

TABLE 2

RESERVES ESTIMATION WITH THE PWD MODEL.

Reserves Estimation	Copula: Independence	Copula: Frank	Copula: Gaussian
Personal	6,464,090	6,511,363	6,423,180
Commercial	490,657	487,904	495,989
Total	6,954,747	6,999,267	6,919,169

copula that supposes independence between the cells. Their model selection is based on a likelihood-based goodness-of-fit measure, namely Akaike's Information Criterion (AIC).

3.2. Empirical illustration

We provide in Tables 1 and 2, the fit statistics and the reserves for the PWD model. Note that even if the results are close to those obtained in Shi and Frees (2011), we do not obtain the same estimates because we have changed the link function of the mean of the gamma distribution to avoid inconsistencies, as explained in Section 2.2.

On the other hand, even if we have chosen a different parametrization, we obtain the same conclusion as their and find that the copula that leads to the smallest AIC is the Gaussian copula. This model generates a reserve of almost 7 million dollars. Interestingly, the dependence parameter obtained for the pairwise model with the Gaussian and the Frank copula is negative, meaning that the model supposes that the two lines of business are negatively correlated.

4. CALENDAR YEAR AND HIERARCHICAL DEPENDENCE

We propose here to further investigate the model of Shi and Frees (2011) to better capture the interactions within and between the runoff triangles of different lines of business. For that purpose, we first propose to consider a dependence construction for the different elements of a diagonal of a given runoff triangle to take into account a CY effect. Second, we add another level of dependence to capture the dependence between the lines of business.

4.1. Calendar year effect

We propose in this section a model that allows a dependence relation within paid claims belonging to a diagonal of a runoff triangle. This reflects a CY effect, more precisely the changes or inflections on paid claims in a CY due to jurisprudence modifications or inflationary trends for example. A CY effect can also highlight the impact of strategic decisions made in a CY such as an incentive to increase payments in a particular CY for all lines of business.

This dependence structure assumes that all cells from the same diagonal are correlated, which implies that the number of cells in the dependence structure is different for each diagonal. Indeed, the number of cells in the dependence structure varies from 1 to t for the tth diagonal, with $t \in \{1, \dots, n\}$, and $t = i + j - 1$. Evidently, the first cell at the top left of the runoff triangle is not linked to any other cell within the triangle.

Such a CY effect has already been analyzed before, for example by Barnett and Zehnwirth (1998) who added a covariate to capture the CY effect. The systematic component of such a model can be written as:

$$\eta_{ij}^{(\ell)} = \zeta^{(\ell)} + \alpha_i^{(\ell)} + \beta_j^{(\ell)} + \Upsilon_t^{(\ell)}, \quad \ell = 1, \dots, L, \tag{4}$$

where $\zeta^{(\ell)}$ is the intercept, $\alpha_i^{(\ell)}$ ($i \in \{1, 2, \dots, I\}$) and $\beta_j^{(\ell)}$ ($j \in \{1, 2, \dots, J\}$) characterize respectively the accident year effect and the development period effect, while $\Upsilon_t^{(\ell)}$ ($t = i + j - 1$) captures the CY effect.

De Jong (2006) modeled the growth rates in cumulative payments in a CY, and Wüthrich (2010) examined the accounting year effect for a single line of business. Wüthrich and Salzmann (2012) used a multivariate Bayes chain-ladder model that allows the modeling of dependence along accounting years within runoff triangles. The authors showed that they are able to derive closed form solutions for the posterior distribution, the claims reserves and the corresponding prediction uncertainty. Kuang *et al.* (2008) have also considered a canonical parametrization with three factors for a single line of business. Each factor represents time scale, in such way the inflation is taken into account. Also, they added an assumption ensuring that the forecasts do not depend on these arbitrary linear trends. They extended this assumption later by combining the canonical parametrization with a non-stationary time series forecasting model in Kuang *et al.* (2011).

In our proposed model, instead of adding an explanatory variable for the CY effect, the dependence relation between the paid claims of a diagonal will be based on a multivariate Archimedean copula. More specifically, the same Archimedean copula with an identical dependence parameter is assumed for each diagonal of a runoff triangle. Hence, all random variables of the same CY

$t = i + j - 1$ and ℓth line of business are included in the vector $\boldsymbol{Y}_{\ell t} = \{Y_{\ell ij} : i + j - 1 = t\}$. The log-likelihood function of this model can be written as:

$$L = \sum_{i=1}^{I} \sum_{j=1}^{I-i+1} \log(f_{ij}) + \sum_{t=2}^{n} \log\ c\left(F_{t-j+1,j}, \ldots, F_{1,t}; \theta\right)_{j=1,\ldots,t}, \tag{5}$$

where f denotes the density of marginal distribution F, and $c(.)$ the probability density function corresponding to the copula distribution function $C(.)$.

The main advantage of the copula approach instead of adding a CY covariate in the mean specification, lies in the fact that the copula approach allows a more general structure of dependence between the observations of a given CY and allows more flexibility. Also, the use of covariates would lead to a great number of parameters to explain the CY effect instead of only one (dependence copula parameter). For example, for two lines of business, we would have 20 parameters instead of 2 (see Equation (4)). This might lead to over-parametrization. Furthermore, the parameter describing a given CY effect, would not have any predictive power, as we cannot use it to compute the lower triangle.

4.2. Line of business dependence

A natural extension to the model behind (5) is to introduce a dependence structure between lines of business based on copulas, more precisely here with the Gaussian copula and hierarchical Archimedean copulas.

Another way to add dependence between lines of business is by modifying Equation (4) and use the same CY covariate for the two lines of business, i.e. $\Upsilon_t = \Upsilon_t^{(1)} = \Upsilon_t^{(2)}$. The correlation induced by common CY effects would then be introduced through the mean specification. Also, as done in Shi *et al.* (2012), in addition to the common CY covariate, a pairwise correlation between the two runoff triangles can be added. This approach has the disadvantage however of adding a new parameter for each diagonal (Υ_t).

4.2.1. *Multivariate Gaussian copula.* We first propose to use the Gaussian copula to capture the dependence within and between runoff triangles. The Gaussian copula which arises from the multivariate normal distribution is the most widely known copula of the elliptical family of copulas. Such a copula allows great flexibility to model dependences simply by modifying its correlation matrix.

