© *The Journal of Risk and Insurance*, 2013, Vol. 80, No. 4, 891–919
DOI: 10.1111/j.1539-6975.2012.01480.x

# Predicting Multivariate Insurance Loss Payments Under the Bayesian Copula Framework

Yanwei Zhang
Vanja Dukic

## Abstract

The literature of predicting the outstanding liability for insurance companies has undergone rapid and profound changes in the past three decades, most recently focusing on Bayesian stochastic modeling and multivariate insurance loss payments. In this article, we introduce a novel Bayesian multivariate model based on the use of parametric copula to account for dependencies between various lines of insurance claims. We derive a full Bayesian stochastic simulation algorithm that can estimate parameters in this class of models. We provide an extensive discussion of this modeling framework and give examples that deal with a wide range of topics encountered in the multivariate loss prediction settings.

## Introduction

In his 1987 article, Bühlmann made the comment that the casualty actuaries have to master the skills of probabilistic thinking in order to deal with a variety of risky situations. Because of the complexity and diversity of risks arising in the property–casualty insurance, probability distributions of unknown quantities are often required for the formal solutions of decision problems. One such important example is the practice of insurance loss reserving. Due to reasons such as late reported claims, judicial proceedings, or schedules of benefits for employer's liability claims, many types of property–casualty insurance claims often have lengthy settlement periods, with liability claims often taking years or even decades to complete. In order to be able to respond to outstanding claims, every insurance company must set aside a provision, known as a loss reserve. The loss reserve is typically a property–casualty insurance company's largest balance sheet liability. Its proper prediction is therefore a matter of vital importance to the company and the research around loss reserve prediction has become a central subject in modern actuarial science.

The loss reserving literature has undergone rapid and profound changes in the past three decades, where a large number of innovative applications of statistical methods

---

Yanwei Zhang is at CNA Insurance Company. Vanja Dukic is at the University of Colorado-Boulder. The first author can be contacted via e-mail: Yanwei.Zhang@cna.com. The authors thank an anonymous referee for his/her many constructive suggestions that have helped improve the contents of this article greatly. The first author is also grateful to Jian Peng at MIT for his generous help while programming the simulation algorithm.

have led to the proliferation of a great variety of stochastic loss reserving models. As the body of the literature continues to grow, some new trends have been observed in recent years. One such important trend is the increasing reliance on Bayesian methods (e.g., de Alba, 2002; Ntzoufras and Dellaportas, 2002; Antonio and Beirlant, 2008; Zhang, Dukic, and Guszcza, 2012). Bayesian analysis has gained some of its popularity due to its inherent advantages in the applied loss forecasting context. For example, the posterior predictive distribution conveys a wealth of information well beyond the uncertainty estimators commonly used in the traditional stochastic loss reserving analysis. This full measure of liability risk can be further incorporated into the determination of economic capital for the purpose of capital allocation and making strategic and tactical decisions. The second trend is developments in the realm of multivariate methods, where the major focus lies in the creation of models that enable integration of correlated multidimensional claim information for simultaneous statistical processing in order to improve, with respect to inference efficiency and prediction and risk assessment accuracy, upon univariate stochastic methods that could potentially fail due to the negligence of the inter-relationship within the multidimensional data. For example, substantial effort (e.g., Braun, 2004; Merz and Wüthrich, 2008; de Jong, 2012; Shi and Frees, 2011) has been devoted to making the extension of standard stochastic reserving models to account for correlations among various lines of insurance, with the aim to depict a more accurate picture of the liability risks faced by the insurance company.

Although considerable advances have been achieved in both of the above fields, Bayesian analysis in the multivariate setting has rarely been found in the loss reserving literature (see Shi and Frees, 2011). It is the intention of this article to fill such a gap by laying out a unified and flexible framework to perform multivariate loss reserving analysis using Bayesian methods. One distinct feature of the proposed framework is the use of parametric copulas (e.g., see Frees and Valdez, 1998) within a Bayesian paradigm to construct multivariate joint distributions. The copula model allows various marginal model structures or multivariate outcomes of mixed types to be specified and modeled simultaneously, yielding a rich family of probability models for analyzing multivariate reserving problems (see, e.g., Shi and Frees, 2011; de Jong, 2012). Associated with the model flexibility in copulas is the increasing difficulty in parameter inference, which, in many situations, would have not been feasible without the use of the Bayesian simulation-based methods. Moreover, the development of the multivariate methods in the Bayesian framework also brings many other advantages to the loss reserving context. For example, compared to existing multivariate loss reserving methods, the Bayesian copula model accounts for all sources of uncertainty from the data and parameters and generates predictive distributions for all quantities of interest. Furthermore, when the collected data have an inherent multilevel structure, hierarchical models can be employed for efficient statistical inference. Although much of the article is within the context of insurance loss reserving, we note that the proposed Bayesian copula framework also has potential applications in the field of economic capital evaluation (Tang and Valdez, 2009), enterprise risk management (CAS, 2003), and dynamic financial analysis (Kaufmann, Gadmer, and Klett, 2001), where Monte Carlo methods are often adopted to generate risk distributions and the aggregation of all risk distributions needs to reflect correlations and portfolio effects.

The organization of the article is as follows. The next section will present a brief survey of the Bayesian and multivariate methods that recently appeared in the literature, point out some of the advantages and shortcomings of these methods, and provide comments motivating the Bayesian copula model to be presented in this article. The third section will introduce the copula method for constructing multivariate distributions and discuss some of the more prominent univariate stochastic loss reserving methods. In the fourth section, we consider numerical examples where copula models along with a variety of marginal methods are implemented within the Bayesian paradigm to deal with a wide range of topics encountered in the multivariate loss reserving setting. We also conduct the standard exercise to compare models using out-of-sample validation data points and recommend the best one under consideration. However, it is to be noted that our emphasis is on statistical inference of the assumed model rather than attempting to search for the best possible one. Finally, the fifth section will provide concluding comments.

## REVIEW OF BAYESIAN AND MULTIVARIATE METHODS IN LOSS RESERVING

### Bayesian Methods in Actuarial Loss Reserving

The application of Bayesian statistics has been found in various areas of actuarial science (e.g., see Makov, 2001), and recently, there has been increasing interest in adapting the Bayesian methodology to loss reserving analyses.

A spate of Bayesian models appeared in the recent loss reserving literature: for example, De Alba (2002) considers a number of specifications that allow the use of measures of exposures or loss counts in addition to loss payment amounts; Ntzoufras and Dellaportas (2002) explore the use of state space modeling and the inclusion of inflation factors; Mildenhall (2005) presents a rich yet tractable Bayesian model of claim count development using the gamma-Poisson and Dirichlet-multinomial conjugate distribution families; de Alba and Nieto-Barajas (2008) propose a model that accounts for various sources of correlation through the use of a correlated Gamma process; Antonio and Beirlant (2008) apply semiparametric smoothing models and employ generalized linear mixed models to build in credibility in the estimation of loss reserves; Peters, Shevchenko, and Wüthrich (2009) examine the loss reserving problem using a Bayesian Tweedie model; Zhang, Dukic, and Guszcza (2012) construct a hierarchical nonlinear parametric model that makes robust estimation with multilevel pooling of information across accident years and companies.

This abundant appearance of Bayesian loss reserving models is due in part to the rapid advances in the theory of Bayesian statistics that is spurred by the use of simulation-based inference, in part to the availability of inexpensive high-speed computing power that makes the construction and estimation of computationally demanding models feasible, but in a larger, and more important part, due to the inherent advantages of Bayesian methodology in the applied loss reserving context; for example, it generates predictive distribution of loss reserves, allows incorporation of expert opinion in a formal way, and minimizes the risk of company management misinterpreting probabilistic statements. See Zhang, Dukic, and Guszcza (2012) for a more detailed account of the advantages of Bayesian loss reserving models.

### Multivariate Loss Reserving

A second distinct trend in the recent loss reserving literature is the increasing interest in methods that cope with multidimensional data. The aim is to understand and utilize the relationship among multivariate outcomes in order to achieve a more efficient and accurate estimation of the outstanding liability and the associated uncertainty measure. Substantial effort has been devoted to creating methods that incorporate structural connections among various types of data in improving model adequacy (e.g., de Alba, 2002; Ntzoufras and Dellaportas, 2002; Quarg and Mack, 2004; Zhang, 2010), to building hierarchical models that reflect within-cohort dependencies and allow the pooling of information across multiple claim histories (e.g., Zhang, Dukic, and Guszcza, 2012), and to a larger extent, to developing models that explicitly account for the contemporaneous correlations among various lines of business (see, e.g., Merz and Wüthrich, 2008a, 2008b). The current article is mainly focused on this last area.

In common practice, the estimation of the loss reserve for a portfolio is usually formed by combining the individual inferences made for each of several subportfolios. The purpose of dividing the portfolio is to make homogeneous subportfolios so that an appropriate univariate stochastic loss reserving model can be chosen for each subportfolio. In general, contemporaneous correlations exist among subportfolios, and independent estimates of individual loss reserve uncertainty cannot be combined directly. In general, contemporaneous correlations exist among the subportfolios and independent estimates of reserve variability cannot be combined directly. To appropriately account for the correlation, a multivariate distribution must be constructed and joint estimation of correlated subportfolios must be carried out.

In addition to appreciating the efficiency gains in parameter estimation, actuarial analyses often place a greater emphasis on the resulting reserve predictions and variability measures. For example, it is generally perceived that subportfolios from different lines of business have positive correlations, and as such, independent applications of the univariate stochastic loss reserving model will lead to underestimation of the reserve variability, depicting a falsely optimistic picture of the economic risks faced by the insurance company. Within this context, several papers (e.g., Braun, 2004; Merz and Wüthrich, 2008b) have extended the univariate chain ladder approach of Mack (1993) to arrive at estimated loss reserves and corresponding variability measures that reflect contemporaneous correlations (see also Merz and Wüthrich, 2008a). Zhang (2010) shows that these methods are equivalent to sequential applications of the feasible generalized least squares. Although these multivariate chain ladder models provide a practical solution to the problem of simultaneous modeling of correlated subportfolios, they inevitably encounter several shortcomings due to the multiplicative structure inherited from the chain ladder model; for example, they lack procedures for model assessment and selection, do not yield the predictive distribution of loss reserves, and are not flexible in choice of model structure—all the existing estimation theories are built exclusively for the multiplicative structure of the chain ladder model.

Distinguished from the multivariate chain ladder models, another approach to the multivariate loss reserving problem relies on parametric copulas (see, e.g., de Jong, 2012; Shi and Frees, 2011). The copula approach overcomes many difficulties within the multivariate chain ladder framework, such as the challenge to generalize to model structures other than the chain ladder. We give a more detailed discussion of the copula framework first, before developing the Bayesian copula framework.

## Multivariate Distribution With Prespecified Marginals

The copula methodology is based on the results of Sklar who proved that a multivariate distribution can be characterized by the set of its marginal distributions and a "copula" function that describes the dependence between the marginal components (uniquely in the case of continuous distributions). Thus, a copula can be used for generating samples from multivariate densities with prespecified marginals and a dependency function.

It is worth noting that copulas are not the sole way to construct multivariate distributions with prespecified marginals. In fact, besides copulas, the methods for generating such multivariate distributions in practice range from simple composition methods (sequentially drawing one variable from its marginal density, followed by the draws of others from conditional densities), to custom algorithms for particular distributions (Lawrance and Lewis, 1981; Schmeiser and Lai, 1982), to multidimensional acceptance/rejection sampling and Markov Chain Monte Carlo (MCMC) methods (Robert and Casella, 2005). An excellent historical overview of the developments in the field of probability distributions with prespecified marginals can be found in the volumes by Dall'Aglio, Kotz, and Salinetti (1991), Rüschendorf, Schweizer, and Taylor (1996), and Schweizer and Sklar (1986).

