Astin Bulletin (2024), **54**, pp. 652–678
doi:10.1017/asb.2024.13

International Actuarial Association
Association Actuarielle Internationale

RESEARCH ARTICLE

Multidimensional credibility: A new approach based on joint distribution function

Limin Wen[1], Wei Liu[1] and Yiying Zhang[2]

[1]School of Mathematics and Statistics, Jiangxi Normal University, Nanchang 330022, China
[2]Department of Mathematics, Southern University of Science and Technology, Shenzhen 518055, China
Corresponding author: Yiying Zhang; Email: zhangyy3@sustech.edu.cn

Received: 29 August 2023; **Revised:** 31 January 2024; **Accepted:** 20 March 2024; **First published online:** 12 April 2024

Keywords: Multidimensional credibility; joint distribution; premium principles; Frobenius norm; consistency; asymptotic normality

JEL codes: G21; G22; G31

Abstract

In the traditional multidimensional credibility models developed by Jewell ((1973) Operations Research Center, pp. 73–77.), the estimation of the hypothetical mean vector involves complex matrix manipulations, which can be challenging to implement in practice. Additionally, the estimation of hyperparameters becomes even more difficult in high-dimensional risk variable scenarios. To address these issues, this paper proposes a new multidimensional credibility model based on the conditional joint distribution function for predicting future premiums. First, we develop an estimator of the joint distribution function of a vector of claims using linear combinations of indicator functions based on past observations. By minimizing the integral of the expected quadratic distance function between the proposed estimator and the true joint distribution function, we obtain the optimal linear Bayesian estimator of the joint distribution function. Using the plug-in method, we obtain an explicit formula for the multidimensional credibility estimator of the hypothetical mean vector. In contrast to the traditional multidimensional credibility approach, our newly proposed estimator does not involve a matrix as the credibility factor, but rather a scalar. This scalar is composed of both population information and sample information, and it still maintains the essential property of increasingness with respect to the sample size. Furthermore, the new estimator based on the joint distribution function can be naturally extended and applied to estimate the process covariance matrix and risk premiums under various premium principles. We further illustrate the performance of the new estimator by comparing it with the traditional multidimensional credibility model using bivariate exponential-gamma and multivariate normal distributions. Finally, we present two real examples to demonstrate the findings of our study.

1. Introduction

Credibility theory plays a crucial role in insurance pricing by allowing insurers to strike a balance between the industry experience and the individual policyholder's claims history. As the pioneering work in modern credibility theory, Bühlmann (1967) developed a Bayesian model for claim amounts that constrains the prediction of future losses to be a linear function of past observations. This model provides the optimal linear Bayesian estimate of future losses, also known as the individual mean or hypothetical mean, under the quadratic loss function. The estimate is expressed as a linear combination of the population mean and individuals' sample mean, with the weight assigned to the sample mean referred to as the *credibility factor*. Bühlmann's Bayesian model for claim amounts laid the foundation for credibility theory and provided a robust framework for estimating future losses based on past observations.

Over time, credibility theory has evolved significantly and found extensive application in nonlife actuarial science in insurance companies (cf. Wen *et al.*, 2009; Gómez-Déniz, 2016; Yan and Song, 2022).

However, with the changing times and the development of the insurance industry and actuarial technologies, actuaries have encountered situations where the factors influencing risks in commercial insurance plans are highly complex (cf. Bauwelinckx and Goovaerts, 1990; Dhaene and Goovaerts, 1996; Yeo and Valdez, 2006; Deresa *et al.*, 2022). Interested readers can refer to some typical situations presented in Chapter 7 of Bühlmann and Gisler (2005). Therefore, theories on multidimensional credibility are in urgent need.

For the first time, Jewell (1973) introduced the concept of *multidimensional credibility*, making a decisive contribution to history of credibility theory. After that, Hachemeister (1975) introduced the credibility theory for regression models. These two works are believed as the pathbreaking contributions generalizing the credibility techniques into higher dimensions (cf. Bühlmann *et al.*, 2003). Besides, Jewell (1974) established *exact multivariate credibility*[1] model under certain conditions when the linear exponential family of likelihoods together with their natural conjugate priors are considered. Since then, credibility estimates for future losses in multivariate scenarios have been widely applied in premium pricing and liability reserve estimation in nonlife actuarial science. For example, Frees (2003) proposed theoretically optimal insurance prices using a multivariate credibility model and demonstrated significant economic differences in premium pricing methods with and without considering the covariance between different components. Englund *et al.* (2008) studied a generalization of the Bühlmann–Straub model, which allows for the age of claims to influence the estimation of future claims. Poon and Lu (2015) investigated a Bühlmann–Straub-type credibility model with dependence structure among risk parameters and conditional spatial cross-sectional dependence. Further generalizations of the multivariate credibility model can be found in Englund *et al.* (2009), Thuring *et al.* (2012), Pechon *et al.* (2021), Gómez-Déniz and Calderín-Ojeda (2018, 2021), Schinzinger *et al.* (2016), and related references.

However, the multivariate credibility estimates proposed by Jewell (1973) have a rather complex form due to the matrix-based credibility factor, making them challenging to understand and implement. Moreover, the credibility factor matrix in the traditional multivariate credibility model involves numerous structural parameters, resulting in an increasing number of unknown parameters as the dimension of the risk vector (denoted by p) increases. Consequently, as the dimension of the risk vector increases, it suffers from the issue of the "curse of dimensionality." In this context, when p is large,[2] these estimates become difficult to apply in actual insurance pricing processes.

To address these issues, this paper proposes a new multivariate credibility approach based on the joint distribution function. The study proceeds in two steps. First, instead of directly estimating the conditional mean of the risk random vector $\mathbf{Y} = (Y^{(1)}, \cdots, Y^{(p)})$ given the risk profile $\mathbf{\Theta}$, we focus on estimating the conditional joint distribution function, denoted as $F\left(y^{(1)}, \cdots, y^{(p)} | \mathbf{\Theta}\right)$, by constraining it to a linear combination of indicator functions labeled on the past claims. Utilizing the principles of credibility theory, the optimal linear Bayesian estimator of the conditional joint distribution function is then derived under the integral quadratic loss function. Second, employing the plug-in method[3], multidimensional credibility estimates for the hypothetical mean, process variance, and risk premiums under various premium principles are obtained. Compared to the traditional multivariate credibility estimator in Jewell (1973), the credibility factor in the proposed approach is no longer a matrix but a scalar. It is shown to be an increasing function of the sample size and satisfies the essential principles of credibility. Moreover, the credibility factor involved in the new estimator is easy to estimate and can be naturally applied in premium pricing for multidimensional risks. Overall, our proposed approach addresses the limitations

[1]The credibility estimator is said to be *exact* if the Bayesian forecast of the mean observation under the same setting is derived to be linear in the data.

[2]Even for $p = 2$, it has been a tedious task; see Example 2.1.

[3]In the context of plug-in method, operating under a parametric model $\mathcal{P} = \left\{P_\gamma ; \gamma \in \Gamma\right\}$, any real-valued characteristic τ of a particular member P_γ can be written as a mapping from the parameter space Γ, that is $\tau : \Gamma \mapsto \mathbb{R}$. If some observations $\mathbf{y}$ come from P_{γ_0} and some estimate $\hat{\gamma} \in \Gamma$ of γ_0 (e.g., by Maximum-Likelihood) has been derived, it is natural to use $\tau(\hat{\gamma})$ as an estimate of $\tau(\gamma_0)$. This method for constructing estimates is commonly referred to as the *plug-in principle*, since we "plug" the estimate $\hat{\gamma}$ into the mapping $\tau(\cdot)$. Relevant applications of this useful statistical method can be found in the literature such as Wu *et al.* (2013), Oliveira *et al.* (2012), and Laloë and Servien (2013).

of traditional multivariate credibility models and provides a more straightforward and computationally efficient solution for credibility estimation in high-dimensional scenarios.

It should be mentioned that, for the case of one dimension $p = 1$, Cai *et al.* (2015) applied the non-parametric method to develop the credibility estimation of distribution functions, which differs in nature compared with the present work. Our proposed multidimensional credibility estimator can be employed to predict premiums for aggregate claims in insurance portfolios, while the results in Cai *et al.* (2015) for the one-dimensional case only deal with a single business line.

The novel contribution of our study is summarized as follows:

(i) We focus on the credibility estimation of multivariate joint distribution functions and plug in the estimator into the hypothetical mean and process covariance matrix, generating a new multidimensional credibility model.

(ii) The traditional multidimensional credibility model introduced in Jewell (1973) suffers from severe difficulty in estimating the process covariance matrix (or even the population covariance matrix), rendering it impractical for use in insurance practice. Our proposed multidimensional credibility estimator overcomes this drawback and has the advantage of computational effectiveness, making it directly applicable to insurance premium prediction problems.

(iii) The proposed estimators can be applied for estimating various premium principles, which is in sharp contrast with the traditional multidimensional credibility that can only estimate the hypothetical mean.

The article is structured as follows: In Section 2, we present some pertinent notations and review the traditional multivariate credibility models. Additionally, we re-obtain the multivariate credibility estimation using the orthogonal projection method. In Section 3, we present the credibility estimation of the conditional joint distribution function and derive the corresponding credibility estimations for the hypothetical mean vector, process covariance matrix, and various premium principles and also study some statistical properties. In Section 4, we compare and demonstrate the performances of our proposed multidimensional credibility estimator and the traditional approach introduced by Jewell (1973). Finally, we analyze two real applications, one from an insurance company mentioned in Bühlmann and Gisler (2005) and the other from the R package "insuranceData." Section 5 concludes the paper. The proofs and derivations of the main results are delegated to the Supplementary Materials.

2. Preliminaries and review of traditional multivariate credibility models

Consider a portfolio of insurance policies with interdependence between risks. The losses associated with these risks are represented by a p-dimensional random vector $\mathbf{Y} = (Y^{(1)}, \ldots, Y^{(p)})'$, where p denotes the number of policies (or business lines). Given $\mathbf{\Theta} = \boldsymbol{\theta}$, the conditional joint distribution function of $\mathbf{Y}$ is denoted as:

$$F(\mathbf{y}|\boldsymbol{\theta}) = \mathbb{P}(\mathbf{Y} \le \mathbf{y}|\mathbf{\Theta} = \boldsymbol{\theta}) = \mathbb{P}\left(Y^{(1)} \le y^{(1)}, \cdots, Y^{(p)} \le y^{(p)}|\mathbf{\Theta} = \boldsymbol{\theta}\right),$$

where $\mathbf{\Theta}$, named as the *risk profile*, encompasses all factors influencing the risk of the insurance policy portfolio. Typically, due to the diversity of risks, $\mathbf{\Theta}$ is assumed to be a p-dimensional random vector, characterized by a certain prior density function $\pi(\cdot)$. The conditional mean (or hypothetical mean) and conditional (co-)variance (or process covariance) of the risk vector $\mathbf{Y}$ are defined as:

$$\boldsymbol{\mu}(\mathbf{\Theta}) = \mathbb{E}(\mathbf{Y}|\mathbf{\Theta}) \quad \text{and} \quad \Sigma(\mathbf{\Theta}) = \mathbb{V}(\mathbf{Y}|\mathbf{\Theta}).$$

For simplicity, let us denote

$$\boldsymbol{\mu}_0 = \mathbb{E}[\boldsymbol{\mu}(\mathbf{\Theta})], \quad T = \mathbb{V}[\boldsymbol{\mu}(\mathbf{\Theta})], \quad \Sigma_0 = \mathbb{E}[\Sigma(\mathbf{\Theta})]. \tag{2.1}$$

We assume that both Σ_0 and T are invertible matrices.[4]

[4]Otherwise, in the subsequent discussions, we can substitute with their generalized inverses.

https://doi.org/10.1017/asb.2024.13 Published online by Cambridge University Press

Since $\boldsymbol{\Theta}$ is unobserved, both $\boldsymbol{\mu}(\boldsymbol{\Theta})$ and $\Sigma(\boldsymbol{\Theta})$ are unobservable random vectors which require estimation based on available information. In the estimation process, we have access to two types of claim information. First, we have the prior information $\pi(\cdot)$ of the risk profile $\boldsymbol{\Theta}$, and second, we have past sample information $\mathbf{Y}_1, \cdots, \mathbf{Y}_n$ obtained by observing the population $\mathbf{Y}$. The approach of combining prior information and sample information to draw statistical inferences about $\boldsymbol{\mu}(\boldsymbol{\Theta})$ and $\Sigma(\boldsymbol{\Theta})$ is known as *Bayesian statistics*. The basic settings of Bayesian statistics models are described as follows.

Assumption 2.1. *Let* $\boldsymbol{\Theta}$ *be a random vector that characterizes the risk parameter of the random vector* $\boldsymbol{Y}$*, governed by a prior distribution function* $\pi(\cdot)$.

Assumption 2.2. *Given* $\boldsymbol{\Theta}$*, the random vector* $\boldsymbol{Y}_1, \cdots, \boldsymbol{Y}_n$ *consists of independent and identically distributed samples, where* $\boldsymbol{Y}_i = \left(Y_i^{(1)}, Y_i^{(2)}, \ldots, Y_i^{(p)}\right)'$*, for* $i = 1, \ldots, n$.

Following the traditional credibility theory, the estimation of $\boldsymbol{\mu}(\boldsymbol{\Theta})$ is constrained to linear functions of the observations. The optimal linear Bayesian estimator of $\boldsymbol{\mu}(\boldsymbol{\Theta})$ under the quadratic loss function can be obtained by solving

$$\min_{\mathbf{b}_0\in\mathbb{R}^p, \mathrm{B}_i\in\mathbb{R}^{p\times p}} \mathbb{E}\left[\left(\boldsymbol{\mu}(\boldsymbol{\Theta}) - \mathbf{b}_0 - \sum_{i=1}^{n} \mathrm{B}_i\mathbf{Y}_i\right)\left(\boldsymbol{\mu}(\boldsymbol{\Theta}) - \mathbf{b}_0 - \sum_{i=1}^{n} \mathrm{B}_i\mathbf{Y}_i\right)'\right]. \tag{2.2}$$

In contrast to the approach presented by Bühlmann and Gisler (2005), we employ the orthogonal projection method to solve (2.2) and derive the optimal linear Bayesian estimate of $\boldsymbol{\mu}(\boldsymbol{\Theta})$. The following lemma demonstrates the connection between the orthogonal projection and the optimal linear Bayesian estimation. The detailed proof can be found in Wen *et al.* (2009).