Let us suppose, for a given CY t, the following set of observations $\mathbf{u}_t = (u_{t-j+1,j}^{(1)}, \ldots, u_{1,t}^{(1)}, u_{t-j+1,j}^{(2)}, \ldots, u_{1,t}^{(2)})_{j=1,\ldots,t}$, with multivariate Gaussian copula density:

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

https://doi.org/10.1017/asb.2015.6 Published online by Cambridge University Press

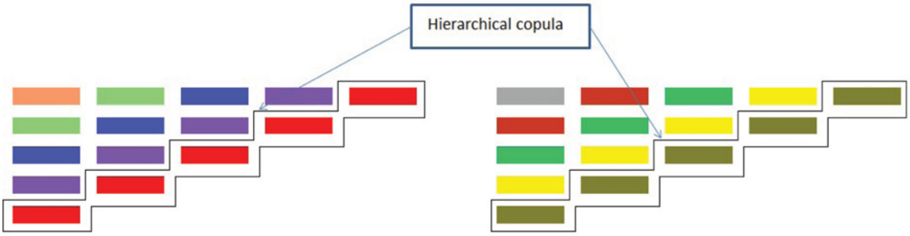

FIGURE 1: Dependence implied by hierarchical dependence. (Color online)

where $\boldsymbol{\xi}_t = (\Phi^{-1}(u^{(1)}_{t-j+1,j}), \ldots, \Phi^{-1}(u^{(1)}_{1,t}), \Phi^{-1}(u^{(2)}_{t-j+1,j}), \ldots, \Phi^{-1}(u^{(2)}_{1,t}))^T_{j=1,\ldots,t}$. The correlation matrix $\boldsymbol{\Sigma}_t$ for the CY t can be represented as a block matrix as follows, given the assumptions of the model:

$$\boldsymbol{\Sigma}_t = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{21} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{12} \end{pmatrix}. \quad (6)$$

In (6), the matrices $\boldsymbol{\Sigma}_{11}$ and $\boldsymbol{\Sigma}_{12}$ are correlation matrices with unit main diagonal and off-diagonal parameters $\theta_{1,1}$ and $\theta_{1,2}$ corresponding to the CY dependence for the first and second line of business respectively. $\boldsymbol{\Sigma}_{21}$ is a matrix filled with parameter $\theta_{2,1}$ representing the dependence between the two lines of business.

Numerical results obtained with the Gaussian copula are presented in the empirical illustration of Section 4.3.

4.2.2. *Hierarchical Archimedean copulas.* Hierarchical Archimedean copulas permit to have different levels of dependence within our framework. We use them here to add another level of dependence to the one proposed in Section 4.1. With this second level of dependence, we capture the dependence between two different runoff triangles in a pairwise manner between corresponding diagonals, instead of between cells. Pairing diagonals instead of cells with a copula has the advantage of being applicable even in a case of missing data in one of the runoff triangles.

The hierarchical approach allows us to visualize the multi-level dependence. Indeed, this dependence structure is illustrated in Figure 1, where a dependence structure between cells of the same CY is supposed as well as a dependence structure between the two lines of business. In the next section, we will also be interested in comparing the hierarchical copula approach with the multivariate Gaussian copula approach, as the latter is often considered as a benchmark model.

The CY effect has not been often studied with more than one line of business. Two recent examples are De Jong (2012), where the CY effect was introduced through the correlation matrix and Shi *et al.* (2012), who used random effects to accommodate the correlation due to accounting year effects within and across runoff triangles. In Shi *et al.* (2012), they work with a Bayesian perspective, using a multivariate lognormal distribution, along with a multivariate Gaussian correlation matrix. The predictive distributions of outstanding payments are generated through Monte Carlo simulations. The CY effect is taken into account through an explanatory variable. A discussion of this paper is suggested in Wüthrich (2012), and where it is also explained that for the method it does not really matter whether we consider incremental or cumulative claims, as long as we have a multivariate Gaussian structure. Also, still with a Bayesian framework, Wüthrich *et al.* (2013) used a multivariate lognormal chain-ladder model and derived predictors and confidence bounds in closed form. Their analytical solutions are such that they allow for any correlation structure. Their models allow a dependence between and within runoff triangles, and for any correlation structure. It has also been shown in this paper that the pairwise dependence form is a rather weak one compared to CY dependence. More recently, Shi (2014) captures the dependencies introduced by various sources, including the common CY effects via the family of elliptical copulas, and use a parametric bootstrapping to quantify the associated reserving variability.

In this paper, to model the complex dependence structure between two runoff triangles, we introduce models based on hierarchical Archimedean copulas. The idea is to use Archimedean copulas at each level, from the lowest (CYs) to the highest (lines of business). Hierarchical Archimedean copulas have first been mentioned in the literature by Joe (1997), and appeared in more details in Savu and Trede (2010). More recently, Okhrin *et al.* (2013) provided a method to estimate multivariate distributions defined through hierarchical Archimedean copulas.

The main advantage of using Archimedean and hierarchical Archimedean copulas is that they can be explicitly defined in terms of a one-dimensional function called the generator of the Archimedean copula. Elliptical copulas, used in Shi (2014), do not possess this nice property; they do not have a closed form. Archimedean copulas are also flexible and allow to model many kinds of dependencies, while Elliptical copulas, have equal lower and upper tail dependence coefficients. In high dimensions, Archimedean copulas are restricted given the exchangeability of the components. This assumption is relaxed with hierarchical Archimedean copulas.

At the lowest level, and for the CY t, we have $2 \times t$ standard uniformly distributed random variables $U^{(1)}_{t-j+1,j}, \ldots, U^{(1)}_{1,t}, U^{(2)}_{t-j+1,j}, \ldots, U^{(2)}_{1,t}$ where j designates the development period ($j = 1, \ldots, t$).

The joint distribution function is evaluated at $\mathbf{u} = (u^{(1)}_{t-j+1,j}, \ldots, u^{(1)}_{1,t}, u^{(2)}_{t-j+1,j}, \ldots, u^{(2)}_{1,t}) \in [0, 1]^{2t}$. Let there be H hierarchy levels indexed by h. For example, the set of elements $\mathbf{u}$ is located at level $h = 0$. At each level $h = 0, \ldots, H$ we have n_h distinct objects with index $k = 1, \ldots, n_h$.

