North American Actuarial Journal

ISSN: 1092-0277 (Print) 2325-0453 (Online) Journal homepage: http://www.tandfonline.com/loi/uaaj20

Credibility Using Copulas

Edward W. Frees & Ping Wang Ph.D

To cite this article: Edward W. Frees & Ping Wang Ph.D (2005) Credibility Using Copulas, North American Actuarial Journal, 9:2, 31-48, DOI: 10.1080/10920277.2005.10596196

To link to this article: http://dx.doi.org/10.1080/10920277.2005.10596196

Credibility Using Copulas

Edward W. Frees* and Ping Wang†

Abstract

Credibility is a form of insurance pricing that is widely used, particularly in North America. The theory of credibility has been called a ''cornerstone'' in the field of actuarial science. Students of the North American actuarial bodies also study loss distributions, the process of statistical inference of relating a set of data to a theoretical (loss) distribution. In this work, we develop a direct link between credibility and loss distributions through the notion of a copula, a tool for understanding relationships among multivariate outcomes.

This paper develops credibility using a longitudinal data framework. In a longitudinal data framework, one might encounter data from a cross section of risk classes (towns) with a history of insurance claims available for each risk class. For the marginal claims distributions, we use generalized linear models, an extension of linear regression that also encompasses Weibull and Gamma regressions. Copulas are used to model the dependencies over time; specifically, this paper is the first to propose using a *t*-copula in the context of generalized linear models. The *t*-copula is the copula associated with the multivariate *t*-distribution; like the univariate *t*-distributions, it seems especially suitable for empirical work. Moreover, we show that the *t*-copula gives rise to easily computable predictive distributions that we use to generate credibility predictors. Like Bayesian methods, our copula credibility prediction methods allow us to provide an entire distribution of predicted claims, not just a point prediction.

We present an illustrative example of Massachusetts automobile claims, and compare our new credibility estimates with those currently existing in the literature.

1. Introduction

Credibility rate making is a technique for predicting future expected claims of a risk class, given past claims of that and related risk classes. This technique has a long history in actuarial science, with fundamental contributions dating back to Mowbray (1914). Whitney (1918) introduced the intuitively appealing concept of using a weighted average of (1) average claims from the risk class and (2) average claims over all risk classes to predict future expected claims. The weight associated with the risk class under consideration is known as the credibility factor.

In part, credibility predictors succeed in practice because they are intuitively appealing. By expressing the predictors as weighted averages, credibility predictors are straightforward in their explanation of actuarial merchandise to consumers. The American Academy of Actuaries published an important piece of evidence on this in the discussion of credibility applications in the Actuarial Standard of Practice Number 25.

In part, credibility predictors succeed because they are known to be the best possible predictors in a broad variety of situations. Bühlmann (1967) described a fundamental model containing latent (unobserved) effects that are common to all claims from a risk class; Bühlmann called these ''structure effects.'' The ''best'' linear unbiased predictors that can be derived from this model turn out to be the

* Edward W. Frees is a Professor with the School of Business, University of Wisconsin, 975 University Avenue, Madison, WI 53706, jfrees@bus.wisc.edu.

† Ping Wang is a Ph.D candidate, Actuarial Science, Risk Management and Insurance, University of Wisconsin-Madison, 3164 Grainger, Madison, WI 53706, pwang4@wisc.edu.

same intuitively appealing linear credibility predictors described above. Bühlmann's basic model formulation extends readily to encompass a large class of models. See Frees, Young, and Luo (1999) for a review that is oriented towards linear regression and longitudinal data models.

In Bühlmann's model formulation, the descriptor ''best'' means minimum mean square error. Although minimizing a mean square error has proven to be very useful in applied statistics, it is well-known that it may not be appropriate for skewed or long-tailed distributions such as commonly encountered in insurance claims analysis.

To account for the entire distribution of claims, a common approach used in credibility is to adopt a Bayesian perspective. Keffer (1929) initially suggested using a Bayesian perspective for experience rating in the context of group life insurance. Subsequently, Bailey (1945, 1950) showed how to derive the linear credibility form from a Bayesian perspective as the mean of a predictive distribution. Several authors have provided useful extensions of this paradigm. Jewell (1974) extended Bailey's results to a broader class of distributions, the natural exponential family, with conjugate prior distributions for the structure variables. Klugman (1992) investigated normal linear hierarchical models; this restricts the class of distributions but allows the analyst to include covariate effects.

In addition to the works cited above, we also note the work of Miller and Hickman (1975) and Pinquet (1997). Miller and Hickman examined credibility in the context of aggregate loss distributions. Pinquet was also interested in automobile claims; he considered collision claims arising from two lines, at fault and no fault coverages. Both of these papers assumed parametric distributions for the claims number and amount distributions and used Bayesian techniques to develop estimators.

In the Bayesian framework, one can explicitly account for the distribution of claims conditional on the latent structure variable (sampling distribution), make a preliminary (prior) assumption about the distribution of the structure variable, and use the data to improve this preliminary assumption (and hence compute the posterior distribution). This new posterior distribution, together with the sampling distribution, can then be used to compute the predictive distribution of a new claim, given prior claims.

In this paper, we will give a frequentist version of a predictive distribution. With this distribution, we will be able to compute the mean, median, or any other measure to summarize the distribution; thus, this aspect is the same as in Bayesian analysis. However, because we are adopting a frequentist perspective, we will not make an assumption concerning the prior distribution of the latent variables. As is well-known, this may be an advantage or disadvantage, depending on the situation.

To model the dependencies among claims within a risk class, we use a copula directly in lieu of a latent variable framework. Although Bühlmann's latent variable framework has proved successful for many applications, a limitation of this approach is that the unobserved variable (structure) is constant over time. This means that dependencies among claims are constant over time; this is a strong assumption in time-series analysis. To illustrate, for our application described below, it will mean that 1998 claims and 1997 claims have the same dependency as between 1998 claims and 1993 claims. It is customary in time-series analysis to assume that dependencies weaken as random variables become further apart in time.

A copula is a tool for understanding relationships among multivariate outcomes; it is a function that links univariate marginals to their full multivariate distribution. Copulas were introduced in 1959 in the context of probabilistic metric spaces. Recently there has been a rapidly developing literature on the statistical properties and applications of copulas, particularly in the enterprise risk management literature, see for example Frees and Valdez (1998), Nelsen (1999), and Embrechts, Lindskog, and McNeil (2001).

This paper extends earlier work by Frees, Young, and Luo (1999, 2001), which showed how to produce credibility predictors for linear longitudinal and panel data models, in two ways. First, we consider a generalized linear model (GLM) for marginal claims distributions. This framework has been applied by actuaries (Haberman and Renshaw, 1996); it allows us to consider long-tailed claims through, for example, a Gamma distribution. Moreover, it also gives a direct method for incorporating covariate (explanatory) variables into credibility estimators for these non-Gaussian situations.

Second, we replace the latent variable method of inducing dependencies with a copula. This direct method of modeling dependencies will allow us to derive models that can be closely fit by the data; an important consideration for applied modeling. An important advantage of the copula approach is that it preserves the shape of marginal distributions. In actuarial applications, we have well-developed methods for estimating marginal distributions; that is, estimating model parameters for each time period in isolation of the others. In this paper, we propose using copula functions to link these period-by-period estimates of distributions, thus preserving all of the standard estimation machinery when developing credibility estimates.

Our approach is to use all of the tools that actuaries (as well as statisticians) use for parametric modeling of the marginal distributions (see, for example, Klugman, Panjer, and Willmot [1998]) but also to connect information in the claims history using theory from copulas. Thus we envision a highly parametric approach to claims rate making. We document several advantages of this new approach compared to the current paradigm in place (as well as some disadvantages). The new approach will be easy to use on a computer in that it is likelihood-based. It should be applicable to a much broader set of problems (such as those listed above), without needing special tools for each problem. We demonstrate that the copula formulation is more flexible than positing a (constant) latent variable. In this paper we compare and contrast the two approaches by examining the Massachusetts automobile claims data set that was used in a previous paper (Frees, 2003) on multivariate credibility (an example where specialized credibility techniques were required). Our intent is to develop a basic theory using this data set as our guide.

The following is an outline for the remainder of the paper. Section 2 lays out the basic stochastic model, including the GLM model for marginal claims and the copula for dependencies over time. Section 3 introduces the Massachusetts automobile claims data and Section 4 shows how to fit this data to our framework. Section 5 summarizes the prediction and Section 6 provides summary and concluding remarks.