Lemma 2.1. *Let* $\begin{pmatrix} \boldsymbol{X}_{p\times 1} \\ \boldsymbol{Y}_{q\times 1}\end{pmatrix}$ *be a random vector with its expectation and covariance matrix given by* $\begin{pmatrix} \boldsymbol{\mu}_X \\ \boldsymbol{\mu}_Y\end{pmatrix}$ *and* $\begin{pmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY}\end{pmatrix}$*, respectively. Then the expected quadratic loss*

$$\mathbb{E}(\boldsymbol{Y} - A - B\boldsymbol{X})\,(\boldsymbol{Y} - A - B\boldsymbol{X})' \tag{2.3}$$

achieves its minimum in the nonnegative definite[5] *sense when*

$$A = \boldsymbol{\mu}_Y - \Sigma_{YX}\Sigma_{XX}^{-1}\boldsymbol{\mu}_X \quad \text{and} \quad B = \Sigma_{YX}\Sigma_{XX}^{-1}. \tag{2.4}$$

In light of Lemma 2.1, the optimal linear Bayesian estimation of the conditional mean vector can be obtained by solving the minimization problem (2.2), which can be expressed as follows:

$$\begin{aligned}\widehat{\boldsymbol{\mu}_{C,n}(\boldsymbol{\Theta})} &= \mathrm{Pro}\big(\boldsymbol{\mu}(\boldsymbol{\Theta})|\, L(\widetilde{\mathbf{Y}}, \mathbf{1})\big) \\ &= \mathbb{E}[\boldsymbol{\mu}(\boldsymbol{\Theta})] + \mathbb{Cov}\big(\boldsymbol{\mu}(\boldsymbol{\Theta}), \widetilde{\mathbf{Y}}\big)\, \mathbb{Cov}^{-1}\big(\widetilde{\mathbf{Y}}, \widetilde{\mathbf{Y}}\big)\big[\widetilde{\mathbf{Y}} - \mathbb{E}\big(\widetilde{\mathbf{Y}}\big)\big],\end{aligned} \tag{2.5}$$

where "Pro" signifies the projection onto the space of random vectors spanned by $\widetilde{\mathbf{Y}} = \left(\mathbf{Y}_1', \mathbf{Y}_2', \cdots, \mathbf{Y}_n'\right)'$ and $L(\widetilde{\mathbf{Y}}, \mathbf{1})$ is the linear subspace spanned by 1 and the components of $\widetilde{\mathbf{Y}}$. By combining Assumptions 2.1 and 2.2 of the Bayesian model for multivariate risk vectors, the following theorem, which was developed in Jewell (1973) and revisited in Bühlmann and Gisler (2005), can be obtained.

Theorem 2.1. *Under Assumptions 2.1 and 2.2, the optimal linear estimation of* $\boldsymbol{\mu}(\boldsymbol{\Theta})$ *is achieved by solving the minimization problem (2.2) and can be expressed as follows:*

$$\widehat{\boldsymbol{\mu}_{C,n}(\boldsymbol{\Theta})} = Z_{C,n}\overline{\boldsymbol{Y}} + \left(I_p - Z_{C,n}\right)\boldsymbol{\mu}_0, \tag{2.6}$$

[5]Henceforth, for two matrixes A and B, we write $A \geq B$ to denote that $A - B \geq 0$, that is, $A - B$ is a nonnegative definite matrix.

https://doi.org/10.1017/asb.2024.13 Published online by Cambridge University Press

where $\overline{\mathbf{Y}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{Y}_i$ *represents the sample mean vector, and the weight matrix*

$$Z_{C,n} = nT(nT + \Sigma_0)^{-1} \tag{2.7}$$

is a $p \times p$ *matrix.*[6] *Here, the capital letter "C" inside the credibility factor* $Z_{C,n}$ *stands for the word "classical."*

As per Theorem 2.1, the optimal linear Bayesian estimator of the vector of hypothetical mean $\boldsymbol{\mu}(\boldsymbol{\Theta})$ is a weighted average of the sample mean $\overline{\mathbf{Y}}$ and the prior mean $\boldsymbol{\mu}_0$. The weight matrix $Z_{C,n}$ has the following properties:

$$\lim_{n\to\infty} Z_{C,n} = I_p \quad \textit{and} \quad \lim_{n\to 0} Z_{C,n} = 0_{p\times p},$$

where I_p is the p-dimensional identity matrix and $0_{p\times p}$ is the zero matrix with p rows and p columns. Note that $Z_{C,n}$ is usually referred to as the *credibility factor matrix*, indicating the influence of the sample mean $\overline{\mathbf{Y}}$ in the optimal linear estimation $\widehat{\boldsymbol{\mu}_{C,n}(\boldsymbol{\Theta})}$.

In practical insurance premium pricing problems, computing the credibility factor matrix $Z_{C,n}$ can be very tedious and challenging since it involves with much more complexity of matrix inversion when p is large, hindering the usage of the classical multidimentional credibility approach for dealing with real problems. The following two examples illustrate this point.

Example 2.1. *Given* $\boldsymbol{\Theta}$, *we assume that the random vectors* $\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_n$ *are independent and identically distributed as Bivariate normal distribution* $N_2(\boldsymbol{\Theta}, \Sigma_0)$, *where* $\mathbf{Y}_i = (Y_i^{(1)}, Y_i^{(2)})$ *is a two-dimensional random vector. The prior distribution of the vector of risk parameters* $\boldsymbol{\Theta}$ *is* $N_2(\boldsymbol{\mu}_0, T)$, *where*

$$T = \begin{pmatrix} \tau_1^2 & \nu_1 \\ \nu_1 & \tau_2^2 \end{pmatrix} \quad \textit{and} \quad \Sigma_0 = \begin{pmatrix} \sigma_1^2 & \nu_2 \\ \nu_2 & \sigma_2^2 \end{pmatrix}.$$

Let $A = nT + \Sigma_0$. *Then the inverse of* A *can be computed using matrix algebra as:*

$$A^{-1} = \frac{A^*}{|A|} = \frac{1}{\left(n\tau_1^2 + \sigma_1^2\right)\left(n\tau_2^2 + \sigma_2^2\right) - (n\nu_1 + \nu_2)^2} \begin{pmatrix} n\tau_2^2 + \sigma_2^2 & -(n\nu_1 + \nu_2) \\ -(n\nu_1 + \nu_2) & n\tau_1^2 + \sigma_1^2 \end{pmatrix}.$$

Using Equation (2.7), we can obtain the expression for $Z_{C,n}$ *as follows:*

$$Z_{C,n} = \begin{pmatrix} \dfrac{n\tau_1^2\left(n\tau_2^2 + \sigma_2^2\right) - n\nu_1\left(n\nu_1 + \nu_2\right)}{\left(n\tau_1^2 + \sigma_1^2\right)\left(n\tau_2^2 + \sigma_2^2\right) - (n\nu_1 + \nu_2)^2} & \dfrac{n\nu_1\sigma_1^2 - n\nu_2\tau_1^2}{\left(n\tau_1^2 + \sigma_1^2\right)\left(n\tau_2^2 + \sigma_2^2\right) - (n\nu_1 + \nu_2)^2} \\[2ex] \dfrac{n\nu_1\sigma_2^2 - n\nu_2\tau_2^2}{\left(n\tau_1^2 + \sigma_1^2\right)\left(n\tau_2^2 + \sigma_2^2\right) - (n\nu_1 + \nu_2)^2} & \dfrac{n\tau_2^2\left(n\tau_1^2 + \sigma_1^2\right) - n\nu_1\left(n\nu_1 + \nu_2\right)}{\left(n\tau_1^2 + \sigma_1^2\right)\left(n\tau_2^2 + \sigma_2^2\right) - (n\nu_1 + \nu_2)^2} \end{pmatrix}. \tag{2.8}$$

As seen in Example 2.1, the computation of $Z_{C,n}$ is manageable when $p = 2$; however, when $p \geq 3$, it becomes much more complex, as shown in the following example.

Example 2.2. *In accordance with Example 2.1, we proceed to compute the matrix of credibility factors for the scenario where* $p = 3$, *wherein the following representations ensue*

$$T = \begin{pmatrix} \tau_1^2 & \nu_1 & \nu_2 \\ \nu_1 & \tau_2^2 & \nu_3 \\ \nu_2 & \nu_3 & \tau_3^2 \end{pmatrix} \quad \textit{and} \quad \Sigma_0 = \begin{pmatrix} \sigma_1^2 & \nu_4 & \nu_5 \\ \nu_4 & \sigma_2^2 & \nu_6 \\ \nu_5 & \nu_6 & \sigma_3^2 \end{pmatrix}.$$

[6]Given that matrices T and Σ_0 are both positive definite matrices, it follows that

$$|Z_{C,n}| = \frac{n|T|}{|nT + \Sigma_0|} > 0,$$

indicating that $Z_{C,n}$ is also a positive definite matrix.

https://doi.org/10.1017/asb.2024.13 Published online by Cambridge University Press

The adjoint matrix of A is determined as:

$$A^* = \begin{pmatrix} \psi_{11} & \psi_{12} & \psi_{13} \\ \psi_{21} & \psi_{22} & \psi_{23} \\ \psi_{31} & \psi_{32} & \psi_{33} \end{pmatrix}, \tag{2.9}$$

where

$$\begin{aligned}
\psi_{11} &= \left(n\tau_2^2 + \sigma_2^2\right)\left(n\tau_3^2 + \sigma_3^2\right) - (\nu_3 + \nu_4)^2, \\
\psi_{12} = \psi_{21} &= (n\nu_2 + \nu_5)(n\nu_3 + \nu_6) - (n\nu_1 + \nu_4)\left(n\tau_3^2 + \sigma_3^2\right), \\
\psi_{13} = \psi_{31} &= (n\nu_1 + \nu_4)(n\nu_3 + \nu_6) - (n\nu_2 + \nu_5)\left(n\tau_2^2 + \sigma_2^2\right), \\
\psi_{22} &= \left(n\tau_1^2 + \sigma_1^2\right)\left(n\tau_3^2 + \sigma_3^2\right) - (n\nu_2 + \nu_5)^2, \\
\psi_{23} = \psi_{32} &= (n\nu_1 + \nu_4)(n\nu_2 + \nu_4) - \left(n\tau_1^2 + \sigma_1^2\right)(n\nu_3 + \nu_6), \\
\psi_{33} &= \left(n\tau_1^2 + \sigma_1^2\right)\left(n\tau_2^2 + \sigma_2^2\right) - (n\nu_1 + \nu_4)^2.
\end{aligned}$$

The determinant of matrix A is formulated as:

$$\begin{aligned}
|A| = {} & \left(n\tau_1^2 + \sigma_1^2\right)\left(\left(n\tau_2^2 + \sigma_2^2\right)\left(n\tau_3^2 + \sigma_3^2\right) - (n\nu_3 + \nu_6)^2\right) \\
& - (n\nu_1 + \nu_4)\left((n\nu_1 + \nu_4)\left(n\tau_3^2 + \sigma_3^2\right) - (n\nu_2 + \nu_5)(n\nu_3 + \nu_6)\right) \\
& + (n\nu_2 + \nu_5)\left((n\nu_1 + \nu_4)(n\nu_3 + \nu_6) - (n\nu_2 + \nu_5)\left(n\tau_2^2 + \sigma_2^2\right)\right).
\end{aligned}$$

The expression for the matrix of credibility factors is provided as follows:

$$Z_{C,n} = \frac{1}{|A|}\begin{pmatrix} Z_{11}^C & Z_{12}^C & Z_{13}^C \\ Z_{21}^C & Z_{22}^C & Z_{23}^C \\ Z_{31}^C & Z_{32}^C & Z_{33}^C \end{pmatrix}, \tag{2.10}$$

where the components are given by:

$$\begin{aligned}
Z_{11}^C &= n\tau_1^2\psi_{11} + n\nu_1\psi_{21} + n\nu_2\psi_{31}, & Z_{12}^C &= n\tau_1^2\psi_{12} + n\nu_1\psi_{22} + n\nu_2\psi_{32}, \\
Z_{13}^C &= n\tau_1^2\psi_{13} + n\nu_1\psi_{23} + n\nu_2\psi_{33}, & Z_{21}^C &= n\nu_1\psi_{11} + n\tau_2^2\psi_{21} + n\nu_3\psi_{31}, \\
Z_{22}^C &= n\nu_1\psi_{12} + n\tau_2^2\psi_{22} + n\nu_3\psi_{32}, & Z_{23}^C &= n\nu_1\psi_{13} + n\tau_2^2\psi_{23} + n\nu_3\psi_{33}, \\
Z_{31}^C &= n\nu_2\psi_{11} + n\nu_3\psi_{21} + n\tau_3^2\psi_{31}, & Z_{32}^C &= n\nu_2\psi_{12} + n\nu_3\psi_{22} + n\tau_3^2\psi_{32}, \\
Z_{33}^C &= n\nu_2\psi_{13} + n\nu_3\psi_{23} + n\tau_3^2\psi_{33}.
\end{aligned}$$

As noted from the above two examples, the traditional multivariate credibility estimation (2.6) may suffer certain shortcomings when compared to Bühlmann's univariate credibility estimation:

(i) **Complexity of computation**: When $p = 3$, the credibility factor matrix $Z_{C,n}$ is already quite complex. For higher dimensions with $p > 3$, the expression of $Z_{C,n}$ involves the inverse of high-order matrices, and there is no general explicit expression. This complexity makes it difficult for practitioners to understand and apply in practical situations.

(ii) **Difficulty in parameter estimation**: The traditional multivariate credibility estimation suffers from the issue of having many unknown structural parameters, which can be difficult to estimate accurately, especially as p increases. The large number of parameters can lead to overfitting and increase computation complexity, making it less useful for real-world applications. Even for $p = 2$, the expression of $Z_{C,n}$ contains six unknown parameters such as $\tau_1^2, \tau_2^2, \sigma_1^2, \sigma_2^2, \nu_1, \nu_2$. Exactly, the *dimensional disaster* problem refers to the challenge of estimating a large number of unknown parameters as the dimension p of the problem increases. In the case of traditional multivariate credibility estimation, the number of unknown parameters increases with

https://doi.org/10.1017/asb.2024.13 Published online by Cambridge University Press

the order $O(p^2+p)$ as p grows, which can quickly become overwhelming and computationally demanding, especially for high-dimensional problems.

(iii) **Lack of applicability under some premium principles**: Traditional multivariate credibility estimation $\widehat{\boldsymbol{\mu}_{C,n}(\boldsymbol{\Theta})}$ is limited in its applicability, particularly when considering some premium principles beyond the expected value principle. Since it only provides an estimation of the conditional mean $\boldsymbol{\mu}(\boldsymbol{\Theta})$, it cannot estimate other functionals of the risk parameter, such as the process covariance matrix $\Sigma(\boldsymbol{\Theta})$.