On the other hand, copula methods are arguably the most widely used algorithm today for generating samples from multivariate densities with prespecified marginals and a dependency function. Copulas have been used in actuarial science for joint estimation of insurance loss payments and loss adjustment expenses (Frees and Valdez, 1998), for modeling the dependence among different claim types (Frees and Valdez, 2008), in economic modeling (Tang and Valdez, 2009), and in loss reserving analyses of individual-level claims data (Zhao and Zhou, 2010) as well as aggregate triangles (Shi and Frees, 2011; de Jong, 2012).

Different from the current copula loss reserving models (e.g., Shi and Frees, 2011), our analysis is carried out in the Bayesian framework so that it enjoys many of the familiar advantages of Bayesian analysis. Although not yet seen in the loss reserving context, Bayesian copula models have found applications in joint survivor analysis (Shemyakin and Youn, 2006) and financial asset allocations (Kang, 2011). The books by Shevchenko (2011) and Böcker (2010) provide a good overview of the implementation of Bayesian copula models in modern risk management practices.

Another important difference between our work and the existing copula loss reserving methodology lies in that we consider flexible and possibly different marginal models such as nonlinear growth models, smoothing models, as well as the commonly used generalized linear models. Furthermore, hierarchical structures are built in whenever appropriate, reflecting the concept of actuarial credibility (Frees, Young, and Luo, 1999) and allowing estimation of group-level effects and variation.

## COPULA AND MARGINAL MODELS

This section reviews some of the statistical concepts used in this article. We first introduce the copula function and then summarize some of the prominent statistical models that have been historically used in the loss reserving analysis. Throughout the article, we will be focusing on the modeling of the incremental losses, which are the actual amounts paid out during a time period. Departing from the convention of denoting the data in a triangular form (see Table A1 in the Appendix), we use the long format representation of the loss reserving data (Table A2), emphasizing the regression-type nature present in most analyses.

The notation we will be using to describe the probability models is as follows. Denote $L$ as the number of margins in the copula model (the number of loss triangles), $N$ the number of data points observed in each marginal model, and $J_l$ the number of mean parameters in the $l$th marginal model. Throughout the article, we use boldface to represent a vector or a matrix, and $'$ to represent matrix transpose. We denote $\boldsymbol{X}$ (or $\boldsymbol{Z}$ in some case) as the design matrix in a linear model, and $\boldsymbol{x}$ as the covariate in the model, discrete or continuous. So, if $\boldsymbol{x}$ is treated continuous, its associated design matrix will be an $N \times 1$ vector, but if $\boldsymbol{x}$ is treated as a categorical variable with $J$ levels, then the design matrix $\boldsymbol{X}$ generated by $\boldsymbol{x}$ will be an $N \times J$ matrix.

### Copula

Formally, a copula is a multivariate distribution function with uniform marginal densities that correspond to the probability integral transformation (PIT) of the individual random variables. More precisely, if $y_1, \ldots, y_L$ are the continuous random variables with the joint distribution function $F$ and marginal distributions $F_1, \ldots, F_L$, then Sklar's theorem establishes that there exists a unique copula function $C : [0,1]^L \to [0,1]$ such that

$$F(y_1, \ldots, y_L) = C(F_1(y_1), \ldots, F_L(y_L)) \quad \forall (y_1, \ldots, y_L) \in \mathbb{R}^L. \tag{1}$$

Hence, the dependence among the original variables is translated, via the copula function, into the dependence on the scale of uniform random variables, and then the desired distributions are obtained via a transformation (often via inexact methods) of these correlated uniform variables. Under standard smoothness conditions, one can recover the joint density by differentiating (1) to obtain

$$f(y_1, \ldots, y_L) = c(F_1(y_1), \ldots F_L(y_L)) \cdot \prod_{i=1}^{L} f_i(y_i), \tag{2}$$

where $c(u_1, \ldots, u_L) = \frac{\partial^L C(u_1, \ldots, u_L)}{\partial u_1 \ldots \partial u_L}$ with $u_i = F_i(y_i)$ is the joint density on the $L$ uniforms. Many copulas are available for constructing multivariate distributions, and in this article, we consider two most frequently used copula families: the Archimedean copulas and the elliptical copulas.

*Archimedean Copula.* An Archimedean copula is constructed through a generator, a convex, and decreasing function $\phi$ with domain $(0,1]$ and range $[0,\infty)$ such that $\phi(1) = 0$, as

**TABLE 1**
Archimedean Copulas

| Copula | $C(u, v)$ | Support of $\alpha$ | Kendall's $\tau$ |
|---|---|---|---|
| Clayton | $(u^{-\alpha} + v^{-\alpha} - 1)^{-\frac{1}{\alpha}}$ | $[-1, \infty)$ | $\frac{\alpha}{\alpha+2}$ |
| Gumbel | $\exp(-[(-\log u)^{\alpha} + (-\log v)^{\alpha}]^{\frac{1}{\alpha}})$ | $[1, \infty)$ | $1 - \frac{1}{\alpha}$ |
| Frank | $\frac{-1}{\alpha} \log\left(1 + \frac{(e^{-\alpha u}-1)(e^{-\alpha v}-1)}{e^{-\alpha}-1}\right)$ | $(-\infty, 0) \cup (0, \infty)$ | $1 - \frac{4}{\alpha} + \frac{4}{\alpha^2}\int_0^{\alpha} \frac{t}{e^t-1}dt$ |

*Note*: Listed are the distribution function, parameter support, and Kendall's $\tau$ for each of the three Archimedean copulas considered in this article. $\alpha$ is the parameter associated with the copula.

$$C(u_1, \ldots, u_L) = \phi^{-1}(\phi(u_1) + \cdots + \phi(u_L)). \tag{3}$$

Table 1 shows, in the bivariate form for simplicity, the three one-parameter Archimedean copulas that are used later in this article. Some detailed discussion of Archimedean copulas can be found in Nelsen (2006).

*Elliptical Copula (Gaussian).* Different in construction from the Archimedean copula, the elliptical copula corresponds to an elliptical distribution (Fang, Kotz, and Ng, 1990). Specifically, it is determined by a multivariate elliptical distribution function $F$ as

$$C(u_1, \ldots, u_L) = F\left(F_1^{-1}(u_1), \ldots, F_L^{-1}(u_L)\right), \tag{4}$$

where $F_i$ is the distribution function for the $i$th margin and $F_i^{-1}$ is the corresponding inverse distribution function. The elliptical copula considered in this article is the Gaussian copula where $F = \Phi_L$, the distribution function for the $L$-variate multivariate normal $N_L(\mathbf{0}, \mathbf{\Gamma})$ with standardized margins, and $F_i = \Phi$, the distribution function for the standard normal variable. Thus, under the Gaussian copula, Equation (4) becomes

$$C(u_1, \ldots, u_L) = \Phi_L(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_L)). \tag{5}$$

The Gaussian copula has much flexibility in the selection of the correlation parameters, and also due to the easy implementation, it has become very popular in many applied fields.

*Kendall's $\tau$.* It is understood that the widely used Pearson's correlation coefficient is affected by nonlinear transformation of scales; thus, it is determined not only by the copula but also by the marginal distributions. On the contrary, Kendall's correlation coefficient ($\tau$) could be expressed only in terms of the copula function, which is indifferent to nonlinear transformation of scales. Thus, we use Kendall's correlation coefficient throughout the article as a measure of association. Kendall's $\tau$ is given in Table 1 for the three Archimedean copulas, and for the multivariate Gaussian copula, it is calculated as

$$\tau(y_i, y_j) = \frac{2}{\pi}\arcsin(\alpha_{ij}), \tag{6}$$

where $\alpha_{ij}$ is the $(i, j)$th element of $\mathbf{\Gamma}$, the covariance matrix of the $L$-variate multivariate normal $N_L(\mathbf{0}, \mathbf{\Gamma})$ with standardized margins.

Copulas provide a convenient way to link several marginal distributions through the imposition of a joint distribution on the individual PITs. To construct a complete joint distribution, one needs to select the appropriate loss reserving model for each margin as well as a proper form of the copula function. The following reviews some of the more popular univariate loss reserving methods that are used later in the data analysis.

### Generalized Linear Models

The use of generalized linear model has been a primary focus in the early stochastic loss reserving literature, largely due to the equivalence of the overdispersed Poisson log-linear model with the popular deterministic chain ladder approach. Here, we briefly summarize the structure of the generalized linear models. England and Verrall (2002) give a detailed overview of their use in insurance loss reserving, and McCullagh and Nelder (1989) present a comprehensive introduction.

In the generalized linear model, the response variable $y_i$s are assumed to be independent and come from one member of the exponential family for $i = 1, \ldots, N$. The mean $\mu_i$ is assumed to relate to some linear predictors through a known link function $\eta$ as

$$\eta(\mu_i) = \boldsymbol{X}_i' \boldsymbol{\beta}, \tag{7}$$

where $\boldsymbol{X}$ is the design matrix and $\boldsymbol{\beta}$ is the vector of mean parameters. A simple but commonly adopted approach is to put the accident year and the development lag as categorical covariates (see Table A2) into the main effect model as

$$\eta(\mu_i) = \text{ accident year} + \text{ development lag}. \tag{8}$$

Throughout the article, the parameter corresponding to the first level of each categorical variable is set to be zero to make the estimation feasible.

### Nonlinear Growth Model

Clark (2003) introduces a stochastic loss reserving model that uses a growth curve to model the loss development process. In such a model, the mean of the cumulative loss in each year is assumed to follow a nonlinear growth curve. Equivalently, one can formulate the model on the basis of incremental losses by taking the first-order difference of the cumulative losses, in which the mean of the incremental loss is assumed to be

$$\mu_i = p_i \cdot \gamma \cdot [G(t_i; \mathbf{\Theta}) - G(t_i - \Delta t_i; \mathbf{\Theta})]. \tag{9}$$

In the above equation, $p_i$ is the (given) premium for the year where $y_i$ is measured, and $\gamma$ denotes the all-year-combined expected ultimate loss ratio. $G(t_i, \mathbf{\Theta})$ is a parametric growth curve that depends on parameters $\mathbf{\Theta}$ and measures the percentage of ultimate losses that have emerged as of time $t_i$. Therefore, $G$ must have the properties that

$G(0; \mathbf{\Theta}) = 0$ and $G(t_i; \mathbf{\Theta}) \to 1$ as $t_i \to \infty$. Denoting $\Delta t_i$ as the time elapsed since the latest loss measurement before $y_i$, then $G(t_i; \mathbf{\Theta}) - G(t_i - \Delta t_i; \mathbf{\Theta})$ measures the percentage of the ultimate losses emerged during the time interval $(t_i - \Delta t_i, t_i)$.

The formulation of the nonlinear growth model has several appealing points; for example, the explicit modeling of the ultimate loss ratios $\gamma$ allows easy incorporation of relevant expert opinions, and the curve performs direct within-model interpolation and extrapolation to produce predictions at any evaluation time. Moreover, hierarchical model structure can be readily incorporated to allow the loss ratio $\gamma$, and in principle the growth curve parameters $\mathbf{\Theta}$, to vary randomly across accident years (see Zhang, Dukic, and Guszcza, 2012).

### Semiparametric Methods

Although intuitively appealing, the parametric growth curve dictated by the above model may not be fulfilled in certain circumstances where no clear nonlinear growth pattern is detected. In such a case, semiparametric models can be of substantial value mostly in that they can retain the essential features of the data while discarding unimportant and noisy details. Compared to the generalized linear model where a tail factor must be supplemented in long-tailed products, the semiparametric smoothing models allow the extrapolation of loss payment to some known point beyond the range of the observed data. Loss reserving applications of these models can be found in Verrall (1996) and Antonio and Beirlant (2008).

To understand the idea of semiparametric linear models, consider the response variable $y_i$ that comes from an exponential family as in generalized linear models, but with the linear predictor replaced by some smoothing functions such that

$$\eta(\mu_i) = g(x_i), \quad \text{for } i = 1, \ldots, N, \tag{10}$$

where $g$ is a smoothing function for the covariate $x_i$. Here, we consider the case where there is only one covariate for ease of formulation, but the generalization of the theory to generalized additive models where multiple smoothing terms are included (Hastie and Tibshirani, 1990) is straightforward (e.g., see Ruppert, Wand, and Carroll, 2003; Antonio and Beirlant, 2008).

