Astin Bulletin (2026), **56**, pp. 1036–1065
doi:10.1017/asb.2026.10094

International Actuarial Association
Association Actuarielle Internationale

RESEARCH ARTICLE

A new credibility estimation of process variance with applications

Limin Wen[1] and Wei Liu[2]

[1]School of Digital Economy Industry, Guangzhou College of Commerce, Guangzhou, Guangdong, China
[2]Research Center of Management Science and Engineering, Jiangxi Normal University, Nanchang, Jiangxi, China
Corresponding author: Limin Wen; Email: lmwen2929@163.com

Received: 7 July 2025; **Revised:** 19 March 2026; **Accepted:** 25 March 2026; **First published online:** 20 April 2026

Keywords: Bayesian estimation; credibility estimation; process variance; hypothetical mean; variance-related premium principles

JEL Classification: G22; C21; G32; C52

Abstract

In classical credibility theory, estimation is typically limited to the hypothetical mean, restricting its use for premium principles that depend on higher-order moments. To address this, we develop a credibility-based framework for estimating the process variance under both known and unknown hypothetical means and apply these estimators to a broad class of variance-related premium principles, including the expected value, variance, standard deviation, and modified-variance principles. The estimators are derived via constrained linear projection techniques, minimizing the mean squared error between the estimator and the true process variance. Explicit formulas are obtained that are optimal among affine transformations of the data. The proposed estimators exhibit desirable statistical properties, including conditional unbiasedness, consistency, mean squared error convergence, and asymptotic normality. Numerical studies demonstrate their favorable convergence behavior, and an empirical analysis with real insurance data highlights their practical relevance. This framework extends Bühlmann's classical credibility theory to second-moment estimation while remaining computationally tractable and requiring only mild moment conditions, without specifying the population or prior distributions.

1. Introduction

In actuarial science, the accurate determination of insurance premiums is of fundamental importance, as setting appropriate premiums directly affects both the financial stability of insurers and the economic welfare of policyholders. Within the field of non-life insurance, credibility theory has long been regarded as a cornerstone of premium calculation, providing a coherent and systematic framework for integrating historical experience with new observations through a credibility factor. By balancing individual experience with collective information, it enhances the accuracy, reliability, and stability of premium estimates and thereby occupies a central position in modern non-life pricing practice. Since its origin in the early twentieth century (Whitney, 1918; Bailey, 1945; Bühlmann, 1967), credibility theory has undergone continuous development and refinement over the past century. Owing to its flexibility in combining prior knowledge with observed data, it has also been widely applied beyond premium determination, including in loss reserving, risk portfolio evaluation, inequality measurement (such as Lorenz curves and Gini coefficients), and bibliometrics (Gisler and Wüthrich, 2008; Pérez-Hornero, *et al*., 2013; Wen *et al*., 2024), demonstrating its broad relevance for risk assessment and predictive modeling.

Classical credibility methods primarily focus on estimating hypothetical means (Bühlmann and Gisler, 2005), which are sufficient for constructing premiums under the net premium and expected value principles. In practice, however, non-life insurers employ a variety of premium principles, including

© The Author(s), 2026. Published by Cambridge University Press on behalf of The International Actuarial Association.

https://doi.org/10.1017/asb.2026.10094 Published online by Cambridge University Press

the variance principle, the standard deviation principle, and the modified variance premium principle (Teugels and Sundt, 2004; Liang and Yuen, 2016; Zhang *et al*., 2016; Sun *et al*., 2017; Yong *et al*., 2025). While the expected value principle is predominantly used in life insurance, the standard deviation principle is more common in non-life insurance, and the variance principle possesses several desirable theoretical properties that have been extensively studied in the literature (Bühlmann, 2007). The coexistence of these premium principles reflects the diversity of actuarial pricing methodologies and highlights the need for a unified framework capable of accommodating different premium calculation rules while fully leveraging the predictive power of credibility theory.

From a practical actuarial perspective, premium calculation must be aligned with the premium principles actually employed in insurance pricing. This consideration has motivated a strand of research that extends credibility theory beyond the classical net premium principle to alternative pricing rules. For instance, Wen *et al*. (2009) derived credibility estimators under generalized weighted premium principles, Pan *et al*. (2008) examined the consistency of credibility premiums under the Esscher principle, and Wen *et al*. (2011) developed credibility premiums corresponding to the exponential principle.

Methodologically, these contributions are rooted in the decision-theoretic foundation of credibility theory (Heilmann, 1989). The premium is reformulated as the solution to an optimization problem under a suitably specified loss function, so that the credibility premium emerges as an optimal linear Bayes estimator with respect to the prescribed loss structure. While this approach is effective for premium principles that admit a natural representation through loss function minimization, it becomes less suitable for pricing rules that cannot be directly characterized in this manner. In particular, variance-related premium principles—such as the variance principle, the standard deviation principle, and the modified variance principle, as well as their unified formulation, the variance-related premium principle (Guerra and de Lourdes Centeno, 2010; Chi, 2012; Zhang *et al*., 2016)—do not readily fit into the classical loss-function-based credibility paradigm.

An important attempt to address this issue is due to Yong *et al*. (2025), who proposed a credibility pricing method under the variance principle by restricting the premium estimator to linear and quadratic functions of the sample observations and deriving the corresponding optimal linear Bayes estimator within this enlarged class. However, because the variance principle combines expectation and variance components measured on different scales, the resulting credibility estimator involves a comparatively complex structure and multiple structural parameters. Moreover, this construction is specifically tailored to the variance principle and does not extend naturally to other variance-related premium principles, such as the standard deviation principle, thereby limiting its general applicability.

Motivated by the aforementioned limitations, the present paper develops a unified framework for credibility estimation of the process variance under variance-related premium principles. In contrast to approaches that embed alternative premium principles into modified loss functions, we adopt a structurally different perspective by decoupling the estimation of the hypothetical mean from that of the process variance and focusing explicitly on the latter. Drawing inspiration from the normal–inverse gamma model with a known mean, we first derive the Bayesian estimator of the process variance, which can be expressed as a weighted combination of the sample variance and the prior variance, where the weights naturally exhibit a credibility interpretation: increasing reliance on sample information as the sample size grows and stronger dependence on prior information when data are limited. When the mean is unknown, we extend the framework by replacing the population mean with the sample mean, leading to a modified credibility estimator that preserves linearity and optimality in expectation. By combining this estimator with the classical credibility estimator of the hypothetical mean, we obtain credibility premiums for a broad class of variance-related premium principles. The resulting formulas are parsimonious, structurally transparent, and readily extendable, thereby enhancing both their theoretical coherence and practical applicability.

The contributions of this paper are fourfold. First, we show that directly applying the classical Bühlmann credibility estimator to process variance is inappropriate, revealing a limitation of existing credibility methods. Second, inspired by Bayesian linearization techniques, we construct new credibility

https://doi.org/10.1017/asb.2026.10094 Published online by Cambridge University Press

estimators for process variance that are valid under both known and unknown hypothetical means. Third, we rigorously analyze the statistical properties of these estimators, including unbiasedness, mean squared error, consistency, and asymptotic normality. Fourth, by combining the proposed estimators of process variance with the classical Bühlmann credibility estimator for hypothetical means, we build a comprehensive framework for estimating risk premiums under a broad class of variance-related premium principles. Compared with existing methods such as Yong *et al*. (2025), our approach separates the estimation of mean and variance, requires fewer structural parameters, ensures numerical stability, and extends naturally to multiple premium principles. The practical performance and applicability of the proposed methodology are further illustrated through extensive simulations and an empirical analysis using real insurance claims data.

The remainder of the paper is organized as follows. Section 2 introduces the variance-related premium principles and their Bayesian formulations, provides a detailed review of classical credibility estimation methods for hypothetical means, and establishes the necessary and sufficient conditions for consistency in risk premium estimation. Section 3 develops credibility estimators for process variance under both known and unknown hypothetical means and rigorously investigates their finite-sample and asymptotic properties. Section 4 integrates the credibility estimators of both hypothetical means and process variance to construct a unified framework for risk premium estimation under variance-related premium principles and demonstrates the practical relevance of the proposed methods through numerical simulations and empirical analysis. Finally, Section 5 concludes the paper by summarizing the main findings and outlining potential directions for future research. To ensure the paper is fully self-contained, all proofs of the theoretical results are provided in a dedicated Appendix A.

2. Credibility estimators for variance-related premium principles

Credibility theory provides a framework for predicting future outcomes of individual risks by combining their own experience with the collective experience of a larger population. A risk X is assumed to depend on an underlying parameter Θ, which may be a scalar or a vector. For example, in automobile liability insurance, Θ may represent the combined effect of driver characteristics (age, gender, and behavior) and vehicle attributes (mileage, region, and engine type).

Since insurers generally lack full knowledge of all risk-influencing factors, Θ is modeled as a random variable or vector with prior distribution $\pi_\Theta(\theta)$. Conditional on $\Theta = \theta$, the observations $X_1, \ldots, X_n$ are assumed to be i.i.d. with distribution $f_\theta(x)$. Under this structure, risk premium estimation naturally fits within the Bayesian framework. The model assumptions are summarized below:

Assumption A1. *All factors influencing the risk X are represented by a random parameter Θ, with prior distribution $\pi_\Theta(\theta)$.*

Assumption A2. *Given $\Theta = \theta$, the samples $X_1, \ldots, X_n$ are i.i.d. with common distribution $f_\theta(x)$.*

Here, the conditional expectation and conditional variance of X are denoted by

$$\mu(\Theta) = \mathbb{E}(X|\Theta), \quad \gamma(\Theta) = \mathrm{Var}(X|\Theta),$$

which are referred to in actuarial science as the hypothetical mean and the process variance, respectively.

A model that satisfies Assumptions A1 and A2 is commonly referred to as a Bayesian model. In this setting, the risk associated with each individual is characterized by an underlying random parameter, and observed outcomes are assumed conditionally independent given this parameter. We focus on the problem of estimating credibility premiums under such a Bayesian framework, which allows the integration of individual claim experience with collective information from the broader population, providing a statistically coherent approach to premium determination.

Traditional credibility theory is typically grounded in the pure premium principle for premium pricing models (Teugels and Sundt, 2004). Essentially, the pure premium principle considers only the expected

https://doi.org/10.1017/asb.2026.10094 Published online by Cambridge University Press

value of the risk, without accounting for its volatility. In non-life insurance actuarial practice, incorporating variance components into the risk loading premium provides a more comprehensive and practical approach. Motivated by this idea, and building upon the variance-related premium principle introduced by Guerra and de Lourdes Centeno (2010), we extend their framework by defining the risk premium as a function of both the hypothetical mean and the process variance. Specifically, we define the risk premium under the variance-related premium principle as follows.

Definition 2.1. *Under the Bayesian model, the risk premium based on the variance-related premium principle can be defined as*

$$R(\Theta) = \mu(\Theta) + g(\mu(\Theta), \gamma(\Theta)), \tag{2.1}$$

where g(x,y) is a continuous bivariate function.

In practice, the risk parameter Θ is typically unobservable, and thus the associated risk premium $R(\Theta)$ is also unknown and must be inferred from available data. To provide a rigorous foundation and clarify the estimation principle, we define the Bayesian premium under the squared loss criterion as follows. Specifically, the Bayesian premium is the measurable function of the sample that minimizes the expected squared loss between the estimator and the true (but unobservable) risk premium $R(\Theta)$ (Bühlmann and Gisler, 2005).

Definition 2.2. *Given a sample* $\mathbf{X} = (X_1, \ldots, X_n)$*, the Bayesian premium of the risk* $R(\Theta)$ *is defined as its posterior expectation:*

$$\widehat{R_B(\Theta)} = \mathbb{E}[R(\Theta) \mid \mathbf{X}].$$

Equivalently, by decomposing $R(\Theta)$ *into its mean and risk loading components, we have*

$$\widehat{R_B(\Theta)} = \mathbb{E}[\mu(\Theta) \mid \mathbf{X}] + \mathbb{E}[g(\mu(\Theta), \gamma(\Theta)) \mid \mathbf{X}].$$

If Θ is a continuous random variable, the Bayesian premium can be equivalently expressed in integral form as

$$\widehat{R_B(\Theta)} = \frac{\int_\Omega \mu(\theta)\, \pi_\Theta(\theta) \prod_{i=1}^n f(x_i \mid \theta)\, d\theta}{\int_\Omega \pi_\Theta(\theta) \prod_{i=1}^n f(x_i \mid \theta)\, d\theta} + \frac{\int_\Omega g(\mu(\theta), \gamma(\theta))\, \pi_\Theta(\theta) \prod_{i=1}^n f(x_i \mid \theta)\, d\theta}{\int_\Omega \pi_\Theta(\theta) \prod_{i=1}^n f(x_i \mid \theta)\, d\theta},$$

where $(x_1, \ldots, x_n)$ denotes the observed sample corresponding to $(X_1, \ldots, X_n)$, and Ω is the support of Θ.