In conclusion, although the classical multivariate credibility estimation has good statistical properties and a mathematically sound form, it is less commonly used in practice due to the challenges of dimension disaster and computational complexity. In this regard, we propose a new multidimensional credibility model based on the credibility estimation of the conditional joint distribution function, which will be the main focus of the next section.

3. Main results

Note that $\boldsymbol{\mu}(\boldsymbol{\Theta})$ and $\boldsymbol{\Sigma}(\boldsymbol{\Theta})$ can be expressed as:

$$\boldsymbol{\mu}(\boldsymbol{\Theta}) = \int_{\mathbb{R}^p} \mathbf{y} dF(\mathbf{y}|\,\boldsymbol{\Theta}) \tag{3.1}$$

and

$$\boldsymbol{\Sigma}(\boldsymbol{\Theta}) = \int_{\mathbb{R}^p} \left[\mathbf{y}-\boldsymbol{\mu}(\boldsymbol{\Theta})\right]\left[\mathbf{y}-\boldsymbol{\mu}\,(\boldsymbol{\Theta})\right]' dF(\mathbf{y}|\,\boldsymbol{\Theta}), \tag{3.2}$$

respectively. Based on these expressions, we shall employ the basic idea of credibility theory to first estimate the conditional joint distribution function $F(\mathbf{y}|\,\boldsymbol{\Theta})$. Subsequently, the credibility estimations for $\boldsymbol{\mu}(\boldsymbol{\Theta})$, $\Sigma(\boldsymbol{\Theta})$, and the general risk premium $R(\boldsymbol{\Theta})$ will be derived using the plug-in technique.

3.1. Credibility estimator for the joint distribution function

Let us define the class of nonhomogeneous linear functions of the conditional joint distribution as:

$$\mathscr{L} = \left\{\alpha_0\,(\mathbf{y}) + \sum_{s=1}^{n} \alpha_s H_s(\mathbf{y})\right\}, \tag{3.3}$$

where $\alpha_0(\mathbf{y})$ is an unspecified multivariate function (in general representing the population information), and

$$H_s(\mathbf{y}) = I_{\{\mathbf{Y}\le\mathbf{y}\}} = \begin{cases} 1, & Y_s^{(1)} \le y^{(1)}, \cdots, Y_s^{(p)} \le y^{(p)} \\ 0, & \text{otherwise} \end{cases},$$

for $\mathbf{y} = \left(y^{(1)}, \cdots, y^{(p)}\right)'$ and $s = 1, \ldots, n$, where I_A represents the indicative function of event A, which is equal to 1 when A occurs, otherwise equal to 0. We aim to solve the following minimization problem:

$$\min_{\widehat{F}(\mathbf{y}|\boldsymbol{\Theta})\in\mathscr{L}} \int_{\mathbb{R}^p} \mathbb{E}\left[\left(F(\mathbf{y}|\,\boldsymbol{\Theta}) - \widehat{F}(\mathbf{y}|\,\boldsymbol{\Theta})\right)^2\right] d\mathbf{y}. \tag{3.4}$$

The solution to problem (3.4) is summarized in the following theorem.

Theorem 3.1. *In the Bayesian model settings for multivariate risks, as described by Assumptions 2.1 and 2.2, the optimal linear Bayesian estimator of the conditional joint distribution function $F(\mathbf{y}|\,\boldsymbol{\Theta})$ solving problem (3.4) is given by:*

$$\widehat{F}(\mathbf{y}|\,\boldsymbol{\Theta}) = Z_{N,n}F_n(\mathbf{y}) + (1-Z_{N,n})F_0(\mathbf{y}), \tag{3.5}$$

https://doi.org/10.1017/asb.2024.13 Published online by Cambridge University Press

where

$$F_n(\mathbf{y}) = \frac{1}{n}\sum_{i=1}^{n} H_i(\mathbf{y}) \tag{3.6}$$

and

$$F_0(\mathbf{y}) = \mathbb{E}\big[F(\mathbf{y}|\,\boldsymbol{\Theta})\big] = \int_{\mathbb{R}^p} F(\mathbf{y}|\,\boldsymbol{\theta})\,\pi(\boldsymbol{\theta})\,d\boldsymbol{\theta}$$

represent the empirical and aggregated joint distribution functions, respectively. The credibility factor $Z_{N,n}$[7] *is defined as:*

$$Z_{N,n} = \frac{n\tau_0^2}{n\tau_0^2 + \sigma_0^2}, \tag{3.7}$$

where

$$\tau_0^2 = \int_{\mathbb{R}^p} \mathbb{V}\big[F(\mathbf{y}|\,\boldsymbol{\Theta})\big]\,d\mathbf{y} \tag{3.8}$$

and

$$\sigma_0^2 = \int_{\mathbb{R}^p} \mathbb{E}\big[F(\mathbf{y}|\,\boldsymbol{\Theta})\,(1 - F(\mathbf{y}|\,\boldsymbol{\Theta}))\big]\,d\mathbf{y}. \tag{3.9}$$

We hereby designate $\widehat{F}(\mathbf{y}|\,\boldsymbol{\Theta})$ *as the credibility estimate of the conditional distribution function* $F(\mathbf{y}|\,\boldsymbol{\Theta})$ *and* $Z_{N,n}$ *as the credibility factor.*

Remark 3.1. Noted that in Equation (3.3), none constraints were imposed on class $\mathscr{L}$ for estimating distribution functions. It is easy to observe that solving problem (3.4) would be quite challenging if we introduce some constraints. Given $\boldsymbol{\Theta}$, due to the exchangeability of $\mathbf{Y}_1, \mathbf{Y}_2, \cdots, \mathbf{Y}_n$, it must hold that $\alpha_1 = \alpha_2 = \cdots = \alpha_n$ for the solution of (3.4). Therefore, the optimization problem (3.4) is actually equivalent to solving:

$$\min_{b\in\mathbb{R}} \int_{\mathbb{R}^p} \mathbb{E}\big[F(\mathbf{y}|\,\boldsymbol{\Theta}) - \alpha_0(\mathbf{y}) - bF_n(\mathbf{y})\big]^2\,d\mathbf{y}. \tag{3.10}$$

Note that

$$\begin{aligned}
&\mathbb{E}\big[F(\mathbf{y}|\,\boldsymbol{\Theta}) - \alpha_0(\mathbf{y}) - bF_n(\mathbf{y})\big]^2\\
&= \mathbb{V}\big[F(\mathbf{y}|\,\boldsymbol{\Theta}) - \alpha_0(\mathbf{y}) - bF_n(\mathbf{y})\big] + \mathbb{E}\big[F(\mathbf{y}|\,\boldsymbol{\Theta}) - \alpha_0(\mathbf{y}) - bF_n(\mathbf{y})\big]\\
&\geq \mathbb{V}\big[F(\mathbf{y}|\,\boldsymbol{\Theta}) - \alpha_0(\mathbf{y}) - bF_n(\mathbf{y})\big],
\end{aligned}$$

where the equality holds if and only if

$$\mathbb{E}\big[F(\mathbf{y}|\,\boldsymbol{\Theta}) - \alpha_0(\mathbf{y}) - bF_n(\mathbf{y})\big] = 0.$$

Therefore, $\alpha_0(\mathbf{y})$ should be chosen such that

$$\alpha_0(\mathbf{y}) = \mathbb{E}\big[F(\mathbf{y}|\,\boldsymbol{\Theta}) - bF_n(\mathbf{y})\big] = aF_0(\mathbf{y}),$$

where $a = 1 - b$. Thus, the optimization problem (3.4) is equivalent to finding the optimal estimate of the conditional distribution function $F(\mathbf{y}|\,\boldsymbol{\Theta})$ within the class:

$$\mathscr{L}^* = \{aF_0(\mathbf{y}) + bF_n(\mathbf{y})\}.$$

To ensure the candidates in $\mathscr{L}^*$ to be reasonable estimates, we need to impose the condition:

$$a \geq 0,\; b \geq 0,\; a + b = 1,$$

[7] Here, the capital letter "N" stands for the word "*new*."

https://doi.org/10.1017/asb.2024.13 Published online by Cambridge University Press

under which $aF_0(\mathbf{y}) + bF_n(\mathbf{y})$ remains a multivariate distribution function. To this regard, we should solve the following constrained optimization problem:

$$\begin{cases} \min\limits_{a,b\in\mathbb{R}} \displaystyle\int_{\mathbb{R}^p} \mathbb{E}\left[F(\mathbf{y}|\,\boldsymbol{\Theta}) - aF_0(\mathbf{y}) - bF_n(\mathbf{y})\right]^2 d\mathbf{y} \\ \text{s.t. } a \geq 0,\ b \geq 0,\ a+b=1 \end{cases}. \tag{3.11}$$

In accordance with Lemma 1 in Section 2 of the supplementary file, one can demonstrate that solving the constrained optimization problem (3.11) is equivalent to dealing with the unconstrained optimization problem (3.4).

Remark 3.2. For the case of one-dimensional claims, that is, $p = 1$, problem (3.4) can be written as:

$$\min_{\alpha_0(\cdot),\alpha_1,\ldots,\alpha_n\in\mathbb{R}} \int_{\mathbb{R}} \mathbb{E}\left[F(y|\,\Theta) - \alpha_0(y) - \sum_{s=1}^{n} \alpha_s I(Y_s \leq y)\right]^2 dy \tag{3.12}$$

yielding the credibility estimation of the univariate conditional distribution function as:

$$\widehat{F}(y|\,\Theta) = Z_{N,n}F_n(y) + (1 - Z_{N,n})F_0(y),$$

which is akin to the credibility estimation of the survival function stated in Equation (2.7) of Cai *et al.* (2015).

Remark 3.3. If the claims have discrete distributions, we only need to modify Equation (3.4) to

$$\min_{\alpha_0(\cdot),\alpha_1,\ldots,\alpha_n\in\mathbb{R}} \sum_{y} \mathbb{E}\left[p(\mathbf{y}|\,\boldsymbol{\Theta}) - \alpha_0(\mathbf{y}) - \sum_{s=1}^{n} \alpha_s I(\mathbf{Y}_s = \mathbf{y})\right]^2,$$

where $p(\mathbf{y}|\,\boldsymbol{\Theta}) = \mathbb{P}(\mathbf{Y} = \mathbf{y}|\,\boldsymbol{\Theta} = \boldsymbol{\theta})$. This yields the corresponding credibility estimation as:

$$\widehat{p}(\mathbf{y}|\,\boldsymbol{\Theta}) = Z_{N,n}p_n\,(\mathbf{y}) + \left(1 - Z_{N,n}\right) p_0\,(\mathbf{y}),$$

where $Z_{N,n}$ shares a resemblance with the results of Theorem 3.1, with the distinction of replacing the multivariate integration by multivariate summation.

It is noteworthy that the new credibility factor $Z_{N,n}$ is a deterministic function of sample size n, but independent of the samples. Additionally, both $F_n(\mathbf{y})$ and $F_0(\mathbf{y})$ conform to the fundamental properties of joint distribution functions. Consequently, it is evident that $\widehat{F}(\mathbf{y}|\,\boldsymbol{\Theta})$ possesses attributes such as nonnegativity, monotonicity, right-continuity for each component, and finite additivity. Moreover, it satisfies the following boundary conditions:

$$\lim_{y^{(1)}\to+\infty,\ldots,y^{(p)}\to+\infty} \left[\widehat{F}(\mathbf{y}|\,\boldsymbol{\Theta})\right] = 1$$

and

$$\lim_{y^{(i)}\to-\infty} \widehat{F}(\mathbf{y}|\,\boldsymbol{\Theta}) = 0 \text{ for some } i \in \{1, 2, \cdots, p\}.$$

Hence, the estimator $\widehat{F}(\mathbf{y}|\,\boldsymbol{\Theta})$ also qualifies as a joint distribution function. As per Theorem 3.1, the estimation $\widehat{F}(\mathbf{y}|\,\boldsymbol{\Theta})$ of the conditional joint distribution function can be elegantly represented as a weighted combination of the empirical joint distribution function $F_n(\mathbf{y})$ and the integrated joint distribution function $F_0(\mathbf{y})$.

The credibility factor $Z_{N,n}$ exhibits an increment with respect to the sample size n, and it fulfills the following asymptotic properties:

$$\lim_{n\to\infty} Z_{N,n} = 1 \quad \text{and} \quad \lim_{n\to 0} Z_{N,n} = 0, \tag{3.13}$$

which enjoys a similar expression with the univariate standard Bühlmann model. Consequently, when the sample size is substantial, greater emphasis is conferred upon the empirical distribution function $F_n(\mathbf{y})$, and vice versa.

https://doi.org/10.1017/asb.2024.13 Published online by Cambridge University Press

3.2. Credibility estimator for the hypothetical mean

Utilizing the credibility estimator $\widehat{F}(\mathbf{y}|\,\boldsymbol{\Theta})$, one can readily derive the credibility estimation for the hypothetical mean vector $\boldsymbol{\mu}(\boldsymbol{\Theta})$ by applying the plug-in technique. As a result, one has the following proposition.

Proposition 3.1. *By substituting the optimal linear estimate $\widehat{F}(\mathbf{y}|\,\boldsymbol{\Theta})$ of the distribution function into Equation (3.1), we obtain the credibility estimate of the conditional mean vector $\boldsymbol{\mu}(\boldsymbol{\Theta})$ as follows:*

$$\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})} = Z_{N,n}\overline{\mathbf{Y}} + (1 - Z_{N,n})\boldsymbol{\mu}_0,$$

where the credibility factor $Z_{N,n}$ is given by Equation (3.7).

It is evident that the estimation $\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}$ can be represented as a linear weighted average of the sample mean $\overline{\mathbf{Y}}$ and the population mean $\boldsymbol{\mu}_0$, where the weights satisfy the property (3.13). This form is very similar with the univariate case for classical Bühlmann model developed in Bühlmann (1967). However, compared with the traditional multivariate credibility estimator given in Theorem 2.1, the credibility factor $Z_{N,n}$ is a scalar, which reduces the computational complexity of traditional multivariate credibility model.