In actuarial applications, the above nonparametric model is often mixed with additional parametric components, for example, the inclusion of the accident year effect, thus becoming a semiparametric model (see Verrall, 1996; Antonio and Beirlant, 2008). There are many possible choices for the smoothing functions, and this article considers the low-rank thin-plate splines that have been shown to exhibit many good numerical properties in Bayesian posterior simulations. This spline can be written as

$$g(x_i; \boldsymbol{\theta}) = \beta_1 + \beta_2 x_i + \sum_{k=1}^{K} b_k |x_i - \kappa_k|^3, \tag{11}$$

where $\boldsymbol{\theta} = (\beta_1, \beta_2, b_1, \ldots, b_K)'$ is the vector of regression parameters and $\kappa_1 < \kappa_2 < \cdots < \kappa_K$ are the positions of the knots. To avoid wiggly fit, it is necessary to impose constraints on the $b$s to penalize the roughness of the fit. It is widely appreciated that this penalized spline model can also be viewed as a particular case of the hierarchical model, in that imposing a penalty on the *b*s in the spline model is equivalent to assuming the *b*s are normally distributed random effects in the hierarchical model. As a result, we can express the above penalized spline model as

$$\eta(\mu_i) = \boldsymbol{X}_i'\boldsymbol{\beta} + \boldsymbol{Z}_i'\boldsymbol{b}, \tag{12}$$

$$b_k \sim N\left(0, \sigma_b^2\right), \tag{13}$$

where $\boldsymbol{X}_i = (1, x_i)', \boldsymbol{\beta} = (\beta_1, \beta_2)'$, and $\boldsymbol{b} = (b_1, \ldots, b_K)'$ is a $K \times 1$ vector, and $\boldsymbol{Z} = \boldsymbol{Z}_K \boldsymbol{\Sigma}_K^{-1/2}$ is an $N \times K$ matrix with the $i$th row of $\boldsymbol{Z}_K$ equal to $(|x_i - \kappa_1|^3, \ldots, |x_i - \kappa_K|^3)$ and the $(l, k)$th entry of $\boldsymbol{\Sigma}_k$ being $|\kappa_l - \kappa_k|^3$. Under this formulation, the Bayesian posterior simulation method can be employed to make parameter inference. See Crainiceanu, Ruppert, and Wand (2005) for details.

## DATA ANALYSIS IN THE BAYESIAN COPULA FRAMEWORK

In this section, the copula model along with the different marginals introduced above is implemented in the Bayesian framework and applied to numerical examples to demonstrate its usefulness and flexibility. Bivariate copulas with marginal generalized linear models are considered first as an extension of this popular univariate stochastic loss reserving methods. Then attention is focused on multivariate copula models admitting various forms of marginal loss reserving methods such as the penalized spline model and the nonlinear model.

The data we analyzed are extracted from historical Schedule P tables (Part 2) in the statutory annual statements that insurance companies are required to report to the National Association of Insurance Commissioners every year. We select one company (Pennsylvania National Insurance Group) to be included in this study and the detailed entries of the data are reported in Table A2 in the Appendix. The accident years available are from 1988 to 1997, resulting in 10 years of historical experience. This covariate is denoted as $x_1$ in Table A2 and the values are converted to the 1–10 scale. Losses are evaluated at 12-month intervals, with the highest available development age being 120 months. This gives rise to the development lag covariate, denoted as $x_2$ and ranging from 1 to 10. Incremental paid losses for these accident years that are reported in calendar years 1998–2006 but up to development lag 10 (i.e., $x_1 + x_2 > 11, x_1, x_2 <= 10$) are also collected and treated as the out-of-sample data set to evaluate the performance of the models considered.[1] Three lines of business, namely, personal auto, commercial auto, and workers' compensation insurance, within the company are included in our analysis, and they are referred to as 1–3, respectively, in specifying the probability model. The collected claim information is the incremental loss and accident year earned premium, denoted as $y$ and $p$, respectively. Thus, we have $L = 3$ lines of business, and for each line of business, we have $N = 55$ data points in the training set.

[1] These entries, along with the code we used for the article, are available in the online supplementary material at: http://www.actuaryzhang.com/publication.html

Additional notation is introduced to formulate the models in the multivariate setting. We impose multiple subscripts on data points, parameters, and functions to indicate the dimensions. The first subscript is often interpreted as the indicator of the margin (line of business in this case). Subsequent subscripts, if any, should be regarded as the row and column indicator, respectively. For example, we use $y_{li}$ to indicate the $i$th observation from the $l$th margin. Thus, the vector $(y_{l1}, \ldots, y_{lN})'$ is the $N \times 1$ vector of the observed incremental losses from the $l$th margin. As another example, given an $N \times J_l$ design matrix $\boldsymbol{X}_l$, its $i$th row is a $J_l \times 1$ vector $\boldsymbol{X}_{li}$, and its $(i, j)$th element is a scalar $X_{lij}$. We denote $M$ as the number of data points to be predicted in each marginal model. These data points, when added together, produce the predicted marginal aggregate loss reserve if tail development is ignored. Note that $M$ is also equal to the number of validation data points available for each line in this analysis, that is, $M = 45$.

As for the distributional form, the lognormal distribution is employed to model incremental losses, which has also been used by many other authors (see, e.g., Ntzoufras and Dellaportas, 2002). In the following, although the term generalized linear model is used when modeling lognormal distributions, we are actually referring to its logarithmic transformation, that is, the normal distribution.

### Example I: Bivariate Copula Model With Marginal Generalized Linear Model

In this example, two insurance lines, the personal auto and commercial auto insurance, will be modeled using bivariate copulas along with marginal generalized linear models. Structure (8) is imposed on both marginal models with identity link.

*Marginal Models.* Given the main effect model in (8), we write the marginal models as

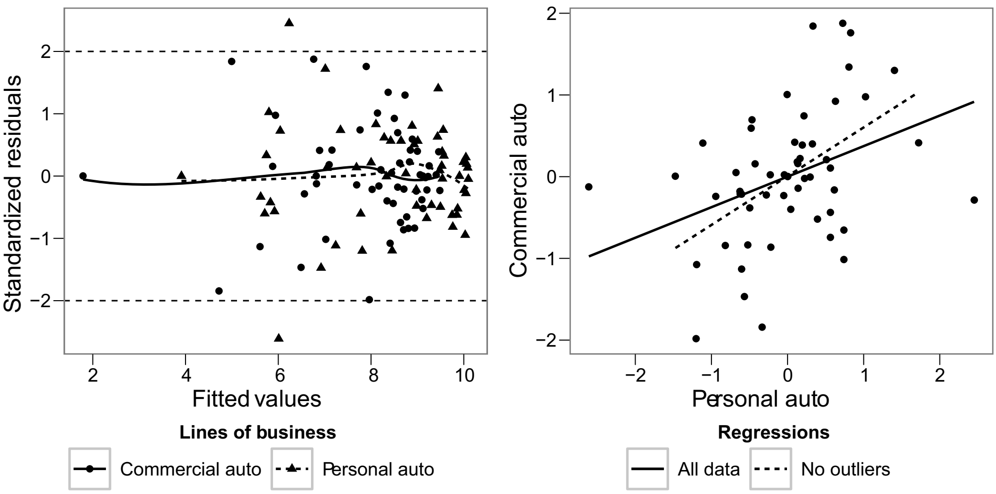
$$\log y_{li} \sim N(\mu_{li}, \sigma_l^2), \tag{14}$$

$$\mu_{li} = \boldsymbol{X}_{li}' \boldsymbol{\beta}_l \quad \text{for } i = 1, \ldots, N;\ l = 1, 2, \tag{15}$$

where $\boldsymbol{X}_{li}$ is the $J_l \times 1$ vector composed of the $i$th observation of the design matrix $\boldsymbol{X}_l$ generated by the main effect model, and $\boldsymbol{\beta}_l$ is the $J_l \times 1$ vector of the parameters in the $l$th margin. Under this specification, we have $J_l = 20$ for $l = 1, 2$.

Such models can be readily estimated using the linear regression technique. The relationship of residuals over the fitted values for each model is investigated graphically in the first plot of Figure 1. Despite two potential outliers outside the range $(-2, 2)$ in the personal auto model, there is no indication of substantial violations of the marginal model assumptions—the LOESS smoothers (Cleveland, 1979) of the residuals are generally flat. However, when the two sets of residuals are compared with each other in the second plot of Figure 1, strong positive relationship is detected, the Kendall's correlation coefficient being estimated at about 0.271. Intuitively, this positive relationship could come from many sources. For example, the automobiles underwritten by this company, whether they are for personal or commercial use, could be involved in the same accident or impacted by the same catastrophic events. Further, loss payments are also likely to obey similar claims settlement policies. When

**FIGURE 1**
Residual Plots From Separate Estimations of the Two Marginal Models




*Note*: The left-hand plot is the standardized residuals over the fitted values, along with the LOESS smoother, for each of the two marginal models. The right-hand plot is a scatterplot of the two sets of residuals. Two regression lines are added to help reveal the positive correlation, one using all the residuals and the other using those falling within the range $(-2, 2)$.

these contributing factors are significant and left out in the main effect model, residual correlation is induced.

*Copula Models.* To capture the positive correlation between the two insurance lines, we now join the two marginal linear models through the use of copula and make simultaneous inference of all parameters. We implement all the Archimedean copulas in Table 1 as well as the Gaussian copula. For example, under the Clayton copula, the joint distribution of the two lines, written in terms of the cumulative distribution function, becomes

$$F(\log y_{1i}, \log y_{2i}) = \left(u_{1i}^{-\alpha} + u_{2i}^{-\alpha} - 1\right)^{-\frac{1}{\alpha}}, \quad \text{for } i = 1, \ldots, N, \tag{16}$$

where

$$u_{li} = \Phi\left(\frac{\log y_{li} - \boldsymbol{X}_{li}'\boldsymbol{\beta}_l}{\sigma_l}\right), \quad \text{for } i = 1, \ldots, N; l = 1, 2, \tag{17}$$

is the PIT of $y_{li}$ under Equations (14) and (15).

For the Gaussian copula, it is to be noted that with Gaussian marginals, it in effect results in a multivariate normal distribution, that is, we have

$$(\log y_{1i}, \log y_{2i})' \sim N_2(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}), \tag{18}$$

where $\boldsymbol{\mu}_i = (\mu_{1i}, \mu_{2i})$ as in (15) and in $\boldsymbol{\Sigma}$, $\Sigma_{11} = \sigma_1^2$, $\Sigma_{22} = \sigma_2^2$ and $\Sigma_{12} = \Sigma_{21} = \alpha\sigma_1\sigma_2$, $\alpha$ being the usual Pearson's correlation parameter.

*Prior Distributions.* Since our analysis is to be implemented in the Bayesian framework, we need to specify the prior distributions for the parameters in the copula and the marginal models to complete the probability model formulation. Here, we will not specify a joint prior distribution for the copula parameter and the marginal model parameters, mainly because it is, in general, very difficult to understand how they are interrelated *a priori*. We treat them independently *a priori*. In this example, we will employ noninformative prior distributions on all parameters so that they are estimated based on the data almost exclusively. Of course, when additional information exists, such as expert opinion about what certain parameters might be, or prior studies illustrating the joint behavior of all parameters, incorporating it by specifying informative joint prior distributions on all parameters would be the Bayesian approach of choice.