Remark 2.1. To compute the Bayesian premium $\widehat{R_B(\Theta)}$ for a given loading function g(x,y), one must specify both the conditional distribution of the claim counts and the prior distribution of the underlying risk parameter. Consider the case where, conditional on $\Theta = \theta$, the observations $X_1, \ldots, X_n$ are independent and identically distributed as Poisson(θ), and the parameter Θ itself follows a Gamma(α, β) distribution. By conjugacy, the posterior distribution of Θ remains Gamma with updated parameters $(\alpha + S, \beta + n)$, where $S = \sum_{i=1}^n X_i$ denotes the total number of observed claims.

Under this specification, the hypothetical mean and process variance satisfy

$$\mu(\Theta) = \gamma(\Theta) = \Theta, \qquad \mathbb{E}[\Theta \mid \mathbf{X}] = \frac{\alpha + S}{\beta + n}.$$

Since the general risk premium is defined in (2.1) and the Bayesian premium is the posterior expectation of $R(\Theta)$, the estimator simplifies to

$$\widehat{R_B(\Theta)} = \mathbb{E}[\Theta \mid \mathbf{X}] + \mathbb{E}[g(\Theta, \Theta) \mid \mathbf{X}].$$

https://doi.org/10.1017/asb.2026.10094 Published online by Cambridge University Press

Table 1. *Examples of g(x, y), corresponding risk premiums, and Bayesian premiums.*

Premium principle	$g(x, y)$	Risk premium $R(\Theta)$	Bayesian premium $\widehat{R_B(\Theta)}$
Net premium	0	$R_0(\Theta) = \mu(\Theta)$	$\frac{\alpha+S}{\beta+n}$
Expected value	$\rho_1 \cdot x$	$R_1(\Theta) = (1+\rho_1)\cdot\mu(\Theta)$	$(1+\rho_1)\,\frac{\alpha+S}{\beta+n}$
Variance premium	$\rho_2 \cdot y$	$R_2(\Theta) = \mu(\Theta) + \rho_2 \cdot \gamma(\Theta)$	$(1+\rho_2)\,\frac{\alpha+S}{\beta+n}$
Standard deviation	$\rho_3 \cdot \sqrt{y}$	$R_3(\Theta) = \mu(\Theta) + \rho_3 \cdot \sqrt{\gamma(\Theta)}$	$\frac{\alpha+S}{\beta+n} + \frac{\rho_3\Gamma\left(\alpha+S+\frac{1}{2}\right)}{(\beta+n)^{1/2}\,\Gamma(\alpha+S)}$
Modified variance	$\rho_4 \cdot \frac{y}{x}$	$R_4(\Theta) = \mu(\Theta) + \rho_4 \cdot \frac{\gamma(\Theta)}{\mu(\Theta)}$	$\frac{\alpha+S}{\beta+n} + \rho_4$

Therefore, the explicit formulas in the last column of Table 1 follow from evaluating the posterior expectation $\mathbb{E}[g(\Theta, \Theta) \mid \mathbf{X}]$ under the Gamma posterior distribution for each specified loading function $g(x, y)$.

The results in Table 1 indicate that, under the Poisson–Gamma model, the Bayesian premiums largely retain the functional form of their classical counterparts. This is because the risk parameter Θ is replaced by its posterior mean, and the loading term is adjusted through the corresponding posterior moments. For the net premium and the expected value premium, this substitution preserves their proportional relationship exactly. For more complex loading functions, such as the standard deviation premium, the Bayesian premiums include additional terms involving ratios of Gamma functions. These closed-form expressions demonstrate the analytical tractability of the Poisson–Gamma framework and clarify how different loading principles are modified under Bayesian estimation.

This remark shows that the Bayesian premium naturally blends individual experience with collective information, reflecting the core idea of credibility theory: assigning optimal weights to the individual and collective components to obtain a stable and equitable premium. In practice, however, the true conditional distribution $f_\theta(x)$ and prior $\pi_\Theta(\theta)$ are rarely known, making the Bayesian premium difficult to compute directly. To address this issue, Bühlmann (1967) proposed restricting the estimator of $\mu(\Theta)$ to a linear function of the observations, leading to the classical linear credibility estimator. This provides a practical and interpretable alternative that preserves the essence of the Bayesian formulation.

By minimizing

$$\mathbb{E}\left[\left(\mu(\Theta) - a_0 - \sum_{i=1}^{n} a_i X_i\right)^2\right],$$

the optimal linear Bayesian estimate of $\mu(\Theta)$ takes the standard credibility form

$$\widehat{\mu_C(\Theta)} = Z_n\,\overline{X}_n + (1 - Z_n)\,\mu_0, \tag{2.2}$$

where $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ is the sample mean, $\mu_0 = \mathbb{E}[\mu(\Theta)]$ is the collective mean, and

$$Z_n = \frac{n\,\mathrm{Var}[\mu(\Theta)]}{n\,\mathrm{Var}[\mu(\Theta)] + \mathbb{E}[\gamma(\Theta)]} \tag{2.3}$$

is the credibility factor. Bühlmann's framework has become a cornerstone of non-life actuarial practice by providing a statistically sound method for combining individual and collective information.

However, classical Bühlmann credibility theory focuses solely on estimating the hypothetical mean $\mu(\Theta)$. In many non-life insurance applications, premium principles depend not only on the mean but also on the process variance, as in variance-related premium principles (cf. (2.1)). Consequently, the traditional approach is insufficient when pricing methods also require accurate estimation of process variance, such as under the variance or standard deviation premium principles.

This paper focuses on the credibility estimation of process variance under the framework of credibility theory, aiming to provide a practical and theoretically grounded method for premium calculation across

https://doi.org/10.1017/asb.2026.10094 Published online by Cambridge University Press

commonly used premium principles. To formalize the concept, we first introduce the definition of the credibility premium.

Definition 2.3. *A credibility estimator systematically combines individual risk experience with collective information to estimate future premiums. Let P_n denote the observed (experience-based) premium from a sample of size n, and P_0 denote the collective (aggregate) premium. The credibility estimator is defined as*

$$\widehat{P} = \xi(n) \cdot P_n + \big(1 - \xi(n)\big) \cdot P_0,$$

where the credibility factor $\xi(n)$ satisfies $0 \le \xi(n) \le 1$, increases with the sample size n, and approaches 1 as $n \to \infty$. This formulation ensures a coherent and stable balance between individual and collective information.

Remark 2.2. The definition above provides a unified representation of credibility estimators across different settings. In particular, the relationship between the individual experience P_n and the collective premium P_0 depends on whether the prior distribution is known:

(1) Prior distribution known (classical credibility): In this case, P_0 corresponds to the fixed prior mean of the risk parameter (Bühlmann 1967; Chen *et al*., 2020). The individual experience P_n is not included in P_0, and the two components can be treated as independent. The credibility estimator is then a linear approximation to the Bayesian premium under the known prior.
(2) Prior distribution unknown (empirical Bayes credibility): Here, P_0 must be estimated from aggregated data, such as other policies in the portfolio or industry data (Mashayekhi, 2002; Bühlmann and Gisler, 2005; Zhang and Wen, 2024). The individual experience P_n may be included in the data used to estimate P_0, inducing statistical dependence between P_n and P_0. The credibility estimator remains a weighted average of the two components, but both are data-driven and potentially correlated.

In practical applications, such as launching a new line of business, the collective premium P_0 may be based on external industry data that exclude the individual insurer's own experience. In these cases, the individual and collective components can again be treated as independent.

This remark clarifies that Definition 2.3 does not assume P_n must be included in P_0; rather, the inclusion or exclusion depends on the specific setting. The subsequent theorems and methodology apply to all these scenarios, with appropriate consideration of independence or dependence between P_n and P_0.

Building on the unified credibility framework presented above, we next apply the classical Bühlmann methodology. Specifically, by constraining the risk premium estimate to be a linear function of the observed sample, we obtain the optimal linear Bayesian estimate of the risk premium, as formalized in the following theorem.

Theorem 2.1. *Under the variance premium principle, the optimal linear Bayesian estimation of the risk premium $R(\Theta)$ is obtained by solving:*

$$\min_{b_0, b_1, \ldots, b_n \in \mathbb{R}} \mathbb{E}\left[\left(R(\Theta) - b_0 - \sum_{i=1}^{n} b_i X_i\right)^2\right],$$

which is given by:

$$\widehat{R_C(\Theta)} = T_n \overline{X}_n + \left(1 - \frac{T_n \mu_0}{R_0}\right) R_0, \tag{2.4}$$

where $R_0 = \mathbb{E}[R(\Theta)]$, $\tau^2 = \mathrm{Var}[\mu(\Theta)]$, $\varphi = \mathrm{Cov}(R(\Theta), \mu(\Theta))$, $\sigma^2 = \mathbb{E}[\mathrm{Var}(X|\Theta)]$, and

$$T_n = \frac{n\varphi}{n\tau^2 + \sigma^2}.$$

Here, $\mathrm{Cov}(R(\Theta), \mu(\Theta))$ indicates the covariance between $R(\Theta)$ and $\mu(\Theta)$.

https://doi.org/10.1017/asb.2026.10094 Published online by Cambridge University Press

The proof of Theorem 2.1 is provided in Appendix A.1. From this result, the optimal linear Bayesian estimator $\widehat{R_C(\Theta)}$ admits a credibility-like weighted representation, in which the factor T_n plays a role analogous to that of a classical credibility factor. However, in contrast to standard credibility estimators, the resulting weights do not necessarily form a convex combination. More importantly, $\widehat{R_C(\Theta)}$ is, in general, not consistent for the true risk premium $R(\Theta)$. Consequently, in most practical settings, this estimator cannot be reliably interpreted as a credibility-based premium.

Theorem 2.2. *The optimal linear Bayesian estimator $\widehat{R_C(\Theta)}$ is consistent with respect to $R(\Theta)$ if and only if there exist constants s and t such that*

$$g\big(\mu(\Theta), \gamma(\Theta)\big) = s\,\mu(\Theta) + t.$$

Theorem 2.2 reveals an important structural implication for the linear Bayesian estimator $\widehat{R_C(\Theta)}$ defined in Theorem 2.1. In order for $\widehat{R_C(\Theta)}$ to be consistent with respect to $R(\Theta)$, the loading function $g(\mu(\Theta), \gamma(\Theta))$ must be linear in the hypothetical mean $\mu(\Theta)$.

This condition implies that the resulting premium principle effectively reduces to the expected premium principle. Consequently, the linear estimator in Theorem 2.1 does not provide a genuine generalization of commonly used premium principles. In particular, many variance-based loading functions widely used in actuarial practice do not satisfy this linearity requirement, and therefore the estimator $\widehat{R_C(\Theta)}$ will generally not be consistent under such premium principles.

This limitation motivates the development of the alternative credibility estimation methods presented in Section 3, which are designed to ensure consistent premium estimation while accommodating more general premium principles. The proof of Theorem 2.2 is provided in Appendix A.2.

Remark 2.3. The structural parameters appearing in Theorem 2.1, namely R_0, τ^2, φ, and σ^2, can be estimated using a fully nonparametric approach. This procedure relies solely on the empirical moments of the observed data and does not require any specific distributional assumptions on the latent risk parameter Θ. Accordingly, it provides a flexible and model-free framework for approximating the underlying structure of the premium model.

Suppose that m independent risk units are observed, with the ith unit containing n_i realizations X_{ij} $(j = 1, \ldots, n_i)$. It should be noted that the credibility estimators described above are not the ones recommended in this paper, as they often fail to yield consistent estimates (c.f. Theorem 2.2). For simplicity, we focus here on estimating the hyperparameters of the balanced model, where $n_1 = n_2 = \cdots = n_m = n$. For a detailed discussion on hyperparameter estimation in unbalanced models, the reader is referred to Bühlmann and Gisler (2005).

For each risk unit i $(i = 1, 2, \ldots, m)$, we define the sample mean and sample variance as

$$\overline{X}_{n_i} = \frac{1}{n}\sum_{j=1}^{n} X_{ij}, \qquad S^2_{n_i} = \frac{1}{n-1}\sum_{j=1}^{n}(X_{ij} - \overline{X}_{n_i})^2.$$

These statistics serve as empirical estimates of the latent parameters $\mu(\Theta_i)$ and $\gamma(\Theta_i)$, respectively, and constitute the fundamental quantities for subsequent nonparametric estimation of the structural parameters.