2. MODELING LONGITUDINAL DATA USING COPULAS

This section outlines the theory part of the paper. Section 2.1 describes the marginal distributions using a GLM framework. Section 2.2 connects the marginals via a copula and Section 2.3 shows how to predict future observations.

2.1 Marginal Distribution

Suppose that there are T_i potential claims for the ith risk class, $\mathbf{Y}_i = (Y_{i1}, Y_{i2}, Y_{i3}, \ldots, Y_{iT_i})'$, and that the corresponding realizations are $\mathbf{y}_i = (y_{i1}, y_{i2}, y_{i3}, \ldots, y_{iT_i})'$, the observed sample. The joint distribution function for the ith risk class is denoted by

$$\mathrm{F}_i(y_{i1}, \ldots, y_{iT_i}) = \mathrm{Prob}(Y_{i1} \le y_{i1}, \ldots, Y_{iT_i} \le y_{iT_i}),$$

with marginal distribution functions

$$\mathrm{F}_{it}(y_{it}) = \mathrm{Prob}(Y_{it} \le y_{it}) = F_{it} = \mathrm{F}(y_{it}, \theta_{it}).$$

We assume independence among risk classes $i = 1, \ldots, n$.

With this notation, we assume that the marginal distribution function F(.) for claims Y_{it} is common up to a systematic component θ_{it} that is known up to K parameters. For applications, we typically work with models such that θ_{it} is a linear function of unknown parameters and use $\theta_{it} = \mathbf{x}_{it}'\boldsymbol{\beta}$, where $\mathbf{x}_{it}$ is a $K \times 1$ vector of known explanatory variables (covariates) and $\boldsymbol{\beta}$ is a $K \times 1$ vector of unknown parameters. The corresponding marginal density (mass) function is $\mathrm{f}_{it}(y_{it}) = f_{it} = \mathrm{f}(y_{it}, \theta_{it})$.

In this paper, we assume that f(.) is from the natural exponential family of distributions. This family encompasses the normal, Poisson and Gamma distributions, as well as others that are important in actuarial applications (Haberman and Renshaw, 1996). Thus, the marginal density (mass) function for the ith risk class at the tth time point can be written as:

$$f(y_{it}, \theta_{it}) = \exp\left(\frac{y_{it}\theta_{it} - b(\theta_{it})}{\phi} + S(y_{it}, \phi)\right).$$

Here, the functions b(.) and S(.,.) are chosen to represent a particular distribution and ϕ is a known dispersion factor. We can also select a function to link the mean component to the systematic component. For illustrative purposes, we focus on the canonical link function so that $b'(\theta_{it}) = E(y_{it})$ and $\theta_{it} = \mathbf{x}_{it}'\boldsymbol{\beta} = g(E\, y_{it})$. This family, with the use of covariates, is commonly known as the *generalized linear model* (GLM) in statistics, see for example, McCullagh and Nelder (1989). The GLM framework is an extension of ordinary linear regression that also encompasses logistic and Poisson regression. Thus, by itself, it is an applied statistical model that is useful in many applications.

2.2 Modeling the Dynamics

Using copulas for the generalized linear model is a natural idea that has been proposed in the biomedical literature; see Meester and MacKay (1994) and Lambert (1996) for early discussions and Lambert and Vandenhende (2002) for a more recent contribution. However the idea is not widely known, perhaps because of the nature of the applications investigated. This is the first such investigation in a social science context.

The joint distribution function of $\mathbf{Y}_i$ can be expressed as a function of the marginal distributions through the copula function

$$F_i(y_{i1}, \ldots, y_{iT_i}) = C(F_{i1}, \ldots, F_{iT_i}),$$

where C is a copula. Thus, the copula allows a fully parametric specification of the probability model, we exploit this specification by using maximum likelihood estimation. We assume independence among risk classes and use the copula to model dependencies over time. Hence, the copula accounts for the *dynamic* aspect of claims behavior.

Let c(.) be the probability density function corresponding to the copula distribution function C(.) (we now assume continuous claims). Thus, the log-likelihood of the ith risk class is

$$l_i = \sum_{t=1}^{T_i} \ln f(y_{it}, \theta_{it}) + \ln c(F_{i1}, F_{i2}, \ldots, F_{iT_i}).$$

For the GLM framework, we have

$$l_i = constant + \sum_{t=1}^{T_i} \frac{y_{it}\mathbf{x}_{it}'\boldsymbol{\beta} - b(\mathbf{x}_{it}'\boldsymbol{\beta})}{\phi} + \ln c(F_{i1}, F_{i2}, \ldots, F_{iT_i}). \tag{2.1}$$

Although not as extensive as with bivariate data ($T = 2$), there are still several options in the choice of a copula function for multivariate data, such as the Archimedean, Markov, and elliptical copulas (see for example, Joe, 1997, Nelson, 1999). We will focus on the copula associated with the multivariate t-distribution, known as the t-copula. This generalization of the normal copula retains many of its desirable properties, including tractability and ease of implementation in simulation studies. Moreover, it has proven popular in the risk management literature recently because of its ability to provide positive large tail dependence. See for example, Embrechts et al. (2001), Venter (2003), and Demarta and McNeil (2004), as well as the discussion in the Appendices.

Appendix B gives the formula for the t-copula density. Substituting the copula density into equation (2.1) gives an expression for the log-likelihood of the ith risk class. Nonlinear optimization subroutines such as NLPNMS and NLPQN from statistics software package SAS provide numerical tools for maximizing the log-likelihood equations. This, together with likelihood expressions, is sufficient for standard parametric estimation. Section 4 provides illustrations.

2.3 Credibility Prediction with Copulas

In Appendix B, we see that the t-copula is parameterized by r, its degrees of freedom, and $\mathbf{\Sigma}$, a correlation matrix. For the $T + 1$ observations from the ith risk class, $(y_{i1}, \ldots, y_{i,T}, y_{i,T+1})$, we may partition this correlation matrix as

$$\mathbf{\Sigma}_{T+1} = \begin{pmatrix} \mathbf{\Sigma}_T & \rho_{T+1\cdot T} \\ \rho'_{T+1\cdot T} & 1 \end{pmatrix}.$$

That is, $\mathbf{\Sigma}_T$ describes that correlations among $(y_{i1}, \ldots, y_{i,T})$ and $\rho_{T+1\cdot T}$ describes the correlation between $y_{i,T+1}$ and $(y_{i1}, \ldots, y_{i,T})$. Using this partition, we may define the conditional variance $\sigma^2_{T+1\cdot T} = 1 - \rho'_{T+1\cdot T}\mathbf{\Sigma}_T^{-1}\rho_{T+1\cdot T}$, where Σ_T^{-1} is the inverse of the correlation matrix Σ_T.

In Appendix C, we show that the density function of the predictive distribution is

$$\mathrm{f}(y_{i,T+1}|y_{i1}, \ldots y_{i,T}) = g_r\left(\frac{\nu_{i,T+1} - \rho'_{T+1\cdot T}\mathbf{\Sigma}_T^{-1}\mathbf{v}_i}{\sigma_{T+1\cdot T}}\right)\frac{\mathrm{f}(y_{i,T+1}, \theta_{i,T+1})}{g_r(\nu_{i,T+1})\sigma_{T+1\cdot T}}, \tag{2.2}$$

where $\nu_{it} = G_r^{-1}(F_{it}(y_{it}))$, $t = 1, \ldots, T + 1$ and $\mathbf{v}_i = (\nu_{i1}, \ldots, \nu_{iT})'$. Here, $F_{it}(y_{it})$ and $f_{it}(y_{it}) = \mathrm{f}(y_{it}, \theta_{it})$, $t = 1, \ldots, T + 1$, are cumulative and density (mass) distribution functions, respectively, of the univariate marginal distribution. Further, G_r is the distribution function of a t-distribution with r degrees of freedom and g_r is the associated density, given by

$$g_r(z) = \frac{\Gamma((r + 1)/2)}{(r\pi)^{1/2}\Gamma(r/2)}\left(1 + \frac{z^2}{r}\right)^{-(r+1)/2}.$$

With the formula for the conditional density in equation (2.2) and the help of computational software, we can easily compute the mean, median, or any percentile of the conditional distribution for the purpose of application or comparison of prediction methods.