For the Archimedean copulas, the following prior distributions are assigned to the parameters in the marginal model:

$$\beta_{lj} \sim N(0, 10^6), \text{ for } j = 1, \ldots, 20; l = 1, 2, \tag{19}$$

$$\sigma_l^2 \sim \text{Inv-Gamma}\,(10^{-3}, 10^{-3}) \text{ for } l = 1, 2. \tag{20}$$

The notation of the inverse gamma distribution means that for a random variable $x \sim \text{Gamma}(a, b)$ such that $E(x) = \frac{a}{b}, \text{Var}(x) = \frac{a}{b^2}$, we have $\frac{1}{x} \sim \text{Inv-Gamma}(a, b)$. We put a uniform prior on the copula parameter, where the bounds are determined so that the prior covers a wide range of the possible values on the support:

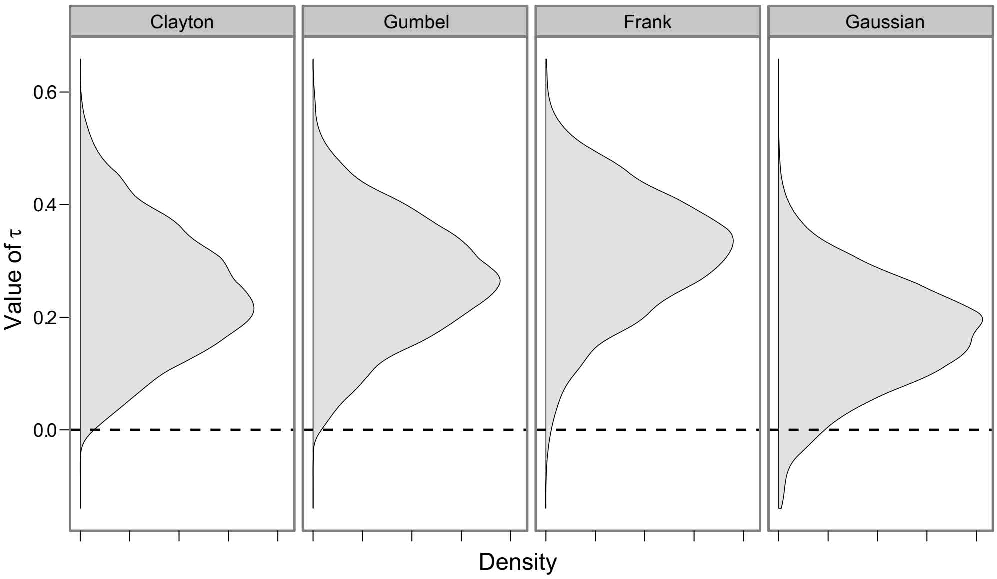
$$\begin{aligned} \text{Clayton: } & \alpha \sim U(0, 100); \\ \text{Gumbel: } & \alpha \sim U(1, 100); \\ \text{Frank: } & \alpha \sim U(-1,000, 1,000). \end{aligned}$$

For the Gaussian copula, we assign the prior distributions of the $\beta$s as in (19), and $\boldsymbol{\Sigma} \sim \text{Inv-Wishart}_3(\boldsymbol{I})$, where $\boldsymbol{I}$ is the $2 \times 2$ identity matrix and the degrees of freedom for the Inverse Wishart are set relatively low, but still high enough to maintain a proper distribution (see Gelman et al., 2004).

*Bayesian Inference.* To estimate the parameters, we simulate the full posterior distribution using Markov Chain Monte Carlo methods (Gelman et al., 2004). The "Metropolis- within-Gibbs" algorithm scheme sequentially samples parameters from their lower dimensional full conditional distributions over many iterations. The full conditionals are derived from the joint posterior density, which is proportional to the product of the prior and the likelihood.

For each of the four copula models presented above, we ran 300,000 iterations in three parallel chains, discarding the burn-in period of the first 250,000 iterations at which point the approximate convergence was achieved (the potential scale reduction factors of Gelman and Rubin (1992) were below 1.1 for all parameters). To reduce the autocorrelation, we used every 50th iteration of each chain. This resulted in 1,000 simulation draws per chain. Based on these simulation draws, we computed

**FIGURE 2**
The Posterior Distribution of the Correlation Coefficient $\tau$ From the Four Bivariate Copula Models




the posterior median and the standard deviation for the parameters in the model, which are included in Table A3 in the Appendix. We also implement the Bayesian estimation for the model where marginals are assumed independent and use it as a baseline model for the purpose of model comparison. This model is referred to as the "Independent" model in the following discussion. Kendall's correlation coefficient is calculated according to the formulas in Table 1, and the posterior median and the standard error are 0.239 (0.111), 0.267 (0.104), 0.320 (0.107), and 0.173 (0.096) for the Clayton, Gumbel, Frank, and Gaussian copulas, respectively. While there is some variation in the estimated Kendall's correlation across the copula models (where the Frank copula reports the largest, 0.320, and the Gaussian copula produces the smallest, 0.173), all these models indicate that statistically significant positive correlation exists between the two lines—one can see from Figure 2, where the posterior distribution of $\tau$ is plotted for each copula model, that the probability of $\tau < 0$ is very small.

*Predictive Distribution.* Given the simulated model parameters from the full posterior distribution, we could make predictions for the 45 out-of-sample data points and compare models by choosing an appropriate loss function. Also, the predictive distribution of the loss reserves can be estimated based on these predicted values.

The predictive simulation for the Gaussian copula is most straightforward via efficient multivariate normal simulation algorithms (see also in the next example). For the Archimedean copulas, we proceed using the following two steps. For the $m$th data point in the out-of-sample set, $m = 1, \ldots, 45$, we generate the $s$th predicted value $y_{lm}^{(s)}$, for $l = 1, 2$ and $s = 1, \ldots, 3,000$, by

**TABLE 2**
Summary of the Loss Reserve Predictions for the Bivariate Models (Numbers Are in 1,000s)

| Statistics | Line | Independent | Clayton | Gumbel | Frank | Gaussian |
|---|---|---|---|---|---|---|
| Median | $R_1$ | 106,537 | 103,674 | 104,613 | 106,676 | 107,930 |
| | $R_2$ | 90,675 | 91,067 | 94,019 | 89,584 | 92,773 |
| | $R_1 + R_2$ | 197,212 | 194,741 | 198,632 | 196,260 | 200,703 |
| Std err | $R_1$ | 17,484 | 18,742 | 16,171 | 16,201 | 21,502 |
| | $R_2$ | 15,272 | 15,820 | 14,884 | 13,664 | 17,902 |
| | $R_1 + R_2$ | 23,465 | 28,376 | 26,301 | 24,673 | 31,333 |
| 99% VaR | $R_1$ | 158,177 | 166,187 | 152,930 | 154,612 | 172,161 |
| | $R_2$ | 135,667 | 135,924 | 134,931 | 127,850 | 147,734 |
| | $R_1 + R_2$ | 268,086 | 283,931 | 274,278 | 262,495 | 295,900 |

1. Simulating the vector of uniforms $(u_{1m}^{(s)}, u_{2m}^{(s)})'$ from the corresponding copula given the simulated copula parameter $\alpha^{(s)}$ and then
2. Producing the value of $\log y_{lm}^{(s)}$ via calculating $\boldsymbol{X}'_{lm}\boldsymbol{\beta}_l^{(s)} + \sigma_l^{(s)}\Phi^{-1}(u_{lm}^{(s)})$, where $\boldsymbol{\beta}_l^{(s)}$ and $\sigma_l^{(s)}$ are the $s$th simulated parameters, and $\boldsymbol{X}_{lm}$ is the design matrix corresponding to the $m$th out-of-sample data point.

In step 1 above, many algorithms are available for the purpose of simulating samples from the Archimedean copulas, for example, the algorithm that sequentially simulates values from the conditional distributions as described in Genest (1987).

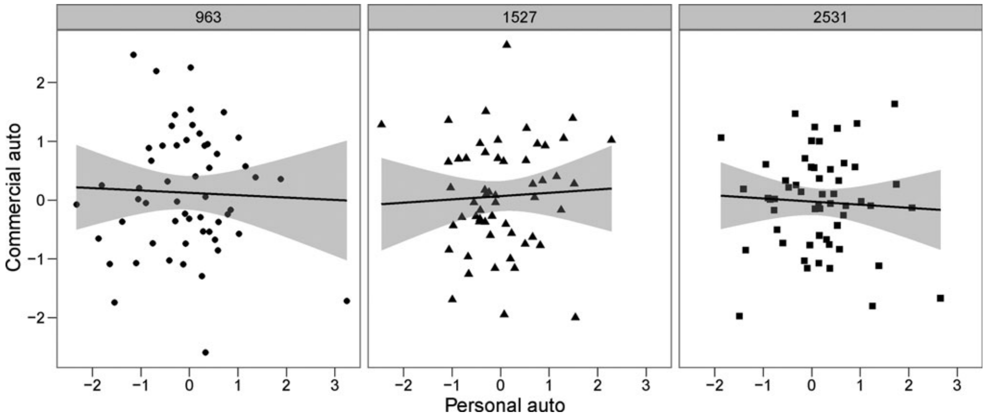
Implementation of the above simulation algorithm results in a sample of 3,000 predictive values for each of the 45 validation points, which can then be used to estimate the predictive loss distributions for loss reserves. That is, for the $l$th margin, the $s$th simulated loss reserve is

$$R_l^{(s)} = \sum_{m=1}^{45} y_{lm}^{(s)}, \tag{21}$$

where $R_l$ represents the loss reserve from the $l$th line. Given these simulated reserves, we calculate such summary statistics as the median, standard deviation, and 99 percent value at risk (VaR). These statistics are reported in Table 2 for each line of business and the portfolio. As a result of properly accounting for the positive correlation, we see that the copula models report larger uncertainty measures for the portfolio reserve ($R_1 + R_2$) than the independent case. The copula models also report larger VaR measures than the independent model except for the Frank copula model.

*Model Comparison.* The simulated predictive values can also be used to examine the model's predictive ability on the out-of-sample validation points. As we are mostly concerned with the model's performance in terms of predictive accuracy, we choose the absolute error as our loss function. Moreover, since the objective of a typical

**FIGURE 3**
The Scatterplot of Residuals From Commercial Auto Over Personal Auto for Three Randomly Selected Samples From the Gaussian Copula Model




*Note*: Also shown is a linear regression line along with standard errors.

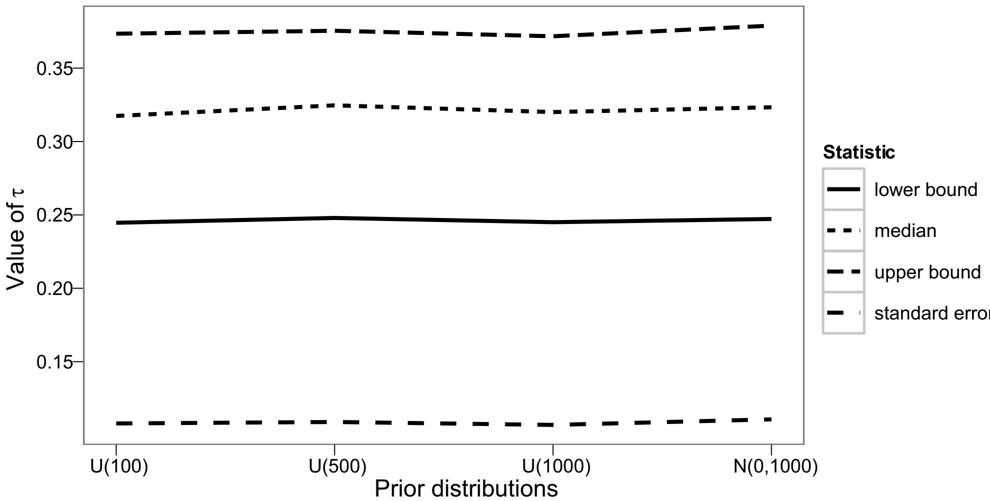
loss reserving analysis is to predict the outstanding loss amounts for the portfolio, we compute the absolute error on the portfolio level, that is, the absolute difference between the median of the predicted portfolio loss reserve and the realized portfolio paid amounts in the out-of-sample periods. The calculated loss statistics are 44,902, 48,806, 47,236, and 51,023 for the Clayton, Gumbel, Frank, and Gaussian copulas, respectively, compared to 47,953 for the independent model. We see that two of the copula models outperform the independent case, where the best performance model, the Clayton copula model, has prediction error about 6 percent lower than that from the independent model. In addition to producing more realistic uncertainty measures, an appropriate copula model also provides the opportunity to improve the predictive performance over the independent model, although, in general, such improvement is not guaranteed.