The structural parameters can then be estimated by

$$\hat{R}_0 = \frac{1}{m}\sum_{i=1}^{m}\big[\overline{X}_{n_i} + g\big(\overline{X}_{n_i}, S^2_{n_i}\big)\big], \quad \hat{\sigma}^2 = \frac{1}{m}\sum_{i=1}^{m} S^2_{n_i}, \quad \hat{\tau}^2 = \frac{1}{m-1}\sum_{i=1}^{m}(\overline{X}_{n_i} - \bar{\mu})^2 - \frac{1}{n}\hat{\sigma}^2,$$

and

$$\hat{\varphi} = \frac{1}{m-1}\sum_{i=1}^{m}\big[\big(\overline{X}_{n_i} + g\big(\overline{X}_{n_i}, S^2_{n_i}\big)\big) - \hat{R}_0\big](\overline{X}_{n_i} - \bar{\mu}),$$

where $\bar{\mu} = \frac{1}{m}\sum_{i=1}^{m}\overline{X}_{n_i}$ represents the overall sample mean across all risk units.

https://doi.org/10.1017/asb.2026.10094 Published online by Cambridge University Press

This nonparametric estimation framework accommodates any continuous specification of the function $g(x, y)$, including $g(x, y) = 0$, $g(x, y) = \rho_1 x$, $g(x, y) = \rho_2 y$, $g(x, y) = \rho_2\sqrt{y}$, and $g(x, y) = \rho \frac{y}{x}$ (see Table 1). This flexibility ensures broad applicability and practical relevance across a wide range of variance-related premium principles.

3. New credibility estimation of process variance

3.1 Known mean case

In the variance-related premium principle, the risk premium is primarily composed of two components: the hypothetical mean $\mu(\Theta)$ and the process variance $\gamma(\Theta)$. While the credibility estimate for the hypothetical mean has been extensively studied (c.f. (2.2)), obtaining a reliable estimate for the process variance component is equally crucial for accurate risk premium calculation.

Let $X_1, \ldots, X_n$ denote independent and identically distributed (i.i.d.) samples conditional on the parameter Θ, and let $\gamma(\Theta) = \mathrm{Var}(X_1|\Theta)$ represent the conditional variance of the population. Under the squared-error loss criterion, the optimal estimator of $\gamma(\Theta)$ is the posterior mean in the Bayesian framework:

$$\widehat{\gamma_B(\Theta)} = \mathbb{E}[\gamma(\Theta) \mid X_1, \ldots, X_n], \tag{3.1}$$

which serves as the Bayesian estimator of the process variance. In specific models, this Bayesian estimator can be expressed in a form equivalent to the classical credibility formula, demonstrating the close connection between Bayesian estimation and credibility theory.

The following example illustrates the Bayesian estimation of the process variance in the normal distribution case, showing how credibility theory can be applied in practical scenarios.

Example 3.1. *Consider the case where, conditional on $\Theta = \theta$, the observations $X_1, \ldots, X_n$ are independent and identically distributed as $N(\mu_0, \theta)$, and the process variance parameter Θ itself follows an inverse gamma distribution* $\mathrm{IG}(\alpha, \beta)$. *The conditional density of X_i given $\Theta = \theta$ is*

$$f_\theta(x) = \frac{1}{\sqrt{2\pi\theta}} \exp\left\{ -\frac{(x-\mu_0)^2}{2\theta} \right\},$$

while the prior density of Θ is

$$\pi_\Theta(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{-(\alpha+1)} e^{-\beta/\theta}.$$

By Bayes' theorem, the posterior distribution of Θ given the observed sample is

$$[\Theta \mid X_1, \ldots, X_n] \sim \mathrm{IG}\Big(\alpha + \frac{n}{2},\ \beta + \frac{1}{2}\sum_{i=1}^{n} (X_i - \mu_0)^2\Big),$$

which belongs to the inverse gamma conjugate family (see, e.g., Gelman et al. 2013). The corresponding Bayesian estimator of the process variance is

$$\widehat{\gamma_B(\Theta)} = \frac{\beta + \frac{1}{2}\sum_{i=1}^{n} (X_i - \mu_0)^2}{\alpha + \frac{n}{2} - 1}.$$

This Bayesian estimator can equivalently be expressed in the standard credibility form:

$$\widehat{\gamma_B(\Theta)} = w_n S_n^2 + (1 - w_n)\gamma_0,$$

where

$$S_n^2 = \frac{1}{n}\sum_{i=1}^{n} (X_i - \mu_0)^2, \quad \gamma_0 = \frac{\beta}{\alpha - 1}, \quad w_n = \frac{n}{n + 2(\alpha - 1)}.$$

https://doi.org/10.1017/asb.2026.10094 Published online by Cambridge University Press

The posterior variance of $\gamma(\Theta)$ is given by

$$\mathrm{Var}[\gamma(\Theta) \mid X_1, \ldots, X_n] = \frac{\left(\beta + \frac{n}{2} S_n^2\right)^2}{\left(\alpha + \frac{n}{2} - 1\right)^2 \left(\alpha + \frac{n}{2} - 2\right)}.$$

Remark 3.1. Note that, for $\Theta \sim \mathrm{IG}(\alpha, \beta)$, the mean and variance are given by

$$\mathbb{E}[\Theta] = \frac{\beta}{\alpha - 1}, \qquad \mathrm{Var}(\Theta) = \frac{\beta^2}{(\alpha - 1)^2(\alpha - 2)}, \quad \alpha > 2.$$

Hence, the posterior variance of the process variance $\gamma(\Theta)$ is well-defined only when $\alpha > 2$.

Furthermore, the marginal distribution of each X_i can be derived by integrating out Θ:

$$\begin{aligned} f_{X_i}(x) &= \int_0^\infty \frac{1}{\sqrt{2\pi\theta}} \exp\left\{ -\frac{(x-\mu_0)^2}{2\theta} \right\} \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{-(\alpha+1)} e^{-\beta/\theta} \, d\theta \\ &= \frac{\Gamma\left(\alpha + \frac{1}{2}\right)}{\Gamma(\alpha)\sqrt{2\pi\beta}} \left[1 + \frac{(x-\mu_0)^2}{2\beta} \right]^{-(\alpha+1/2)}. \end{aligned}$$

This is precisely a Student's t-distribution with 2α degrees of freedom, location parameter μ_0, and scale parameter $\sqrt{\beta/\alpha}$, that is,

$$X_i \sim t_{2\alpha}\left(\mu_0, \sqrt{\beta/\alpha}\right).$$

This derivation explicitly shows how the heavy-tailed Inverse-Gamma prior for Θ induces a Student's t-marginal distribution for X_i, providing robustness in modeling and inference.

In Example 3.1, the weight w_n plays a central role in determining the posterior mean estimator $\widehat{\Theta}_B$ of the process variance parameter Θ. It quantifies the relative influence of the sample variance S_n^2 on $\widehat{\Theta}_B$ compared to the prior mean $\mathbb{E}[\Theta] = \beta/(\alpha - 1)$. It is straightforward to verify that $w_n \to 1$ as $n \to \infty$ and $w_n \to 0$ as $n \to 0$. Consequently, for large sample sizes, the Bayesian estimator $\widehat{\Theta}_B$ is primarily driven by the observed data through S_n^2, making it largely data-driven. In contrast, for small n, w_n is smaller, and the prior mean exerts greater influence, reflecting the more significant role of prior information when observations are limited.

The weight w_n therefore balances the contributions of the observed data and the prior information, allowing $\widehat{\Theta}_B$ to smoothly transition from reliance on the prior under scarce data conditions to reliance on the sample information as the number of observations increases. This property ensures a robust estimation procedure that accounts for both data variability and prior knowledge (Bühlmann, 1967; Bühlmann and Gisler, 2005; Zhang, 2023).

In practical insurance applications, the conditional distribution of the population and the distribution of the risk parameter are often unknown, which generally prevents exact Bayesian computations. While the example in Example 3.1 is specific and admits an explicit solution only under conjugate assumptions, it provides useful insight: Bayesian estimation of conditional parameters can be interpreted as a weighted combination of the sample variance and the prior mean of the process variance, closely aligning with the principles of traditional credibility theory.

To better utilize the principles of credibility theory for obtaining credibility estimates of process variance $\gamma(\Theta)$, we assume that the hypothetical mean $\mu(\Theta)$ is known, that is, $\mu(\Theta) = \mu_0$, and solve the following minimization problem:

$$\min_{a_0, a_1, \ldots, a_n \in \mathbb{R}} \mathbb{E}\left[\left(\gamma(\Theta) - a_0 - \sum_{i=1}^n a_i \, (X_i - \mu_0)^2 \right)^2 \right]. \tag{3.2}$$

The objective of the optimization problem (3.2) is to construct a linear estimator for the process variance $\gamma(\Theta)$ by minimizing the expected squared deviation between $\gamma(\Theta)$ and a linear combination

https://doi.org/10.1017/asb.2026.10094 Published online by Cambridge University Press

of the squared deviations $(X_i - \mu_0)^2$. This formulation is equivalent to the classical credibility problem applied to the transformed variable $Y_i = (X_i - \mu_0)^2$, where the estimation focus is shifted from the original observations X_i to their squared deviations.

To characterize the statistical structure of this problem, let $Y_i = (X_i - \mu_0)^2$ for $i = 1, \ldots, n$. Then, $Y_1, \ldots, Y_n$ are conditionally independent and identically distributed given Θ, with

$$\mathbb{E}(Y_i|\Theta) = \gamma(\Theta), \quad \text{Var}(Y_i|\Theta) = k_4(\Theta) - \gamma^2(\Theta),$$

where $k_4(\Theta) = \mathbb{E}[(X_1 - \mu_0)^4|\Theta]$. Taking expectations with respect to Θ yields

$$\mathbb{E}[\text{Var}(Y_i|\Theta)] = k_4 - \lambda - \gamma_0^2 = v,$$

where $k_4 = \mathbb{E}[k_4(\Theta)]$, $\lambda = \text{Var}[\gamma(\Theta)]$ and $\gamma_0 = \mathbb{E}[\gamma(\Theta)]$.

According to the standard methods in credibility theory (Bühlmann, 1967; Bühlmann and Gisler, 2005), the optimal linear credibility estimator in this setting is a weighted combination of the sample variance and the collective variance, with the weight determined by the relative magnitudes of the within-group and between-group variances. Consequently, we obtain the following corollary.

Corollary 3.1. *Under Assumptions A1 and A2 of the Bayesian model, if $\mu(\Theta) = \mu_0$, the optimal estimator of the process variance $\gamma(\Theta)$ obtained by solving (3.2) is given by*

$$\widehat{\gamma_C(\Theta)} = w_n^* S_n^2 + \left(1 - w_n^*\right) \gamma_0, \tag{3.3}$$

where $\gamma_0 = \mathbb{E}[\gamma(\Theta)]$ denotes the collective (a priori) variance, $S_n^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \mu_0)^2$ is the sample variance, and

$$w_n^* = \frac{n\lambda}{n\lambda + v} \tag{3.4}$$

is the credibility factor.

This corollary establishes a Bayesian–credibility framework for estimating the process variance $\gamma(\Theta)$. The result shows that the optimal estimator is a weighted average of the sample variance S_n^2 and the collective variance γ_0, with the credibility factor w_n^* reflecting the relative importance of the observed data and the prior information. Specifically, w_n^* depends on the ratio between the between-group variance λ and the within-group variance v and increases with the sample size n. Hence, as more data become available, the estimator increasingly relies on the empirical sample variance, while for smaller samples, the estimation remains more influenced by the prior collective information.

This balance between prior belief and observed data lies at the core of credibility theory, providing a natural Bayesian interpretation of the credibility premium principle. Moreover, the formulation ensures that the estimator $\widehat{\gamma_C(\Theta)}$ is unbiased and efficient under the assumed conditional independence and distributional conditions, highlighting its practical relevance in actuarial science and risk management.

Remark 3.2. In Example 3.1, under the normal model with known mean and an inverse-gamma prior, the Bayesian estimator of the process variance admits the linear representation

$$\widehat{\gamma_B(\Theta)} = w_n S_n^2 + (1 - w_n)\gamma_0,$$

which coincides exactly with the credibility estimator in Corollary 3.1. This agreement is structural rather than accidental. In the normal model with known mean, the variance parameter Θ is itself the scale parameter, and the likelihood depends on Θ only through the quadratic sufficient statistic $\sum_{i=1}^n (X_i - \mu_0)^2$. The inverse-gamma prior is conjugate with respect to this quadratic structure, so posterior updating simply modifies the hyperparameters via functions of $\sum_{i=1}^n (X_i - \mu_0)^2$. Consequently, the posterior mean $\mathbb{E}[\Theta \mid X_1, \ldots, X_n]$ is an affine function of S_n^2, yielding an exact credibility-type formula.

In contrast, for a general one-parameter exponential family with conjugate prior, the process variance $\gamma(\Theta) = \text{Var}(X_1 \mid \Theta)$ is typically a nonlinear function of the structural parameter Θ. Although conjugacy ensures analytical tractability of the posterior distribution of Θ, it does not guarantee that the posterior mean $\mathbb{E}[\gamma(\Theta) \mid X_1, \ldots, X_n]$ can be expressed as a linear function of quadratic statistics such as S_n^2. Unless

https://doi.org/10.1017/asb.2026.10094 Published online by Cambridge University Press

the variance parameter coincides directly with the canonical scale parameter and the sufficient statistic has a purely quadratic form, such an exact credibility representation generally fails.