3. MASSACHUSETTS AUTOMOBILE CLAIMS

To illustrate our proposed procedures, this article considers automobile bodily injury liability claims from a sample of $n = 29$ Massachusetts towns described in Frees (2003). For this coverage, we consider annual data from $T = 6$ years, 1993–98, inclusive. To mitigate the effect of time trends, claims amounts have been rescaled to adjust for the effects of inflation. Specifically, all claims are in 1991 dollars, using the Consumer Price Index (CPI) for the rescaling factor. We study the behavior of average claims per unit of exposure (the pure premium), defined to be the total claim amount divided by the amount of exposure, for each town, and each year.

We first present summary statistics of the claims data in Section 3.1. Section 3.2 examines the marginal claims distribution and Section 3.3 introduces explanatory variables.

3.1 Descriptive Statistics

Table 1 displays the descriptive statistics for average claims (AC) by year. For instance, the mean of average claims in 1993 is $133.00 and the standard deviation for the same period is $31.59. This table suggests that the claims distribution appears to be stable over time.

Table 1
Descriptive Statistics of Average Claims (in dollars)

	AC1993	AC1994	AC1995	AC1996	AC1997	AC1998
Mean	133.00	129.03	143.38	141.17	142.94	134.37
Median	131.57	131.45	138.76	149.00	144.73	131.96
Standard deviation	31.59	32.63	38.28	39.28	36.22	32.85
Minimum	80.03	42.74	61.04	66.20	61.68	74.89
Maximum	212.46	209.52	238.22	201.99	248.75	191.05

Table 2
Claims Correlations

	AC1993	AC1994	AC1995	AC1996	AC1997	AC1998
AC1993	1.000					
AC1994	0.811	1.000				
AC1995	0.731	0.668	1.000			
AC1996	0.754	0.670	0.680	1.000		
AC1997	0.761	0.626	0.875	0.745	1.000	
AC1998	0.645	0.573	0.648	0.711	0.674	1.000

Table 2 displays correlations of claims among the six years. Clearly the multivariate average claims variables are not independent. For example, the correlation coefficient between the average claims of 1993 and 1994 is 0.811. The smallest correlation coefficient, 0.573, occurs between 1994 and 1998. We will demonstrate how copulas can be employed to model these relationships in Section 4.

3.2 Marginal Claims Distributions

To validate our prediction methods in Section 4, the observations for year 1998 are reserved as the ''hold-out'' sample. This leaves $T = 5$ years of observations for each town at our disposal for parameter estimation.

To obtain intuitive knowledge of the distribution of the average claims, several probability distribution (q-q or quantile-quantile) plots were produced and presented in Figure 1. These plots compare empirical quantiles to quantiles from an estimated parametric model. There are two fitted Weibull distributions. One sets the position parameter θ at zero while estimating the scale parameter and shape parameter. The other estimates all three parameters simultaneously. Table 3 reports three goodness-of-fit statistics that assess the relation between the empirical distribution and the estimated parametric distribution. A large p-value indicates a nonsignificant difference between the two. Table 3 gives the results of goodness-of-fit for all candidate distributions.

Both probability plots and p-values indicate that, except for the exponential distribution, all hypothesized distributions provide a reasonable fit for the average claims variable. The variable of interest, average claims, is equal to the sum of all claim amounts divided by the number of exposures, which is

Figure 1
Q-Q Plots of Six Marginal Claims Distributions

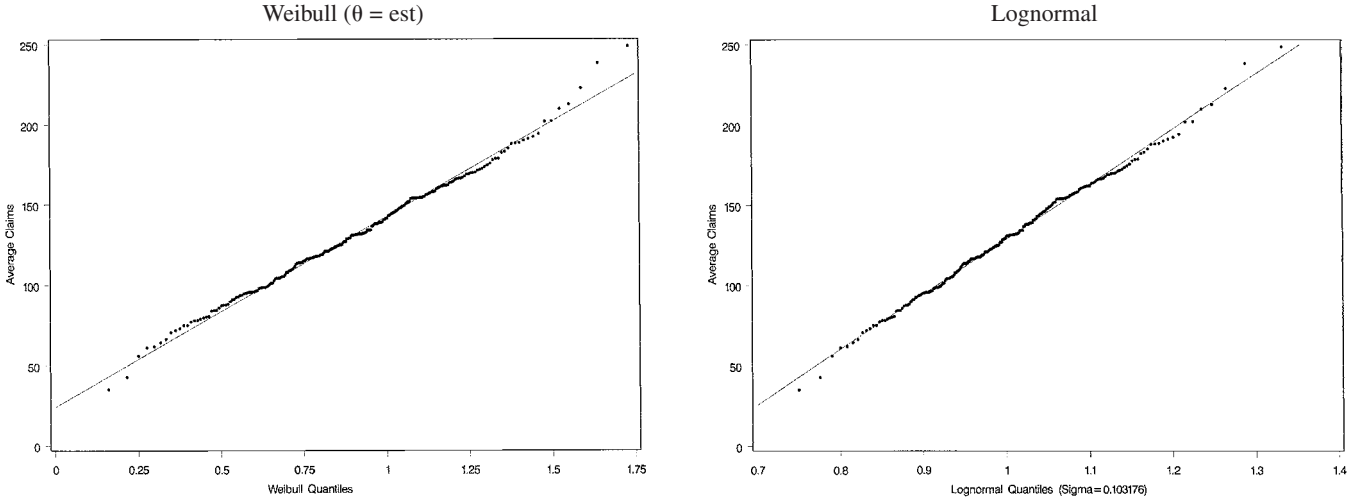
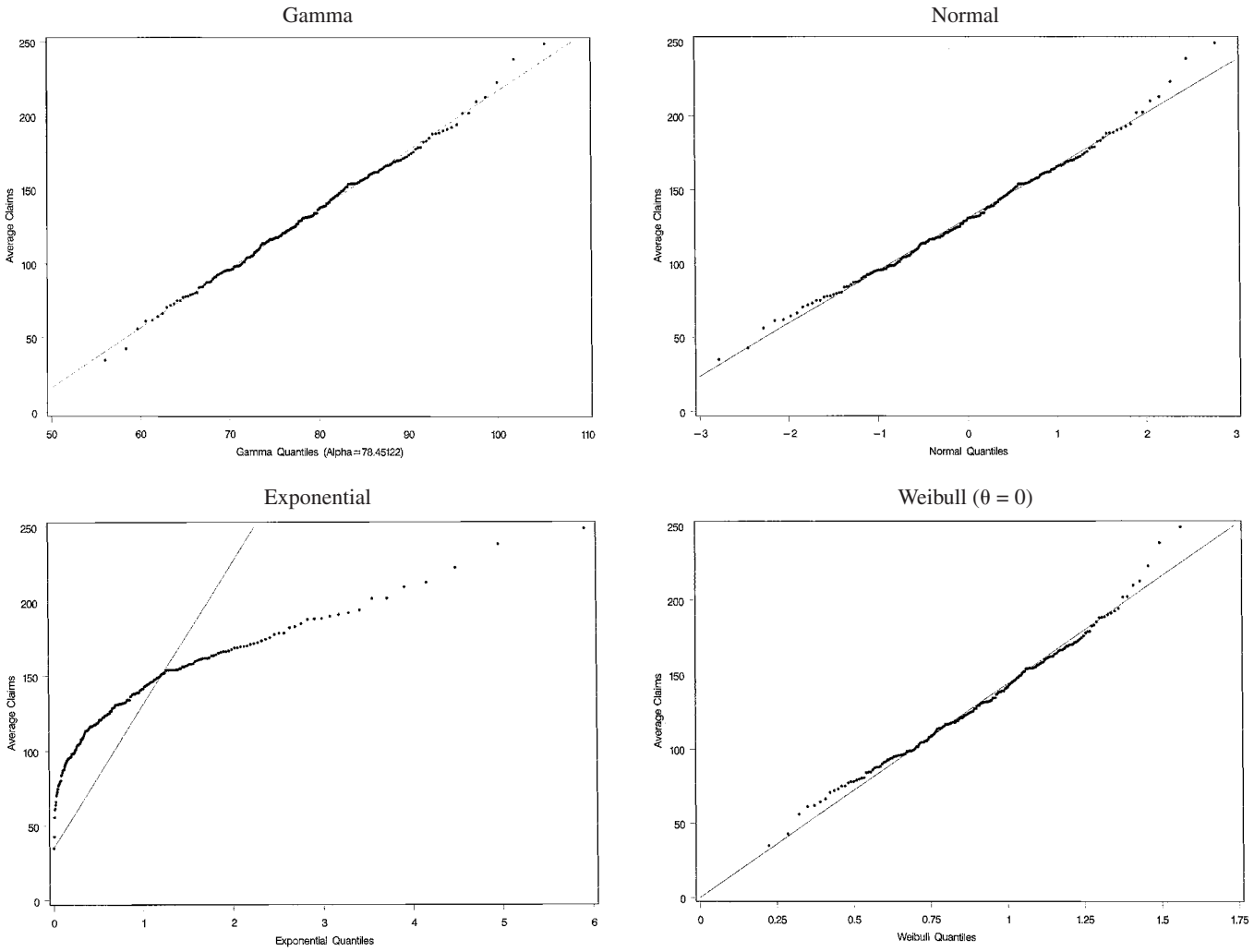