At level $h = 1$, the $u^{(1)}_{t-j+1,j}, \ldots, u^{(1)}_{1,t}, u^{(2)}_{t-j+1,j}, \ldots, u^{(2)}_{1,t}$ are grouped into n_1 ordinary multivariate Archimedean copulas $C_{1,k}, k = 1, \ldots, n_1$ (in our case with two lines of business, we have $n_1 = 2$), of the form

$$C_{1,k}(\mathbf{u}_{1,k}) = \phi^{-1}_{1,k}\left(\sum_{\mathbf{u}_{1,k}} \phi_{1,k}(\mathbf{u}_{1,k})\right),$$

where $\phi_{1,k}$ denotes the generator of the copula $C_{1,k}$. Let $\mathbf{u}_{1,k}$ denote the set of elements of $u^{(k)}_{t-j+1,j}, \ldots, u^{(k)}_{1,t}$ belonging to the copula $C_{1,k}$ for $k = 1, \ldots, n_1$, which represents the elements of a given CY for a single line of business ℓ. At this level only, k corresponds to ℓ. In our model, we have three levels, i.e. $H = 2$. At the highest level, we have a single object ($n_2 = 1$), which is the hierarchical Archimedean copula $C_{2,1}$, that aggregates the multivariate Archimedean copulas of the previous level, and can be represented as

$$C_{2,k}(\mathbf{C}_{2,k}) = \phi^{-1}_{2,k}\left(\sum_{\mathbf{C}_{2,k}} \phi_{2,k}(\mathbf{C}_{2,k})\right),$$

where $\phi_{2,k}$ denotes the generator of the copula $C_{2,k}$ and $\mathbf{C}_{2,k}$ represents the set of all copulas from level $h = 1$ entering copula $C_{2,k}$ for $k = 1, \ldots, n_2$.

Obviously, there are numerous conditions to be satisfied for the existence of a hierarchical Archimedean copula. The number of copulas must decrease at each level, i.e. $n_h < n_{h-1}$, as well as the degree of dependence, i.e. $\theta_{h+1,k'} < \theta_{h,k}$ for all $h = 0, \ldots, H$ and $k = 1, \ldots, n_h, k' = 1, \ldots, n_{h+1}$ such that $C_{h,k} \in C_{h+1,k'}$ where $\theta_{h,k}$ is the parameter belonging to the generator $\phi_{h,k}$. This means that for runoff triangles, elements of a same line of business can have a higher degree of dependence than elements of different lines of business. Mathematically, the conditions that have to be verified by a hierarchical Archimedean copula are summarized as follows:

1. All inverse generator functions $\phi^{-1}_{h,k}$ are completely monotone.
2. The composite $\phi_{h+1,k'} \circ \phi^{-1}_{h,k}$ are convex functions for all $h = 0, \ldots, H$ and $k = 1, \ldots, n_h, k' = 1, \ldots, n_{h+1}$ such that $C_{h,k} \in C_{h+1,k'}$.

In our application, we will limit the number of levels to three, and the number of lines of business to two. This means that we will have at the highest level ($h = 2$), one (hierarchical) bivariate Archimedean copula between lines of business, and for $h = 1$, two (ordinary) multivariate Archimedean copula within a runoff triangle.

As an illustration, let us consider a dependence structure between two runoff triangles for the second CY. The resulting hierarchical Archimedean copula has the following analytical form

$$\begin{aligned}C_{2,1} &= C_{2,1}(u^{(1)}_{2,1}, u^{(1)}_{1,2}, u^{(2)}_{2,1}, u^{(2)}_{1,2})\\ &= C_{2,1}(C_{1,1}(u^{(1)}_{2,1}, u^{(1)}_{1,2}), C_{1,2}(u^{(2)}_{2,1}, u^{(2)}_{1,2}))\\ &= \phi^{-1}_{2,1}\Big(\phi_{2,1}\circ\phi^{-1}_{1,1}[\phi_{1,1}(u^{(1)}_{2,1}) + \phi_{1,1}(u^{(1)}_{1,2})]\\ &\quad + \phi_{2,1}\circ\phi^{-1}_{1,2}[\phi_{1,2}(u^{(2)}_{2,1}) + \phi_{1,2}(u^{(2)}_{1,2})]\Big).\end{aligned}$$

This hierarchical Archimedean copula will be applied to each CY, with the dataset described in Section 3.2. The CY t takes values from 1 to 10 because the runoff triangles both have 10 diagonals, i.e. $I = J = 10$. The resulting hierarchical Archimedean copula for our model has the following general analytical form:

$$\begin{aligned}C_{2,1} &= C_{2,1}(u^{(1)}_{t-j+1,j}, \ldots, u^{(1)}_{1,t}, u^{(2)}_{t-j+1,j}, \ldots, u^{(2)}_{1,t})\\ &= C_{2,1}(C_{1,1}(u^{(1)}_{t-j+1,j}, \ldots, u^{(1)}_{1,t}), C_{1,2}(u^{(2)}_{t-j+1,j}, \ldots, u^{(2)}_{1,t}))\\ &= \phi^{-1}_{2,1}\Big(\phi_{2,1}\circ\phi^{-1}_{1,1}[\phi_{1,1}(u^{(1)}_{t-j+1,j}) + \cdots + \phi_{1,1}(u^{(1)}_{1,t})]\\ &\quad + \phi_{2,1}\circ\phi^{-1}_{1,2}[\phi_{1,2}(u^{(2)}_{t-j+1,j}) + \cdots + \phi_{1,2}(u^{(2)}_{1,t})]\Big).\end{aligned}$$