Nevertheless, this lack of exact equivalence does not diminish the relevance of the credibility estimator. The estimator derived from (3.2) is the optimal predictor within the class of linear estimators under squared-error loss. Hence, even when the Bayesian estimator is not exactly linear, the credibility estimator can be interpreted as its best linear approximation, depending only on low-order moment quantities (γ_0, λ, v). This moment-based structure provides robustness and computational simplicity, making the credibility approach a practically attractive alternative when full distributional specification is uncertain or when conjugacy does not preserve linearity.

A systematic characterization of the distribution–prior combinations under which the Bayesian estimator of the process variance admits an exact credibility-type linear representation remains an interesting open problem beyond the scope of the present paper.

Remark 3.3. The credibility estimator $\widehat{\gamma_C(\Theta)}$ is unconditionally unbiased for γ_0 in the sense that

$$\mathbb{E}\big[\widehat{\gamma_C(\Theta)}\big] = w_n^* \,\mathbb{E}\big[S_n^2\big] + (1 - w_n^*)\gamma_0 = w_n^*\gamma_0 + (1 - w_n^*)\gamma_0 = \gamma_0 = \mathbb{E}\big[\gamma(\Theta)\big].$$

Its unconditional mean squared error can be written as

$$\begin{aligned}\mathrm{MSE}\big(\widehat{\gamma_C(\Theta)}\big) = (w_n^*)^2\,\mathbb{E}\big[(S_n^2 - \gamma(\Theta))^2\big] + (1 - w_n^*)^2\,\mathbb{E}\big[(\gamma_0 - \gamma(\Theta))^2\big] \\ + 2w_n^*(1 - w_n^*)\,\mathbb{E}\big[\big(S_n^2 - \gamma(\Theta)\big)(\gamma_0 - \gamma(\Theta))\big].\end{aligned}$$

Since $\mathbb{E}[S_n^2 \mid \Theta] = \gamma(\Theta)$, the cross term vanishes: $\mathbb{E}[(S_n^2 - \gamma(\Theta))(\gamma_0 - \gamma(\Theta))] = 0$. Hence, the unconditional MSE simplifies to

$$\mathrm{MSE}\big(\widehat{\gamma_C(\Theta)}\big) = (w_n^*)^2\,\mathbb{E}\big[\mathrm{Var}\big(S_n^2 \mid \Theta\big)\big] + (1 - w_n^*)^2\lambda = \lambda(1 - w_n^*) = w_n^*\frac{v}{n},$$

where v denotes the variance of S_n^2 and λ the prior variance of $\gamma(\Theta)$. It follows that $\mathrm{MSE}\big(\widehat{\gamma_C(\Theta)}\big) \to 0$ as $n \to \infty$.

Furthermore, we can prove the consistency and asymptotic normality of the credibility estimation.

Theorem 3.1. *Given* Θ, *the estimator* $\widehat{\gamma_C(\Theta)}$ *is consistent and asymptotically normal. Specifically, the following results hold:*

(1) $\widehat{\gamma_C(\Theta)} \xrightarrow{a.s.} \gamma(\Theta)$ *as* $n \to \infty$;

(2) $\sqrt{n}\left[\widehat{\gamma_C(\Theta)} - \gamma(\Theta)\right] \xrightarrow{d} \mathcal{N}\left(0, \sigma_\gamma^2(\Theta)\right)$ *as* $n \to \infty$, *where* $\sigma_\gamma^2(\Theta) = k_4(\Theta) - \gamma^2(\Theta)$ *is the conditional asymptotic variance.*

The proof is provided in Appendix A.3. This theorem establishes the large-sample properties of $\widehat{\gamma_C(\Theta)}$ under a fixed risk parameter Θ. The almost sure consistency ensures convergence to the true process variance, while asymptotic normality provides a basis for statistical inference in large samples. Together, these results justify the use of the proposed credibility estimator in practical applications where the sample size is moderate to large.

As discussed in Remark 3.2, we consider a hierarchical model in which, conditional on $\Theta = \theta$, the observations $X_1, \ldots, X_n$ are i.i.d. $N(\mu_0, \theta)$.

To explore the performance of the credibility estimator beyond conjugacy, we now replace the inverse-gamma prior with a non-conjugate specification. The results show that, even without a conjugate prior, the credibility estimator remains well-defined and closely approximates the Bayesian estimator, providing a practical and accurate alternative in such settings.

Example 3.2. *Consider a hierarchical model in which, conditional on* $\Theta = \theta$, *the observations* $X_1, \ldots, X_n$ *are i.i.d.* $N(\mu_0, \theta)$, *and the prior distribution of* Θ *is* Gamma(α, β) *with* $\alpha, \beta > 0$.

https://doi.org/10.1017/asb.2026.10094 Published online by Cambridge University Press

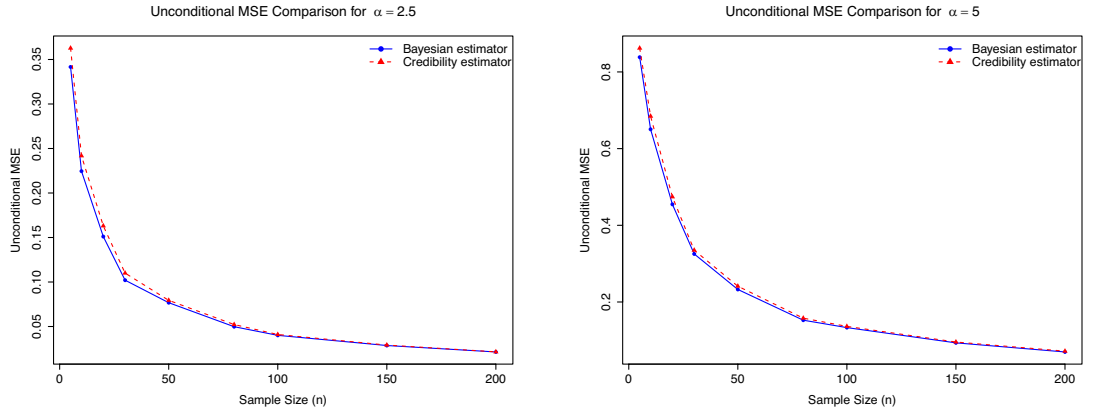


Figure 1. *Unconditional mean squared error of the Bayesian and credibility estimators for different sample sizes.*

Under this model, we have $\gamma(\Theta) = \Theta$. *Moreover,*

$$\mathbb{E}\big[(X_i - \mu_0)^2 \mid \Theta\big] = \Theta, \qquad \mathrm{Var}\big((X_i - \mu_0)^2 \mid \Theta\big) = 2\Theta^2.$$

It follows that the structural parameters in the credibility framework are

$$\gamma_0 = \mathbb{E}[\Theta] = \frac{\alpha}{\beta}, \quad \lambda = \mathrm{Var}(\Theta) = \frac{\alpha}{\beta^2}, \quad v = \mathbb{E}[2\Theta^2] = \frac{2\alpha(1+\alpha)}{\beta^2}.$$

The resulting credibility estimator is

$$\widehat{\gamma_C(\Theta)} = w_n^* S_n^2 + (1 - w_n^*)\gamma_0, \quad w_n^* = \frac{n}{n + 2(1+\alpha)}.$$

Since the Gamma prior is not conjugate to the normal likelihood with known mean, the posterior distribution of Θ *is not available in closed form. Consequently, the Bayesian estimator* $\widehat{\gamma_B(\Theta)} = \mathbb{E}[\Theta \mid X_1, \ldots, X_n]$ *is no longer an affine function of* S_n^2 *and must be evaluated numerically.*

In contrast, the credibility estimator remains explicit and depends only on the low-order moments (γ_0, λ, v). *To quantify the discrepancy between the Bayesian and credibility approaches, we consider the unconditional mean squared error (MSE),*

$$\mathrm{MSE}\big(\widehat{\gamma_v(\Theta)}\big) = \mathbb{E}\Big[\big(\widehat{\gamma_v(\Theta)} - \gamma(\Theta)\big)^2\Big], \quad v \in \{B, C\}, \tag{3.5}$$

which is evaluated via Monte Carlo simulation.

Figure 1 reports the simulated MSEs under different sample sizes. For all values of n, the Bayesian estimator exhibits a slightly lower MSE than the credibility estimator. However, the difference remains modest. As the sample size increases, the gap further diminishes and becomes negligible for moderate to large samples (e.g., $n \geq 100$*), indicating that the credibility estimator provides an accurate approximation to the Bayesian estimator in such settings.*

To further investigate the distributional behavior of the Bayesian and credibility estimators at a fixed value of $\Theta = 1.5$, *we examine QQ-plots for two representative sample sizes,* $n = 30$ *and* $n = 100$. *These plots allow us to visually assess how closely the empirical distributions of the two estimators align and to evaluate their approximate normality.*

Figure 2 shows that the points are approximately aligned along the diagonal for both sample sizes, indicating that the Bayesian and credibility estimators are nearly normally distributed. Mild deviations are observed for $n = 30$, *reflecting finite-sample differences, while for* $n = 100$ *the points lie closer to the diagonal, demonstrating increasingly similar distributions and supporting their asymptotic normality.*

https://doi.org/10.1017/asb.2026.10094 Published online by Cambridge University Press

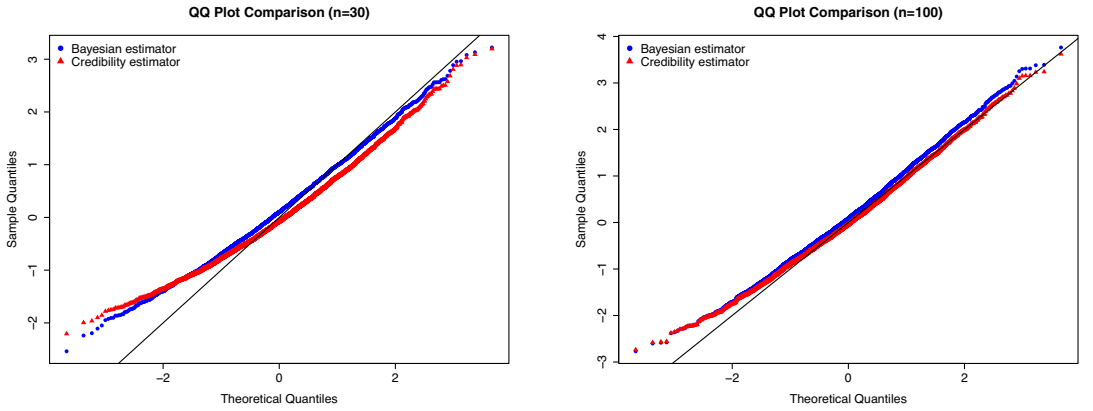


Figure 2. *QQ-plots of the Bayesian and credibility estimators at* $\Theta = 1.5$ *for* $n = 30$ *(left) and* $n = 100$ *(right).*

Example 3.2 highlights several important features arising in the non-conjugate setting. First, while the credibility estimator of the process variance retains an explicit closed-form expression, the corresponding Bayesian estimator is no longer analytically tractable and must be evaluated numerically, for instance via simulation-based methods.

Second, despite the loss of conjugacy, the credibility estimator continues to provide a close approximation to the Bayesian estimator. This is supported by the simulation results, which show that the two estimators have very similar mean squared errors, with only minor differences even for moderate sample sizes.

Finally, from a distributional perspective, both estimators exhibit asymptotic normality. As the sample size n increases, their empirical distributions become increasingly similar, further confirming that the credibility estimator serves as an accurate and computationally efficient proxy for the Bayesian estimator in large samples.

3.2 Unknown mean case

In the previous subsections, we considered the estimation of the process variance $\gamma(\Theta)$ under the assumption that the hypothetical mean $\mu(\Theta)$ is known. However, in many practical situations, this assumption does not hold, and the hypothetical mean must be estimated jointly with the process variance. This joint estimation increases the complexity of the risk premium evaluation $R(\Theta)$, since both the mean and the variance must be inferred from the available data.

Nevertheless, the Bühlmann credibility estimator $\widehat{\mu_C(\Theta)}$ in (2.2) already provides an optimal estimator of the hypothetical mean when the prior distribution is unknown. Therefore, in the context of variance-related premium principles, we will not further address the estimation of $\mu(\Theta)$. Instead, we directly substitute its Bühlmann credibility estimate when required. Our primary focus remains on developing a credibility estimator for the process variance $\gamma(\Theta)$.

Different from the previous subsection, when the hypothetical mean is unknown, the key idea is to replace the term $(X_i - \mu_0)^2$ with its centered counterpart $(X_i - \overline{X}_n)^2$, where $\overline{X}_n$ is the sample mean. To formalize this approach, we introduce the following quantities:

$$u_4(\Theta) = \mathbb{E}\big[(X_1 - \mu(\Theta))^4 \big| \Theta\big], \qquad \delta_n(\Theta) = \mathrm{Var}\big[\tilde{S}_n^2 \big| \Theta\big], \qquad \delta_n = \mathbb{E}[\delta_n(\Theta)],$$

where $\tilde{S}_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$ denotes the unbiased sample variance. Here, the quantity λ appearing later in Theorem 3.2 refers to the between-group variability of the process variance $\gamma(\Theta)$, as defined previously. The above notation summarizes the conditional fourth central moment and the conditional

https://doi.org/10.1017/asb.2026.10094 Published online by Cambridge University Press

and unconditional variability of $\tilde{S}_n^2$, which play a key role in the derivation of the credibility estimator for the unknown mean case.