*Residuals Revisited.* We have seen that the copula model has the potential to enhance predictive accuracy; however, we still need to investigate whether the copula models correctly reflect the residual correlation. We examine the residual relationship for the Gaussian copula, where the standardized residuals can be readily calculated as

$$(r_{1i}, r_{2i})' = \boldsymbol{B}^{-1} \cdot (\log y_{1i} - \boldsymbol{X}_{1i}'\boldsymbol{\beta}_1, \log y_{2i} - \boldsymbol{X}_{2i}'\boldsymbol{\beta}_2)'. \tag{22}$$

In the above, $r_{li}$ denotes the $i$th standardized residual in the $l$th line, and $\boldsymbol{B}$ is the Cholesky factor of $\boldsymbol{\Sigma}$ such that $\boldsymbol{\Sigma} = \boldsymbol{B}\boldsymbol{B}'$. These pairs of standardized residuals are then plotted against each other in Figure 3 to check for positive relationships. As this is a Bayesian analysis, thousands of residuals will be generated. The plot in Figure 3 shows three randomly selected realizations. We see that the strong positive relationship as in Figure 1 has disappeared, indicating that the model correctly captures the contemporaneous correlations.

**FIGURE 4**
The Median, Standard Deviation, and Bounds of the 50% Credible Interval of $\tau$ in Different Settings of Prior Distributions




*Sensitivity Analysis.* In Bayesian analysis, it is typically the case that a multitude of prior distributions is available, and we need to assess the sensitivity of the Bayesian posterior estimates to various prior distribution assumptions. We do this for the Frank copula. Specifically, we examine another three alternative priors $U(-100, 100)$, $U(-500, 500)$, and $N(0, 10^3)$ for the parameter $\tau$ and see if changing the bounds of the uniform prior or the form of the prior distribution will affect the posterior estimates of $\tau$. The posterior median, standard deviation, and lower and upper bounds of the 50 percent intervals of $\tau$ using different priors are plotted in Figure 4, along with the one using the present prior. We see that the posterior distribution of $\tau$ is fairly consistent across these different scenarios of prior distributions, indicating that our estimate is robust to the alternative relatively vague prior specifications.

## Example II: Gaussian Copula With Semiparametric Regression and Nonlinear Model

The previous example demonstrates that a proper bivariate copula model that accounts for correlations can improve upon the independent counterpart in terms of model adequacy and prediction accuracy. In this section, we will consider an example where three insurance lines are jointly modeled, each marginal chosen from a particular class of the standard loss reserving models that would reflect the inherent nature of the loss reporting process unique to each line. One such feature of particular interest here is the tail of the insurance line, which is the length of time period required for all claims to proceed to final settlement. In the case of "short-tailed" insurance products such as property damage or motor collision, the majority of claims are settled within the first several years. As such, it suffices to estimate the loss reserve for each accident year, assuming that there is no further claims reported beyond the range of the observed data (10 years), as in the generalized linear models in the previous example. In general, however, the loss payments can extend well beyond this observed range to some future time. It is therefore necessary to employ models such as penalized splines or nonlinear curves that enable predictions of multiple years ahead.

Further, as opposed to the previous example, only the Gaussian copula is implemented here, mainly because the Gaussian copula admits flexible correlation structures and when combined with Normal marginals, it leads to a multivariate Normal distribution for which the posterior simulation could be much more efficient.

Specifically, a third line, the workers' compensation insurance, is added to the previous example. Thus, we have $L = 3$ and the $i$th observation is now a $3 \times 1$ vector $(y_{1i}, y_{2i}, y_{3i})'$. The model structures chosen for $y_1$ through $y_3$ are the semiparametric spline regression, the generalized linear model, and the nonlinear growth curve model, respectively. Besides the heuristic reason of illustrating the flexibility of the copula model, our choice of incorporating three different forms of marginal models is also determined via inspecting the pattern in the data, for example, looking at the reported losses in the 10th lag in Table A2 or the last column in its triangular representation such as Table A1. While it is reasonably safe to assume that there is no tail development in commercial auto, hence the justification of maintaining the use of the generalized linear model, the tails must be explicitly accounted for in the other two lines. The choice of semiparametric regression for personal auto and growth curve for workers' compensation is largely based on the duration of the tail development. For workers' compensation that has long-lasting claim reporting periods, the parametric nonlinear model may be preferred to the penalized spline model that, due to the heavy reliance of extrapolation on the inferred pattern from the data, is likely to produce unrealistic predictions for multiple years ahead. On the other hand, for the personal auto insurance that is shorter tailed, the implicit assumption of the parametric nonlinear model that loss reporting will only terminate at an infinite unknown time point is not intuitively appealing, so we adopt the semiparametric regression model. The choice of the marginal models depends largely on the analyst's judgment besides the tail behavior of the type of claims under consideration, and one might prefer a different combination from the above, for example, the use of consistent model structures such as penalized splines across the three lines. In that case, one can formalize the selection process and choose the best model according to the error of predictions criterion as in "Model Comparison" section. This point is not pursued further here, and in the following, we will proceed with the above marginal models.

Similar to the bivariate Gaussian copula model, the use of lognormal marginals results in a multivariate normal distribution on the logarithmic scale, that is,

$$(\log y_{1i}, \log y_{2i}, \log y_{3i})' \sim N_3(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}), \tag{23}$$

where $\boldsymbol{\mu}_i = (\mu_{1i}, \mu_{2i}, \mu_{3i})'$ is a vector composed of the $i$th (log) mean for each line, and $\boldsymbol{\Sigma}$ is a $3 \times 3$ covariance matrix. We assign the prior distribution of $\boldsymbol{\Sigma}$ to be $\boldsymbol{\Sigma} \sim$ Inv-Wishart$_4(\boldsymbol{I})$ similar to the bivariate case. In the following, we specify the mean model and prior distributions for the associated parameters.

*Personal Auto Insurance.* As in Equation (10), we specify only one smoother in the mean structure using the development lag ($x_2$ in Table A2), allowing a smooth estimation of the development year effect. Identity link is used. Four knots are chosen and the positions of the knots are set at $\kappa_1 = 2, \kappa_2 = 4, \kappa_3 = 6,$ and $\kappa_4 = 8$. An additional accident year effect that varies randomly with mean 0 is also specified in the mean model, allowing the estimate of a particular year to deviate from the overall mean. Thus, this leads to

$$\mu_{1i} = \boldsymbol{X}_{1i}'\boldsymbol{\beta}_1 + \boldsymbol{Z}_i'\boldsymbol{b} + a_k, \quad \text{with } k = x_{1i}, \quad \text{for } i = 1, \ldots, 55, \tag{24}$$

where $\boldsymbol{X}_{1i} = (1, x_{2i})'$, $\boldsymbol{\beta}_1 = (\beta_{11}, \beta_{12})'$, $\boldsymbol{b} = (b_1, \ldots, b_4)'$, and $\boldsymbol{Z}$ constructed in the same fashion as in Equation (12). Normal distributions are used for $a_k$, the deviation of the $k$th year from the grand mean, and for $b_l$, the coefficient associated with the $l$th knot:

$$\begin{aligned} a_k &\sim N\big(0, \sigma_\alpha^2\big), \quad \text{for } k = 1, \ldots, 10, \\ b_l &\sim N\big(0, \sigma_b^2\big), \quad \text{for } l = 1, \ldots, 4. \end{aligned}$$

We specify noninformative priors: for example, $\beta_{1j} \sim N(0, 10^6)$ for $j = 1, 2$, $\sigma_a^2 \sim$ Inv-Gamma$(10^{-3}, 10^{-3})$, and $\sigma_b^2 \sim$ Inv-Gamma$(10^{-3}, 10^{-3})$.

*Commercial Auto Insurance.* The model in the commercial auto line remains unchanged from the previous example, that is,

$$\mu_{2i} = \boldsymbol{X}_{2i}'\boldsymbol{\beta}_2, \quad \text{for } i = 1, \ldots, 55, \tag{25}$$

where $\boldsymbol{X}_{2i}$ is the design matrix generated by including the accident year and the development the lag as categorical covariates. Noninformative priors are used for the $\beta$s as $\beta_{2j} \sim N(0, 10^6)$ for $j = 1, \ldots, 20$.

*Workers' Compensation Insurance.* The nonlinear growth curve model is implemented here, where the parametric form of the growth curve is chosen to be the log-logistic curve:

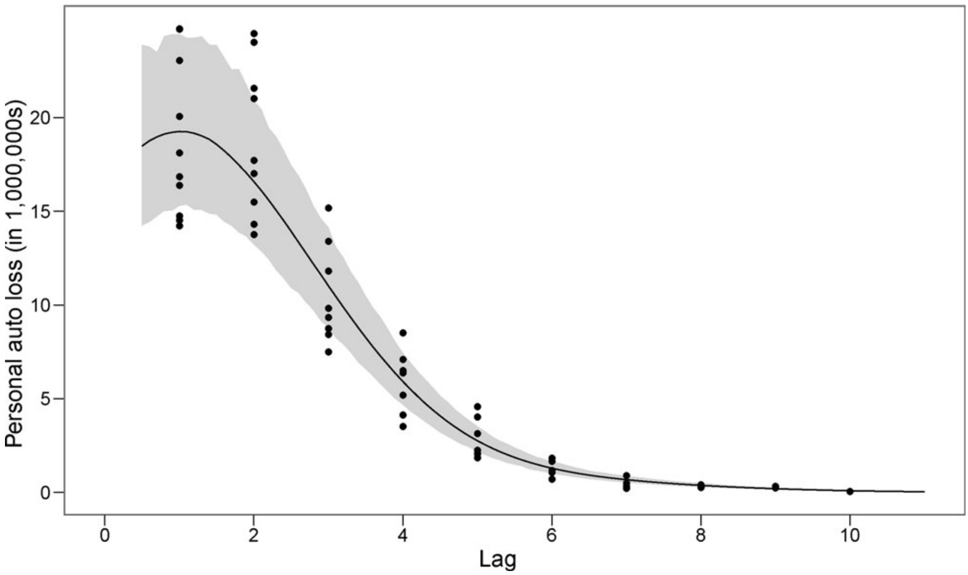
$$G(t_i; \omega, \theta) = \frac{t_i^\omega}{t_i^\omega + \theta^\omega}, \; \omega > 0, \theta > 0. \tag{26}$$

Also, as in Zhang, Dukic, and Guszcza (2012), the expected ultimate loss ratios are allowed to vary by accident year, which leads to the following mean model:

$$\exp(\mu_{3i}) = p_{3i} \cdot \gamma_k \cdot [G(t_i; \omega, \theta) - G(t_i - 12; \omega, \theta)] \quad \text{with } k = x_{1i}, \tag{27}$$

$$\log \gamma_k \sim N\big(\log \gamma_0, \sigma_\gamma^2\big), \quad \text{for } k = 1, \ldots, 10. \tag{28}$$

**FIGURE 5**
The Plot of the Company-Level Smoother Along With the 50% Prediction Interval for Personal Auto



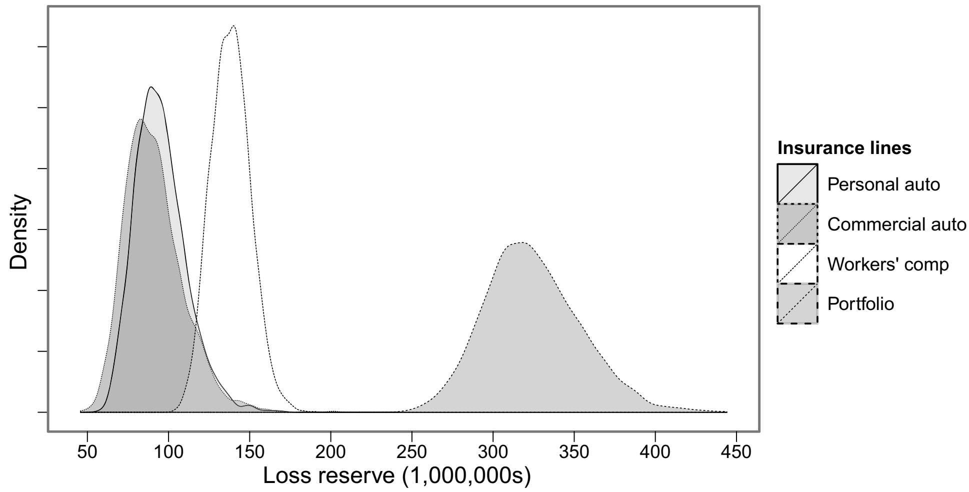

Since our model is on the logarithmic scale, we make the transformation $\exp(\mu_{3i})$ such that the losses on the original scale approximately follow a nonlinear growth curve. We also have $t_i = x_{2i} \cdot 12$ and $\Delta t_i = 12$ for $i = 1, \ldots, 55$, as the data in schedule P are reported annually. The expected ultimate loss ratios $\gamma_k$ are modeled on the logarithmic scale since they must be positive. The parameter $\gamma_0$ measures the company-level average ultimate loss ratio across all years, and $\sigma_\gamma$ corresponds to the variation of accident-year-level ultimate loss ratios on the logarithmic scale.