Finally, the log-likelihood function of the hierarchical model can be written as follows:

$$\begin{aligned}L = &\sum_{\ell=1}^{2}\sum_{i=1}^{I}\sum_{j=1}^{I-i+1}\ln(f^{(\ell)}_{ij}) + \sum_{t=2}^{n}\{\ln c_{1,1}\left(F^{(1)}_{t-j+1,j}, \ldots, F^{(1)}_{1,t}; \theta_{1,1}\right)_{j=1,\ldots,t}\\ &+ \ln c_{1,2}\left(F^{(2)}_{t-j+1,j}, \ldots, F^{(2)}_{1,t}; \theta_{1,2}\right)_{j=1,\ldots,t}\\ &+ \ln c_{2,1}\Big(C_{1,1}(F^{(1)}_{t-j+1,j}, \ldots, F^{(1)}_{1,t}; \theta_{1,1}),\\ &\quad C_{1,2}(F^{(2)}_{t-j+1,j}, \ldots, F^{(2)}_{1,t}; \theta_{1,2}); \theta_{2,1}\Big)_{j=1,\ldots,t}\}.\end{aligned} \tag{7}$$

The simpler form of hierarchical dependence is to suppose a product copula between the two runoff triangles, meaning independence between lines of business. In this situation, the log-likelihood of the model is simply $L = L^{(1)} + L^{(2)}$, where $L^{(\ell)}$, $\ell = 1, 2$ is simply the log-likelihood obtained by (5). Of course, it is very easy to extend this model to more than two lines of business.

TABLE 3

FIT STATISTICS OF INDEPENDENT CALENDAR YEAR DEPENDENCE MODEL.

	Copula-Estimates and Standard Errors							
	Gaussian		Frank		Clayton		Gumbel	
$\theta_{1,1}$	0.6091	(0.1366)	8.1249	(1.4959)	2.2680	(0.4463)	2.7267	(0.6360)
$\theta_{1,2}$	0.7634	(0.0983)	14.0604	(1.8180)	2.9748	(0.5743)	4.6339	(1.0385)
Log-Lik.	391.5		404.7		403.9		400.4	
AIC	−699.0		−725.4		−723.9		−716.8	
BIC	−585.6		−612.0		−610.4		−603.4	

TABLE 4

RESERVES ESTIMATION OF INDEPENDENT CALENDAR YEAR DEPENDENCE MODEL.

	Copula			
Reserves Estimation	Gaussian	Frank	Clayton	Gumbel
Personal	6,175,574	6,015,229	6,158,971	6,676,692
Commercial	751,725	901,641	662,929	769,428
Total	6,927,299	6,916,870	6,821,900	7,446,120

4.3. Empirical illustration

Hierarchical models based on different copulas have been applied to the runoff triangles used in Section 3.2. For this model, the CY dependence has been modeled with five different copulas (product, Frank, Gumbel, Clayton and Gaussian). In our empirical study, we first use a model that supposes independence between lines of business, i.e. a product copula between runoff triangles. We call this model ICYD, for independence CY dependence. Fit statistics as long as dependence parameters of this model are shown in Table 3, while the estimated reserves are presented in Table 4. In terms of AIC, we observe that all Archimedean copula models offer a better fit than the multivariate Gaussian copula. Note that a CY dependence with a product copula within and between the two lines of business is simply a cell-by-cell modeling. The empirical results of this simple model have already been given in Section 3.2, for the PWD model with a product copula.

Three copulas (Frank, Clayton and Gumbel) have been considered in a hierarchical model to investigate dependence between the two lines of business. The same copula is used for each level, meaning for example that if a Frank copula is chosen within a runoff triangle, then it is also used between the business lines. This is due to the convexity condition on hierarchical Archimedean copulas. We call this model HCYD, for hierarchical CY dependence. The fit statistics and the reserves obtained for this model are shown in Tables 5 and 6 respectively. To compare the degree of dependence between different copulas, we also provide

TABLE 5

PARAMETER ESTIMATION OF THE CALENDAR YEAR DEPENDENCE MODEL BETWEEN LOB'S.

	Copula-Estimates and Standard errors							
	Gaussian		Frank		Clayton		Gumbel	
$\theta_{1,1}$	0.5989	(0.1396)	8.6648	(1.2668)	1.9343	(0.8738)	3.1292	(0.2962)
$\theta_{1,2}$	0.7586	(0.0997)	16.3327	(2.1025)	2.3357	(0.7781)	3.4505	(0.5443)
$\theta_{2,1}$	−0.1727	(0.2384)	6.0999	(3.0045)	0.2217	(0.2981)	1.8052	(0.5253)
ρ_S	−0.1651		0.7166		0.1491		0.6186	
τ_K	−0.1105		0.5193		0.0997		0.4460	
LogLik.	391.7		406.2		404.2		404.9	
AIC	−697.4		−726.4		−722.4		−723.7	
BIC	−581.3		−610.3		−606.3		−607.7	

TABLE 6

RESERVES ESTIMATION OF THE CALENDAR YEAR DEPENDENCE MODEL BETWEEN LOB'S.

	Copula			
Reserves Estimation	Gaussian	Frank	Clayton	Gumbel
Personal	6,103,937	6,253,043	6,263,024	7,144,724
Commercial	740,747	1,095,438	720,353	1,222,431
Total	6,844,684	7,348,481	6,983,377	8,367,155

the two non-linear association measures Spearman's rho ρ_S and Kandall's tau τ_K for the four copulas, see Table 5. We notice that the Clayton copula captures a smaller dependence than the Frank and Gumbel copulas, whose association measures are higher. Indeed, the Clayton family is characterized by a lower tail dependence. Also, once again, the hierarchical Archimedean copula models offer a better fit than the multivariate Gaussian copula as shown by the values of the log-likelihood function. The Frank copula offers the best fit between all HCYD models. Additionally, by looking at the values of the AIC, the hierarchical model with a Frank copula provides the best fit between all models studied in this paper. As previously mentioned, the choice of the PWD model was based on the AIC criterion.

4.3.1. *Analysis of dependence.* It is interesting to note that, unlike the pairwise model of Shi and Frees (2011) and the multivariate Gaussian copula model which generate negative dependence, hierarchical models generate positive dependence between loss triangles with our dataset. For the multivariate Gaussian copula, we can observe that the parameter $\theta_{2,1}$ is not statistically significant (estimate of −0.1727 with standard error of 0.2384), which means that this negative association might therefore be misleading. The hierarchical bivariate copula $C_{2,1}$ is not restricted, and allows for positive and negative dependence.