Figure 1
(continued)



at least three thousands for each town-year in our sample. Theoretically, the central limit theorem suggests that average claims be approximately normally distributed and thus have thin-tails. Because all of our hypothesized distributions are capable of fitting thin-tailed distributions, it is not surprising that they fit well. To illustrate the procedures proposed, we chose the Gamma distribution as the fitted marginal distribution of average claims. The Gamma distribution also has the flexibility to allow for long-tail claims distributions.

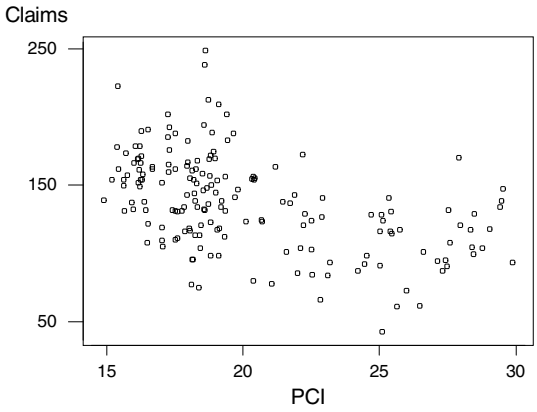
Table 3
***p*-values of Goodness-of-Fit**

	Kolmogorov-Smirnov	Cramer-von Mises	Anderson-Darling
Exponential	<0.001	<0.001	<0.001
Gamma	>0.250	>0.500	>0.500
Lognormal	>0.250	>0.500	>0.500
Normal	>0.150	>0.250	>0.250
Weibull (θ = 0)	N/A	>0.250	>0.150
Weibull (θ est)	>0.500	>0.500	>0.250

3.3 Explanatory Variables

In our study, two explanatory variables, per capita income (PCI) and logarithmic population per square mile (PPSM) were identified as being related to the response variable, claims. Population estimates of the towns in Massachusetts for years 1993 through 1998 were prepared by MISER/State Data Center of Massachusetts and available at http:///www.umass.edu/miser. Data of per capita income were constructed as follows. We first collected information about per capita income in 1989 for each sampled town from the 1990 census report. We then found county estimates for median household income on the website of U.S. Census Bureau at http:///www.census.gov/hhes/www/saipe/stcty for years 1990 through 1999. The annual increase rate of median household income in a county was used to proxy that of per capita income of towns in the same county, so per capita income for each town can be estimated. Finally all income data were deflated to 1991 dollars using the CPI index.

Figure 2
Claims versus Per Capita Income (in thousands, PCI)



Figures 2 and 3 reveal the association of AC with PCI and with PPSM, respectively. Figure 2 suggests that lower bodily injury claims are associated with higher income. A reasonable explanation may be that when more money is at disposal people are more willing to settle disputes by themselves to avoid the penalty of increased premium that is associated with claim payments by the insurance company. In towns of high population density, Figure 3 indicates average claims tend to be high, which may be a consequence of the more frequent occurrences of losses and claims.

Figure 3
Claims versus Population per Square Mile (in logarithmic units, PPSM).

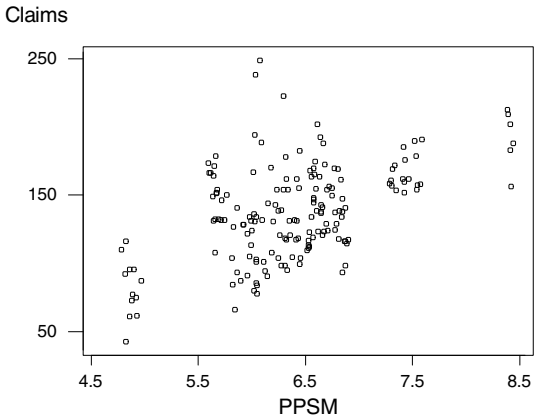


Table 4
Regression of Claim versus PCI and PPSM

	Coefficients	Standard Error	*t*-statistic	*p*-value
Intercept	76.344	23.357	3.27	0.0014
PCI	−4.123	0.569	−7.25	<.0001
PPSM	22.604	3.013	7.50	<.0001
R-square			0.458	

The relationships of average claims with income and population density are supported by a regression analysis whose results are displayed in Table 4. For instance, every $1,000 increase in per capita income is associated with $4.12 decrease in average claim amount per exposure, i.e., pure premium. Every 2.72 persons increase in population density (equivalent to increase of 1 in logarithmic population per square mile) is associated with $22.60 increase in average claim amount. Moreover both covariates are significant at 1% level.

4. INFERENCE USING COPULAS

In this section we show how copulas can be applied to incorporate the dependence structure. Section 4.1 discusses parameterizations of the *t*-copula. To demonstrate the application of copulas in modeling and prediction, we divide our analysis into two stages. Section 4.2 summarizes estimation that does not involve explanatory variables while concentrating on estimation of the copula and marginal distribution parameters. Section 4.3 incorporates the explanatory variables parameters.

4.1 Parameters of the *t*-Copula

For inference, one needs to estimate the *t*-copula parameters given by the correlation matrix $\mathbf{\Sigma}$ and the degrees of freedom r.

We will compare four different structures of $\mathbf{\Sigma}$: the identity, exchangeable, $AR(1)$, and band Toeplitz. Essentially, these four choices capture different aspects of the (time-series) correlation structure. One well-known fact of *t*-copulas is that the identity correlation matrix implies a type of zero correlation although not independence among observations. (From Appendix A, this is because of the common denominator, χ_r^2, used in the definition of the multivariate *t*-distribution.) However, when r tends to infinity, the *t*-copula tends to a normal copula in which case the identity matrix does imply independent components (because χ_r^2/r tends to a constant).

The exchangeable structure, also known as ''compound symmetry'' or ''uniform correlation'' in longitudinal data models (see for example, Frees 2004), is most closely aligned with traditional credibility models. Bühlmann (1967) posited a latent ''structure'' variable that is common to each claim within a risk class; this structure variable induces dependencies among claims within a risk class that does not vary over time.

In contrast, the $AR(1)$ structure is a traditional time-series representation of temporal relationships; this structure implies that the claim experience of current year poses diminishing influence on claims of the following years. A band Toeplitz structure is adopted when we assume that the claims have a Markovian characteristic; current year claims only affects claims of next several, say, l, years. In linear time-series analysis, this structure corresponds to the ''moving-average'' model.

Because the time dimension consists of $T = 5$ years, the four different structures of $\mathbf{\Sigma}$ can be expressed as:

$$\boldsymbol{\Sigma}_I = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \qquad \boldsymbol{\Sigma}_{EX} = \begin{pmatrix} 1 & \rho & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho & \rho \\ \rho & \rho & 1 & \rho & \rho \\ \rho & \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & \rho & 1 \end{pmatrix},$$

$$\boldsymbol{\Sigma}_{AR} = \begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 & \rho^4 \\ \rho & 1 & \rho & \rho^2 & \rho^3 \\ \rho^2 & \rho & 1 & \rho & \rho^2 \\ \rho^3 & \rho^2 & \rho & 1 & \rho \\ \rho^4 & \rho^3 & \rho^2 & \rho & 1 \end{pmatrix}, \text{ or } \qquad \boldsymbol{\Sigma}_T = \begin{pmatrix} 1 & \rho_1 & \rho_2 & 0 & 0 \\ \rho_1 & 1 & \rho_1 & \rho_2 & 0 \\ \rho_2 & \rho_1 & 1 & \rho_1 & \rho_2 \\ 0 & \rho_2 & \rho_1 & 1 & \rho_1 \\ 0 & 0 & \rho_2 & \rho_1 & 1 \end{pmatrix}. \tag{4.1}$$

In expression (4.1) of band Toeplitz matrix, we use a band of $l = 2$.