This leads to the following theorem, which generalizes the credibility estimation of the process variance to the case of an unknown hypothetical mean.

Theorem 3.2. *Under the Bayesian model with unknown mean $\mu(\Theta)$, the optimal linear estimator of the process variance $\gamma(\Theta)$ is obtained by solving*

$$\min_{a_0,a_1,\ldots,a_n\in\mathbb{R}} \mathbb{E}\left[\left(\gamma(\Theta) - a_0 - \sum_{i=1}^{n} a_i (X_i - \overline{X}_n)^2\right)^2\right]. \tag{3.6}$$

The solution is given by the credibility-weighted estimator

$$\widetilde{\gamma_C(\Theta)} = t_n^* \, \tilde{S}_n^2 + (1 - t_n^*) \, \gamma_0, \tag{3.7}$$

where $\tilde{S}_n^2$ is the unbiased sample variance of the observations, $\gamma_0 = \mathbb{E}[\gamma(\Theta)]$ denotes the collective (a priori) variance, $\lambda = \mathrm{Var}[\gamma(\Theta)]$ is the between-group variance of the process variance, and $t_n^* = \frac{\lambda}{\lambda+\delta_n}$ is the credibility factor balancing the individual sample variance and the collective variance.

The proof of Theorem 3.2 is provided in Appendix A.4. This result extends the classical credibility framework to the case where the hypothetical mean is unknown. By using the centered observations $(X_i - \overline{X}_n)^2$, the estimator retains its interpretability and computational simplicity, while effectively adjusting for the uncertainty introduced by estimating the mean. As in the known-mean case, the credibility factor t_n^* reflects the relative weight assigned to the observed sample variance and the collective (prior) variance, ensuring a balanced integration of empirical data and prior structure.

Remark 3.4. By evaluating the conditional expectation and fourth-order moment of $\tilde{S}_n^2$ given Θ under the hierarchical framework (Benhamou, 2018), we have

$$\delta_n(\Theta) = \mathrm{Var}\left(\tilde{S}_n^2 \mid \Theta\right) = \frac{u_4(\Theta)}{n} - \frac{n-3}{n(n-1)} \gamma^2(\Theta),$$

and its unconditional expectation

$$\delta_n = \mathbb{E}[\delta_n(\Theta)] = \frac{u_4}{n} - \frac{(n-3)(\lambda+\gamma_0^2)}{n(n-1)} = \frac{1}{n}(u_4 - \lambda - \gamma_0^2) + \frac{2(\lambda+\gamma_0^2)}{n(n-1)}.$$

Consequently, the credibility factor can be expressed as

$$t_n^* = \frac{\lambda}{\lambda + \delta_n} = \frac{n}{n + \kappa_n},$$

where $\kappa_n = (u_4 - \lambda - \gamma_0^2)/\lambda + 2\,(\lambda + \gamma_0^2)/[\lambda(n-1)]$ is a constant of order $O(1)$ with respect to n.

This form explicitly shows the dependence of t_n^* on the sample size n: t_n^* increases with n and approaches 1 as $n \to \infty$. Intuitively, as the sample size grows, the credibility estimator gives more weight to the sample component and less to the prior, whereas for small n, the prior receives more weight, consistent with standard credibility theory.

Remark 3.5. The credibility estimation $\widetilde{\gamma_C(\Theta)}$ is an unconditionally unbiased estimator of γ_0 in the sense of $\mathbb{E}\left[\widetilde{\gamma_C(\Theta)}\right] = \mathbb{E}\left[\gamma(\Theta)\right] = \gamma_0$. In addition, the unconditional mean square error of $\widetilde{\gamma_C(\Theta)}$ is given by:

$$\mathrm{MSE}\left(\widetilde{\gamma_C(\Theta)}\right) = \mathbb{E}\left[\left(\widetilde{\gamma_C(\Theta)} - \gamma(\Theta)\right)^2\right] = \frac{\lambda\delta_n}{\lambda+\delta_n} = \lambda\left(1 - t_n^*\right) = t_n^* \delta_n,$$

and we have $\mathrm{MSE}\left(\widetilde{\gamma_C(\Theta)}\right) = O\left(\frac{1}{n}\right) \to 0$ as $n \to \infty$.

https://doi.org/10.1017/asb.2026.10094 Published online by Cambridge University Press

The remark addresses two fundamental properties of the credibility estimator $\widetilde{\gamma_C(\Theta)}$. This property ensures that, on average, the estimator neither overestimates nor underestimates the true parameter value γ_0. The MSE formula shows that as the sample size n increases, the MSE decreases at a rate proportional to $1/n$, asymptotically approaching zero as $n \to \infty$. This indicates that the estimator's accuracy improves with larger sample sizes, which is a desirable characteristic.

Example 3.3. *Suppose that, conditional on the parameter $\Theta = \theta$, the samples $X_1, \ldots, X_n$ are independent and identically distributed, each following a normal distribution $N(\mu(\theta), \gamma(\theta))$ with conditional density*

$$f_\theta(x) = \frac{1}{\sqrt{2\pi\gamma(\theta)}} \exp\left[-\frac{(x-\mu(\theta))^2}{2\gamma(\theta)}\right], \quad x \in \mathbb{R}.$$

The parameter Θ is assumed to be a random variable following an unknown prior density $\pi_\Theta(\theta)$.

According to Theorem 3.2, the credibility estimation (3.7) still holds under this scenario. In this case, the credibility estimator provides the best linear approximation to the Bayesian estimator of the process variance $\gamma(\Theta)$, thereby offering a practical and robust alternative when the exact prior distribution is unavailable.

Due to the fact that the excess kurtosis of a normal distribution is zero, we have

$$u_4(\Theta) = 3\gamma^2(\Theta),$$

and thus

$$\delta_n(\Theta) = \frac{2\gamma^2(\Theta)}{n-1}.$$

Accordingly, the credibility factor t_n^* can be simplified as

$$t_n^* = \frac{\lambda}{\lambda + 2(\lambda+\gamma_0^2)/(n-1)} = \frac{(n-1)\lambda}{(n-1)\lambda + 2(\lambda+\gamma_0^2)}.$$

Therefore, the credibility estimation of $\gamma(\Theta)$ can be expressed as

$$\widetilde{\gamma_C(\Theta)} = \frac{(n-1)\lambda\,\tilde{S}_n^2}{(n-1)\lambda + 2(\lambda+\gamma_0^2)} + \frac{2(\lambda+\gamma_0^2)\,\gamma_0}{(n-1)\lambda + 2(\lambda+\gamma_0^2)}. \tag{3.8}$$

This formulation highlights an important property of credibility estimation: even without explicit assumptions about the prior distribution, the estimator $\widetilde{\gamma_C(\Theta)}$ maintains a functional form that is closely aligned with the Bayesian estimator. Consequently, it provides a robust and interpretable approximation that effectively balances empirical data information with collective-level variance characteristics.

This example demonstrates how credibility estimation can be effectively applied to normally distributed data when the prior distribution is unknown. By leveraging the properties of the normal distribution and simplifying the problem, the example derives a credibility factor and estimator that appropriately balance the influence of sample data and prior information. Notably, because the kurtosis under the normal distribution is zero, the credibility estimator is elegantly simplified, ensuring that the estimation process remains both efficient and accurate.

Remark 3.6. In general, when the conditional distribution of the risk variable X given Θ has zero excess kurtosis (i.e., a kurtosis equal to 3), it follows that

$$u_4(\Theta) = \mathbb{E}\left[(X-\mu(\Theta))^4 \,\middle|\, \Theta\right] = 3\gamma^2(\Theta),$$

where $\gamma(\Theta) = \mathrm{Var}[X \mid \Theta]$. Under this condition, the credibility estimator can be simplified to the form given in (3.8).

While the normal distribution is the most well-known and canonical example exhibiting zero excess kurtosis, it is not the only one. Other symmetric distributions, including certain parameterizations of the generalized normal (exponential power) distribution, may also have zero excess kurtosis. This highlights

https://doi.org/10.1017/asb.2026.10094 Published online by Cambridge University Press

that the simplification of the credibility estimator based on the kurtosis property is applicable beyond the normal case. For further discussion on distributions with zero excess kurtosis, see, for example, Bühlmann (2007).

Overall, this remark emphasizes that the key condition for the simplified form of the credibility estimator is the kurtosis of the conditional distribution rather than normality per se, thereby extending the applicability of the result to a broader class of symmetric distributions.

Theoretically, regardless of whether the hypothetical mean is known, the Bayesian estimator of the process variance is optimal under the squared loss function. Specifically, the Bayesian estimator of $\gamma(\Theta)$ is given by

$$\widetilde{\gamma_B(\Theta)} = \mathbb{E}\big[\gamma(\Theta) \mid X_1, X_2, \ldots, X_n\big] = \int_{\Omega} \gamma(\theta)\, \pi(\theta \mid X_1, X_2, \ldots, X_n)\, d\theta, \tag{3.9}$$

where Ω denotes the support of Θ. However, the posterior density $\pi(\theta \mid X_1, X_2, \ldots, X_n)$ is typically unknown because the conditional distribution $f_\theta(x)$ or the prior distribution $\pi_\Theta(\theta)$ is unspecified. Consequently, the Bayesian estimator $\widetilde{\gamma_B(\Theta)}$ cannot be obtained in closed form.

However, with a sufficiently large sample size, we can establish the asymptotic properties of both the Bayesian and credibility estimators. Before presenting these results, it is useful to recall a fundamental structural feature of credibility estimators: they are typically expressed as a convex combination of a constant (representing prior or collective information) and a moment-based estimator computed from the sample. As the sample size increases, the weight placed on the sample-based component tends to one, while the weight placed on the constant term diminishes. This property implies that credibility estimators automatically recover the classical moment estimators in large samples, which in turn guarantees their consistency and asymptotic normality.

More importantly, credibility estimators are designed to perform well in small samples. By optimally combining prior information with limited empirical data, they achieve a smaller mean squared error than purely sample-based estimators when n is small. Thus, credibility estimation not only satisfies the desirable large-sample properties of consistency and asymptotic Gaussianity but also offers strong small-sample efficiency, which is a central motivation for the development of credibility theory.

Theorem 3.3. *Under the Bayesian models and in the case of an unknown mean, consider the credibility estimator* $\widetilde{\gamma_C(\Theta)}$ *and the Bayesian estimator* $\widetilde{\gamma_B(\Theta)}$ *of the process variance* $\gamma(\Theta)$*. Then, as* $n \to \infty$*, the following properties hold:*

(1) Both estimators are consistent:

$$\widetilde{\gamma_C(\Theta)} \xrightarrow{a.s.} \gamma(\Theta), \quad \widetilde{\gamma_B(\Theta)} \xrightarrow{a.s.} \gamma(\Theta).$$

(2) The credibility estimator is asymptotically normal:

$$\sqrt{n}\Big[\widetilde{\gamma_C(\Theta)} - \gamma(\Theta)\Big] \xrightarrow{d} \mathcal{N}\big(0, \sigma_\upsilon^2(\Theta)\big),$$

where $\sigma_\upsilon^2(\Theta) = u_4(\Theta) - \gamma^2(\Theta)$.

Details of the proof can be found in Appendix A.5. Theorem 3.3 confirms that both the credibility and Bayesian estimators possess desirable large-sample properties. Consistency ensures convergence to the true process variance, while asymptotic normality provides a foundation for statistical inference in settings with sufficiently large sample sizes.

To provide a concise summary of the similarities and differences between the Bayesian and credibility estimators introduced in Theorems 3.3, we include the following compact comparison table. This table highlights key features, including consistency, use of prior information, optimality, model dependence, closed-form availability, and typical MSE behavior.

https://doi.org/10.1017/asb.2026.10094 Published online by Cambridge University Press

Table 2. *Comparison of Bayesian and credibility estimators of the process variance* $\gamma(\Theta)$.

Feature	Bayesian $\widetilde{\gamma_B(\Theta)}$	Credibility $\widetilde{\gamma_C(\Theta)}$
Consistency	Yes; converges to $\gamma(\Theta)$	Yes; converges to $\gamma(\Theta)$
Prior/structural info	Uses prior information	Uses structural parameters
Optimality	Minimises expected squared loss over all functions	Minimises loss over sample-summary class
Model dependence	Requires full likelihood and prior; may lack closed form	Depends on low-order moments; always closed form
Closed-form	Only in certain special conjugate cases	Always available
MSE	Typically smaller	Typically larger

This table allows the reader to clearly visualize the relationship between the two estimators. It demonstrates that while both $\widetilde{\gamma_C(\Theta)}$ and $\widetilde{\gamma_B(\Theta)}$ are consistent and asymptotically valid, the Bayesian estimator generally achieves smaller mean squared errors due to its optimality among all measurable functions and its model-dependent flexibility. Overall, the table provides a concise and accessible illustration of the theoretical properties discussed above, thereby enhancing both the readability and practical understanding of the estimators.