We observe that the parameter $\theta_{2,1}$ is not statistically significant for the Clayton copula model (estimate of 0.2217 with standard error of 0.2981). However, it is significantly greater than 0 for the Gumbel and the Frank copulas, which favors models with positive dependence, as opposed to the conclusion of Shi and Frees (2011). This analysis highlights the fact that the choice of the models can lead to different conclusions for the dependence structure. This was also well illustrated in Figure 4 of Shi *et al.* (2012).

When we incorporate a CY correlation within the lines of business (level 1), the residual dependence becomes positive. Intuitively, this can be explained by the trends and common effects that are detected with the introduction of the proposed dependence structure but not with the chain-ladder coefficients. In a given CY, exogenous common factors such as inflation, interest rates, jurisprudence or strategic decisions such as the acceleration of the payments for the entire portfolio can have simultaneous impacts on all lines of business of a given sector, such as the two lines of business considered in the present paper. These effects may as well result in trends in the development period parameters. It is important to note that these trends can detect both positive and negative associations.

Finally, a hierarchical copula model requires a higher degree of dependence for variates linked at a lower level than those linked at a higher level. In our context, this means that the degree of dependence within lines of business should be greater than between lines of business, as illustrated in Figure 1. One can observe in Table 5 that this condition is respected with a dependence parameter $\theta_{2,1}$ lower than the dependence parameters $\theta_{1,1}$ and $\theta_{1,2}$.

4.3.2. *Fitted values.* It can be interesting to analyse how the new models proposed in this paper adjust the data. We examined fitted values $\widehat{y}_{i,j}$, for $i + j \leq n$, with the HCYD model using a Frank copula. The fitted values are compared to observed loss ratios in Figure 2. We observe that despite the fact the HCYD model offers the best fit, it generates outliers. We analyzed in details this situation and observed that these outliers come from the last CY of the commercial line. For this CY, smaller incremental amounts are observed, compared to the previous CYs. We observed that the individual development factors of the last CY are all smaller than the median development factor for each development period. This could be due to a policy reform, where it was decided to accelerate payments, causing a decrease in the payments for the last diagonal. HCYD model allowed us to clearly identify these outliers. For this dataset used as an illustration, we think that other methods can be used to evaluate the impact of the possible administrative change. For example, new methods attributing weights to outliers could be investigated. But even in this situation, a generalization of the ICYD and the HCYD should be used.

When we look at the difference in the reserves between the models, this can be explained in part by analyzing the values of the estimated parameters. One of the most sensitive parameters is $\alpha_{10}^{(\ell)}$, which comes from the last accident year.

https://doi.org/10.1017/asb.2015.6 Published online by Cambridge University Press

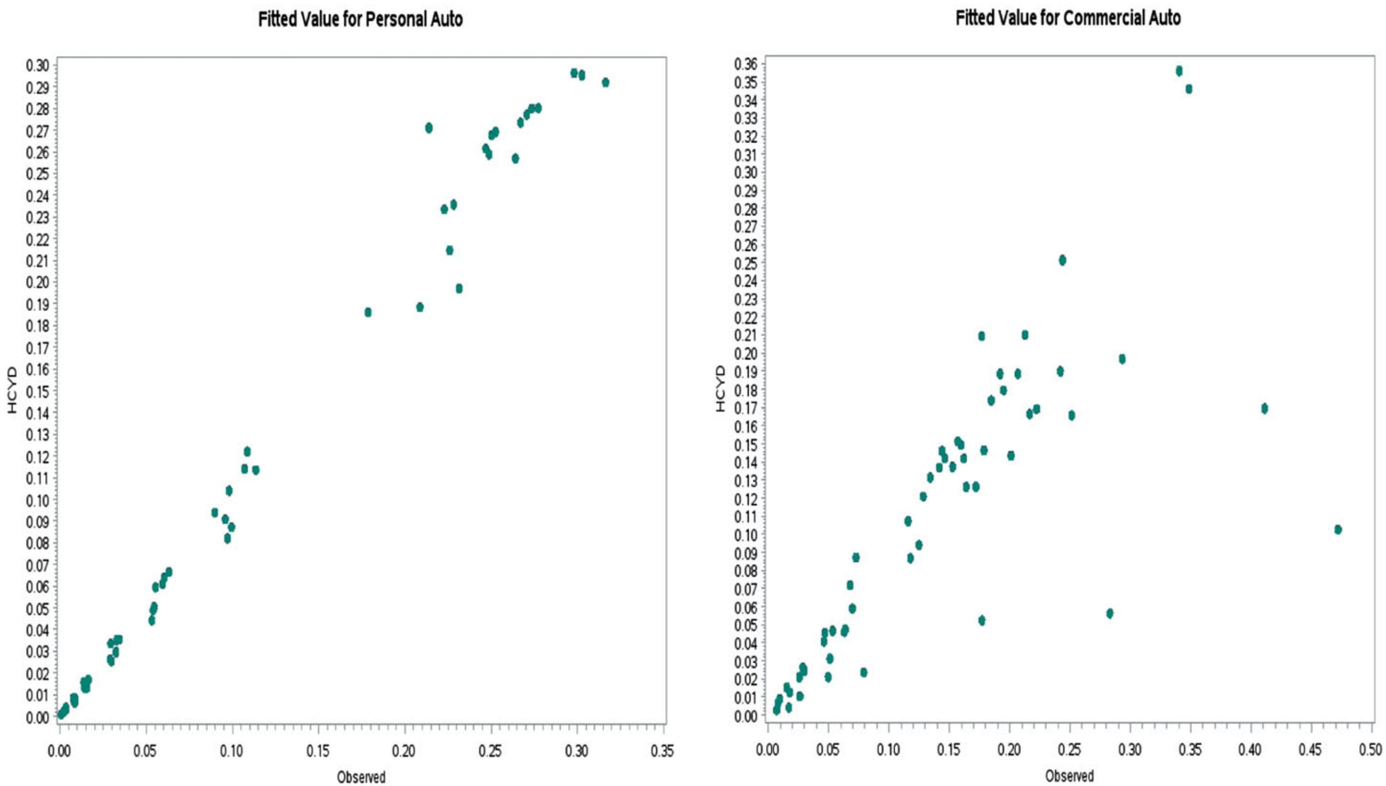

FIGURE 2: Scatter plots of fitted value between the observed data and HCYD model for the personal auto and commercial auto lines. (Color online)

$\beta_{10}^{(\ell)}$, from the most recent development period, significantly drives the reserves. Given that the idea behind a model with two explanatory variables, accident year and development period, is the chain-ladder model, one can compare or relate the sensitivity of the parameters $\alpha_{10}^{(\ell)}$ and $\beta_{10}^{(\ell)}$ to the sensitivity of the top right and bottom left parts of the triangle with the chain-ladder model.