We specify noninformative priors for the parameters, for example, $\log \gamma_0 \sim N(0, 10^6), \log \omega \sim N(0, 10^6), \log \theta \sim N(0, 10^6)$, and $\sigma_\gamma^2 \sim$ Inv-Gamma $(10^{-3}, 10^{-3})$.

*Inference.* Given the above probability model, MCMC algorithms are employed to draw posterior samples for the parameters in the model. We ran 300,000 iterations in three parallel chains, discarding the burn-in period of the first 200,000 iterations and use every 100th iteration of each chain. This resulted in 1,000 simulation draws per chain. Based on these simulation draws, we compute the posterior median and the standard deviation for the parameters in the model, as shown in Table A4 in the Appendix.

Several interesting statistics can be derived from these estimates. First, under the estimated $\beta$s and $b$s in personal auto, it can be shown that the expected incremental loss, approximated by the smoother, is a decreasing function of $x_2$ when $x_2 > 10$ (this is more easily seen in Figure 5). This allows us to approximate the time required for all claims to be settled, using the time at which the projected incremental loss

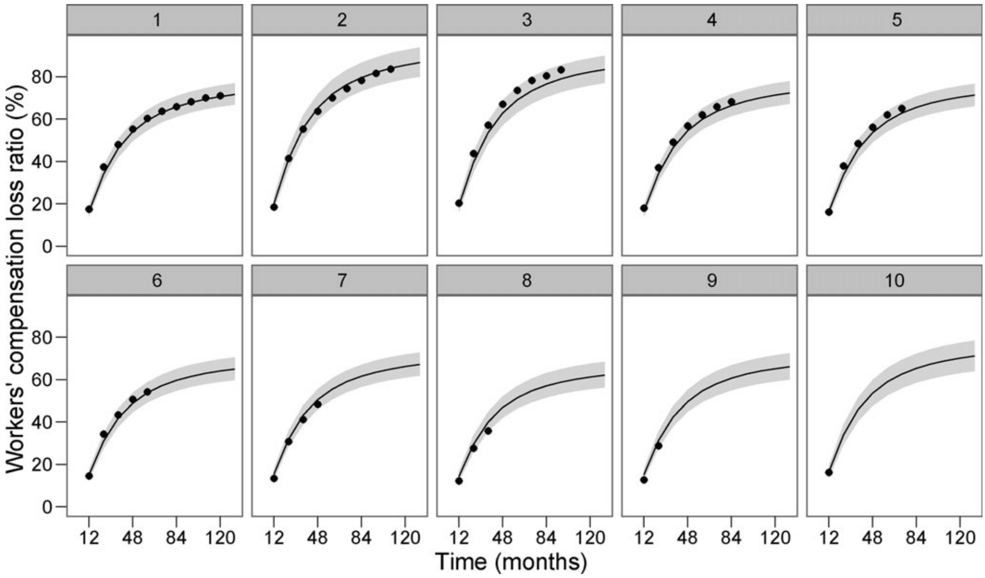
**FIGURE 6**
Predictive Distribution of the Loss Reserves for Each Line of Business and the Portfolio




falls below a given threshold, say $y_1 < 5$. Such a calculation shows that the average claims settlement period in personal auto is about 13 years for this company. Second, read directly from the table, the average loss ratio for workers' compensation is about 77.0 percent, and the variation of the loss ratios across different accident years approximately corresponds to $0.770 \times 0.135 = 0.104$ on the original scale using the delta method. Third, the average emergence percentage of the ultimate claim as of 10 years (120 months) in workers' compensation is about 89.8 percent using the log-logistic growth curve and using the median estimate of $\omega$ and $\theta$. Lastly, estimate of correlations between the three lines can be derived using the estimate of the covariance matrix. For example, the Kendall's correlation coefficient between personal auto and commercial auto is $\tau_{12} = \frac{2}{\pi} \arcsin(\frac{\boldsymbol{\Sigma}_{12}}{\sqrt{\boldsymbol{\Sigma}_{11} \cdot \boldsymbol{\Sigma}_{22}}}) = 0.184$. Similar calculations lead to the estimates $\tau_{13} = 0.088$ and $\tau_{23} = 0.0012$, the Kendall's correlation coefficient between personal auto and workers' compensation, and between commercial auto and workers' compensation, respectively.

*Predictive Distribution, Loss Reserve Calculation, and Model Comparison.* The predictive simulation in the Gaussian copula model here is much more straightforward, compared to the Archimedean copula models. To proceed, one computes the posterior mean $\boldsymbol{\mu}_m^{(s)}$ for each of the validation data points $m = 1, \ldots, 45$ and for each of the simulated sample $s = 1, \ldots 3{,}000$, then the logarithm of the predicted values can be drawn from a multivariate normal distribution with mean $\boldsymbol{\mu}_m^{(s)}$ and covariance matrix $\boldsymbol{\Sigma}^{(s)}$. Then, the loss reserve for the $l$th line and the $s$th sample is just $\sum_{m=1}^{45} y_{lm}^{(s)}$. The estimates for the loss reserves as of 10 years are 93,544 (15,702), 88,959 (17,891), and 138,082 (12,060) for personal auto, commercial auto, and workers' compensation, respectively, where the numbers in the parentheses are the standard errors. The

**Figure 7**
Lattice Plot of the Workers' Compensation Data (Cumulative Loss Ratios) by Accident Year, Together With the Fitted Curve and 50% Prediction Intervals




*Note*: Panel labels indicate accident years.

corresponding posterior distributions for each line and the portfolio are shown in Figure 6. We see that due to the use of the lognormal distribution, all these distributions are skewed to the right, and commercial auto is more skewed than personal auto. Further, as the semiparametric model and the growth curve model allow predictions of loss reserves beyond 10 years, we also calculate the estimated loss reserves as of 13 years for personal auto and those as of 20 years for workers' compensation for illustrating purpose. The estimates are 94,447 (16,183) and 180,169 (15,333), respectively.

Using the validation data points, the loss statistics on the portfolio level is 32,901, resulting in a substantial reduction (about 25 percent) in prediction error compared to the Clayton copula model, the best performance model in the bivariate case. A closer look at the model's performance on individual lines of business reveals that this substantial improvement is mainly due to the replacement of the generalized linear model with the hierarchical semiparametric model in the personal auto insurance. Thus, one might also want to examine how changing the model structure for commercial auto affects the prediction accuracy. The company-level estimate for the median smoother along with the 50 percent prediction intervals is plotted in Figure 5. We see that overall trend in the data is sufficiently captured by the smoother with four knots, indicating that the previous use of generalized linear model is likely to be overparameterized and consequently has inferior predictive performance. For the workers' compensation insurance, we evaluate the goodness of fit of the produced curve based on the cumulative loss ratio, that is, the sum of incremental losses divided by the accident year premium. The fitted curve and the corresponding 50 percent prediction interval for each year are plotted together with the observed data in Figure 7. We see that the growth curves describe the observed patterns fairly well overall, although it slightly underestimates accident year 3 where the data fall on the upper edge of the 50 percent intervals.

## CONCLUSIONS

The goal of this article is to introduce Bayesian copula modeling into the field of multivariate loss reserving. Supported via the Bayesian computational machinery, the complexity of models that can be entertained with this framework is vast, as we show via two examples, where we present the Bayesian copula model, computational scheme, model checking and comparison procedures, as well as posterior predictive inference. Since in this article, our emphasis is on statistical inference of the assumed model rather than model selection, we only conduct a simple yet intuitive exercise to compare models using portfolio-level out-of-sample validation information. In the case where model selection is the primary interest, the deviance-based model selection criteria can be applied to the Bayesian copula models (see Silva and Lopes, 2008).

In our data analysis, we take relatively simple but still reasonable marginal models because our main focus in this article is on the modeling of contemporaneous correlations and the investigation of whether multivariate copula models could improve on their independent counterparts. As a result, the models we considered may not be "optimal" ones—in specifying our marginal models, we have left out the examination of other potentially useful factors such as inflation, calendar year effect, and so on. When necessary, one should examine such additional factors in the choice of appropriate marginals to be used in the copula model and use the Bayesian computational scheme to make inference. In general, we advocate the constructive approach, where simpler models are the starting point, and additional complexity is added in stages as necessary. As another result, we do not compare our copula results to other existing methods such as the (multivariate) chain ladder. The copula model may or may not outperform the (multivariate) chain ladder model, depending, to some extent, on whether the marginal model we used is better or worse. This comparison would not provide evidence for the use of copula one way or the other.

It is our hope that through the detailed examples, we have been able to elucidate the power and feasibility of Bayesian copula modeling for multivariate loss reserving. It is also our hope that the proposed Bayesian copula framework would be adopted to other insurance fields such as economic capital evaluation and enterprise risk management, where there is often great need for both account of correlation and generation of posterior risk distributions. Indeed, as the financial credit crisis highlighted the importance of uncertainty measures, there has been growing interest in the financial service field on the use of Bayesian framework as more effective risk management tools. The combination of Bayesian methods and copula models thus provides a natural way to better understand the uncertainty of risk profiles in dependence modeling. Readers interested at such topics would find helpful the books by Shevchenko (2011) and Böcker (2010).