Although Bayesian estimation typically has a smaller mean squared error than credibility estimation, credibility estimation is more practical. Even when the population and sample distributions are unknown, credibility estimation still provides explicit expressions. In contrast, Bayesian estimation heavily relies on the specific forms of the population and sample distributions. Moreover, even if these distributions are known, Bayesian estimation may not yield a closed-form solution and often requires numerical integration methods, such as Markov Chain Monte Carlo (MCMC), for computation. Therefore, in practical applications, credibility estimation is generally preferred over Bayesian estimation.

4. Numerical illustrations and applications

4.1 Numerical comparisons

The variance-related premium principle has significant applications in insurance and risk management. By considering both the expected value and the variance of the risk, this principle provides a more comprehensive evaluation of risk compared to traditional pure premium principles. This section discusses practical applications of the variance-related premium principle, highlighting how credibility estimation can be effectively used to estimate risk premiums.

Definition 4.1. *Under the Bayesian models, the credibility estimation of the risk premium* $R(\Theta)$ *by the "plug-in" principle is defined by:*

$$\widetilde{R_C(\Theta)} = \widetilde{\mu_C(\Theta)} + g\left(\widetilde{\mu_C(\Theta)}, \widetilde{\gamma_C(\Theta)}\right) \tag{4.1}$$

where $\widetilde{\mu_C(\Theta)}$ *and* $\widetilde{\gamma_C(\Theta)}$ *are the credibility estimators for the hypothetical mean* $\mu(\Theta)$ *and the process variance* $\gamma(\Theta)$ *given in (2.2) and (3.7), respectively.*

The plug-in credibility estimator defined above provides a unified framework for constructing risk premium estimates under various premium principles by substituting the credibility estimators of the hypothetical mean and process variance into the functional form of the premium principle. This formulation not only preserves the interpretability of the classical credibility approach but also enables its extension to a broader class of variance-related premium principles. To highlight the generality and

https://doi.org/10.1017/asb.2026.10094 Published online by Cambridge University Press

practical advantages of this framework, it is instructive to compare it with the method proposed by Yong *et al*. (2025), which focuses specifically on the variance premium principle.

Remark 4.1. Yong et al. (2025) investigated credibility estimation under the classical variance premium principle, where the risk premium is expressed as

$$R(\Theta) = \mu(\Theta) + \rho\gamma(\Theta).$$

Specifically, they considered a class of transformed quadratic functions of the sample,

$$\Phi = \left\{ a_0 + \sum_{i=1}^{n} a_i X_i + \sum_{i=1}^{n} b_i X_i^2 + \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij} X_i X_j \;\middle|\; a_i, b_i, c_{ij} \in \mathbb{R} \right\},$$

and obtained the credibility estimator by minimizing the expected squared error:

$$\min_{\widehat{R(\Theta)}\in\Phi} \mathbb{E}\left[\left(\mu(\Theta) + \rho\gamma(\Theta) - \widehat{R(\Theta)} \right)^2 \right].$$

This provides a consistent estimate for the variance-based risk premium and represents a significant extension of Bühlmann's classical credibility theory. Notably, as mentioned on page 6 of Yong et al. (2025), their credibility estimation involves at least 14 structural parameters, which are generally non-trivial to estimate. Furthermore, because the two components $\mu(\Theta)$ and $\gamma(\Theta)$ are generally dependent, an exact computation of the variance of $R(\Theta)$ is challenging.

Building on this foundation, the framework proposed in this paper introduces several key features. First, we separate the estimation of the hypothetical mean $\mu(\Theta)$ and the process variance $\gamma(\Theta)$. Since the mean and variance have different dimensions, estimating them separately is both necessary and statistically reasonable. The process variance is constructed as a linear combination of transformed sample data, minimizing squared error in a manner analogous to Bühlmann's approach for the mean. This separation enhances analytical tractability and stability.

Second, our method requires only a small number of easily interpretable hyperparameters. As shown in Table 6 of Subsection 4.2, the framework involves only five structural parameters, all of which can be readily estimated using an empirical Bayes approach. This parsimonious structure facilitates practical implementation and reliable estimation even with limited data, while maintaining flexibility to accommodate different types of premium functions.

Finally, by combining the separately estimated $\mu(\Theta)$ and $\gamma(\Theta)$ through a plug-in variance-related premium function

$$R(\Theta) = \mu(\Theta) + g(\mu(\Theta), \gamma(\Theta)),$$

our framework naturally extends to a broader class of premium principles, including the variance, standard deviation, and modified variance principles. This approach preserves desirable statistical properties such as conditional unbiasedness, mean squared error convergence, consistency, and asymptotic normality, which can be established analytically.

Both the Bayesian estimator in Definition 2.1 and the credibility estimator in (4.1) share a common analytical structure. Each consists of a *location term*, representing the estimated expected loss, and a *risk loading term*, determined by the function $g(x, y)$, which captures the adjustment for higher-order uncertainty.

The Bayesian estimator computes these components using posterior expectations of the model parameters, providing an exact posterior risk premium under the assumed prior-likelihood framework. In contrast, the credibility estimator replaces these posterior quantities with their credibility estimates of the hypothetical mean and process variance given in (2.2) and (3.7). This yields a practical, data-driven approximation that preserves the Bayesian interpretation while avoiding explicit prior specification and posterior computation.

Thus, both estimators belong to the same family, distinguished primarily by how expectations are evaluated: the Bayesian estimator serves as the theoretical benchmark, whereas the credibility estimator

https://doi.org/10.1017/asb.2026.10094 Published online by Cambridge University Press

offers an operational alternative suited for actuarial applications with limited data or a preference for computational simplicity.

Table 3 summarizes the explicit forms of both estimators for several common premium principles, highlighting their structural coherence and the credibility estimator's role as a natural extension of the Bayesian approach.

The variance-related premium principle extends traditional actuarial pricing methodologies by incorporating not only the expected loss but also the associated uncertainty, typically represented by the variance. The estimators summarized in Table 3 offer a practical means of implementing this generalized principle. To demonstrate their application, we present a numerical example in which the latent risk variable follows a Gamma distribution, while the observed claim amounts are conditionally exponential given the risk parameter.

Example 4.1. *Let $\Theta = \theta$ denote the underlying risk parameter, and assume that $X_1, \ldots, X_n$ are conditionally i.i.d.* Exp(θ) *with density*

$$f_\theta(x) = \theta e^{-\theta x}, \qquad x > 0.$$

Hence,

$$\mu(\Theta) = \mathbb{E}(X \mid \Theta) = \Theta^{-1}, \qquad \gamma(\Theta) = \mathrm{Var}(X \mid \Theta) = \Theta^{-2}.$$

The prior distribution is taken as $\Theta \sim \Gamma(\alpha, \beta)$ with density

$$\pi_\Theta(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta}, \qquad \theta > 0,\ \alpha > 4,$$

where $\alpha > 4$ ensures the existence of the required higher-order moments. Under the Exponential-Gamma hierarchy, the marginal distribution of X_i is

$$f_{X_i}(x) = \alpha\beta^\alpha (x+\beta)^{-(\alpha+1)}, \qquad x > 0,$$

that is, a Pareto (Type II)/Lomax distribution (Lemonte and Cordeiro, 2013), exhibiting the heavy-tailed behavior induced by the mixing. Moreover, $\mu(\Theta)$ follows an inverse-gamma(α, β) distribution.

Given the sample $\mathbf{X} = (X_1, \ldots, X_n)$ with mean $\overline{X}_n$, conjugacy implies

$$[\Theta \mid \mathbf{X}] \sim \Gamma(\alpha + n, \beta + n\overline{X}_n),$$

so that the Bayesian estimators of the hypothetical mean and process variance are

$$\mathbb{E}[\mu(\Theta) \mid \mathbf{X}] = \frac{\beta + n\overline{X}_n}{\alpha + n - 1}, \qquad \mathbb{E}[\gamma(\Theta) \mid \mathbf{X}] = \frac{(\beta + n\overline{X}_n)^2}{(\alpha + n - 1)(\alpha + n - 2)}.$$

Consequently, all premium principles expressed in terms of $(\mu(\Theta), \gamma(\Theta))$ admit closed Bayesian forms. For example, the variance premium estimator is

$$\widetilde{R_2^B(\Theta)} = \mathbb{E}[\mu(\Theta) \mid \mathbf{X}] + \rho_2\, \mathbb{E}[\gamma(\Theta) \mid \mathbf{X}].$$

The prior moments are given by

$$\mu_0 = \frac{\beta}{\alpha - 1}, \quad \tau^2 = \frac{\beta^2}{(\alpha-1)^2(\alpha-2)}, \quad \gamma_0 = \frac{\beta^2}{(\alpha-1)(\alpha-2)}, \quad \lambda = \frac{2\beta^4(2\alpha-5)}{(\alpha-1)^2(\alpha-2)^2(\alpha-3)(\alpha-4)}.$$

For exponential claims, the fourth conditional central moment leads to

$$u_4(\Theta) = \frac{9}{\Theta^4}, \quad \delta_n(\Theta) = \frac{2(4n-3)}{\Theta^4 n(n-1)}, \quad \delta_n = \frac{2(4n-3)\,\beta^4}{n(n-1)(\alpha-1)(\alpha-2)(\alpha-3)(\alpha-4)}.$$

The associated credibility factors for the hypothetical mean and process variance are

$$Z_n = \frac{n\tau^2}{n\tau^2 + \gamma_0} = \frac{n}{n + \alpha - 1}, t_n^* = \frac{\lambda}{\lambda + \delta_n} = \frac{1}{1 + \frac{(4n-3)}{n(n-1)} \cdot \frac{(\alpha-1)(\alpha-2)}{2\alpha-5}}.$$

https://doi.org/10.1017/asb.2026.10094 Published online by Cambridge University Press

https://doi.org/10.1017/asb.2026.10094 Published online by Cambridge University Press

Table 3. *Bayesian and credibility estimators of variance-related risk premiums.*

Premium principle	$g(x, y)$	Bayesian estimator	Credibility estimator
Net	$g(x,y)=0$	$\widetilde{R_0^B(\Theta)}=\mathbb{E}\left[\mu(\Theta)\vert\mathbf{X}\right]$	$\widetilde{R_0^C(\Theta)}=\widetilde{\mu_C(\Theta)}$
Expected value	$g(x,y)=\rho_1\cdot x$	$\widetilde{R_1^B(\Theta)}=(1+\rho_1)\mathbb{E}\left[\mu(\Theta)\vert\mathbf{X}\right]$	$\widetilde{R_1^C(\Theta)}=(1+\rho_1)\cdot\widetilde{\mu_C(\Theta)}$
Variance	$g(x,y)=\rho_2\cdot y$	$\widetilde{R_2^B(\Theta)}=\mathbb{E}\left[\mu(\Theta)\vert\mathbf{X}\right]+\rho_2\cdot\mathbb{E}\left[\gamma(\Theta\mid\mathbf{X})\right]$	$\widetilde{R_2^C(\Theta)}=\widetilde{\mu_C(\Theta)}+\rho_2\cdot\widetilde{\gamma_C(\Theta)}$
Standard deviation	$g(x,y)=\rho_3\cdot\sqrt{y}$	$\widetilde{R_3^B(\Theta)}=\mathbb{E}\left[\mu(\Theta)\vert\mathbf{X}\right]+\rho_3\cdot\mathbb{E}\left[\sqrt{\gamma(\Theta)}\mid\mathbf{X}\right]$	$\widetilde{R_3^C(\Theta)}=\widetilde{\mu_C(\Theta)}+\rho_3\cdot\sqrt{\widetilde{\gamma_C(\Theta)}}$
Modified variance	$g(x,y)=\rho_4\cdot\frac{y}{x}$	$\widetilde{R_4^B(\Theta)}=\mathbb{E}\left[\mu(\Theta)\vert\mathbf{X}\right]+\rho_4\cdot\mathbb{E}\left[\gamma(\Theta)/\mu(\Theta)\mid\mathbf{X}\right]$	$\widetilde{R_4^C(\Theta)}=\widetilde{\mu_C(\Theta)}+\rho_4\cdot\frac{\widetilde{\gamma_C(\Theta)}}{\widetilde{\mu_C(\Theta)}}$

Table 4. *Conditional MSEs of Bayesian and credibility estimators.*

	$\Theta = 0.6$			$\Theta = 1.0$		
Estimator	$n = 10$	$n = 50$	$n = 200$	$n = 10$	$n = 50$	$n = 200$
$\text{CMSE}_{B,2}$	0.5013	0.1385	0.0375	0.0970	0.0337	0.0092
$\text{CMSE}_{C,2}$	0.5440	0.1550	0.0427	0.0847	0.0336	0.0098
$\text{CMSE}_{B,3}$	0.3420	0.0871	0.0230	0.0832	0.0290	0.0080
$\text{CMSE}_{C,3}$	0.3651	0.0924	0.0242	0.0741	0.0287	0.0083
$\text{CMSE}_{B,4}$	0.3493	0.0877	0.0230	0.0836	0.0290	0.0080
$\text{CMSE}_{C,4}$	0.3878	0.1034	0.0276	0.0688	0.0309	0.0094

These expressions provide all quantities required to compute both the Bayesian and credibility estimators for the considered premium principles. Their accuracy can be assessed through the conditional mean squared error (CMSE),

$$\text{CMSE}_{k,i}\left[\widetilde{R_i^k(\Theta)}\right] = \mathbb{E}\left[\left(\widetilde{R_i^k(\Theta)} - R_i(\Theta)\right)^2 \middle| \Theta\right],$$

and the unconditional mean squared error (UMSE),

$$\text{UMSE}_{k,i}\left[\widetilde{R_i^k(\Theta)}\right] = \mathbb{E}\left[\left(\widetilde{R_i^k(\Theta)} - R_i(\Theta)\right)^2\right],$$

where $k = B$ or C and $i = 0, 1, 2, 3, 4$ correspond to the net, expected value, variance, standard deviation, and modified variance premium principles.