The reserve's difference between models can also be explained by the disproportional size of the two sub-portfolios. The dominating size of the personal auto line in the insurance portfolio can drive the results, as it was stated in one of the conclusions of Wüthrich (2013): *Strong portfolio growth or decrease needs a careful analysis because change of volume often conceals other effects*. This conclusion would be even stronger for a risk capital analysis. This difference of reserves has also been noticed previously in Table 9 of Shi and Frees (2011) when they compare the prediction of unpaid losses of the insurance portfolio from the copula model with various existing approaches.

Other interesting avenues are also raised in Meyers (2013), where it is indicated that such situations could be explained by the fact that the insurance loss environment is too dynamic to be captured in a single stochastic loss reserve model. This may also be due to the data used to calibrate the model in which can be missing crucial information. Examples of such changes could include changes in the way the underlying business is conducted, such as changes in claim processes or changes in the direct/ceded/assumed reinsurance composition of the claim values in triangles.

5. Predictive distribution

In practice, actuaries are interested in knowing the uncertainty of the reserve. A parametric technique, the bootstrap, not only provides such information but most importantly lets one determine the entire predictive distribution, rarely obtained for non-Bayesian models. The predictive distribution notably allows assessment of risk capital for an insurance portfolio. Bootstrapping is also ideal from a practical point of view, because it avoids the complex theoretical calculations and can easily be implemented. Moreover, it tackles the potential model overfitting, typically encountered in loss reserving problems, due to the small sample size.

The bootstrap technique is increasingly popular in loss reserving, and allows a wide range of applications. It was first introduced in a loss reserving context with a distribution-free approach by Lowe (1994). For a multivariate loss reserving analysis, Kirschner *et al.* (2008) used a synchronized parametric bootstrap to model dependence between correlated lines of business, and Taylor and McGuire (2007) extended this result to a generalized linear model context. Shi and Frees (2011) and more recently Shi (2014) have also performed a parametric bootstrap to incorporate the uncertainty in parameter estimates, while modeling dependence between loss triangles using copulas.

5.1. Parametric bootstrap

The parametric bootstrap allows us to obtain the whole distribution of the reserves. We follow the same bootstrap algorithm of Taylor and McGuire (2007), and summarized in Shi and Frees (2011).

5.1.1. *Copula simulation.* The first step of the parametric bootstrap is to generate pseudo-responses of normalized incremental paid losses $y^{*(\ell)}_{ij,r}$, for i, j such that $i + j \leq I$ and $\ell = 1, 2$. We know that $y^{*(\ell)}_{ij,r} = F^{(-1)(\ell)}(u^{(\ell)}_{ij}, \widehat{\mu}^{(\ell)}_{ij}, \widehat{\gamma}^{(\ell)})$, with $\widehat{\mu}^{(\ell)}_{ij}$ and $\widehat{\gamma}^{(\ell)}$ already estimated. Therefore, a technique to generate the realizations of the copula $u^{(\ell)}_{ij}$, with $\ell = 1, 2$ should be used.

Given that the Frank copula generates the best fit for many models in this paper, we have decided to focus on this copula for the bootstrap. Below we study the Frank PWD model, the Frank independent model (a Frank copula is chosen for each CY dependence, and a product-copula between lines of business), and the Frank hierarchical model.

To generate a multivariate Frank copula, we follow the method based on the inversion of the Laplace transform, an idea that can be traced back to Marshall and Olkin (1988).

The above cited algorithms allow us to generate the set of realizations $\mathbf{u}^{(1)}_{1,1}$ and $\mathbf{u}^{(2)}_{1,2}$ for the first level of hierarchy (CY level at $h = 1$) from the ordinary multivariate Archimedean copulas $C_{1,1}$ and $C_{1,2}$, for a given CY t and development period j ($j = 1, \ldots, t$), with $\mathbf{u}^{(1)}_{1,1} = (u^{(1)}_{t-j+1,j}, \ldots, u^{(1)}_{1,t})$ and $\mathbf{u}^{(2)}_{1,2} = (u^{(2)}_{t-j+1,j}, \ldots, u^{(2)}_{1,t})$. To generate realizations with a Frank copula at the highest level of the hierarchy (line of business level at $h = 2$), we propose to use a new algorithm. We will start from the last step where uniform realizations of a Frank copula are generated for a given CY (CY level), and transform them into two sets of standard exponential variables, with $\mathbf{Z}^{(\ell)}_{2,1} = -\ln(\mathbf{U}^{(\ell)}_{1,\ell})$, for $\ell = 1, 2$. Thus, the algorithm for a hierarchical model with a Frank copula between the lines of business is as follows:

1. Generate realizations of random variables $\mathbf{Z}^{(\ell)}_{2,1} = -\ln(\mathbf{U}^{(\ell)}_{1,\ell})$, with $\mathbf{Z}_{2,\ell} \sim$ Exp(1), for $\ell = 1, 2$.
2. Generate a realization $y_{2,1}$ of a logarithmic random variable $Y_{2,1}$, with parameter $\eta_{2,1} = 1 - \exp(-\theta_{2,1})$.
3. Calculate the set $\mathbf{U}^{(\ell)}_{2,1} = \frac{\ln(1-\eta_{2,1}\exp(-\frac{\mathbf{Z}^{(\ell)}_{2,1}}{Y}))}{\ln(1-\eta_{2,1})}$, for $\ell = 1, 2$.

Consequently, we have obtained the set of realizations $\mathbf{u}^{(1)}_{2,1}$ and $\mathbf{u}^{(2)}_{2,1}$ for the second level of hierarchy (business line level at $h = 2$) from the hierarchical Archimedean copula $C_{2,1}$.

TABLE 7

BOOTSTRAP BIAS FOR VARIOUS MODELS.