## APPENDIX

**TABLE A1**

Triangular Representation of the Incremental Paid Losses (in 1,000s) in Commercial Auto From Table 1

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 5,407 | 9,015 | 4,641 | 3,384 | 1,695 | 1,262 | 1,425 | 373 | 241 | 6 |
| 2 | 6,279 | 8,752 | 6,172 | 4,494 | 2,110 | 919 | 447 | 202 | 69 | |
| 3 | 7,256 | 8,667 | 4,778 | 4,262 | 2,884 | 1,427 | 889 | 493 | | |
| 4 | 5,028 | 5,317 | 4,697 | 3,795 | 2,871 | 1,100 | 657 | | | |
| 5 | 5,712 | 6,097 | 6,389 | 3,802 | 4,306 | 862 | | | | |
| 6 | 7,413 | 9,385 | 7,772 | 5,850 | 3,383 | | | | | |
| 7 | 10,868 | 12,337 | 7,966 | 8,531 | | | | | | |
| 8 | 10,143 | 14,193 | 8,070 | | | | | | | |
| 9 | 9,596 | 12,235 | | | | | | | | |
| 10 | 9,076 | | | | | | | | | |

**TABLE A2**

Modeling Data Set in the Study

| $x_1$ (Year) | $x_2$ (Lag) | $y_1$ | $y_2$ | $y_3$ | $p_1$ | $p_2$ | $p_3$ |
|---|---|---|---|---|---|---|---|
| 1 | 1 | 16,864 | 5,407 | 10,167 | 62,467 | 42,874 | 58,278 |
| 1 | 2 | 15,508 | 9,015 | 11,583 | 62,467 | 42,874 | 58,278 |
| 1 | 3 | 9,341 | 4,641 | 6,217 | 62,467 | 42,874 | 58,278 |
| 1 | 4 | 3,537 | 3,384 | 4,305 | 62,467 | 42,874 | 58,278 |
| 1 | 5 | 1,853 | 1,695 | 2,912 | 62,467 | 42,874 | 58,278 |
| 1 | 6 | 1,184 | 1,262 | 1,920 | 62,467 | 42,874 | 58,278 |
| 1 | 7 | 500 | 1,425 | 1,322 | 62,467 | 42,874 | 58,278 |
| 1 | 8 | 308 | 373 | 1,316 | 62,467 | 42,874 | 58,278 |
| 1 | 9 | 338 | 241 | 1,089 | 62,467 | 42,874 | 58,278 |
| 1 | 10 | 50 | 6 | 600 | 62,467 | 42,874 | 58,278 |
| 2 | 1 | 14,528 | 6,279 | 11,123 | 59,821 | 38,829 | 60,268 |
| 2 | 2 | 17,727 | 8,752 | 13,884 | 59,821 | 38,829 | 60,268 |
| 2 | 3 | 8,747 | 6,172 | 8,349 | 59,821 | 38,829 | 60,268 |
| 2 | 4 | 4,149 | 4,494 | 5,070 | 59,821 | 38,829 | 60,268 |
| 2 | 5 | 2,252 | 2,110 | 3,766 | 59,821 | 38,829 | 60,268 |
| 2 | 6 | 715 | 919 | 2,657 | 59,821 | 38,829 | 60,268 |
| 2 | 7 | 325 | 447 | 2,303 | 59,821 | 38,829 | 60,268 |
| 2 | 8 | 261 | 202 | 2,002 | 59,821 | 38,829 | 60,268 |
| 2 | 9 | 255 | 69 | 1,226 | 59,821 | 38,829 | 60,268 |
| 3 | 1 | 14,241 | 7,256 | 13,924 | 62,968 | 43,001 | 68,462 |
| 3 | 2 | 13,763 | 8,667 | 16,065 | 62,968 | 43,001 | 68,462 |
| 3 | 3 | 7,512 | 4,778 | 9,102 | 62,968 | 43,001 | 68,462 |
| 3 | 4 | 5,207 | 4,262 | 6,767 | 62,968 | 43,001 | 68,462 |
| 3 | 5 | 2,068 | 2,884 | 4,515 | 62,968 | 43,001 | 68,462 |
| 3 | 6 | 1,674 | 1,427 | 3,211 | 62,968 | 43,001 | 68,462 |
| 3 | 7 | 219 | 889 | 1,467 | 62,968 | 43,001 | 68,462 |
| 3 | 8 | 421 | 493 | 2,008 | 62,968 | 43,001 | 68,462 |

*(Continued)*

**TABLE A2**
Continued

| $x_1$ (Year) | $x_2$ (Lag) | $y_1$ | $y_2$ | $y_3$ | $p_1$ | $p_2$ | $p_3$ |
|---|---|---|---|---|---|---|---|
| 4 | 1 | 14,765 | 5,028 | 12,468 | 64,453 | 41,840 | 69,254 |
| 4 | 2 | 14,323 | 5,317 | 13,171 | 64,453 | 41,840 | 69,254 |
| 4 | 3 | 8,426 | 4,697 | 8,346 | 64,453 | 41,840 | 69,254 |
| 4 | 4 | 6,513 | 3,795 | 5,383 | 64,453 | 41,840 | 69,254 |
| 4 | 5 | 3,144 | 2,871 | 3,594 | 64,453 | 41,840 | 69,254 |
| 4 | 6 | 1,067 | 1,100 | 2,636 | 64,453 | 41,840 | 69,254 |
| 4 | 7 | 913 | 657 | 1,596 | 64,453 | 41,840 | 69,254 |
| 5 | 1 | 16,395 | 5,712 | 12,626 | 71,185 | 44,525 | 78,387 |
| 5 | 2 | 17,038 | 6,097 | 17,119 | 71,185 | 44,525 | 78,387 |
| 5 | 3 | 9,826 | 6,389 | 8,244 | 71,185 | 44,525 | 78,387 |
| 5 | 4 | 6,381 | 3,802 | 6,045 | 71,185 | 44,525 | 78,387 |
| 5 | 5 | 4,037 | 4,306 | 4,519 | 71,185 | 44,525 | 78,387 |
| 5 | 6 | 1,839 | 862 | 2,383 | 71,185 | 44,525 | 78,387 |
| 6 | 1 | 18,136 | 7,413 | 14,694 | 82,793 | 50,923 | 100,705 |
| 6 | 2 | 21,582 | 9,385 | 19,850 | 82,793 | 50,923 | 100,705 |
| 6 | 3 | 13,415 | 7,772 | 9,053 | 82,793 | 50,923 | 100,705 |
| 6 | 4 | 8,519 | 5,850 | 7,520 | 82,793 | 50,923 | 100,705 |
| 6 | 4 | 8,519 | 5,850 | 7,520 | 82,793 | 50,923 | 100,705 |
| 6 | 5 | 4,583 | 3,383 | 3,457 | 82,793 | 50,923 | 100,705 |
| 7 | 1 | 24,727 | 10,868 | 16,503 | 100,826 | 56,601 | 123,551 |
| 7 | 2 | 24,037 | 12,337 | 21,475 | 100,826 | 56,601 | 123,551 |
| 7 | 3 | 15,181 | 7,966 | 12,974 | 100,826 | 56,601 | 123,551 |
| 7 | 4 | 7,105 | 8,531 | 8,785 | 100,826 | 56,601 | 123,551 |
| 8 | 1 | 24,749 | 10,143 | 14,906 | 98,358 | 54,609 | 122,257 |
| 8 | 2 | 24,501 | 14,193 | 18,904 | 98,358 | 54,609 | 122,257 |
| 8 | 3 | 11,830 | 8,070 | 10,065 | 98,358 | 54,609 | 122,257 |
| 9 | 1 | 23,063 | 9,596 | 11,852 | 76,653 | 47,204 | 92,720 |
| 9 | 2 | 21,035 | 12,235 | 14,859 | 76,653 | 47,204 | 92,720 |
| 10 | 1 | 20,083 | 9,076 | 11,668 | 71,326 | 42,412 | 72,154 |

*Note*: $y$ and $p$ represent the incremental paid losses and accident year earned premium (in 1,000s), and 1–3 indicate the personal auto, commercial auto, and worker's compensation insurance, respectively.

**TABLE A3**
Estimated Parameters for the Bivariate Models Along With the Univariate Estimates

| Parameter | Independent | Clayton | Gumbel | Frank | Gaussian |
|---|---|---|---|---|---|
| $\beta_{1,1}$ | 9.555(0.120) | 9.626(0.133) | 9.552(0.114) | 9.557(0.113) | 9.559(0.141) |
| $\beta_{1,2}$ | −0.117(0.119) | −0.163(0.122) | −0.137(0.122) | −0.127(0.106) | −0.127(0.137) |
| $\beta_{1,3}$ | −0.028(0.122) | −0.115(0.133) | −0.015(0.117) | −0.015(0.125) | −0.027(0.142) |
| $\beta_{1,4}$ | 0.195(0.127) | 0.112(0.142) | 0.219(0.127) | 0.171(0.124) | 0.194(0.150) |
| $\beta_{1,5}$ | 0.296(0.134) | 0.233(0.141) | 0.314(0.136) | 0.295(0.140) | 0.301(0.158) |
| $\beta_{1,6}$ | 0.474(0.144) | 0.392(0.154) | 0.505(0.136) | 0.465(0.138) | 0.478(0.167) |
| $\beta_{1,7}$ | 0.534(0.157) | 0.454(0.165) | 0.552(0.151) | 0.508(0.158) | 0.530(0.182) |
| $\beta_{1,8}$ | 0.500(0.171) | 0.430(0.201) | 0.513(0.165) | 0.485(0.185) | 0.492(0.206) |

*(Continued)*

**TABLE A3**
Continued

| Parameter | Independent | Clayton | Gumbel | Frank | Gaussian |
|---|---|---|---|---|---|
| $\beta_{1,9}$ | 0.440(0.203) | 0.326(0.225) | 0.432(0.205) | 0.452(0.201) | 0.445(0.237) |
| $\beta_{1,10}$ | 0.362(0.273) | 0.264(0.287) | 0.313(0.264) | 0.346(0.253) | 0.344(0.323) |
| $\beta_{1,11}$ | 0.014(0.114) | 0.029(0.115) | −0.005(0.113) | 0.024(0.111) | 0.011(0.137) |
| $\beta_{1,12}$ | −0.554(0.122) | −0.559(0.125) | −0.561(0.118) | −0.532(0.114) | −0.552(0.142) |
| $\beta_{1,13}$ | −1.102(0.126) | −1.116(0.135) | −1.103(0.127) | −1.095(0.122) | −1.100(0.153) |
| $\beta_{1,14}$ | −1.747(0.132) | −1.718(0.141) | −1.771(0.131) | −1.755(0.127) | −1.747(0.158) |
| $\beta_{1,15}$ | −2.507(0.142) | −2.584(0.153) | −2.481(0.136) | −2.508(0.150) | −2.518(0.167) |
| $\beta_{1,16}$ | −3.517(0.155) | −3.573(0.164) | −3.503(0.154) | −3.529(0.149) | −3.516(0.182) |
| $\beta_{1,17}$ | −3.725(0.172) | −3.732(0.174) | −3.731(0.173) | −3.743(0.159) | −3.735(0.201) |
| $\beta_{1,18}$ | −3.810(0.203) | −3.826(0.208) | −3.843(0.236) | −3.802(0.194) | −3.820(0.239) |
| $\beta_{1,19}$ | −5.640(0.272) | −5.741(0.296) | −5.671(0.238) | −5.640(0.243) | −5.638(0.321) |
| $\beta_{2,1}$ | 8.775(0.133) | 8.790(0.137) | 8.730(0.116) | 8.763(0.142) | 8.767(0.150) |
| $\beta_{2,2}$ | −0.269(0.128) | −0.293(0.133) | −0.273(0.125) | −0.254(0.138) | −0.269(0.149) |
| $\beta_{2,3}$ | 0.059(0.135) | 0.045(0.146) | 0.056(0.129) | 0.074(0.138) | 0.066(0.153) |
| $\beta_{2,4}$ | −0.203(0.142) | −0.207(0.147) | −0.155(0.131) | −0.182(0.147) | −0.192(0.161) |
| $\beta_{2,5}$ | −0.067(0.148) | −0.104(0.156) | 0.000(0.137) | −0.038(0.163) | −0.064(0.173) |
| $\beta_{2,6}$ | 0.208(0.158) | 0.184(0.167) | 0.249(0.151) | 0.208(0.159) | 0.212(0.184) |
| $\beta_{2,7}$ | 0.478(0.175) | 0.445(0.178) | 0.518(0.160) | 0.469(0.191) | 0.490(0.193) |
| $\beta_{2,8}$ | 0.446(0.190) | 0.448(0.208) | 0.483(0.186) | 0.460(0.195) | 0.458(0.218) |
| $\beta_{2,9}$ | 0.400(0.229) | 0.352(0.217) | 0.432(0.212) | 0.438(0.207) | 0.410(0.260) |
| $\beta_{2,10}$ | 0.345(0.306) | 0.316(0.292) | 0.412(0.292) | 0.320(0.292) | 0.353(0.347) |
| $\beta_{2,11}$ | 0.227(0.128) | 0.244(0.140) | 0.228(0.115) | 0.227(0.129) | 0.231(0.147) |
| $\beta_{2,12}$ | −0.131(0.137) | −0.142(0.146) | −0.104(0.119) | −0.126(0.136) | −0.131(0.155) |
| $\beta_{2,13}$ | −0.366(0.142) | −0.354(0.155) | −0.342(0.125) | −0.376(0.145) | −0.359(0.161) |
| $\beta_{2,14}$ | −0.812(0.151) | −0.792(0.154) | −0.784(0.138) | −0.831(0.161) | −0.804(0.170) |
| $\beta_{2,15}$ | −1.677(0.154) | −1.693(0.173) | −1.651(0.148) | −1.676(0.162) | −1.670(0.183) |
| $\beta_{2,16}$ | −2.012(0.176) | −1.997(0.186) | −1.931(0.164) | −2.021(0.175) | −2.007(0.197) |
| $\beta_{2,17}$ | −2.892(0.188) | −2.880(0.203) | −2.851(0.184) | −2.897(0.198) | −2.895(0.217) |
| $\beta_{2,18}$ | −3.780(0.226) | −3.867(0.237) | −3.703(0.178) | −3.771(0.237) | −3.777(0.253) |
| $\beta_{2,19}$ | −6.979(0.309) | −7.080(0.325) | −6.999(0.288) | −6.998(0.242) | −6.975(0.348) |
| $\sigma_1$ | 0.241(0.030) | 0.252(0.033) | 0.244(0.030) | 0.241(0.030) | 0.284(0.033) |
| $\sigma_2$ | 0.269(0.033) | 0.280(0.036) | 0.276(0.034) | 0.274(0.034) | 0.308(0.036) |
| $\alpha$ | | 0.627(0.436) | 1.365(0.210) | 3.149(1.288) | 0.268(0.144) |

*Note*: Standard deviation is reported in parentheses.