Given the expression of the credibility estimation $\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}$, we can calculate the conditional expectation of $\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}$ as:

$$\mathbb{E}\left[\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}\,\middle|\,\boldsymbol{\Theta}\right] = Z_{N,n}\mathbb{E}\left(\overline{\mathbf{Y}}\,\middle|\,\boldsymbol{\Theta}\right) + (1 - Z_{N,n})\boldsymbol{\mu}_0 = Z_{N,n}\boldsymbol{\mu}(\boldsymbol{\Theta}) + (1 - Z_{N,n})\boldsymbol{\mu}_0.$$

By using the formula for double expectation, we can further find

$$\mathbb{E}\left[\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}\right] = \mathbb{E}\left[\mathbb{E}\left(\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}\,\middle|\,\boldsymbol{\Theta}\right)\right] = Z_{N,n}\mathbb{E}[\boldsymbol{\mu}(\boldsymbol{\Theta})] + (1 - Z_{N,n})\boldsymbol{\mu}_0 = \boldsymbol{\mu}_0.$$

In the average sense, $\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}$ is an unconditional unbiased estimation of $\mathbb{E}[\boldsymbol{\mu}(\boldsymbol{\Theta})]$, which is described as the following proposition.

Proposition 3.2. *The credibility estimation $\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}$ is an unbiased estimator of the hypothetical mean vector $\boldsymbol{\mu}(\boldsymbol{\Theta})$, that is,*

$$\mathbb{E}\left[\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}\right] = \mathbb{E}[\boldsymbol{\mu}(\boldsymbol{\Theta})] = \boldsymbol{\mu}_0.$$

As the estimation of $\boldsymbol{\mu}(\boldsymbol{\Theta})$, the mean quadratic error matrix of the credibility estimator $\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}$ is given below.

Proposition 3.3. *The mean quadratic error matrix of the credibility estimator $\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}$ is given by:*

$$\mathbb{E}\left[\left(\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})} - \boldsymbol{\mu}(\boldsymbol{\Theta})\right)\left(\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})} - \boldsymbol{\mu}(\boldsymbol{\Theta})\right)'\right] = \frac{Z_{N,n}^2}{n}\Sigma_0 + (1 - Z_{N,n})^2 T.$$

Moreover, as $n \to \infty$, the mean quadratic error matrix tends to a $p \times p$ zero matrix:

$$\mathbb{E}\left[\left(\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})} - \boldsymbol{\mu}(\boldsymbol{\Theta})\right)\left(\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})} - \boldsymbol{\mu}(\boldsymbol{\Theta})\right)'\right] \to 0_{p\times p},$$

and hence

$$\mathbb{E}\left[\left|\left|\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})} - \boldsymbol{\mu}(\boldsymbol{\Theta})\right|\right|_{\xi}^2\right] = \mathbb{E}\left[\boldsymbol{\xi}'\left(\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})} - \boldsymbol{\mu}(\boldsymbol{\Theta})\right)\left(\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})} - \boldsymbol{\mu}(\boldsymbol{\Theta})\right)'\boldsymbol{\xi}\right] \to 0.$$

where $||\cdot||_{\xi}^2$ represents the weighted Frobenius norm,[8] *$\boldsymbol{\xi} = \left(\xi_1, \xi_2, \cdots, \xi_p\right)'$ represents the weight vector, and $\xi_i \geq 0$, for $i = 1, 2, \cdots, p$ such that $\sum_{i=1}^{p} \xi_i = 1$.*

[8] See details in Belitskii and Lyubich (2013). Henceforth, we shall simplify the notation as the weighted F-norm in the following text.

https://doi.org/10.1017/asb.2024.13 Published online by Cambridge University Press

3.3. Credibility estimator for the process variance

Note that

$$\Sigma(\boldsymbol{\Theta})=\int_{\mathbb{R}^p}\left[\left(\mathbf{y}-\int_{\mathbb{R}^p}\mathbf{y}dF(\mathbf{y}|\,\boldsymbol{\Theta})\right)\left(\mathbf{y}-\int_{\mathbb{R}^p}\mathbf{y}dF(\mathbf{y}|\,\boldsymbol{\Theta})\right)'\right]dF(\mathbf{y}|\,\boldsymbol{\Theta})\,.$$

We can first achieve the credibility estimator of the conditional variance based on $\widehat{F}(\mathbf{y}|\,\boldsymbol{\Theta})$ and then obtain the credibility estimator of $\Sigma(\boldsymbol{\Theta})$. For this purpose, let

$$U_n=\frac{1}{n}\sum_{i=1}^{n}\mathbf{Y}_i\mathbf{Y}_i' \quad \text{and} \quad U_0=\int_{\mathbb{R}^p}\mathbf{y}\mathbf{y}'dF_0(\mathbf{y})\,.$$

The following proposition can be reached.

Proposition 3.4. *Under the Bayesian model settings stated in Assumptions 2.1 and 2.2, replacing $F(\boldsymbol{y}|\,\boldsymbol{\Theta})$ with $\widehat{F}(\boldsymbol{y}|\,\boldsymbol{\Theta})$ in Equation (3.2), the credibility estimation of the conditional covariance matrix $\Sigma(\boldsymbol{\Theta})$ is given by:*

$$\widehat{\Sigma_{N,n}(\boldsymbol{\Theta})}=\omega_{1,n}\Sigma_n+\omega_{2,n}\Sigma_0+(1-\omega_{1,n}-\omega_{2,n})M_0,$$

where $\Sigma_0=\mathbb{E}[\Sigma(\boldsymbol{\Theta})]$, $\Sigma_n=\frac{1}{n}\sum_{i=1}^{n}\left(\boldsymbol{Y}_i-\overline{\boldsymbol{Y}}\right)\left(\boldsymbol{Y}_i-\overline{\boldsymbol{Y}}\right)'$ is the sample covariance matrix, and

$$\omega_{1,n}=Z_{N,n}^2,\quad \omega_{2,n}=(1-Z_{N,n})^2,\quad M_0=\frac{1}{2}\left(U_n-\overline{\boldsymbol{Y}}\boldsymbol{\mu}_0'-\boldsymbol{\mu}_0\overline{\boldsymbol{Y}}'+U_0\right).$$

Notably, the credibility estimation of the conditional covariance matrix is consisting of three parts. According to Proposition 3.4 and property (3.13), it holds that $\omega_{1,n}\to 1$ and $\omega_{2,n}\to 0$ as $n\to\infty$. Therefore, when n is large, the credibility estimate $\widehat{\Sigma_{N,n}(\boldsymbol{\Theta})}$ assigns more weight to sample covariance matrix Σ_n. Conversely, more weight is assigned to Σ_0 and M_0 when the sample size is smaller.

3.4. Statistical properties of the estimators

The consistency and asymptotic normality of $\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}$ and $\widehat{\Sigma_{N,n}(\boldsymbol{\Theta})}$ are given in the following two theorems.

Theorem 3.2. *Given $\boldsymbol{\Theta}$, the estimators $\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}$ and $\widehat{\Sigma_{N,n}(\boldsymbol{\Theta})}$ are strongly consistent with respect to the conditional mean vector $\boldsymbol{\mu}(\boldsymbol{\Theta})$ and conditional covariance matrix $\Sigma(\boldsymbol{\Theta})$, respectively. That is, when $n\to\infty$, we have*

$$\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}\to\boldsymbol{\mu}(\boldsymbol{\Theta}),\ a.s.\quad \text{and}\quad \widehat{\Sigma_{N,n}(\boldsymbol{\Theta})}\to\Sigma(\boldsymbol{\Theta}),\ a.s.$$

Theorem 3.3. *Given $\boldsymbol{\Theta}$, the estimator $\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}$ is asymptotically normal when $n\to\infty$, that is,*

$$\sqrt{n}\left(\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}-\boldsymbol{\mu}(\boldsymbol{\Theta})\right)\xrightarrow{L}N(0,\Sigma(\boldsymbol{\Theta})),$$

where "$\xrightarrow{L}$" means convergence in law/distribution.

Theorems 3.2 and 3.3 demonstrate that $\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}$ and $\widehat{\Sigma_{N,n}(\boldsymbol{\Theta})}$ serve as good estimators of the conditional mean vector $\boldsymbol{\mu}(\boldsymbol{\Theta})$ and conditional covariance matrix $\Sigma(\boldsymbol{\Theta})$. Both of them can be plugged in any continuous statistical functionals of these two quantities.

Next, we proceed with a theoretical comparison on the mean of the weighted F-norms between our new credibility estimation $\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}$ and the traditional credibility estimation $\widehat{\boldsymbol{\mu}_{C,n}(\boldsymbol{\Theta})}$. To simplify the analysis, we first assume that $p=2$ and introduce the following notations:

$$\boldsymbol{\mu}(\boldsymbol{\Theta})=\mathbb{E}[\mathbf{Y}_1|\boldsymbol{\Theta}]=\begin{pmatrix}\mu_1(\Theta)\\ \mu_2(\Theta)\end{pmatrix},\quad \Sigma(\boldsymbol{\Theta})=\mathbb{V}[\mathbf{Y}_1|\boldsymbol{\Theta}]=\begin{pmatrix}\sigma_1^2(\Theta) & \nu_2(\Theta)\\ \nu_2(\Theta) & \sigma_2^2(\Theta)\end{pmatrix}$$

https://doi.org/10.1017/asb.2024.13 Published online by Cambridge University Press

and

$$\boldsymbol{\mu}_0 = \mathbb{E}(\boldsymbol{\mu}(\boldsymbol{\Theta})) = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad T = \mathbb{V}(\boldsymbol{\mu}(\boldsymbol{\Theta})) = \begin{pmatrix} \tau_1^2 & \nu_1 \\ \nu_1 & \tau_2^2 \end{pmatrix}, \quad \Sigma_0 = \mathbb{E}(\Sigma(\boldsymbol{\Theta})) = \begin{pmatrix} \sigma_1^2 & \nu_2 \\ \nu_2 & \sigma_2^2 \end{pmatrix}.$$

where $\nu_1 = \rho_1 \tau_1 \tau_2$, $\nu_2 = \rho_2 \sigma_1 \sigma_2$, and

$$\rho_1 = \frac{\mathbb{C}\mathrm{ov}(\mu_1(\Theta), \mu_2(\Theta))}{\sqrt{\mathbb{V}(\mu_1(\Theta))\,\mathbb{V}(\mu_2(\Theta))}}, \quad \rho_2 = \frac{\mathbb{E}\left[\mathbb{C}\mathrm{ov}\left(Y_1^{(1)}, Y_1^{(2)} \middle| \boldsymbol{\Theta}\right)\right]}{\sqrt{\mathbb{E}\left[\mathbb{V}\left(Y_1^{(1)} \middle| \Theta_1\right)\right] \mathbb{E}\left[\mathbb{V}\left(Y_1^{(2)} \middle| \Theta_2\right)\right]}}. \tag{3.14}$$

Here, ρ_1 and ρ_2 represent the two correlation coefficients denoting the inter-dependence among risks and the intra-dependence within each risk, respectively.

Proposition 3.5. *Consider $p = 2$ and $\boldsymbol{\xi} = (\xi_1, \xi_2)'$. Based on Theorem 2.1 and Proposition 3.1, the mean of the weighted F-norms for the error of estimates $\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}$ and $\widehat{\boldsymbol{\mu}_{C,n}(\boldsymbol{\Theta})}$ can be expressed as:*

$$\mathbb{E}\left[\left|\left|\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})} - \boldsymbol{\mu}(\boldsymbol{\Theta})\right|\right|_{\boldsymbol{\xi}}^2\right] = \frac{n\tau_0^4\left(\xi_1\sigma_1^2 + \xi_2\sigma_2^2\right)}{\left(n\tau_0^2 + \sigma_0^2\right)^2} + \frac{\sigma_0^4\left(\xi_1\tau_1^2 + \xi_2\tau_2^2\right)}{\left(n\tau_0^2 + \sigma_0^2\right)^2} \tag{3.15}$$

and

$$\mathbb{E}\left[\left|\left|\widehat{\boldsymbol{\mu}_{C,n}(\boldsymbol{\Theta})} - \boldsymbol{\mu}(\boldsymbol{\Theta})\right|\right|_{\boldsymbol{\xi}}^2\right] = \Delta_1 + \Delta_2 + \Delta_3 + \Delta_4, \tag{3.16}$$

respectively, where

$$\Delta_1 = \frac{n\delta_1\left(\xi_1\left(\tau_1^2\delta_2 - \nu_1\delta_3\right)^2 + \xi_2\left(\nu_1\sigma_1^2 - \nu_2\tau_1^2\right)^2\right)}{\left(\delta_1\delta_2 - \delta_3^2\right)^2},$$

$$\Delta_2 = \frac{n\delta_2\left(\xi_1\left(\tau_2^2\delta_1 - \nu_1\delta_3\right)^2 + \xi_2\left(\nu_1\sigma_2^2 - \nu_2\tau_2^2\right)^2\right)}{\left(\delta_1\delta_2 - \delta_3^2\right)^2},$$

$$\Delta_3 = \frac{1}{(-2\nu_1)^{-1}\left(\delta_1\delta_2 - \delta_3^2\right)^2} \times \Big[n\nu_1\sigma_1^2\sigma_2^2(\xi_2\delta_1 + \xi_1\delta_2) - n\nu_2\left(\xi_2\tau_1^2\sigma_2^2\delta_1\right.$$
$$\left.+\xi_1\tau_2^2\sigma_1^2\delta_2 - \delta_3\left(\nu_2(\xi_2\tau_1^2 + \xi_1\tau_2^2) - \nu_1\left(\xi_2\sigma_1^2 + \xi_1\sigma_2^2\right)\right)\right)\Big],$$

$$\Delta_4 = \frac{\xi_2\left(\sigma_2^2 - n\tau_2^2\right)\delta_1\tau_2^2 + \xi_1\left(\sigma_1^2 - n\tau_1^2\right)\delta_2\tau_1^2 - (\nu_2 - n\nu_1)\left(\xi_1\tau_1^2 + \xi_2\tau_2^2\right)\delta_3}{\delta_1\delta_2 - \delta_3^2}$$

and

$$\delta_1 = n\tau_1^2 + \sigma_1^2, \quad \delta_2 = n\tau_2^2 + \sigma_2^2, \quad \delta_3 = n\nu_1 + \nu_2.$$

Compared to the traditional multivariate credibility estimation $\widehat{\boldsymbol{\mu}_{C,n}(\boldsymbol{\Theta})}$, the mean of the weighted F-norms for the error of the new credibility estimation $\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}$ is more straightforward to compute. The main reason is that the credibility factor $Z_{C,n}$ in the traditional model is a matrix, while in our estimation $\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}$, the corresponding credibility factor $Z_{N,n}$ is a scalar. This scalar form simplifies the computation process and enhances user-friendliness and interpretability.