The number of parameters to be evaluated depends on the matrix structure adopted. For example, when either the exchangeable or $AR(1)$ structure is employed, two parameters determine the t-copula, r and ρ. In addition, there are parameters associated with the fitted marginal distribution, Gamma. Generally a two-parameter Gamma distribution has shape parameter α and scale parameter γ. This means we have to evaluate four parameters when the structure of $\boldsymbol{\Sigma}$ is assumed to be exchangeable or $AR(1)$.

There are at least three approaches to dealing with degrees of freedom, r. One is to treat r in the same way as other parameters and estimate it using maximum likelihood by treating it as a continuous variable. Specifically, r and other parameters are estimated altogether by maximum likelihood, as in Lambert and Vandenhende (2002). Another approach is to estimate other parameters for selected values of r, essentially treating it as discrete. We report results using both approaches. A third approach is to assume that r is known, either as infinity resulting in a normal copula or as a finite value when based on a ''degrees of freedom'' argument, such as is customary in the univariate case.

4.2 Estimation Without Explanatory Variables

We assume that the marginal claims distribution can be modeled by a two-parameter Gamma distribution with density function

$$f(y;\, \alpha,\, \gamma) = \frac{y^{\alpha-1}}{\gamma^{\alpha}\, \Gamma(\alpha)} \exp\left(-\frac{y}{\gamma}\right).$$

So that this can be expressed as a member of the exponential family, one chooses $\theta = -1/(\alpha\, \gamma)$ and $\phi = 1/\alpha$. We also assume that the dependence structure can be modeled by a t-copula with r degrees of freedom. With these assumptions, the log-likelihood function for town i over T_i years can be expressed as

$$\begin{aligned} l_i(\alpha,\, \gamma) = {} & -T_i \alpha \log \gamma - T_i \log \Gamma(\alpha) + (\alpha - 1) \sum_{t=1}^{T_i} \log y_{it} - \frac{1}{\gamma} \sum_{t=1}^{T_i} y_{it} \\ & + \log \Gamma\left(\frac{T_i + r}{2}\right) + (T_i - 1) \log \Gamma\left(\frac{r}{2}\right) - T_i \log \Gamma\left(\frac{r+1}{2}\right) - \frac{1}{2} \log |\boldsymbol{\Sigma}| \\ & + \frac{r+1}{2} \sum_{t=1}^{T_i} \log\left(1 + \frac{v_{it}^2}{r}\right) - \frac{r + T_i}{2} \log\left(1 + \frac{1}{r}\, \mathbf{v}_i' \boldsymbol{\Sigma}^{-1} \mathbf{v}_i\right), \end{aligned} \tag{4.2}$$

where $\mathbf{v}_i = (v_{i1},\, v_{i2},\, \ldots,\, v_{iT_i})'$ and $v_{it} = G_r^{-1}(F_{it}(y_{it};\, \alpha,\, \gamma))$ for $t = 1, \ldots, T_i$.

At this stage of analysis, we assume that the Gamma distribution parameters are the same for all towns in our sample, so there are only four parameters, ρ, α, γ, and r, to be estimated for exchangeable and $AR(1)$ models, three for the identity and five for Toeplitz. For estimation, we employ maximum likelihood. Table 5 displays values of Akaike Information Criteria (AIC) over several choices of r to compare goodness of fit. For this criterion, smaller values of AIC mean a better fit. The entry $r =$

Table 5
Akaike Information Criterion (AIC) by Correlation Matrix (Σ) and Degrees of Freedom (*r*).

Degrees of Freedom (*r*)	Correlation Matrix (Σ)			
	Identity	Exchangeable	*AR*(1)	Toeplitz (*l* = 2)
2	1,426.36	1,357.13	1,379.83	1,389.08
4	1,427.61	1,348.68	1,372.37	1,382.30
6	1,431.07	1,346.39	1,370.61	1,380.86
8	1,434.13	1,345.63	1,370.20	1,380.65
10	1,436.58	1,345.40	1,370.23	1,380.79
12	1,438.54	1,345.38	1,370.42	1,381.04
14	1,440.13	1,345.46	1,370.65	1,381.32
16	1,441.45	1,345.56	1,370.90	1,381.58
18	1,442.56	1,345.69	1,371.14	1,381.83
20	1,443.51	1,345.81	1,371.36	1,382.05
22	1,444.32	1,345.93	1,371.57	1,382.25
24	1,445.04	1,346.04	1,371.76	1,382.43
26	1,445.66	1,346.15	1,371.94	1,382.60
28	1,446.22	1,346.25	1,372.10	1,382.75
30	1,446.72	1,346.34	1,372.25	1,382.88
45	1,449.25	1,346.86	1,373.05	1,383.60
60	1,450.67	1,347.18	1,373.53	1,384.01
120	1,453.02	1,347.75	1,374.36	1,384.70
1,000	1,455.37	1,348.34	1,375.22	1,385.40
10,000	1,455.68	1,348.42	1,375.33	1,385.49
∞ (normal)	1,455.72	1,348.43	1,375.34	1,385.51

Note: No covariates are used in the model fitting.

infinity is also added to compare the fitness of *t*-copula with normal copula. Table 5 shows that the *t*-copula provides a better fit than the normal model regardless of the choice of the correlation matrix.

Because of this, in Table 6 we only report results using *t*-copula. We can see from Table 6 that all parameters (except *r*) are statistically significant. Our major interest is the significance of correlation coefficients; they are all statistically significant which provides strong evidence that the correlation structure is not independent.

Table 6
Maximum Likelihood Estimation Results by Correlation Matrix (Σ).

Parameter	Correlation Matrix (Σ)			
	Identity	Exchangeable	*AR*(1)	Toeplitz
ρ_1	NA	0.744 (0.067)	0.710 (0.063)	0.478 (0.081)
ρ_2	NA	NA	NA	0.402 (0.080)
γ	14.410 (2.766)	13.242 (2.912)	14.123 (2.662)	16.728 (2.645)
α	10.106 (2.044)	10.393 (2.311)	9.759 (1.849)	8.233 (1.319)
r	2.494 (0.897)	11.236 (8.225)	8.685 (5.338)	7.823 (4.752)
AIC	1,427.978	1,347.377	1,372.180	1,382.644

Notes: Standard errors are reported in parentheses. No covariates are used in the model fitting.

4.3 Estimation with Explanatory Variables

The discussion in Section 3.3 suggests that PCI and PPSM are useful predictors of claims. With Gamma distributed claims, we use a canonical link function so that $\theta_{it} = \mathbf{x}'_{it}\boldsymbol{\beta} = \beta_0 + \beta_1 PCI_{it} + \beta_2 PPSM_{it}$. As is customary in generalized linear models, we assume that the scale parameter $\phi = 1/\alpha$ is constant over towns and years while the shape parameter varies through the relation $\theta_{it} = -1/(\alpha\gamma_{it})$.

Now the parameters to be determined include β_0, β_1, and β_2. Again, we first compare the *t*-copula and normal copula under different correlation structures, as reported in Table 7. Values of AIC indicate,

Table 7
Akaike Information Criterion (AIC) by Correlation Matrix (Σ) and Degrees of Freedom (*r*).