Model	Copula Reserve	Bootstrap Reserve	Bias	Std Error
Frank PWD Model	6,999,267	7,042,276	0.61%	352,825
Frank Independent Model	6,916,870	7,011,597	1.35%	565,001
Frank Hierarchical Model	7,348,481	7,291,707	0.77%	886,364

5.1.2. *Bias and MLE.* The maximum likelihood estimation technique is known to be asymptotically unbiased. In practice, we work with a finite number of observations, particularly with runoff triangles. Indeed, in our empirical illustrations, only 55 observations have been used in each triangle. Consequently, regardless of the number of simulations, our estimation is done each time on limited datasets of 55 observations.

The impact of the bias on the estimation has been analyzed. Recently, the lognormal MLE bias has been studied in Johnson *et al.* (2011), along with the gamma and Weibull distributions. Consequently, inter alia, a bias is necessarily observed in the bootstrapping procedure. In our empirical illustration, the bootstrap bias obtained for various models is exhibited in Table 7.

5.2. Reserve indications

In Table 7, we exhibit the bootstrap results for all three models (PWD model, CY dependence model with independence between lines of business and the hierarchical model). We show a histogram of the reserve distribution, with the corresponding percentiles. Figure 3 is important and useful for actuaries, when they want to select a reserve at a desired level of conservatism. The Frank independent model captures the dependence within a runoff triangle but not between the two lines of business. This explains why the prediction error is smaller than the Frank hierarchical model but greater than the Frank PWD model. It has been shown in Wüthrich *et al.* (2013) that the CY modeling is more performant than the PWD modeling. One sees for example that the PWD models can underestimate the variability because they implicitly assume an independence between accident years. The introduction of accounting year dependence may substantially increase the prediction uncertainty (see Table 7). Note that to compute the mean square error of prediction, the process uncertainty must be added to this prediction error (see England and Verrall (2002)).

Note that to obtain the lower triangle (step 3 of the Bootstrap procedure described in Section 5.1.1), we can either calculate the projected mean for each cell of the lower triangle, as shown in this paper (projected mean approach), or generating (by simulation) each cell of the lower triangle starting from the new estimates obtained for each bootstrap sample. The second approach (the simulation based approach) offers a wider range of possible reserves, and will consequently have a larger standard error. This second approach can be

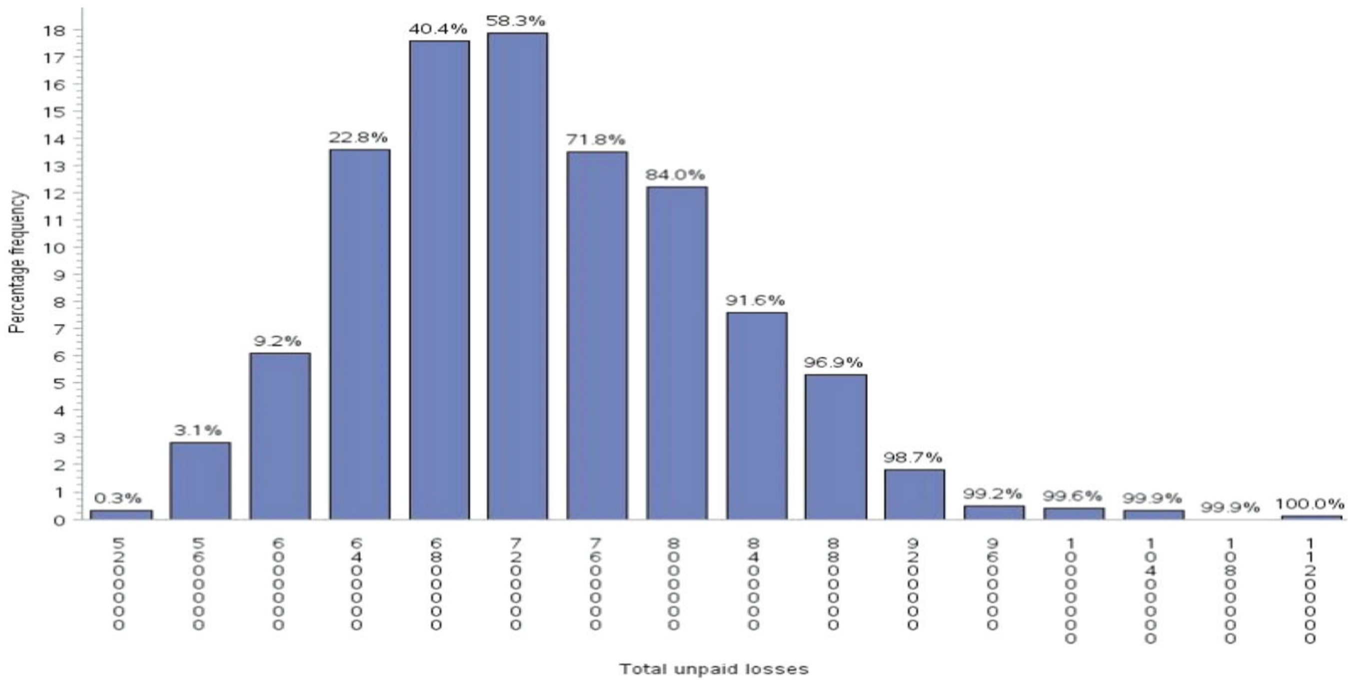

FIGURE 3: Percentiles of total unpaid losses (in millions)-complete hierarchical model. (Color online)

https://doi.org/10.1017/asb.2015.6 Published online by Cambridge University Press

particularly interesting from a capital risk standpoint where extreme loss events have to be considered. Both bootstrap approaches (projected mean approach and simulation based approach) are relevant information for property-casualty insurers.

6. Conclusion

In this paper, we have studied different approaches to model the dependence between loss triangles using multivariate copulas. If losses in different lines of business are correlated, aggregate reserves must reflect this dependence. To allow a complex dependence relation, we propose the use of new models using hierarchial Archimedean copulas. To illustrate the model, an empirical illustration was performed using the same data as the one used by Shi and Frees (2011). Despite the fact that the commercial line generates outliers with ICYD and HCYD models, we identified this situation by the inflexion of the last CY. Based on the AIC and on the BIC, we show that the hierarchical Archimedean models provide a better fit than PWD models.