Indeed, according to the result of Proposition 3.5, when $p > 2$, the mean of the weighted F-norms for the error of $\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}$ can be calculated as:

$$\mathbb{E}\left[\left|\left|\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})} - \boldsymbol{\mu}(\boldsymbol{\Theta})\right|\right|_{\boldsymbol{\xi}}^2\right] = \mathbb{E}\left[\xi_1\left(\widehat{\mu_{N,n}^{(1)}(\Theta)} - \mu_1(\Theta)\right)^2\right] + \cdots + \mathbb{E}\left[\xi_p\left(\widehat{\mu_{N,n}^{(p)}(\Theta)} - \boldsymbol{\mu}_p(\Theta)\right)^2\right]$$
$$= \frac{n\tau_0^4\left(\xi_1\sigma_1^2 + \cdots + \xi_p\sigma_p^2\right)}{\left(n\tau_0^2 + \sigma_0^2\right)^2} + \frac{\sigma_0^4\left(\xi_1\tau_1^2 + \cdots + \xi_p\tau_p^2\right)}{\left(n\tau_0^2 + \sigma_0^2\right)^2}.$$

The traditional multivariate credibility estimation $\widehat{\boldsymbol{\mu}_{C,n}(\boldsymbol{\Theta})}$ minimizes problem (2.2), and thus, in the sense that the matrix $A - B \leq 0$ is equivalent to $B - A$ being a positive definite matrix, it is evident that

$$\mathbb{E}\left[\left(\widehat{\boldsymbol{\mu}_{C,n}(\boldsymbol{\Theta})} - \boldsymbol{\mu}(\boldsymbol{\Theta})\right)\left(\widehat{\boldsymbol{\mu}_{C,n}(\boldsymbol{\Theta})} - \boldsymbol{\mu}(\boldsymbol{\Theta})\right)'\right] \leq \mathbb{E}\left[\left(\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})} - \boldsymbol{\mu}(\boldsymbol{\Theta})\right)\left(\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})} - \boldsymbol{\mu}(\boldsymbol{\Theta})\right)'\right].$$

In other words, $\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}$ is dominated by $\widehat{\boldsymbol{\mu}_{C,n}(\boldsymbol{\Theta})}$ in terms of the mean squared loss, which will be illustrated in Section 4 by two numerical examples. However, it would be interesting to see that our method perform competitively with the traditional one when the sample size is moderate or large.

For the case of $p = 2$, a comparison of the mean of the weighted F-norms for the error for both credibility estimations leads to the following conclusion.

Proposition 3.6. *In Equation (3.14), if $\rho_1 = \rho_2 = 0$ (i.e., $\nu_1 = \nu_2 = 0$), then we have*

$$\mathbb{E}\left[\left|\left|\widehat{\boldsymbol{\mu}_{C,n}(\boldsymbol{\Theta})} - \boldsymbol{\mu}(\boldsymbol{\Theta})\right|\right|_{\boldsymbol{\xi}}^2\right] \leq \mathbb{E}\left[\left|\left|\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})} - \boldsymbol{\mu}(\boldsymbol{\Theta})\right|\right|_{\boldsymbol{\xi}}^2\right]. \tag{3.17}$$

In accordance with Proposition 3.6, it can be readily demonstrated that when $p > 2$ and $\rho_1 = \cdots = \rho_p = 0$, the p-dimensional random vector $\mathbf{Y}$ demonstrates pairwise independence among its components. The mean of the weighted F-norm for the error in traditional multivariate credibility estimation is smaller than that for the error in the new estimator. However, when the correlation coefficients $\rho_i \neq 0$ for some $i = 1, \cdots, p$, the expression for the mean of the weighted F-norm for the error in traditional credibility estimation becomes significantly intricate, rendering a direct comparison of their magnitudes. In the next section dedicated to numerical simulations, we will employ Monte Carlo methods to perform the comparison.

3.5. Application in various premium principles

In nonlife insurance, premium calculation principles are real functionals mapping risk variables to the set of nonnegative real numbers. For a p-dimensional risk $\mathbf{Y}$ with conditional distribution function $F(\mathbf{y}|\ \boldsymbol{\Theta})$, consider the aggregate risk in for such insurance portfolio:

$$Z = a_1Y^{(1)} + \cdots + a_pY^{(p)} = \mathbf{a}'\mathbf{Y},$$

where $\mathbf{Y} = \left(Y^{(1)}, Y^{(2)}, \cdots, Y^{(p)}\right)'$ and $\mathbf{a} = \left(a_1, \cdots, a_p\right)'$ are the weights for these business lines such that $a_i \geq 0$, $i = 1, 2, \cdots, p$. Guerra and Centeno (2010) introduced variance-related premium principle[9] of risk Z with its expression given by:

$$H_Z(\boldsymbol{\Theta}) = (1 + \lambda_1)\,\mathbb{E}(Z|\boldsymbol{\Theta}) + g(\mathbb{V}(Z|\boldsymbol{\Theta})) = (1 + \lambda_1)\,\mathbf{a}'\boldsymbol{\mu}\,(\boldsymbol{\Theta}) + g\big(\mathbf{a}'\Sigma(\boldsymbol{\Theta})\,\mathbf{a}\big)\,, \tag{3.18}$$

where $\lambda_1 \geq 0$ signifies the nonnegative safety loading factor for the expectation and $g: \mathbb{R}_+ \mapsto \mathbb{R}_+$ is an increasing function such that $g(0) = 0$.

[9]Here, we replace the pure expectation in the expression introduced in Guerra and Centeno (2010) with the expected value premium principle.

https://doi.org/10.1017/asb.2024.13 Published online by Cambridge University Press

By substituting $\boldsymbol{\mu}(\boldsymbol{\Theta})$ and $\Sigma(\boldsymbol{\Theta})$ with their credibility estimators $\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}$ and $\widehat{\Sigma_{N,n}(\boldsymbol{\Theta})}$, respectively, a credibility estimation of $H_Z(\boldsymbol{\Theta})$ is achieved as:

$$\widehat{H_Z(\boldsymbol{\Theta})} = (1+\lambda_1)\,\mathbf{a}'\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})} + g\left(\mathbf{a}'\widehat{\Sigma_{N,n}(\boldsymbol{\Theta})}\mathbf{a}\right).$$

It is easy to show that the estimation $\widehat{H_Z(\boldsymbol{\Theta})}$ is consistent estimator for risk premium $H_Z(\boldsymbol{\Theta})$ according to Theorem 3.2 and the continuity of function g. Based on the premium principle given in Equation (3.18), we obtain credibility estimates of risk premiums under some commonly used actuarial premium principles as follows:

- If we set $g(x) = 0$ and $\lambda_1 \geq 0$, then

$$\widehat{H_Z^{(1)}(\boldsymbol{\Theta})} = (1+\lambda_1)\,\mathbf{a}'\left[Z_{N,n}\overline{\mathbf{Y}} + (1-Z_{N,n})\boldsymbol{\mu}_0\right] \tag{3.19}$$

  is the credibility estimation of risk premium under the expected value premium principle.
- If we set $g(x) = \lambda_2 \cdot x$ and $\lambda_1 = 0$, where $\lambda_2 \geq 0$ signifies the nonnegative safety loading factor for the variance term, then

$$\begin{aligned}\widehat{H_Z^{(2)}(\boldsymbol{\Theta})} &= \mathbf{a}'\left[Z_{N,n}\overline{\mathbf{Y}} + (1-Z_{N,n})\boldsymbol{\mu}_0\right]\\ &\quad + \lambda_2\cdot\mathbf{a}'\left[\omega_{1,n}\Sigma_n + \omega_{2,n}\Sigma_0 + (1-\omega_{1,n}-\omega_{2,n})M_0\right]\mathbf{a}\end{aligned} \tag{3.20}$$

  is the credibility estimation of risk premium under the variance premium principle.
- If we set $g(x) = \lambda_2 \cdot \sqrt{x}$ and $\lambda_1 = 0$, then we derive the credibility estimation of risk premium under the standard deviation premium principle as:

$$\begin{aligned}\widehat{H_Z^{(3)}(\boldsymbol{\Theta})} &= \mathbf{a}'\left[Z_{N,n}\overline{\mathbf{Y}} + (1-Z_{N,n})\boldsymbol{\mu}_0\right]\\ &\quad + \lambda_2\cdot\sqrt{\mathbf{a}'\left(\omega_{1,n}\Sigma_n + \omega_{2,n}\Sigma_0 + (1-\omega_{1,n}-\omega_{2,n})M_0\right)\mathbf{a}}.\end{aligned} \tag{3.21}$$

Of course, we can also obtain the credibility estimation of the exponential premium principle charged for Z as follows:

- By plug in the new credibility estimation (3.5), the credibility estimation of risk premium under the exponential premium principle is given by:

$$\begin{aligned}\widehat{H_Z^{(4)}(\boldsymbol{\Theta})} &= \frac{1}{\beta}\log\left[\int_{\mathbb{R}^p} e^{\beta\mathbf{a}'\mathbf{y}}d\widehat{F}(\mathbf{y}|\,\boldsymbol{\Theta})\right]\\ &= \frac{1}{\beta}\log\left[Z_{N,n}\int_{\mathbb{R}^p} e^{\beta\mathbf{a}'\mathbf{y}}dF_n(\mathbf{y}) + \left(1-Z_{N,n}\right)\int_{\mathbb{R}^p} e^{\beta\mathbf{a}'\mathbf{y}}dF_0(\mathbf{y})\right]\\ &= \frac{1}{\beta}\log\left[Z_{N,n}L_n + (1-Z_{N,n})L_0\right],\end{aligned} \tag{3.22}$$

  where $\beta > 0$, $L_n = \frac{1}{n}\sum_{i=1}^n \exp(\beta\mathbf{a}'\mathbf{Y}_i)$, and $L_0 = \mathbb{E}\left[\exp(\beta\mathbf{a}'\mathbf{Y}_1)\right]$.

Clearly, compared to traditional multivariate credibility estimates, our estimations have broader applicability, as they can be naturally applied to most premium principles in nonlife insurance and provide corresponding credibility estimators based on the credibility estimate of conditional distribution function.

4. Numerical simulations and case analysis

In this section, we first conduct two numerical simulations to compare the conditional and unconditional mean of the weighted F-norms for the errors of the traditional estimator and our proposed estimator. We then calculate predicted values of various premium principles by applying the new credibility estimator to two insurance scenarios.

4.1. Bivariate exponential-gamma model

Assuming that, for a given $\mathbf{\Theta}=\boldsymbol{\theta}$, where $\mathbf{\Theta}=(\Theta_1,\Theta_2)'$ and $\boldsymbol{\theta}=(\theta_1,\theta_2)'$, $\mathbf{Y}_1,\mathbf{Y}_2,\cdots,\mathbf{Y}_n$ are independently and identically distributed according to the bivariate exponential distribution, with the cumulative joint distribution function given by:

$$F\left(y^{(1)},y^{(2)}\,\middle|\,\mathbf{\Theta}=\boldsymbol{\theta}\right)=1-\exp\left\{-\frac{y^{(1)}}{\theta_1}\right\}-\exp\left\{-\frac{y^{(2)}}{\theta_2}\right\}+\exp\left\{-\frac{y^{(1)}}{\theta_1}-\frac{y^{(2)}}{\theta_2}-\frac{\rho_2}{\theta_1\theta_2}y^{(1)}y^{(2)}\right\},$$

and the corresponding probability density function is

$$f\left(y^{(1)},y^{(2)}\,\middle|\,\mathbf{\Theta}=\boldsymbol{\theta}\right)=\left[-\frac{\rho_2}{\theta_1\theta_2}+\left(\frac{1}{\theta_1}+\frac{\rho_2}{\theta_1\theta_2}y^{(2)}\right)\left(\frac{1}{\theta_2}+\frac{\rho_2}{\theta_1\theta_2}y^{(1)}\right)\right]e^{-\frac{y^{(1)}}{\theta_1}-\frac{y^{(2)}}{\theta_2}-\frac{\rho_2}{\theta_1\theta_2}y^{(1)}y^{(2)}},$$

where $y^{(1)},y^{(2)}\geq 0$ and $\rho_2\in[-1,1]$.

Further, we assume that θ_1 and θ_2 are independent of each other and follow gamma prior with density functions:

$$\pi_{\Theta_1}(\theta_1)=\frac{\beta_1^{\alpha_1}}{\Gamma(\alpha_1)}\theta_1^{\alpha_1-1}e^{-\beta_1\theta_1},\qquad \theta_1,\alpha_1,\beta_1>0,$$

and

$$\pi_{\Theta_2}(\theta_2)=\frac{\beta_2^{\alpha_2}}{\Gamma(\alpha_2)}\theta_2^{\alpha_2-1}e^{-\beta_2\theta_2},\qquad \theta_2,\alpha_2,\beta_2>0$$

respectively.

Set $\alpha_1=2.4$, $\alpha_2=3.2$, $\beta_1=5.4$, $\beta_2=2.8$, and $\rho_2=0.2$. By utilizing Equations (2.1), (3.8), and (3.9) and using the Monte Carlo integration method,[10] the values for various structural parameters are approximately obtained as follows:

$$\boldsymbol{\mu}_0\approx\begin{pmatrix}0.0263\\0.0424\end{pmatrix},\quad T\approx\begin{pmatrix}0.0015&0\\0&0.0017\end{pmatrix},\quad \Sigma_0\approx\begin{pmatrix}0.0133&0.0043\\0.0043&0.0355\end{pmatrix},$$

and

$$\tau_0^2\approx 0.0026,\qquad \sigma_0^2\approx 0.0387.$$

Furthermore, the corresponding conditional mean of the weighted F-norm for the error are obtained as:

$$\mathrm{MSE}_{k,n}(\boldsymbol{\theta})=\mathbb{E}\left[\left|\left|\widehat{\boldsymbol{\mu}_{k,n}(\mathbf{\Theta})}-\boldsymbol{\mu}(\mathbf{\Theta})\right|\right|_{\boldsymbol{\xi}}^2\,\middle|\,\mathbf{\Theta}=\boldsymbol{\theta}\right],\quad k=C\ or\ N. \tag{4.1}$$

Taking $\boldsymbol{\xi}=(0.4,0.6)'$ and different values of (θ_1,θ_2) and n, the simulation is repeated 10,000 times, and the simulation results are given in Table 1.

[10]The Monte Carlo integration method is a numerical integration technique based on random sampling. Its main idea involves generating a large number of random samples, calculating the values of these samples on the function, and then estimating the integral value through the average of these function values. Unlike traditional deterministic numerical integration methods, Monte Carlo methods demonstrate significant advantages when dealing with high-dimensional, complex, or integrals that cannot be solved analytically. Relevant literature can be found in the work by Robert *et al.* (1999). In this study, we stochastically generated 100,000 samples over the integration domain for the approximate computation of structural parameters.