Degrees of Freedom (*r*)	Correlation Matrix (Σ)			
	Identity	Exchangeable	*AR*(1)	Toeplitz
2	1,354.24	1,335.35	1,345.43	1,342.84
4	1,349.48	1,328.06	1,338.84	1,335.44
6	1,349.64	1,326.41	1,337.66	1,333.76
8	1,350.62	1,326.06	1,337.67	1,333.39
10	1,351.66	1,326.13	1,338.01	1,333.43
12	1,352.61	1,326.33	1,338.44	1,333.60
14	1,353.44	1,326.57	1,338.86	1,333.81
16	1,354.16	1,326.82	1,339.26	1,334.02
18	1,354.79	1,327.05	1,339.63	1,334.22
20	1,355.34	1,327.27	1,339.96	1,334.40
22	1,355.83	1,327.47	1,340.26	1,334.56
24	1,356.25	1,327.65	1,340.53	1,334.71
26	1,356.64	1,327.82	1,340.77	1,334.85
28	1,356.98	1,327.97	1,340.99	1,334.97
30	1,357.29	1,328.11	1,341.20	1,335.08
45	1,358.89	1,328.86	1,342.26	1,335.67
60	1,359.80	1,329.30	1,342.88	1,336.01
120	1,361.34	1,330.06	1,343.94	1,336.58
1,000	1,362.90	1,330.84	1,345.02	1,337.15
10,000	1,363.11	1,330.94	1,345.17	1,337.23
∞ (normal)	1,363.13	1,330.96	1,345.18	1,337.24

Note: Covariates are used in the model fitting.

as in the case of no covariates, that the *t*-copula fits our sample data better than the normal copula. Another observation about Table 7 is that the AIC for each of the models with covariates present is less than that of corresponding models without covariates; this shows evidence that covariates provide useful explanatory information.

Table 8 displays the results for models with covariates, using the *t*-copula and different matrix structures. Once again we observe that the correlation coefficients for exchangeable, *AR*(1) and Toeplitz are strong statistically greater than zero (*p*-values less than 1%), indicating the value of the dependence structure. The coefficients of the explanatory variables, β_1 and β_2, are also significant at the 1% level. This substantiates the hypothesis that lower average income and higher population density are associated with higher bodily injury claims.

Table 8
Maximum Likelihood Estimation Results by Correlation Matrix (Σ).

Parameter	Correlation Matrix (Σ)			
	Identity	Exchangeable	*AR*(1)	Toeplitz
ρ_1	NA	0.442 (0.105)	0.395 (0.097)	0.275 (0.096)
ρ_2	NA	NA	NA	0.372 (0.084)
β_0	1.374 (0.861)	1.416 (1.300)	1.714 (1.088)	1.836 (0.947)
β_1	−0.142 (0.029)	−0.143 (0.037)	−0.147 (0.033)	−0.137 (0.028)
β_2	1.022 (0.204)	1.049 (0.269)	0.987 (0.217)	0.881 (0.169)
α	27.340 (4.370)	26.309 (4.616)	27.125 (4.326)	29.017 (4.013)
r	4.686 (1.983)	8.405 (5.200)	6.863 (3.679)	19.230 (22.438)
AIC	1,351.33	1,328.057	1,339.592	1,336.331

Note: Standard errors are reported in parentheses. Covariates are used in the model fitting.

5. PREDICTION WITH COPULAS

In Section 4, using data for years 1993 through 1997, we estimated parameters associated with the *t*-copula and the coefficients of explanatory variables per capita income and population density. Now we can predict the pure premium for 1998, the major interest of this study, and compare predictions using copulas to that of standard existing approaches, namely, full credibility and Bühlmann credibility.

On the one hand, when full credibility is granted to past observations, the predicted value of the next period equals the mean of prior observations. Specifically, the predicted value of the $(T + 1)$-*st* period for the ith risk class is

$$\bar{y}_i = \frac{1}{T}\sum_{t=1}^{T} y_{it}.$$

On the other hand, Bühlmann credibility suggests that prediction for the ith risk class can ''borrow'' information of other risk classes as well as using its own past experience. Here the credibility predictor is given by

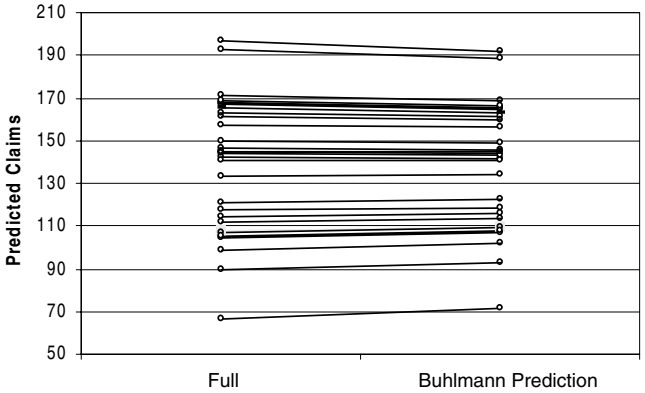
$$\bar{y}_{i,cred} = z \cdot \bar{y}_i + (1 - z)\bar{y},$$

where $\bar{y}$ is the overall mean and $z = T/(T + \sigma_\varepsilon^2/\sigma_\alpha^2)$ is the credibility factor. Refer to Frees, Young, and Luo (2001) for details. There is less spread in the credibility predictions compared to the mean for each risk class; the credibility predictor for a risk class is ''shrunk'' to the overall mean $\bar{y}$. The shrinkage effect of the Bühlmann credibility predictor, applied to our sample data, is shown in Figure 4. From this figure, we see that for our data set, the estimated value of σ_α^2 is large relative to σ_ε^2 so that the credibility factor is close to one, indicating near full credibility.

For copula-based credibility, we used the parameter estimates in Section 4 and 1998 covariate values to estimate the predictive density, given in equation (2.2). For this predictive density, we computed the mean of the distribution that we refer to as the ''copula credibility predictor.'' Predictions using copula credibility were made with an exchangeable correlation matrix of *t*-copula. Figure 5 summarizes the predictions for our data set.

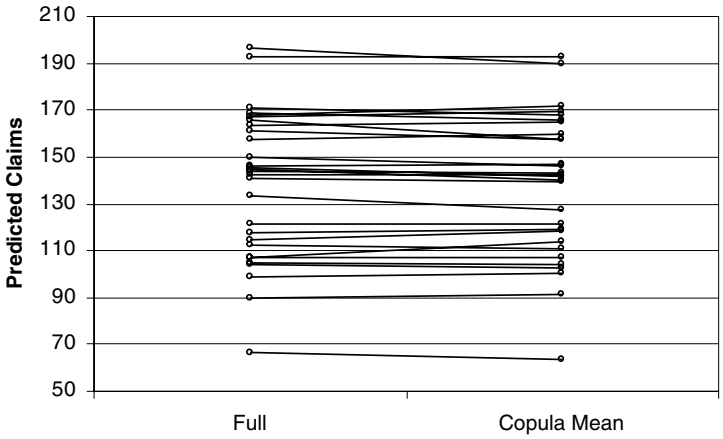
Figure 5 shows that copula credibility predictors also have a mild shrinkage effect. On the left-hand side of each panel, values of full credibility prediction vary over a broader range than values on the right-hand side do, where copula credibility predictions are displayed. Not surprisingly, because these copula credibility predictors use covariates, Figure 5 displays some crossover among lines. Unlike linear predictor theory, we do not know of a broad statistical principle that would ensure a shrinkage effect.

Figure 4
Shrinkage Effect of Bühlmann Credibility



Notes: Each town is connected by a line. The left-hand vertical axis displays the full credibility predictions. The right-hand vertical axis displays the predictions using Bühlmann's credibility.

Figure 5
Shrinkage Effect of Copula Credibility Predictors

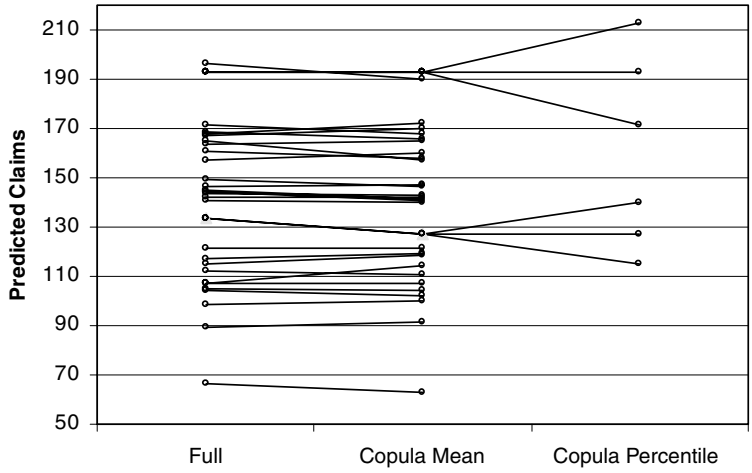


Notes: Each town is connected by a line. The left-hand vertical axis displays the full credibility predictions. The right-hand vertical axis displays the predictions using the copula credibility. The graph suggests that the copula credibility predictors preserve the shrinkage effect.