For the parameter choices $\alpha = 6$, $\beta = 4$, $\rho_1 = \rho_2 = 0.2$, and $\rho_3 = \rho_4 = 0.3$, Monte Carlo simulations in R are conducted to compute the conditional mean squared errors (conditional MSEs, CMSEs) for risk levels $\Theta = 0.6$ and $\Theta = 1.0$ across sample sizes $n = 10, 50, 200$. The resulting conditional MSEs are reported in Table 4.[1] *Here, we only display the results for $i = 2, 3, 4$ because the premiums for $i = 0$ and $i = 1$ coincide under both Bayesian and credibility approaches and including them would be redundant.*

Furthermore, to evaluate the unconditional mean squared errors (unconditional MSEs, UMSEs), values of Θ are randomly generated from the Gamma prior. Two different parameter settings (α, β) are considered, and numerical simulations are conducted to compute the unconditional MSEs for each credibility estimator under the various premium principles. The resulting unconditional MSEs are reported in Table 5.

From Tables 4 to 5, it is evident that both conditional MSEs and unconditional MSEs decline substantially as the sample size n increases. The errors are consistently smaller for larger values of Θ, reflecting the higher precision attainable for less volatile risks. Moreover, while the Bayesian and credibility estimators exhibit almost identical performance for large sample sizes, the Bayesian estimator generally achieves slightly lower mean squared errors when the sample size is small.

Finally, we examine the asymptotic behavior of the credibility estimator $\widetilde{R_i^C(\Theta)}$ for $\alpha = 6$, $\beta = 4$, $\rho_1 = \rho_2 = 0.2$, $\rho_3 = \rho_4 = 0.3$, and $\Theta = 1.4$. Figure 3 presents kernel density estimates of $\widetilde{R_i^C(\Theta)}$ for sample sizes $n = 5$ and $n = 50$, with dashed lines indicating the theoretical premium values. As the sample size increases from $n = 5$ to $n = 50$, the empirical distributions become more symmetric and more concentrated around the theoretical values, providing clear numerical evidence in favor of the estimator's asymptotic normality.

[1] Since

$$\widehat{R_0^B(\Theta)} = \widetilde{R_0^C(\Theta)} = \frac{\beta + n\overline{X}_n}{\alpha + n - 1}, \quad \widehat{R_1^B(\Theta)} = \widetilde{R_1^C(\Theta)} = (1 + \rho_1)\frac{\beta + n\overline{X}_n}{\alpha + n - 1},$$

the net and expected value premiums coincide under both approaches; the tables therefore report only $i = 2, 3, 4$.

https://doi.org/10.1017/asb.2026.10094 Published online by Cambridge University Press

Table 5. *Unconditional MSEs of Bayesian and credibility estimators.*

	$\alpha = 5, \beta = 3$			$\alpha = 6, \beta = 4$		
Estimator	$n = 10$	$n = 50$	$n = 200$	$n = 10$	$n = 50$	$n = 200$
$\text{UMSE}_{B,2}$	0.1248	0.0294	0.0112	0.1259	0.0399	0.0103
$\text{UMSE}_{C,2}$	0.1320	0.0326	0.0135	0.1349	0.0456	0.0119
$\text{UMSE}_{B,3}$	0.0902	0.0215	0.0067	0.0897	0.0263	0.0069
$\text{UMSE}_{C,3}$	0.0935	0.0224	0.0070	0.0926	0.0274	0.0073
$\text{UMSE}_{B,4}$	0.0901	0.0215	0.0067	0.0896	0.0263	0.0069
$\text{UMSE}_{C,4}$	0.1058	0.0258	0.0081	0.1026	0.0311	0.0084

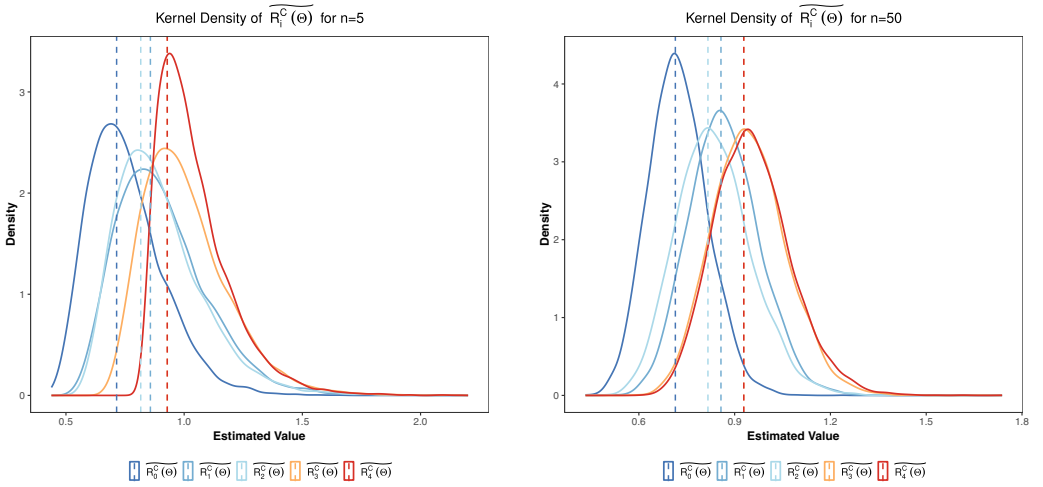


Figure 3. *Kernel density estimates of credibility estimators for different sample sizes.*

Under this hierarchical specification, the marginal distribution of X_i follows a Pareto (Type II) or Lomax distribution, which naturally arises from the exponential–gamma conjugate structure. Furthermore, since $\mu(\Theta) = 1/\Theta$, the transformed random variable $\mu(\Theta)$ follows an inverse Gamma distribution with parameters (α, β) (Klugman *et al.*, 2019). This alternative representation provides valuable insight into the mean–variance relationship inherent in the exponential–gamma model.

4.2 Practical applications

To illustrate the implementation of the proposed credibility framework, we use the dataset `fremotor2freq9907u`[2] available in the `CASdatasets` package in `R` (Rizzo, 2019; Dutang and Charpentier, 2024). Each policy is treated as a separate credibility unit associated with a latent risk parameter Θ_i $(i = 1, \ldots, K)$. Conditional on Θ_i, the observed annual claim counts $\{X_{ij}\}_{j=1}^{n_i}$ are assumed i.i.d. with conditional distribution $f_{\theta_i}(x)$. Additional implementation details, including data processing, structural parameter estimation, and bootstrap approximation of δ_{n_i} and λ, are provided in Appendix A.6.

[2] The dataset contains annual claim counts from French motor insurance policies between 1999 and 2007. Policies with only one year of observations were removed to ensure $(n_i - 1) \neq 0$ for parameter estimation. The resulting sample consists of 71,392 policies, each observed for at least two years, with realistic heterogeneity in the number of observations per policy.

https://doi.org/10.1017/asb.2026.10094 Published online by Cambridge University Press

Table 6. *Structural parameter estimates.*

Parameter	$\widehat{\gamma_0}$	$\widehat{\tau^2}$	$\widehat{\delta}$	$\widehat{\lambda}$	$\widehat{\mu_0}$
Estimate	0.1680	0.0585	0.2345	0.1801	0.1635

Table 7. *Estimated premiums for seven sampled policies under different premium principles.*

Policy ID	PN148469	PN191262	PN105635	PN62080	PN76584	PN38328	PN6536
Net premium principle	0.7122	0.4206	0.2139	0.1867	0.1656	0.2502	0.1352
Expected value ($\rho_1 = 2.0574$)	2.1774	1.2859	0.6538	0.5708	0.5064	0.7648	0.4133
Variance principle ($\rho_2 = 3.5843$)	2.0351	1.4360	1.0377	0.8798	0.7639	1.0642	0.6048
Standard deviation ($\rho_3 = 1.0982$)	1.3793	1.0051	0.7404	0.6696	0.6143	0.7735	0.5327
Modified variance ($\rho_4 = 0.5862$)	1.0160	0.8154	0.8439	0.7939	0.7563	0.7823	0.7033

Based on the sample, we estimate the five structural parameters required for the credibility model, namely μ_0, γ_0, τ^2, δ_{n_i}, and λ. Following Bühlmann and Gisler (2005, p. 94), μ_0 is the portfolio-level mean, γ_0 is the within-policy variance, and τ^2 is the between-policy variance.

The resulting unbiased estimates of the structural parameters are summarized in Table 6. The policy-specific dispersion parameters $\widehat{\delta}_{n_i}$ can be obtained from these estimates as described in Appendix A.6, but are not listed here for brevity.

Given the estimates in Table 6, we next compute credibility estimates of risk premiums using the sample. At this stage, the structural parameters are treated as known, and the resulting credibility estimators correspond to empirical Bayes estimators in the sense of Robbins (1964), with actuarial developments discussed in Mashayekhi (2002).

Before calculating premiums, we determine the safety loading coefficients ρ_i $(i = 1, \ldots, 4)$, which link risk-based estimates to actual premium levels under different principles. Following Yong *et al.* (2025), these coefficients are calibrated via a value-at-risk (VaR) approach at confidence level $p = 0.9$. Let Y denote the aggregate claim amount. The loadings are defined by

$$\begin{aligned}
&\text{(i) Expected value premium principle:} && (1+\rho_1)\,\mathbb{E}[Y] = \mathrm{VaR}_p(Y),\\
&\text{(ii) Variance premium principle:} && \mathbb{E}[Y] + \rho_2\,\mathrm{Var}(Y) = \mathrm{VaR}_p(Y),\\
&\text{(iii) Standard deviation principle:} && \mathbb{E}[Y] + \rho_3\,\sqrt{\mathrm{Var}(Y)} = \mathrm{VaR}_p(Y),\\
&\text{(iv) Modified variance principle:} && \mathbb{E}[Y] + \rho_4\,\tfrac{\mathrm{Var}(Y)}{\mathbb{E}[Y]} = \mathrm{VaR}_p(Y),
\end{aligned}$$

where $\mathbb{E}[Y]$, $\mathrm{Var}(Y)$, and $\mathrm{VaR}_p(Y)$ are estimated using, respectively, the sample mean $\overline{X}_i$, the sample variance $\widetilde{S}^2_{n_i}$, and the empirical p-quantile of policy i computed from the sample $(i = 1, 2, \ldots, K)$. This VaR-based calibration ensures that, under each premium principle, the premiums cover aggregate claims at the specified confidence level.

Using the estimated structural parameters and calibrated safety loadings ρ_i, we then compute the credibility estimator in (4.1) under the net, expected value, variance, standard deviation, and modified variance principles. The sample still contains a large number of policies, so it is not feasible to report results for all contracts. For illustration, we select seven policies from the sample and report their credibility-adjusted premium estimates in Table 7.

https://doi.org/10.1017/asb.2026.10094 Published online by Cambridge University Press

Table 7 shows the resulting premiums under the different principles. As expected, the net premium corresponds to the pure risk expectation, while the other principles incorporate additional loadings through the parameters ρ_i. The variance and modified variance principles produce higher premiums for policies with more volatile claim experience, reflecting their sensitivity to underlying risk heterogeneity and the ability of the credibility-adjusted framework to differentiate premiums accordingly.

5. Concluding remarks

This paper develops a novel framework for estimating process variance, extending the classical Bühlmann credibility theory. Unlike traditional approaches that focus primarily on the hypothetical mean, the proposed method directly addresses the variability of the underlying risk process, capturing higher-order moments and portfolio heterogeneity. This enables the derivation of explicit credibility-adjusted risk premiums under multiple premium principles, including net premium, expected value, variance, standard deviation, and modified variance principles.

By embedding credibility theory within a Bayesian framework, we address some of the limitations associated with conventional Bayesian risk premium estimation. The methodology seamlessly integrates observed claim experience with the underlying portfolio risk structure, producing premiums that reflect both individual policy history and collective characteristics. Calibration of safety loading factors in a data-driven manner ensures that the resulting premiums are consistent with solvency requirements and responsive to varying levels of risk exposure.