Table 1. *Conditional mean of the weighted F-norm for the error for different values of* (θ_1, θ_2) *and n.*

(θ_1, θ_2)	$\text{MSE}_{k,n}(\boldsymbol{\theta})$	$n = 10$	$n = 20$	$n = 50$	$n = 100$	$n = 200$	$n = 500$
(0.1,0.7)	$\text{MSE}_{C,n}(\boldsymbol{\theta})$	0.1238	0.0734	0.0262	0.0101	0.0037	0.0009
	$\text{MSE}_{N,n}(\boldsymbol{\theta})$	0.0981	0.0528	0.0172	0.0066	0.0026	0.0007
(0.3,1.0)	$\text{MSE}_{C,n}(\boldsymbol{\theta})$	0.2774	0.1645	0.0589	0.0225	0.0080	0.0021
	$\text{MSE}_{N,n}(\boldsymbol{\theta})$	0.2172	0.1164	0.0379	0.0145	0.0055	0.0017
(0.5,1.3)	$\text{MSE}_{C,n}(\boldsymbol{\theta})$	0.4978	0.2933	0.1068	0.0403	0.0144	0.0039
	$\text{MSE}_{N,n}(\boldsymbol{\theta})$	0.3906	0.2075	0.0691	0.0259	0.0098	0.0030
(0.7,1.6)	$\text{MSE}_{C,n}(\boldsymbol{\theta})$	0.7792	0.4603	0.1651	0.0632	0.0227	0.0060
	$\text{MSE}_{N,n}(\boldsymbol{\theta})$	0.6124	0.3268	0.1065	0.0407	0.0155	0.0047

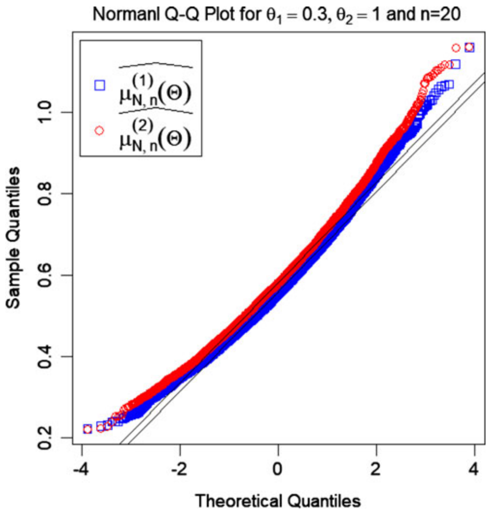

Figure 1. *Q-Q plot for* $\theta_1 = 0.3, \theta_2 = 1.0$.

Based on the simulation results presented in Table 1, some observations can be made as follows:

- When the sample size n is given, as the θ_1 and θ_2 increase, the conditional mean of the weighted F-norms for the errors of $\widehat{\boldsymbol{\mu}_{C,n}(\boldsymbol{\Theta})}$ and $\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}$ also increase.
- For a given set of (θ_1, θ_2), both estimators exhibit a decreasing trend in their conditional mean of the weighted F-norms for the errors as the sample size n increases, indicating that they both satisfy the consistency property.
- Note that the true means of Θ_1 and Θ_2 are 4/9 and 8/7, respectively. For various settings of (θ_1, θ_2) taking values around $(4/9, 8/7)$, $\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}$ has lower values of conditional mean of the weighted F-norms compared with $\widehat{\boldsymbol{\mu}_{C,n}(\boldsymbol{\Theta})}$. As the sample size grows, the difference becomes negligible. This demonstrates great performances of our proposed estimators.

To verify the asymptotic normality of $\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}$, we take different values of (θ_1, θ_2) and n and repeat the simulation 10,000 times, where we continue to set $\alpha_1 = 2.4$, $\alpha_2 = 3.2$, $\beta_1 = 5.4$, $\beta_2 = 2.8$, and $\rho_2 = 0.2$. The corresponding Q-Q plots are shown in Figures 1–6.

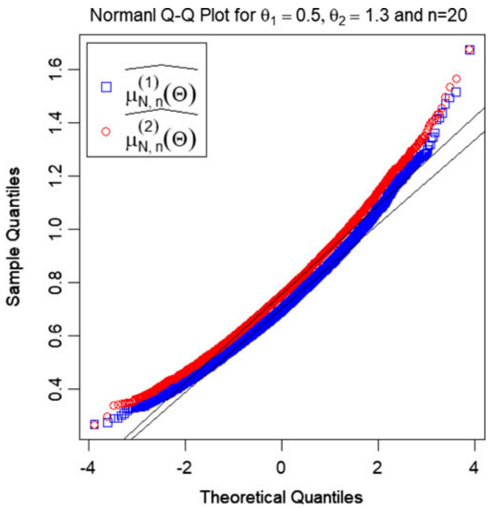

Figure 2. *Q-Q plot for* $\theta_1 = 0.5, \theta_2 = 1.3$.

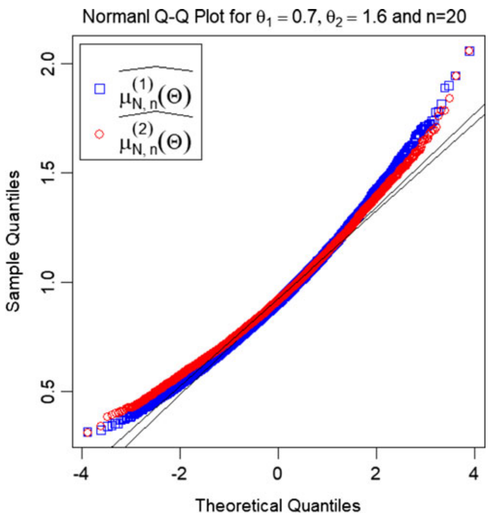

Figure 3. *Q-Q plot for* $\theta_1 = 0.7, \theta_2 = 1.6$.

Figures 1–3. display the Q-Q plots for different values of (θ_1, θ_2) with a fixed sample size of $n = 20$. It is evident that the variations in (θ_1, θ_2) have minimal impact on the asymptotic normality of $\widehat{\mu^{(1)}_{N,n}(\Theta)}$ and $\widehat{\mu^{(2)}_{N,n}(\Theta)}$ when the sample size remains constant, where $\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})} = \left(\widehat{\mu^{(1)}_{N,n}(\Theta)}, \widehat{\mu^{(2)}_{N,n}(\Theta)}\right)'$.

On the other hand, Figures 4–6 show the Q-Q plots for $\theta_1 = 0.5$ and $\theta_2 = 1.3$ with sample sizes of 5, 30, and 50, respectively. It is apparent that the Q-Q plot of the estimation $\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}$ obtained in this paper gradually approaches a straight line as the sample size increases, indicating that $\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}$ fulfills with asymptotic normality.

4.2. Multivariate normal-normal model

Given the joint distribution function of the sample and the prior distribution $\pi(\cdot)$, we can in general obtain the values of τ_0^2 and σ_0^2 according to Equations (3.8) and (3.9). However, since there is no

https://doi.org/10.1017/asb.2024.13 Published online by Cambridge University Press

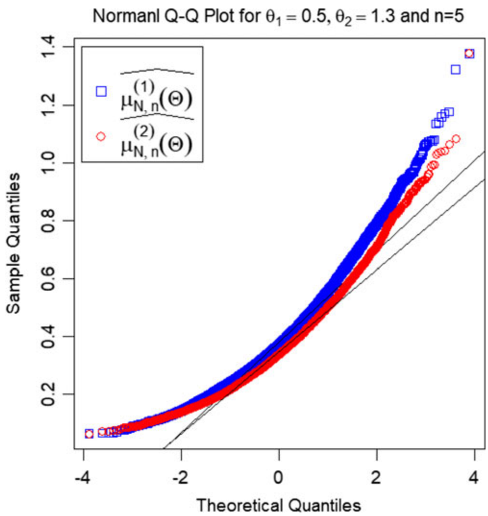

Figure 4. *Q-Q plot for* $n = 5$.

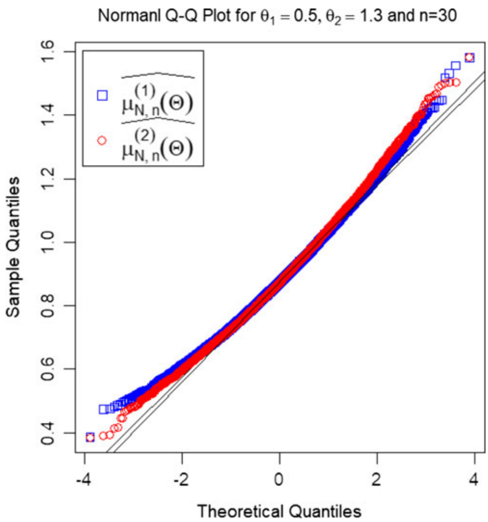

Figure 5. *Q-Q plot for* $n = 30$.

explicit cumulative joint distribution function for multivariate normal distribution, we need to estimate the structure parameters τ_0^2 and σ_0^2 with moment method.

First, we generate a sample of $\mathbf{\Theta}$ randomly with multivariate normal distribution $N_2(\boldsymbol{\mu}_0, T)$ for the sample size m. Then, based on this sample, we randomly generate m groups of samples with a sample size of n with multivariate normal distribution $N_3(\mathbf{\Theta}, \Sigma_0)$ and then obtain the estimations of τ_0^2 and σ_0^2 as:

$$\widehat{\tau_0^2} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{m-1}\sum_{i=1}^{m}\left(F_n^{(i)}\left(y^{(1)}, y^{(2)}, y^{(3)}\right) - \overline{F_n(y^{(1)}, y^{(2)}, y^{(3)})}\right)^2 dy^{(1)}dy^{(2)}dy^{(3)} \tag{4.2}$$

and

$$\widehat{\sigma_0^2} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{m}\sum_{i=1}^{m}\left(F_n^{(i)}\left(y^{(1)}, y^{(2)}, y^{(3)}\right)\left(1 - F_n^{(i)}\left(y^{(1)}, y^{(2)}, y^{(3)}\right)\right)\right) dy^{(1)}dy^{(2)}dy^{(3)}, \tag{4.3}$$

https://doi.org/10.1017/asb.2024.13 Published online by Cambridge University Press

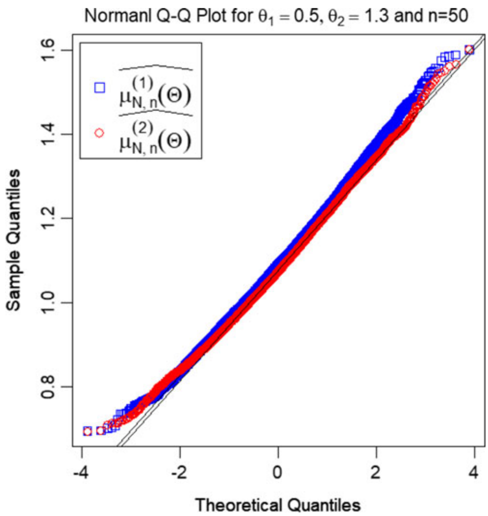

Figure 6. *Q-Q plot for* $n = 50$.

where $F_n^{(i)}\left(y^{(1)}, y^{(2)}, y^{(3)}\right)$ represents the empirical joint distribution function of the i-th sample, $i = 1, \ldots, m$, and

$$\overline{F_n(y^{(1)}, y^{(2)}, y^{(3)})} = \frac{1}{m}\sum_{i=1}^{m} F_n^{(i)}\left(y^{(1)}, y^{(2)}, y^{(3)}\right).$$

Based on the properties of multivariate normal distribution and Example 2.2, the matrices are defined as follows:

$$T = \begin{pmatrix} \tau_1^2 & \rho_1\tau_1\tau_2 & \rho_2\tau_1\tau_3 \\ \rho_1\tau_1\tau_2 & \tau_2^2 & \rho_3\tau_2\tau_3 \\ \rho_2\tau_1\tau_3 & \rho_3\tau_2\tau_3 & \tau_3^2 \end{pmatrix}, \quad \Sigma_0 = \begin{pmatrix} \sigma_1^2 & \rho_4\sigma_1\sigma_2 & \rho_5\sigma_1\sigma_3 \\ \rho_4\sigma_1\sigma_2 & \sigma_2^2 & \rho_6\sigma_2\sigma_3 \\ \rho_5\sigma_1\sigma_3 & \rho_6\sigma_2\sigma_3 & \sigma_3^2 \end{pmatrix},$$

and let $\boldsymbol{\mu}_0 = (1.4, 0.8, 1.2)'$. By varying the elements of matrices T and Σ_0 with different values, the unconditional mean of the weighted F-norms for the errors of $\widehat{\boldsymbol{\mu}_{C,n}(\boldsymbol{\Theta})}$ and $\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}$ are computed, where the corresponding formulas are given by:

$$\text{MSE}_{k,n} = \mathbb{E}\left[\left|\left|\widehat{\boldsymbol{\mu}_{k,n}(\boldsymbol{\Theta})} - \boldsymbol{\mu}(\boldsymbol{\Theta})\right|\right|_{\boldsymbol{\xi}}^2\right], \quad k = C \ or \ N,$$

where we set $\boldsymbol{\xi} = (0.3, 0.5, 0.2)'$. Upon conducting 1000 repetitions of the procedure, the corresponding means have been computed and reported in Table 2.