With the proposed models, we can derive analytically the value of the reserve. However, to obtain the distribution of the reserve, in taking into account the estimation errors of all parameters, we proposed a bootstrap method where an algorithm to simulate a hierarchical copula model is proposed. From this bootstrap method, we can observe, as expected, that the dependence assumption between lines of business increase significantly the prediction uncertainty.

These new models that use hierarchical copula theory constitute a new way to model the dependence structures of runoff triangles. Those models are promising tools to better take into account dependencies within and between business lines. Indeed, this approach can easily be generalized to more than two lines of business because hierarchical Archimedean copulas are flexible and allow more refined possible dependence constructions. Because of their flexibility, hierarchical copula models should also be considered in other areas of actuarial science.

Acknowledgements

The authors wish to thank Mario V. Wüthrich and the anonymous reviewers for their valuable comments and suggestions to improve the original manuscript. Anas Abdallah gratefully acknowledge the financial support of Cooperators General Insurance Company and Mitacs Accelerate program. Jean-Philippe Boucher and Hélène Cossette would like to thank the financial support from the Natural Sciences and Engineering Research Council of Canada.

Note

1. ORSA: Own Risk and Solvency Assessment.

References

BARNETT, G. and ZEHNWIRTH, B. (1998) Best estimates for reserves. *Casualty Actuarial Society Forum (Fall)*, 1–54.

BRAUN, C. (2004) The prediction error of the chain ladder method applied to correlated run-off triangles. *ASTIN Bulletin*, **34**(2), 399–434.

BREHM, P. (2002) Correlation and the aggregation of unpaid loss distributions. *Casualty Actuarial Society Forum (Fall)*, 1–23.

DE JONG, P. (2006) Forecasting runoff triangles. *North American Actuarial Journal*, **10**(2), 28–38.

DE JONG, P. (2012) Modeling dependence between loss triangles. *North American Actuarial Journal*, **16**(1), 74–86.

ENGLAND, P. and VERRALL, R. (2002) Stochastic claims reserving in general insurance. *British Actuarial Journal*, **8**(3), 443–518.

HOFERT, M., MÄCHLER, M. and MCNEIL, A. J. (2012) Likelihood inference for Archimedean copulas in high dimensions under known margins. *Journal of Multivariate Analysis*, **110**, 133–150.

JOE, H. (1997) *Multivariate Models and Dependence Concepts*. London: Chapman and Hall.

JOHNSON, P.H.J., QI, Y. and CHUEH, Y. (2011) Bias-corrected maximum likelihood estimation in actuarial science. *Working paper*.

KIRSCHNER, G., KERLEY, C. and ISAACS, B. (2008) Two approaches to calculating correlated reserve indications across multiple lines of business. *Variance*, **2**(1), 15–38.

KUANG, D., NIELSEN, B. and NIELSEN, J. (2008) Forecasting with the age-period-cohort model and the extended chain-ladder model. *Biometrika*, **95**(4), 987–991.

KUANG, D., NIELSEN, B. and NIELSEN, J. (2011) Forecasting in an extended chain-ladder-type model. *Journal of Risk and Insurance*, **78**(2), 345–359.

LOWE, J. (1994) A practical guide to measuring reserve variability using bootstrapping, operational times and a distribution-free approach. *General Insurance Convention, Institute of Actuaries and Faculty of Actuaries.*

MARSHALL, A.W. and OLKIN, I. (1988) Families of multivariate distributions. *Journal of the American Statistical Association*, **83**(403), 834–841.

MCNEIL, A. and NEŠLEHOVÁ, J. (2009) Multivariate Archimedean copulas, d-monotone functions and ℓ_1-norm symmetric distributions. *The Annals of Statistics* **37**(5B), 3059–3097.

MEYERS, G. (2013) *Stochastic loss reserving using Bayesian MCMC models*. ASTIN Colloquium May 23, 2013.

OKHRIN, O., OKHRIN, Y. and SCHMID, W. (2013) On the structure and estimation of hierarchical Archimedean copulas. *Journal of Econometrics* **173**(2), 189–204.

SAVU, C. and TREDE, M. (2010) Hierarchies of Archimedean copulas. *Quantitative Finance*, **10**(3), 295–304.

SHI, P. (2014) A copula regression for modeling multivariate loss triangles and quantifying reserving variability. *ASTIN Bulletin* **44**(1), 85–102.

SHI, P., BASU, S. and MEYERS, G. (2012) A Bayesian log-normal model for multivariate loss reserving. *North American Actuarial Journal* **16**(1), 29–51.

SHI, P. and FREES, E. (2011) Dependent loss reserving using copulas. *ASTIN Bulletin*, **41**(2), 449–486.

TAYLOR, G. and MCGUIRE, G. (2007) A synchronous bootstrap to account for dependencies between lines of business in the estimation of loss reserve prediction error. *North American Actuarial Journal*, **11**(3), 70–88.

WÜTHRICH, M. (2010) Accounting year effects modeling in the stochastic chain ladder reserving method. *North American Actuarial Journal*, **14**(2), 235–255.

WÜTHRICH, M. (2012) Discussion of "A Bayesian log-normal model for multivariate loss reserving" by Shi-Basu-Meyers. *North American Actuarial Journal*, **16**(3), 398–401.

WÜTHRICH, M. (2013) *Calendar year dependence modeling in run-off triangles*. ASTIN Colloquium, May 21–24, The Hague.

WÜTHRICH, M., MERZ, M. and HASHORVA, E. (2013) Dependence modelling in multivariate claims run-off triangles. *Annals of Actuarial Science*, **7**(1), 3–25.

WÜTHRICH, M. and SALZMANN, R. (2012) Modeling accounting year dependence in runoff triangles. *European Actuarial Journal*, **2**(2), 227–242.

ANAS ABDALLAH
École d'Actuariat
Université Laval
Quebec City (Quebec)
E-Mail: anas.abdallah.1@ulaval.ca

JEAN-PHILIPPE BOUCHER (Corresponding author)
Département de mathématiques
UQAM Montreal (Quebec)
E-Mail: boucher.jean-philippe@uqam.ca

HÉLÈNE COSSETTE
École d'Actuariat
Université Laval
Quebec City (Quebec)
E-Mail: helene.cossette@act.ulaval.ca