**TABLE A4**
Summary of the Bayesian Inference for the Trivariate Gaussian Copula Model

| Personal Auto | | Commercial Auto | | | | Workers' Comp | | Covariance | |
|---|---|---|---|---|---|---|---|---|---|
| $\beta_{11}$ | 11.620(0.989) | $\beta_{2,1}$ | 8.814(0.154) | $\beta_{2,11}$ | 0.195(0.154) | $\omega$ | 1.510(0.067) | $\boldsymbol{\Sigma}_{11}$ | 0.109 |
| $\beta_{12}$ | −0.752(0.194) | $\beta_{2,2}$ | −0.252(0.145) | $\beta_{2,12}$ | −0.135(0.150) | $\theta$ | 28.490(1.334) | $\boldsymbol{\Sigma}_{12}$ | 0.029 |
| $b_1$ | −0.038(0.033) | $\beta_{2,3}$ | 0.059(0.154) | $\beta_{2,13}$ | −0.375(0.160) | $\gamma_0$ | 0.770(0.047) | $\boldsymbol{\Sigma}_{13}$ | 0.009 |
| $b_2$ | 0.055(0.049) | $\beta_{2,4}$ | −0.222(0.159) | $\beta_{2,14}$ | −0.855(0.174) | $\sigma_\gamma$ | 0.135(0.061) | $\boldsymbol{\Sigma}_{22}$ | 0.098 |
| $b_3$ | −0.005(0.046) | $\beta_{2,5}$ | −0.092(0.170) | $\beta_{2,15}$ | −1.724(0.179) | | | $\boldsymbol{\Sigma}_{23}$ | 1.2E − 4 |

*(Continued)*

**TABLE A4**
Continued

| Personal Auto | | Commercial Auto | | | | Workers' Comp | Covariance | |
|---|---|---|---|---|---|---|---|---|
| $b_4$ | 0.025(0.033) | $\beta_{2,6}$ | 0.154(0.189) | $\beta_{2,16}$ | −1.961(0.200) | | $\Sigma_{33}$ | 0.043 |
| $\sigma_a$ | 0.182(0.092) | $\beta_{2,7}$ | 0.406(0.197) | $\beta_{2,17}$ | −2.934(0.224) | | | |
| $\sigma_b$ | 0.059(0.044) | $\beta_{2,8}$ | 0.391(0.221) | $\beta_{2,18}$ | −3.974(0.283) | | | |
| | | $\beta_{2,9}$ | 0.345(0.259) | $\beta_{2,19}$ | −6.914(0.347) | | | |
| | | $\beta_{2,10}$ | 0.288(0.343) | | | | | |

*Note*: Numbers in parentheses are posterior estimate of standard errors.

## REFERENCES


Antonio, K., and J. Beirlant, 2008, Issues in Claims Reserving and Credibility: A Semiparametric Approach With Mixed Models, *Journal of Risk and Insurance*, 75(3): 643-676.

Böcker, K., (ed.) 2010, *Rethinking Risk Measurement and Reporting—Volumes I and II Uncertainty, Bayesian Analysis and Expert Judgement* (London: Risk Books).

Braun, C., 2004, The Prediction Error of the Chain Ladder Method Applied to Correlated Run-Off Triangles, *ASTIN Bulletin*, 34(2): 399-423.

Bühlmann, H., 1987, Actuaries of the Third Kind? *ASTIN Bulletin*, 17(2): 137–138.

Casualty Actuarial Society (CAS)—Enterprise Risk Management Committee, 2003, Overview of Enterprise Risk Management. World Wide Web: http://www.casact.org/research/erm/overview.pdf.

Clark, D. R., 2003, LDF Curve-Fitting and Stochastic Reserving: A Maximum Likelihood Approach. *Casualty Actuarial Society Forum*. World Wide Web: http://www.casact.org /pubs/forum/03fforum/03ff041.pdf.

Cleveland, W. S., 1979, Robust Locally Weighted Regression and Smoothing Scatterplots, *Journal of the American Statistical Association*, 74(368): 829-836.

Crainiceanu, C. M., D. Ruppert, and M. P. Wand, 2005, Bayesian Analysis for Penalized Spline Regression Using WinBUGS, *Journal of Statistical Software*, 14: 1-24.

Dall'Aglio, G., S. Kotz, and G. Salinetti, 1991, *Advances in Probability Distributions With Given Marginals: Beyond the Copulas* (Rome, Italy: Universitá degli studi di Roma La Sapienza).

de Alba, E., 2002, Bayesian Estimation of Outstanding Claims Reserves, *North American Actuarial Journal*, 6(4): 1-20.

de Alba, E., and L. E. Nieto-Barajas, 2008, Claims Reserving: A Correlated Bayesian Model, *Insurance: Mathematics and Economics*, 43: 368-376.

de Jong, P., 2012, Modeling Dependence Between Loss Triangles, *North American Actuarial Journal*, 15(1): 74-86.

England, P. D., and R. J. Verrall, 2002, Stochastic Claims Reserving in General Insurance (With Discussion), *British Actuarial Journal*, 8: 443-544.

Fang, K. T., S. Kotz, and K. W. Ng, 1990, *Symmetric Multivariate and Related Distributions*, (London: Chapman and Hall).

Frees, E. W., and E. A. Valdez, 1998, Understanding Relationships Using Copulas, *North American Actuarial Journal*, 2(1): 1-25.

Frees, E. W., and E. A. Valdez, 2008, Hierarchical Insurance Claims Modeling, *Journal of the American Statistical Association*, 103(484): 1457-1469.

Frees, E. W., V. R. Young, and Y. Luo, 1999, A Longitudinal Data Analysis Interpretation of Credibility Models, *Insurance: Mathematics and Economics*, 24(3): 229-247.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin, 2004, *Bayesian Data Analysis* (London: Chapman and Hall).

Gelman, A., and D. B. Rubin, 1992, Inference From Iterative Simulation Using Multiple Sequences (With Discussion), *Statistical Science*, 7: 457-511.

Genest, C., 1987, Frank's Family of Bivariate Distributions, *Biometrika*, 74: 549-555.

Hastie, T., and R. Tibshirani, 1990, *Generalized Additive Models* (London: Chapman and Hall).

Kang, L., 2011, Asset Allocation in a Bayesian Copula Garch Framework: An Application to the "Passive Funds Versus Active Funds" Problem, *Journal of Asset Management*, 12: 45-66.

Kaufmann, R., A. Gadmer, and R. Klett, 2001, Introduction to Dynamic Financial Analysis, *ASTIN Bulletin*, 31(1): 213-249.

Lawrance, A., and P. Lewis, 1981, A New Autoregressive Time Series Model in Exponential Variables (NEAR(1)), *Advances in Applied Probability*, 13: 826-845.

Mack, T., 1993, Distribution-Free Calculation of the Standard Error, *ASTIN Bulletin*, 23(2): 213-225.

Makov, U. E., 2001, Principal Applications of Bayesian Methods in Actuarial Science: A Perspective, *North American Actuarial Journal*, 5(4): 53-56.

McCullagh, P., and J. Nelder, 1989, *Generalized Linear Models* (Boca Raton, FL: Chapman and Hall).

Merz, M., and M. Wüthrich, 2008a, *Stochastic Claims Reserving Methods in Insurance* (Hoboken, NJ: John Wiley & Sons).

Merz, M., and M. Wüthrich, 2008b, Prediction Error of the Multivariate Chain Ladder Reserving Method, *North American Actuarial Journal*, 12(2): 175-197.

Mildenhall, S., 2005, A Multivariate Bayesian Claim Count Development Model With Closed Form Posterior and Predictive Distributions, Preprint.

Nelsen, R. B., 2006, *An Introduction to Copulas* (New York: Springer).

Ntzoufras, I., and P. Dellaportas, 2002, Bayesian Modelling of Outstanding Liabilities Incorporating Claim Count Uncertainty, *North American Actuarial Journal*, 6(1): 113-128.

Peters, G. W., P. V. Shevchenko, and M. V. Wüthrich, 2009, Model Uncertainty in Claims Reserving Within Tweedie's Compound Poisson Models, *ASTIN Bulletin*, 39(1): 1-33.

Quarg, G., and T. Mack, 2004, Munich Chain Ladder, *Blatter DGVFM*, Band XXVI, 4: 597-630.

Robert, C., and G. Casella, 2005, *Monte Carlo Statistical Methods* (New York: Springer-Verlag).

Ruppert, D., M. Wand, and R. Carroll, 2003, *Semiparametric Regression* (Cambridge, UK: Cambridge University Press).

Rüschendorf, L., B. Schweizer, and M. Taylor, eds., 1996, Distributions With Fixed Marginals and Related Topics, *IMS Lecture Notes—Monograph Series*, Vol. 28. (Hayward, CA: Institute of Mathematical Statistics).

Schmeiser, B., and R. Lai, 1982, Bivariate Gamma Random Vectors, *Operations Research*, 30(2): 355-374.

Schweizer, B., and A. Sklar, 1986, Probability Distributions With Given Margins: Note on a Paper by Finch and Groblicki, *Foundations of Physics*, 16(10): 1061-1064.

Shemyakin, A. E., and H. Youn, 2006, Copula Models of Joint Last Survivor Analysis, *Applied Stochastic Models in Business and Industry*, 22(2): 211-224.

Shevchenko, P. V., 2011, *Modelling Operational Risk Using Bayesian Inference* (Berlin: Springer).

Shi, P., and E. W. Frees, 2011, Dependent Loss Reserving Using Copulas, *ASTIN Bulletin*, 41(2): 449-486.

Silva, R. S., and H. F. Lopes, 2008, Copula, Marginal Distributions and Model Selection: A Bayesian Note, *Statistics and Computing*, 18(3): 313-332.

Tang, A., and E. A. Valdez, 2009, Economic Capital and the Aggregation of Risks Using Copulas. World Wide Web: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1347675.

Verrall, R. J., 1996, Claims Reserving and Generalized Additive Models, *Insurance: Mathematics and Economics*, 19: 31-43.

Zhang, Y., 2010, A General Multivariate Chain Ladder Model, *Insurance: Mathematics and Economics*, 46(3): 588-599.

Zhang, Y., V. Dukic, and J. Guszcza, 2012, A Bayesian Nonlinear Model for Forecasting Insurance Loss Payments, *Journal of the Royal Statistical Society A*, 175: 637-656.

Zhao, X., and X. Zhou, 2010, Applying Copula Models to Individual Claim Loss Reserving Methods, *Insurance: Mathematics and Economics*, 46: 290-299.