Table 2 reports the mean value of $\text{MSE}_{k,n}$ after undergoing 1000 repetitions. Additionally, in Case 1, the matrices T and Σ_0 are assumed to hold the following values:

$$T = \begin{pmatrix} 0.2^2 & 0.006 & 0.008 \\ 0.006 & 0.3^2 & 0.012 \\ 0.008 & 0.012 & 0.4^2 \end{pmatrix}, \quad \Sigma_0 = \begin{pmatrix} 0.4^2 & 0.06 & 0.024 \\ 0.06 & 1.5^2 & 0.09 \\ 0.024 & 0.09 & 0.6^2 \end{pmatrix}, \tag{4.4}$$

where $\rho_i = 0.1$, for $i = 1, 2, \ldots, 6$. In Case 2 and Case 3, we set $\rho_i = 0.5$ and 0.9 for $i = 1, \cdots, 6$, respectively, while keeping other values unchanged. In Case 4 and Case 5, we take $(\tau_1, \tau_2, \tau_3) = (0.4, 0.6, 0.8)$ and $(0.8, 2.0, 1.2)$, respectively, while other values remain unchanged. In Case 6 and Case 7, we consider $(\sigma_1, \sigma_2, \sigma_3) = (0.6, 1.7, 0.8)$ and $(0.8, 1.9, 1.0)$, respectively, while other values remain unchanged. The following observations can be made based on the simulation results presented in Table 2:

https://doi.org/10.1017/asb.2024.13 Published online by Cambridge University Press

Table 2. *Comparison of the unconditional means for different cases.*

Case	$\text{MSE}_{k,n}$	$n=10$	$n=20$	$n=50$	$n=100$	$n=200$	$n=500$
Case 1	$\text{MSE}_{C,n}$	0.0431	0.0305	0.0171	0.0105	0.0056	0.0024
	$\text{MSE}_{N,n}$	0.0987	0.0577	0.0241	0.0130	0.0063	0.0025
Case 2	$\text{MSE}_{C,n}$	0.0398	0.0290	0.0165	0.0095	0.0054	0.0023
	$\text{MSE}_{N,n}$	0.0988	0.0590	0.0250	0.0120	0.0060	0.0025
Case 3	$\text{MSE}_{C,n}$	0.0175	0.0131	0.0093	0.0062	0.0043	0.0020
	$\text{MSE}_{N,n}$	0.1089	0.0602	0.0248	0.0114	0.0062	0.0024
Case 4	$\text{MSE}_{C,n}$	0.0812	0.0475	0.0217	0.0117	0.0062	0.0024
	$\text{MSE}_{N,n}$	0.1051	0.0562	0.0239	0.0122	0.0063	0.0024
Case 5	$\text{MSE}_{C,n}$	0.1192	0.0590	0.0255	0.0128	0.0064	0.0024
	$\text{MSE}_{N,n}$	0.1200	0.0591	0.0254	0.0127	0.0063	0.0023
Case 6	$\text{MSE}_{C,n}$	0.0633	0.0369	0.0210	0.0135	0.0072	0.0032
	$\text{MSE}_{N,n}$	0.1543	0.0839	0.0325	0.0179	0.0081	0.0033
Case 7	$\text{MSE}_{C,n}$	0.0687	0.0429	0.0262	0.0159	0.0097	0.0043
	$\text{MSE}_{N,n}$	0.1912	0.1250	0.0426	0.0219	0.0116	0.0046

- According to the simulation outcomes in Cases 1–3, $\text{MSE}_{C,n}$ diminishes with escalating correlation coefficients. In contrast, the influence on $\text{MSE}_{N,n}$ remains minimal.
- The results obtained from simulations in Cases 1, 4, and 5 reveal a noteworthy trend: with all other factors unchanged, the value of $\text{MSE}_{C,n}$ increases with the values of (τ_1, τ_2, τ_3). On the contrary, the impact on $\text{MSE}_{N,n}$ is particularly pronounced in the context of smaller sample sizes; however, as the sample size grows, this effect diminishes. Moreover, as the parameters (τ_1, τ_2, τ_3) increase, the gap between the values of $\text{MSE}_{C,n}$ and $\text{MSE}_{N,n}$ gradually narrows. Furthermore, with larger values of n, situations may arise where $\text{MSE}_{N,n}$ becomes smaller than $\text{MSE}_{C,n}$, which is in favor of the potential application of our proposed estimator.
- Based on the simulation outcomes from Case 1, Case 6, and Case 7, both $\text{MSE}_{C,n}$ and $\text{MSE}_{N,n}$ exhibit an increasingness in $(\sigma_1, \sigma_2, \sigma_3)$. Besides, $\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}$ converges more quickly as the sample size increases, and its performance becomes comparable to that of the traditional credibility estimator $\widehat{\boldsymbol{\mu}_{C,n}(\boldsymbol{\Theta})}$.

According to Equations (4.1) and (4.4), Q-Q plots are implemented for different values of $(\theta_1, \theta_1, \theta_3)$ and n, which are depicted in Figures 7–12. It is readily seen that the simulation results based on the multivariate normal-normal model exhibit similarities to those under the bivariate exponential-gamma model. In particular irrespective of the specific values of $(\theta_1, \theta_2, \theta_3)$, the asymptotic normality of $\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}$ remains largely unaffected by changes in the sample size. However, when $(\theta_1, \theta_2, \theta_3)$ is fixed, an interesting trend emerges as the sample size increases: the Q-Q plot of the estimation $\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}$ becomes progressively closer to a straight line. This observation indicates that $\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}$ satisfies asymptotic normality, bolstering its practicability as a new credibility estimator.

4.3. Case analysis

In this part, we shall adopt two explicit insurance scenarios to show the performances of our proposed estimator. At the meantime, the predicted values based on the traditional credibility method will be also reported for the first application under some specified premium principles.

https://doi.org/10.1017/asb.2024.13 Published online by Cambridge University Press

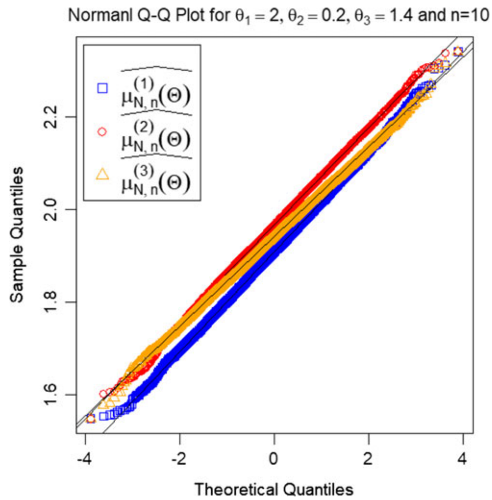

Figure 7. *Q-Q plot for* $\theta_1 = 2, \theta_2 = 0.2, \theta_3 = 1.4$.

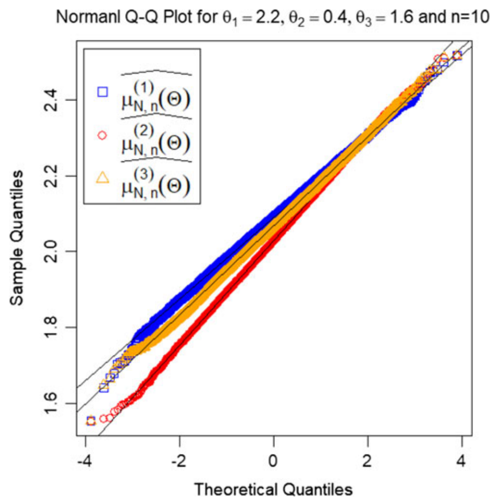

Figure 8. *Q-Q plot for* $\theta_1 = 2.2, \theta_2 = 0.4, \theta_3 = 1.6$.

4.3.1. Application I

To demonstrate the specific application of the new estimators presented in this paper, the data on page 117 of Bühlmann and Gisler (2005) from some fire insurance business are used. Table 3 shows the related data for five portfolios of insurance policies in five policy years, where $Y_{ij}^{(1)}$ and $Y_{ij}^{(2)}$ represent the total loss and loss rate[11] of the i-th policy portfolio in the j-th year, respectively.

In order to obtain the credibility estimator, the structure parameters $\boldsymbol{\mu}_0$, T, Σ_0, τ_0^2, and σ_0^2 need to be estimated based on the dataset. With moment estimation method, the corresponding estimations of $\boldsymbol{\mu}_0$, T, and Σ_0 are given by:

$$\widehat{\boldsymbol{\mu}}_0 = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\mathbf{Y}_{ij}, \quad \widehat{\Sigma}_0 = \frac{1}{m(n-1)}\sum_{i=1}^{m}\sum_{j=1}^{n}\left(\mathbf{Y}_{ij} - \overline{\mathbf{Y}}_i\right)\left(\mathbf{Y}_{ij} - \overline{\mathbf{Y}}_i\right)',$$

[11] Note that the number of policyholders in each group is not fixed for those five time periods, and thus they are not comonotonic. Therefore, it is reasonable to model the joint distribution of the total loss and the loss rate.

https://doi.org/10.1017/asb.2024.13 Published online by Cambridge University Press

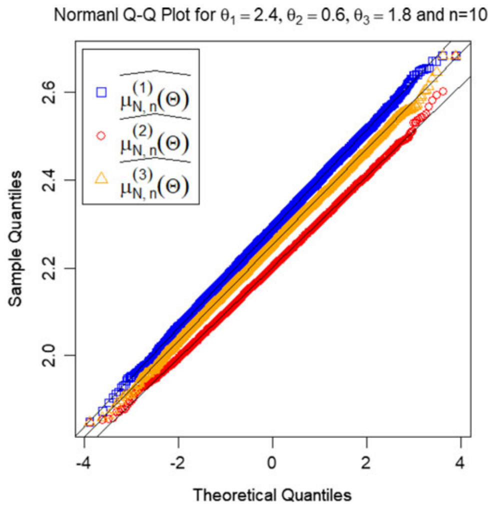

Figure 9. *Q-Q plot for* $\theta_1 = 2.4, \theta_2 = 0.6, \theta_3 = 1.8$.

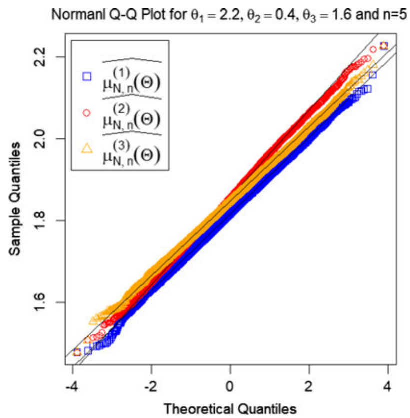

Figure 10. *Q-Q plot for* $n = 5$.

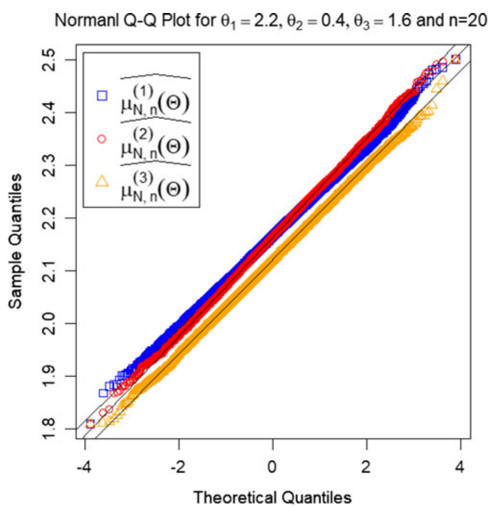

Figure 11. *Q-Q plot for* $n = 20$.

https://doi.org/10.1017/asb.2024.13 Published online by Cambridge University Press

Table 3. *Annual loss data related to fire insurance.*

Risk groups	$\mathbf{Y}_{ij}$	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$
Group 1	$Y_{1j}^{(1)}$	0.583	1.100	0.262	0.837	1.630
	$Y_{1j}^{(2)}$	0.80	1.40	0.30	0.88	1.60
Group 2	$Y_{2j}^{(1)}$	0.099	1.298	0.326	0.463	0.895
	$Y_{2j}^{(2)}$	0.06	0.72	0.16	0.20	0.38
Group 3	$Y_{3j}^{(1)}$	1.433	0.496	0.699	1.742	1.038
	$Y_{3j}^{(2)}$	1.80	0.60	0.80	1.90	1.10
Group 4	$Y_{4j}^{(1)}$	1.765	4.145	3.121	4.129	3.358
	$Y_{4j}^{(2)}$	0.56	1.20	0.84	1.07	0.80
Group 5	$Y_{5j}^{(1)}$	0.040	0	0.169	1.018	0.044
	$Y_{5j}^{(2)}$	0.10	0.00	0.40	2.40	0.10

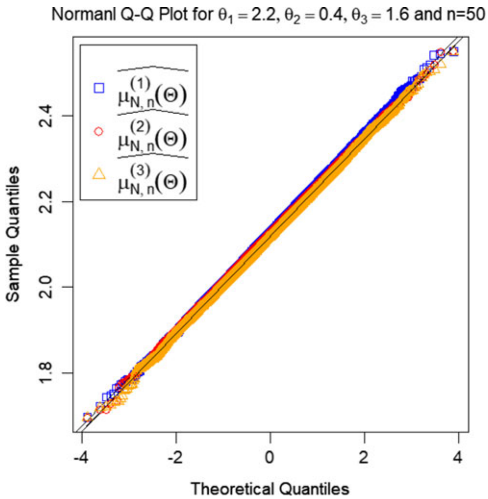

Figure 12. *Q-Q plot for* $n = 50$.

and

$$\widehat{T} = \frac{1}{n(m-1)}\left[\sum_{i=1}^{m} n\left(\overline{\mathbf{Y}}_i - \widehat{\boldsymbol{\mu}}_0\right)\left(\overline{\mathbf{Y}}_i - \widehat{\boldsymbol{\mu}}_0\right)' - (m-1)\widehat{\Sigma}_0\right],$$

where $m = n = 5$, and the estimations of τ_0^2 and σ_0^2 are given by (4.2) and (4.3), respectively. Based on the dataset in Table 3, those estimators can be calculated as $\widehat{\tau_0^2} = 0.5450$, $\widehat{\sigma_0^2} = 0.9591$, and

$$\widehat{\boldsymbol{\mu}}_0 = \begin{pmatrix} 1.2276 \\ 0.8068 \end{pmatrix}, \quad \widehat{T} = \begin{pmatrix} 1.3669 & 0.0864 \\ 0.0864 & 0.0607 \end{pmatrix}, \quad \widehat{\Sigma}_0 = \begin{pmatrix} 0.3795 & 0.2692 \\ 0.2692 & 0.3547 \end{pmatrix}.$$

Based on these results and Equations (3.19)–(3.22), the corresponding credibility estimations under the new estimator and the classical method based on various premium principles are shown in

Table 4. *Our credibility estimations based on various premium principles.*

	$\mathbf{a}=(1,0)'$				$\mathbf{a}=(0,1)'$			
Risk groups	$\widehat{H_Z^{(1)}(\boldsymbol{\Theta})}$	$\widehat{H_Z^{(2)}(\boldsymbol{\Theta})}$	$\widehat{H_Z^{(3)}(\boldsymbol{\Theta})}$	$\widehat{H_Z^{(4)}(\boldsymbol{\Theta})}$	$\widehat{H_Z^{(1)}(\boldsymbol{\Theta})}$	$\widehat{H_Z^{(2)}(\boldsymbol{\Theta})}$	$\widehat{H_Z^{(3)}(\boldsymbol{\Theta})}$	$\widehat{H_Z^{(4)}(\boldsymbol{\Theta})}$
Group 1	1.1667	1.1727	1.1916	1.1792	1.1361	1.0394	1.0959	1.0191
Group 2	0.9304	0.9919	1.0033	1.0050	0.5219	0.5121	0.5710	0.4955
Group 3	1.3435	1.3095	1.3331	1.3095	1.3527	1.2477	1.2973	1.2196
Group 4	3.3158	3.4721	3.1757	3.1608	1.0456	0.9248	0.9846	0.9108
Group 5	0.6091	0.7694	0.7583	0.7865	0.7846	0.8848	0.8893	0.8847