We conjecture that it may be because all of the independent risk classes are being used to estimate common parameters.

Figure 6 augments Figure 5 by adding prediction percentiles for two selected towns. For each of the two towns, Figure 6 shows the 25th and 75th percentiles of the predictive distribution, as well as the mean (for this data set, the mean and median of the predictive distribution are close). For more skewed data sets, the median may be more appropriate. The percentiles provide the actuary with a range of reliability for assessing the copula credibility predictor for rate making purposes. This figure emphasizes that copula credibility predictors share a desirable property with Bayesian credibility; namely the ability to provide a full predictive distribution of future claims.

Figure 6
Distributions of Copula Credibility Predictor



Notes: Each town is connected by a line. On the horizontal axis, ''Full'' indicates prediction using full credibility. Copula Mean reports the predicted mean using copula credibility. Copula Percentile displays the predicted 25% percentile, mean, and 75% percentile for two towns.

Table 9
Comparisons of Sum of Squared Prediction Error

Full Credibility	Bühlmann Credibility	Copula Credibility		
		AR(1)	Exchangeable	Toeplitz
15,700.8	14,916.4	14,437.5	14,255.6	15,265.4

Another examination of the usefulness of copula credibility is displayed in Table 9, the sum of squared prediction errors (SSPE). The error of prediction is defined as the difference between the actual average claim of 1998 and the predicted value using different methods. Table 9 shows that the SSPE of copula credibility methods are less than that of full credibility and the SSPE of the $AR(1)$ and exchangeable correlation copula are less than that of Bühlmann predictors; this suggests that copula credibility deserves a position in the toolbox of actuaries. This result is not surprising in that the copula credibility predictors use the information in the distribution of claims, the dynamic dependencies, and the associated explanatory variables.

6. SUMMARY AND CONCLUDING REMARKS

Credibility estimators are designed to predict claims for a risk class, given prior claims from a risk class and claims from other risk classes. In the traditional linear random effects setting, one models dependencies among claims through latent random quantities known as structure variables; predictors of claims are minimum mean square error among the class of all linear unbiased predictors. This paper considers claims distributions that may be skewed so that the mean square error criterion may not be suitable. Claims distributions are modeled parametrically; this allows one to calculate predictive distributions for future claims given past claims. With the predictive distribution, one can compute means, medians, or any other measure to summarize the predictive distribution.

Computing predictive distributions is an exercise well-known to Bayesian enthusiasts of credibility. From this perspective, it is traditional in credibility theory to assume a prior distribution for the structure variables and use posterior distributions to compute predictive distributions. Instead, in this paper we directly use a copula to model dependences among claims for a risk class. In this way we need not make assumptions about prior distribution. Moreover, we need not assume that the common latent variable induces an exchangeable structure among claims; we can and do investigate time-series models of claims.

For this paper, we used claims from a Gamma family and provided the necessary theoretical underpinnings for the exponential family of distributions that also includes the normal and Weibull distributions. Although any parametric family of copulas fits within the framework described here, this paper explores the advantages of the t-copula. We find that this is a desirable dependence structure, at least for the bodily injury liability automobile claims data investigated here. We hope to explore the robustness of the choice of marginal distributions, covariates, and the copula in subsequent work.

For our data, we compared the copula-based credibility predictors and found that they performed well compared to traditional credibility estimators. They even demonstrated the well-known ''shrinkage'' characteristic that actuaries find appealing for traditional estimators. This may not be a general characteristic of copula-based credibility predictors; we are not aware of general conditions under which an actuary could expect that these new predictors would possess this characteristic. We leave this a problem for future investigation.

We do not anticipate that current available credibility rate making techniques will disappear; to illustrate, the proposed structure does not enjoy the intuitively appealing linear credibility theory formula that actuarial students learn. However, we illustrated procedures that are easy to implement in today's computing environment and that should be applicable in a broad set of circumstances. We did this by positing a stochastic model of insurance claims and developing algorithms for producing cred-

ibility forecasts based on this model. We showed how the algorithms work with real data and compared our new procedures to existing methods.

APPENDIX

THE MULTIVARIATE *t*-COPULA

In this appendix, we collect properties about the *t*-copula that actuaries will find useful for GLM modeling. Important references include Johnson and Kotz (1972), Embrechts, Lindskog, and McNeil (2001) and Venter (2003).

APPENDIX A

THE MULTIVARIATE *t*-DISTRIBUTION

Suppose $(N_1, \ldots, N_T)'$ has a joint standardized multivariate normal distribution with correlation matrix $\boldsymbol{\Sigma}$. Further suppose that χ_r^2 has a chi-square distribution with r degrees of freedom and is independent of $(N_1, \ldots, N_T)$. Then, the joint distribution of $Z_t = N_t(\chi_r^2/\sqrt{r})^{-1}$, $t = 1, \ldots, T$ constitutes a multivariate *t*-distribution with r degrees of freedom. One property of this distribution is that each marginal distribution is a *t*-distribution with r degrees of freedom, denoted by G_r. Moreover, subsets have the same family as the joint. Thus, if we assume that $(Z_1, \ldots, Z_{T+1})$ has a multivariate *t*-distribution, then $(Z_1, \ldots, Z_T)$ also has a multivariate *t*-distribution. The joint probability density function of $(Z_1, \ldots, Z_T)'$ is

$$\mathrm{f}_{\mathrm{z}}(\mathbf{z};\, r,\, \boldsymbol{\Sigma}) = \frac{\Gamma((r+T)/2)}{(r\pi)^{T/2}\Gamma(r/2)|\boldsymbol{\Sigma}|^{1/2}} \left(1 + \frac{1}{r}\,\mathbf{z}'\boldsymbol{\Sigma}^{-1}\mathbf{z}\right)^{-(r+T)/2}, \tag{A.1}$$

where $\mathbf{z} = (z_1, \ldots, z_T)'$. See, for example, Johnson and Kotz (1972).

Conditional distributions can be derived in a straightforward manner. Suppose that the correlation matrix associated with $(N_1, \ldots, N_{T+1})'$ is given by

$$\boldsymbol{\Sigma}_{T+1} = \begin{pmatrix} \boldsymbol{\Sigma}_T & \rho_{T+1\cdot T} \\ \rho'_{T+1\cdot T} & 1 \end{pmatrix}.$$

Then, from standard multivariate normal theory, we have that $N_{T+1}|\{N_1, \ldots, N_T\}$ is normal with mean $\rho'_{T+1\cdot T}\boldsymbol{\Sigma}_T^{-1}(N_1, \ldots, N_T)'$ and variance $\sigma^2_{T+1\cdot T} = 1 - \rho'_{T+1\cdot T}\boldsymbol{\Sigma}_T^{-1}\,\rho_{T+1\cdot T}$. Thus, $Z_{T+1}|\{Z_1, \ldots, Z_T, \chi_r^2\}$ is also normal with mean $\rho'_{T+1\cdot T}\boldsymbol{\Sigma}_T^{-1}(Z_1, \ldots, Z_T)'$ and variance $\sigma^2_{T+1\cdot T}(\chi_r^2/\sqrt{r})^{-2}$. Integrating over the distribution of χ_r^2, we have that $Z_{T+1}|\{Z_1, \ldots, Z_T\}$ is equal in distribution to $\rho'_{T+1\cdot T}\boldsymbol{\Sigma}_T^{-1}(Z_1, \ldots, Z_T)' + \sigma_{T+1\cdot T}\, t_r$, where t_r is a *t*-distributed random variable with r degrees of freedom. Thus, the conditional density function is

$$\mathrm{f}_{\mathrm{z}}(z_{T+1}|\mathbf{z}) = \frac{1}{\sigma_{T+1\cdot T}}\, g_r\left(\frac{z_{T+1} - \rho'_{T+1\cdot T}\boldsymbol{\Sigma}_T^{-1}\mathbf{z}}{\sigma_{T+1\cdot T}}\right), \tag{A.2}$$

where $g_r(.)$ is the probability density function of a *t*-distribution with r degrees of freedom.