Table 5. *Our credibility estimations based on various premium principles.*

	$\mathbf{a}=(0.5,0.5)'$			
Risk groups	$\widehat{H_Z^{(1)}(\boldsymbol{\Theta})}$	$\widehat{H_Z^{(2)}(\boldsymbol{\Theta})}$	$\widehat{H_Z^{(3)}(\boldsymbol{\Theta})}$	$\widehat{H_Z^{(4)}(\boldsymbol{\Theta})}$
Group 1	1.1514	1.0702	1.2253	1.0503
Group 2	0.7261	0.7210	0.7719	0.7014
Group 3	1.3481	1.2419	1.2920	1.2180
Group 4	2.1807	2.0542	2.0557	1.9640
Group 5	0.6969	0.7698	0.7938	0.7469

Table 6. *The traditional credibility estimation based on the expected value premium principle.*

Risk groups	Group 1	Group 2	Group 3	Group 4	Group 5
$\widehat{H_Z^{(1)}(\boldsymbol{\Theta})}$	0.9754	0.9875	1.0958	3.8874	0.4195

Tables 4–6, where we set $\lambda=0.2$, $\rho=0.24$, and $\beta=0.54$ in those four premium principles, and the estimates of U_0 and L_0 are given by:

$$\widehat{U_0}=\frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\mathbf{Y}_{ij}\mathbf{Y}_{ij}' \text{ and } \widehat{L_0}=\frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\exp\left(\beta\mathbf{a}'\mathbf{Y}_{ij}\right). \tag{4.5}$$

The values in Tables 4 through 6 illustrate both of the two credibility estimators of risk premium derived for different weight matrix $\mathbf{a}$ and different premium principles. Specifically, Tables 4 and 5 display the predictions using the novel credibility estimates $\widehat{\boldsymbol{\mu}_{N,n}(\boldsymbol{\Theta})}$ and $\widehat{\Sigma_{N,n}(\boldsymbol{\Theta})}$, whereas in Table 6, owing to the challenge of estimating the conditional covariance matrix $\Sigma(\boldsymbol{\Theta})$ within traditional multivariate credibility method, we can only report the performance of the expected value premium principle, where the expression for $\widehat{H_Z^{(1)}(\boldsymbol{\Theta})}$ based on the traditional multivariate credibility method is

$$\widehat{H_Z^{(1)}(\boldsymbol{\Theta})}=(1+\lambda_1)\,\mathbf{a}'\left[Z_{C,n}\overline{\mathbf{Y}}+(I_p-Z_{C,n})\boldsymbol{\mu}_0\right].$$

It can be noted that the outcomes predicted by the new estimator are slightly different with those calculated from the traditional method with respect to $\widehat{H_Z^{(1)}(\boldsymbol{\Theta})}$. While for the other three types of premium principles, our method can be easily applied, but the traditional multidimensional credibility is hard to put to use.

https://doi.org/10.1017/asb.2024.13 Published online by Cambridge University Press

Table 7. *Credibility estimations for the aggregate risk under various premium principles.*

Premium principle	Estimated value	Safety loading coefficient
Net($\lambda_1 = 0$)	39.6350	0
Expected value($\lambda_1 = 0.2$)	46.1367	0.1640
Variance($\lambda_2 = 0.01$)	50.8007	0.2817
Standard deviation($\lambda_2 = 0.2$)	45.4767	0.1474
Exponential($\beta = 0.8$)	46.7492	0.1795

4.3.2. Application II

To illustrate the practical application of the newly proposed credibility estimates obtained in this paper in higher dimensions, we use the "ClaimsLong" dataset[12] from the "insuranceData" package in R software; see Wolny-Dominiak and Trzesiok (2022). In this dataset, we treat the drivers' age as the risk profile and categorize it into six classes, that is, 1, 2, 3, 4, 5, 6, with 1 stands for the youngest and 6 stands for the oldest. We use $\mathbf{Y}_i = \left(Y_i^{(1)}, \cdots, Y_i^{(6)}\right)'$ to model the claim severities for these six classes in year i, and we assume a payout of 1000 yuan for each claim. Consequently, $\mathbf{Y}_i$ represents the claim amount in year i, calculated as the number of claims in year i divided by the total number of policies and then multiplied by 1000. To obtain estimates of the structural parameters, we employ the Bootstrap method to resample the dataset. After obtaining m sets of data, we estimate the structural parameters using the same method as the previous case and the estimated values are obtained as $\widehat{\tau_0^2} = 12,219.09$, $\widehat{\sigma_0^2} = 33,728.27$,

$$\widehat{\boldsymbol{\mu}}_0 = (19.3750, 38.6250, 48.0500, 46.7222, 25.0972, 15.8972)',$$

And

$$\widehat{\Sigma}_0 = \begin{pmatrix} 422.0484 & 812.1451 & 996.7712 & 970.2049 & 528.0890 & 342.2654 \\ 812.1451 & 1584.5024 & 1943.0102 & 1893.6903 & 1027.0936 & 660.1979 \\ 996.7712 & 1943.0102 & 2409.5705 & 2339.4193 & 1266.2069 & 814.4658 \\ 970.2049 & 1893.6903 & 2339.4193 & 2275.9546 & 1231.5569 & 792.0600 \\ 528.0890 & 1027.0936 & 1266.2069 & 1231.5569 & 668.0973 & 430.0422 \\ 342.2654 & 660.1979 & 814.4658 & 792.0600 & 430.0422 & 279.8111 \end{pmatrix}.$$

With these values and Equations (3.19)–(3.22), the corresponding credibility estimations of the aggregate risk under various premium principles are reported in Table 7, where the weight vector is set as $\mathbf{a} = (0.1, 0.2, 0.2, 0.2, 0.2, 0.1)'$ and the estimates of U_0 and L_0 are given by Equation (4.5). For the sake of comparisons, we define the expression for the safety loading coefficient as follows:

$$\text{Safety loading coefficient} = \frac{\text{Various premium} - \text{Net premium}}{\text{Net premium}}.$$

However, the traditional multidimensional credibility estimator developed by Jewell (1973) is challenging to apply to this dataset due to the computational complexity involved in calculating the credibility factor matrix, which requires inverting a high-dimensional matrix with six rows and six columns. Although the obtained credibility factor in this study necessitates the calculation of integrals involving multivariate functions, in practical applications, two approaches can be employed to address the integration computation issue. First, if samples from multiple policy contracts are available, the estimation of credibility factors can be obtained based on samples resembling policy contracts outlined in

[12]This is a simulated dataset, based on the car insurance dataset. There are 40,000 policies over 3 years, giving 120,000 records.

https://doi.org/10.1017/asb.2024.13 Published online by Cambridge University Press

Application II. The second approach involves utilizing Markov Chain Monte Carlo methods to numerically integrate the multivariate function. In comparison with traditional multivariate credibility methods, we believe that the outcomes of this study hold greater practical applicability.

5. Conclusions

This study addresses the computational intricacies associated with the conventional multivariate credibility estimation procedure, specifically in relation to the computation of the credibility factor matrix. To address these challenges, we propose an innovative approach based on the conditional joint distribution function. Moreover, we extend the methodology to include the estimation of the conditional variance matrix and the computation of risk premiums using various premium principles. Through numerical simulations, we verify the cohesiveness and asymptotic normality of the newly formulated estimators. Finally, we demonstrate the practical applicability of these estimators using real insurance company data, highlighting their effectiveness in calculating risk premiums across a range of premium principles. The results consistently indicate the estimated magnitudes of risk groups, regardless of the premium principles considered.

The novel multivariate credibility estimation method proposed in this study simplifies the computation of the credibility factor while maintaining specific structural parameters. In typical situations, the conditional joint distribution function of the sample and the prior distribution is often unknown. In future research, we plan to apply the approach of credibility estimation using joint distributions to address credibility prediction problems in multivariate regression models. Another promising extension would be generalizing the current study to the case of the multidimensional Bühlmann–Straub model.

Acknowledgments. The authors are very grateful for the insightful and constructive comments and suggestions from the editor and three anonymous reviewers, which have greatly improved the presentation of this manuscript. Limin Wen thanks the financial support from Jiangxi Provincial Natural Science foundation (No. 20224BCD41001). Wei Liu thanks the Jiangxi Provincial Graduate Innovation Fund (No. YC2022-s285). Yiying Zhang acknowledges the financial support from the National Natural Science Foundation of China (No. 12101336, 72311530687), GuangDong Basic and Applied Basic Research Foundation (No. 2023A1515011806), and Shenzhen Science and Technology Program (No. JCYJ20230807093312026).

Competing interests. No potential competing interests were reported by the authors.

Supplementary material. To view supplementary material for this article, please visit https://doi.org/10.1017/asb.2024.13

References

Bauwelinckx, T. and Goovaerts, M.J. (1990) On a multilevel hierarchical credibility algorithm. *Insurance: Mathematics and Economics*, **9**, 221–228.

Belitskii, G.R. and Lyubich, Y.I. (2013) *Matrix Norms and their Applications*, Vol. **36**. Birkhäuser.

Bühlmann, H. (1967) Experience rating and credibility. *ASTIN Bulletin: The Journal of the IAA*, **4**, 199–207.

Bühlmann, H. and Gisler, A. (2005) *A Course in Credibility Theory and its Applications*. Springer Netherlands.

Bühlmann, H., Gisler, A. and Kollöffel, D. (2003). Multidimensional credibility applied to estimating the frequency of big claims. *ASTIN Colloquium International Actuarial Association*, 1–25.

Cai, X.Q., Wen, L.M., Wu, X.Y. and Zhou, X. (2015) Credibility estimation of distribution functions with applications to experience rating in general insurance. *North American Actuarial Journal*, **19**, 311–335.

Deresa, N.W., Van Keilegom, I. and Antonio, K. (2022) Copula-based inference for bivariate survival data with left truncation and dependent censoring. *Insurance: Mathematics and Economics*, **107**, 1–21.

Dhaene, J. and Goovaerts, M.J. (1996). Dependency of risks and stop-loss order. *ASTIN Bulletin: The Journal of the IAA*, **26**, 201–212.

Englund, M., Guillén, M., Gustafsson, J., Nielsen, L.H. and Nielsen, J.P. (2008) Multivariate latent risk: A credibility approach. *ASTIN Bulletin: The Journal of the IAA*, **38**, 137–146.

Englund, M., Gustafsson, J., Nielsen, J.P. and Thuring, F. (2009) Multidimensional credibility with time effects: An application to commercial business lines. *Journal of Risk and Insurance*, **76**, 443–453.

Frees, E.W. (2003). Multivariate credibility for aggregate loss models. *North American Actuarial Journal*, **7**, 13–37.

Gómez-Déniz, E. (2016). Bivariate credibility bonus–malus premiums distinguishing between two types of claims. *Insurance: Mathematics and Economics*, **70**, 117–124.

Gómez-Déniz, E. and Calderín-Ojeda, E. (2018) Multivariate credibility in bonus–malus systems distinguishing between different types of claims. *Risks*, **6**, 1–11.

Gómez-Déniz, E. and Calderín-Ojeda, E. (2021) A priori ratemaking selection using multivariate regression models allowing different coverages in auto insurance. *Risks*, **9**, 1–28.

Guerra, M. and Centeno, M.D.L. (2010) Optimal reinsurance for variance related premium calculation principles1. *ASTIN Bulletin: The Journal of the IAA*, **40**, 97–121.

Hachemeister, C. (1975) Credibility for regression models with application to trend (reprint). In *Credibility: Theory and Applications* (ed. P. Kahn), pp. 307–48. New York: Academic Press, Inc.

Jewell, W.S. (1973) Multidimensional credibility. Operations Research Center, pp. 73–77.

Jewell, W.S. (1974) Exact multidimensional credibility. *Bulletin of Swiss Association of Actuaries*, **74**, 193–214.

Laloë, T. and Servien, R. (2013). Nonparametric estimation of regression level sets using kernel plug-in estimation. *Journal of the Korean Statistical Society*, **42**, 301–311.

Oliveira, M., Crujeiras, R.M. and Rodríguez-Casal, A. (2012) A plug-in rule for bandwidth selection in circular density estimation. *Computational Statistics & Data Analysis,* **56**, 3898–3908.

Pechon, F., Denuit, M. and Trufin, J. (2021). Home and motor insurance joined at a household level using multivariate credibility. *Annals of Actuarial Science*, **15**, 82–114.

Poon, J. and Lu, Y. (2015) A spatial cross-sectional credibility model with dependence among risks. *North American Actuarial Journal*, **19**, 289–310.

Robert, C.P., Casella, G. and Casella, G. (1999) *Monte Carlo Statistical Methods*, Vol. **2**. New York: Springer.

Schinzinger, E., Denuit, M. and Christiansen, M. (2016) A multivariate evolutionary credibility model for mortality improvement rates. *Insurance: Mathematics and Economics*, **69**, 70–81.

Thuring, F., Nielsen, J.P., Guillén, M. and Bolancé, C. (2012) Selecting prospects for cross-selling financial products using multivariate credibility. *Expert Systems with Applications*, **39**, 8809–8816.

Wen, L.M., Wu, X.Y. and Zhou, X. (2009) The credibility premiums for models with dependence induced by common effects. *Insurance: Mathematics and Economics*, **44**, 19–25.

Wolny-Dominiak, A. and Trzesiok, M. (2022) Package ‘insuranceData’.

Wu, B., Yao, Q. and Zhu, S. (2013) Estimation in the presence of many nuisance parameters: Composite likelihood and plug-in likelihood. *Stochastic Processes and their Applications*, **123**, 2877–2898.

Yan, Y.J. and Song, K.-S. (2022) A general optimal approach to Bühlmann credibility theory. *Insurance: Mathematics and Economics*, **104**, 262–282.

Yeo, K.L. and Valdez, E.A. (2006) Claim dependence with common effects in credibility models. *Insurance: Mathematics and Economics*, **38**, 609–629.

https://doi.org/10.1017/asb.2024.13 Published online by Cambridge University Press