APPENDIX B

THE *t*-COPULA

We are now ready to define the *multivariate t-copula*, a function defined for all $(u_1, u_2, \ldots, u_T) \in [0,1]^T$ by

$$\mathrm{C}(u_1, \ldots, u_T) = \mathrm{F}_{\mathrm{z}}(G_r^{-1}(u_1), \ldots, G_r^{-1}(u_T)),$$

where F_{z} is the distribution function associated with the probability density function f_{z}. From equation (A.1), the corresponding probability density function is

$$\mathrm{c}(u_1, \ldots, u_T) = \mathrm{f}_{\mathbf{Z}}(G_r^{-1}(u_1), \ldots, G_r^{-1}(u_T)) \prod_{t=1}^{T} \frac{1}{g_r(G_r^{-1}(u_t))}, \tag{B.1}$$

where $g_r(.)$ is the probability density function associated with G_r, that is, a t-distribution with r degrees of freedom.

The conditional density function

$$\mathrm{c}(u_{T+1}|u_1, \ldots, u_T) = \mathrm{f}_{\mathbf{Z}}(G_r^{-1}(u_{T+1})|G_r^{-1}(u_1), \ldots, G_r^{-1}(u_T)) \frac{1}{g_r(G_r^{-1}(u_{T+1}))}, \tag{B.2}$$

can be evaluated using equation (A.2).

APPENDIX C

PREDICTIVE DENSITY

The joint density function is given by

$$\mathrm{f}_i(y_{i1}, \ldots, y_{i,T}) = \mathrm{c}(F_{i1}, F_{i2}, \ldots, F_{iT_i}) \prod_{t=1}^{T_i} \mathrm{f}(y_{it}, \theta_{it}).$$

where $f_{it}(y_{it}) = f_{it} = \mathrm{f}(y_{it}, \theta_{it})$ and $F_{it}(y_{it}) = F_{it}$ is the corresponding distribution function. Thus, using equations (A.2) and (B.2), the predictive distribution is

$$\begin{aligned}
\mathrm{f}(y_{i,T+1}|y_{i1}, \ldots y_{i,T}) &= \frac{f_i(y_{i1}, \ldots, y_{i,T}, y_{i,T+1})}{f_i(y_{i1}, \ldots, y_{i,T})} = \frac{\mathrm{c}(F_{i1}, \ldots, F_{i,T+1})}{\mathrm{c}(F_{i1}, \ldots, F_{iT})} \mathrm{f}(y_{i,T+1}, \theta_{i,T+1}) \\
&= \mathrm{f}_{\mathbf{Z}}(v_{i,T+1}|v_{i1}, \ldots, v_{iT}) \frac{1}{g_r(v_{i,T+1})} \mathrm{f}(y_{i,T+1}, \theta_{i,T+1}) \\
&= g_r\left(\frac{v_{i,T+1} - \rho'_{T+1\cdot T}\boldsymbol{\Sigma}_T^{-1}\mathbf{v}_i}{\sigma_{T+1\cdot T}}\right) \frac{\mathrm{f}(y_{i,T+1}, \theta_{i,T+1})}{g_r(v_{i,T+1})\sigma_{T+1\cdot T}}
\end{aligned}$$

where $v_{it} = G_r^{-1}(F_{it}(y_{it}))$, $t = 1, \ldots, T+1$ and $\mathbf{v}_i = (v_{i1}, \ldots, v_{iT})'$.

ACKNOWLEDGMENTS

The Society of Actuaries (through the Committee on Knowledge Extension Research), the National Science Foundation (Grant Number SES-0436274), and the Fortis Health Insurance Professorship (to the first author) provided funding to support this research. We are grateful for this support.

REFERENCES

BAILEY, ARTHUR. 1945. A Generalized Theory of Credibility. *Proceedings of the Casualty Actuarial Society* 32.

———. 1950. Credibility Procedures: LaPlace's Generalization of Bayes' Rule and the Combination of Collateral Knowledge with Observed Data. *Proceedings of the Casualty Actuarial Society* 37: 7–23.

BÜHLMANN, HANS. 1967. Experience Rating and Credibility. *ASTIN Bulletin* 4: 199–207.

DEMARTA, STEFANO, AND ALEXANDER J. MCNEIL. 2004. The *t*-Copula and Related Copulas. Working paper. Online at www.math.ethz.ch/%7Emcneil/pub_list.html.

EMBRECHTS, PAUL, FILIP LINDSKOG, AND ALEXANDER MCNEIL. 2001. *Modeling Dependence with Copulas and Applications to Risk Management*. Online at www.risklab.ch/ftp/papers/DependenceWithCopulas.pdf.

FREES, EDWARD W. 2003. Multivariate Credibility for Aggregate Loss Models. *North American Actuarial Journal* 7(1): 13–37.

———. 2004. *Longitudinal and Panel Data: Analysis and Applications for the Social Sciences*. Cambridge: Cambridge University Press.

FREES, EDWARD W., AND EMILIANO VALDEZ. 1998. Understanding Relationships Using Copulas. *North American Actuarial Journal* 2(1): 1–25.

FREES, EDWARD W., VIRGINIA R. YOUNG, AND YU LUO. 1999. A Longitudinal Data Analysis Interpretation of Credibility Models. *Insurance: Mathematics and Economics*. 24: 229–48.

———. 2001. Case Studies using Panel Data Models. *North American Actuarial Journal* 4(4): 24–42.

HABERMAN, STEVEN, AND ARTHUR E. RENSHAW. 1996. Generalized Linear Models and Actuarial Science. *The Statistician* 45(4): 407–36.

HICKMAN, JAMES C., AND LINDA HEACOX. 1999. Credibility Theory: The Cornerstone of Actuarial Science. *North American Actuarial Journal* 3(2): 1–8.

JEWELL, WILLIAM S. 1974. Credible Means are Exact Bayesian for Exponential Families. *Astin Bulletin*. 8(1): 77–90.

JOE, HARRY. 1997. *Multivariate Models and Dependence Concepts*. New York: Chapman & Hall.

JOHNSON, NORMAN L., AND SAMUEL KOTZ. 1972. *Distributions in Statistics: Continuous Multivariate Distributions*. New York: Wiley.

KEFFER, RALPH. 1929. An Experience Rating Formula. *Transactions of the Actuarial Society of America* 30: 130–39.

KLUGMAN, STUART A. 1992. *Bayesian Statistics in Actuarial Science*. Boston, Mass.: Kluwer.

KLUGMAN, STUART A., HARRY H. PANJER, AND GORDON WILLMOT. 1998. *Loss Models: From Data to Decisions*. New York: Wiley.

LAMBERT, PHILIPPE. 1996. Modelling Irregularly Sample Profiles of Non-negative Dog Triglyceride Responses Under Different Distributional Assumptions. *Statistics in Medicine* 15: 1695–1708.

LAMBERT, PHILIPPE, AND FRANCOIS VANDENHENDE. 2002. A Copula-Based Model for Multivariate Non-normal Longitudinal Data. *Statistics in Medicine* 21: 3197–3217.

MCCULLAGH, P., AND J. A. NELDER. 1989. *Generalized Linear Models*. 2nd ed. London: Chapman and Hall.

MEESTER, STEVEN G., AND JOCK MACKAY. 1994. A Parametric Model for Cluster Correlated Categorical Data. *Biometrics* 50(4): 954–63.

MILLER, ROBERT B., AND HICKMAN, JAMES C. 1975. Insurance Credibility Theory and Bayesian Estimation. In *Credibility—Theory and Applications*, edited by Paul M. Kahn. New York: Academic Press.

MOWBRAY, ALBERT H. 1914. How Extensive a Payroll Exposure is Necessary to Give a Dependable Pure Premium. *Proceedings of the Casualty Actuarial Society* 1: 24–30.

NELSEN, ROGER B. 1999. *An Introduction to Copulas*. Lecture Notes in Statistics 139, Springer-Verlag.

PINQUET, JEAN. 1997. Allowance for Cost of Claims in Bonus-Malus Systems. *ASTIN Bulletin* 27: 33–57.

VENTER, GARY G. 2003. Quantifying Correlated Reinsurance Exposures with Copulas. *Casualty Actuarial Society Forum* 215–29. Online at www.casact.org/pubs/forum/03spforum/03spf215.pdf.

WHITNEY, ALBERT W. 1918. The Theory of Experience Rating. *Proceedings of the Casualty Actuarial Society* 4.

Discussions on this paper can be submitted until October 1, 2005. The authors reserve the right to reply to any discussion. Please see the Submission Guidelines for Authors on the inside back cover for instructions on the submission of discussions.