The proposed approach is practical, computationally tractable, and operates under mild assumptions without requiring full specification of population or prior distributions. Rigorous theoretical analysis confirms that the estimators possess desirable statistical properties, while empirical results based on real insurance data demonstrate their robustness and practical relevance. The framework effectively differentiates premiums according to policy-specific risk profiles and premium principles, supporting more accurate and informed pricing decisions.

Looking forward, the methodology can be extended to multidimensional risk settings, incorporate hierarchical or nonparametric priors for greater flexibility, and be combined with modern machine learning techniques to improve both estimation accuracy and computational efficiency. Applications beyond traditional insurance pricing, such as portfolio optimization or systemic risk assessment, may also benefit from this framework, illustrating its broad practical relevance.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/asb.2026.10094.

Acknowledgements. The authors would like to express their sincere gratitude to the anonymous reviewers for their insightful comments and constructive suggestions. Their careful reading and valuable feedback have greatly enhanced the rigor, clarity, and overall quality of this manuscript. The authors also acknowledge the financial support from the National Social Science Fund of China under the Later Stage Project (Grant No. 25FTJB002).

Competing interests. No potential competing interest was reported by the authors.

References

Bailey, A.L. (1945) A generalized theory of credibility. *Proceedings of the Casualty Actuarial Society*, **32**, 13–20.

Benhamou, E. (2018) A few properties of sample variance. *arXiv preprint.* https://arxiv.org/abs/1809.03774.

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

Bühlmann, H. (2007). *Mathematical Methods in Risk Theory*, 2nd edn. Berlin: Springer.

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

Chen, Y., Cheung, K.C., Choi, H.M.C. and Yam, S.C.P. (2020) Evolutionary credibility risk premium. *Insurance: Mathematics and Economics*, **93**, 216–229.

Chi, Y. (2012). Optimal reinsurance under variance related premium principles. *Insurance: Mathematics and Economics*, **51**, 310–321.

https://doi.org/10.1017/asb.2026.10094 Published online by Cambridge University Press

Dutang, C. and Charpentier, A. (2024) *CASdatasets: Insurance datasets.* R package version 1.2-0. doi: 10.57745/P0KHAG.
Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A. and Rubin, D.B. (2013) *Bayesian Data Analysis*, 3rd edn. Boca Raton: CRC Press.
Gisler, A. and Wüthrich, M.V. (2008) Credibility for the chain ladder reserving method. *ASTIN Bulletin: The Journal of the IAA*, **38**, 565–600.
Guerra, M. and de Lourdes Centeno, M. (2010) Optimal reinsurance for variance related premium calculation principles. *ASTIN Bulletin: The Journal of the IAA*, **40**, 97–121.
Heilmann, W.-R. (1989) Decision theoretic foundations of credibility theory. *Insurance: Mathematics and Economics*, **8**, 77–95.
Kapadia, A.S., Chan, W. and Moyé, L.A. (2017) *Mathematical Statistics with Applications.* Boca Raton: CRC Press.
Klugman, S.A., Panjer, H.H. and Willmot, G.E. (2019) *Loss Models: From Data to Decisions*, 5th edn. Hoboken, NJ: Wiley.
Lemonte, A.J. and Cordeiro, G.M. (2013) An extended lomax distribution. *Statistics*, **47**, 800–816.
Liang, Z.B. and Yuen, K.C. (2016) Optimal dynamic reinsurance with dependent risks: variance premium principle. *Scandinavian Actuarial Journal*, **3**, 18–36.
Mashayekhi, M. (2002) On asymptotic optimality in empirical Bayes credibility. *Insurance: Mathematics and Economics*, **31**, 285–295.
Pan, M.L., Wang, R.M. and Wu, X.Y. (2008) On the consistency of credibility premiums regarding Esscher principle. *Insurance: Mathematics and Economics*, **42**, 119–126.
Pérez-Hornero, P., Arias-Nicolás, J.P., Pulgarín, A.A. and Pulgarín, A. (2013) An annual JCR impact factor calculation based on Bayesian credibility formulas. *Journal of Informetrics*, **7**, 1–9.
Rizzo, M.L. (2019) *Statistical Computing with R*. Boca Raton, FL: Chapman and Hall/CRC.
Robbins, H. (1964) The empirical Bayes approach to statistical decision problems. *Annals of Mathematical Statistics*, **35**, 1–20.
Schmidt, K.D. (1990). Convergence of Bayes and credibility premiums. *ASTIN Bulletin: The Journal of the IAA*, **20**, 167–172.
Sun, Z.Y., Zheng, X.X. and Zhang, X. (2017) Robust optimal investment and reinsurance of an insurer under variance premium principle and default risk. *Journal of Mathematical Analysis and Applications*, **446**, 1666–1686.
Teugels, J.L. and Sundt, B. (Eds.) (2004) *Encyclopedia of Actuarial Science*. Chichester, UK: Wiley.
Wen, L., Wu, X. and Zhao, X. (2009) The credibility premiums under generalized weighted loss functions. *Journal of Industrial and Management Optimization*, **5**, 893–910.
Wen, L.M., Cao, R.Y. and Zhang, Y. (2024) Novel credibility approaches for Lorenz curve and Gini coefficient estimation. *Statistics*, **58**, 795–826.
Wen, L.M., Wang, W. and Wang, J. L. (2011) The credibility premiums for exponential principle. *Acta Mathematica Sinica, English Series*, **27**, 2217–2228.
Whitney, A.W. (1918) Theory of experience rating. *Proceedings of the Casualty Actuarial Society*, **4**, 274–292.
Yong, Y., Zeng, P. and Zhang, Y. (2025) Credibility theory for variance premium principle. *North American Actuarial Journal*, **29**, 11–43.
Zhang, X., Meng, H. and Zeng, Y. (2016) Optimal investment and reinsurance strategies for insurers with generalized mean–variance premium principle and no-short selling. *Insurance: Mathematics and Economics*, **67**, 125–132.
Zhang, Y. (2023) The comparison and optimality analysis of insurance premium calculation principles. *Journal of Jiangxi Normal University*, **41**, 15–24.
Zhang, Y. and Wen, L. (2024) Experience rating of risk premium for Esscher premium principle. *Communications in Statistics-Theory and Methods*, **53**, 8659–8687.

Appendix A

A.1 The Proof of Theorem 2.1

Proof. For notational convenience, let

$$\Phi = \mathbb{E}\Big[\big(R(\Theta) - b_0 - \sum_{i=1}^{n} b_i X_i\big)^2\Big].$$

Since $X_1, \ldots, X_n$ are conditionally independent and identically distributed given Θ (i.e., conditionally exchangeable), all coefficients b_i can be taken equal. This simplifies the original minimization problem from $n+1$ variables to just two:

$$b_1 = b_2 = \cdots = b_n, \quad \text{and} \quad \min_{b_0,b_1} \mathbb{E}\Big[\big(R(\Theta) - b_0 - b_1 \sum_{i=1}^{n} X_i\big)^2\Big] = \min_{b_0,b_1} \mathbb{E}\Big[\big(R(\Theta) - b_0 - nb_1\overline{X}_n\big)^2\Big],$$

where $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$.

Taking the derivative of Φ with respect to b_0 and setting it to zero yields the optimal intercept:

$$b_0 = \mathbb{E}[R(\Theta)] - nb_1\mathbb{E}[X_i] = R_0 - nb_1\mu_0.$$

https://doi.org/10.1017/asb.2026.10094 Published online by Cambridge University Press

Substituting this expression back into Φ and differentiating with respect to b_1, we obtain

$$\frac{\partial \Phi}{\partial b_1} = n\,\mathrm{Cov}(R(\Theta), X_i) - n^2 b_1\,\mathrm{Var}(X_i) - n(n-1)b_1\,\mathrm{Cov}(X_i, X_j) = 0, \quad i \neq j.$$

Here, we use the standard Bühlmann variance decomposition, which separates the variance into individual and common components:

$$\mathrm{Var}(X_i) = \sigma^2 + \tau^2, \quad \mathrm{Cov}(X_i, X_j) = \tau^2, \quad \mathrm{Cov}(R(\Theta), X_i) = \varphi.$$

Solving for b_1 gives

$$\widehat{b}_1 = \frac{\varphi}{\sigma^2 + n\tau^2}.$$

Finally, we have

$$\widehat{R_C(\Theta)} = \widehat{b}_0 + n\widehat{b}_1\overline{X}_n = R_0 - n\widehat{b}_1\mu_0 + n\widehat{b}_1\overline{X}_n = T_n\overline{X}_n + \left(1 - \frac{T_n\mu_0}{R_0}\right)R_0.$$

This completes the proof. □

A.2 The Proof of Theorem 2.2

Proof. We first prove the sufficiency part. Suppose there exist constants s and t such that $g(\gamma(\Theta)) = s\mu(\Theta) + t$, then we have

$$R(\Theta) = \mu(\Theta) + \rho \cdot g(\gamma(\Theta)) = \mu(\Theta) + \rho(s\mu(\Theta) + t) = (1 + \rho s)\mu(\Theta) + \rho t.$$

Thus, we have $R_0 = (1 + \rho s)\mu_0 + \rho t$, $\varphi = (1 + \rho s)\tau^2$, and

$$T_n = \frac{(1 + \rho s)n\tau^2}{n\tau^2 + \sigma^2} = (1 + \rho s)Z_n,$$

where Z_n is given by (2.3). The credibility estimation of $R_C(\Theta)$ is simplified to

$$\widehat{R_C(\Theta)} = T_n\overline{X}_n + R_0 - T_n\mu_0 = (1 + \rho s)Z_n\overline{X}_n + (1 + \rho s)\mu_0(1 - Z_n) + \rho t.$$

By the Strong Law of Large Numbers, we have $\overline{X}_n \to \mu(\Theta)$ almost surely as $n \to \infty$. Given that $Z_n \to 1$ and $1 - Z_n \to 0$, the strong consistency for $\widehat{R_C(\Theta)}$ is given by

$$\widehat{R_C(\Theta)} \to (1 + \rho s)\mu(\Theta) + \rho t = R(\Theta).$$

Next, we prove the necessity part. Suppose $\widehat{R_C(\Theta)}$ is consistent with respect to $R(\Theta)$. We denote $\mathrm{Cov}\left(\mu(\Theta), g(\gamma(\Theta))\right) = \eta$ and $\mathbb{E}\left[g(\gamma(\Theta))\right] = \lambda$ for convenience. Due to $R_0 = \mathbb{E}\left[R(\Theta)\right] = \mu_0 + \lambda$ and

$$\varphi = \mathrm{Cov}\left(R(\Theta), R(\Theta)\right) = \mathrm{Cov}\left(\mu(\Theta) + g\left(\gamma(\Theta)\right), \mu(\Theta)\right) = \tau^2 + \eta,$$

we have

$$T_n = \frac{n\varphi}{n\tau^2 + \sigma^2} = \frac{\varphi}{\tau^2 + \sigma^2/n} \to 1 + \frac{\eta}{\tau^2}.$$

Therefore, we have

$$\begin{aligned}\widehat{R_C(\Theta)} &\to \left(1 + \frac{\eta}{\tau^2}\right)\mu(\Theta) + \mu_0 + \lambda - \left(1 + \frac{\eta}{\tau^2}\right)\mu_0 \\ &= \left(1 + \frac{\eta}{\tau^2}\right)\mu(\Theta) + \lambda - \frac{\eta}{\tau^2}\mu_0 \\ &= s\mu(\Theta) + t,\end{aligned}$$

where $s = \left(1 + \frac{\eta}{\tau^2}\right)$ and $t = \lambda - \frac{\eta}{\tau^2}\mu_0$ are constants. This completes the proof. □

https://doi.org/10.1017/asb.2026.10094 Published online by Cambridge University Press

A.3 The Proof of Theorem 3.1

Proof. Note that as $n \to \infty$, $w_n^* = \frac{n\lambda}{n\lambda+v} \to 1$. From the strong law of large numbers, we have $S_n^2 \to \mathbb{E}(Y_1|\Theta) = \gamma(\Theta)$, almost surely (a.s.). Thus, it follows that:

$$\widehat{\gamma_C(\Theta)} = w_n^* \cdot S_n^2 + \left(1 - w_n^*\right) \gamma_0 \to \gamma(\Theta), a.s..$$

Additionally, from the central limit theorem, we have

$$\sqrt{n}(S_n^2 - \gamma(\Theta)) \xrightarrow{d} \mathcal{N}(0, \sigma_\gamma^2(\Theta)).$$

On the other hand, we obtain that

$$\sqrt{n}\left[\widehat{\gamma_C(\Theta)} - \gamma(\Theta)\right] - \sqrt{n}(S_n^2 - \gamma(\Theta)) = \sqrt{n}\left(1 - w_n^*\right)\left(\gamma_0 - S_n^2\right) = O_p\left(\frac{1}{\sqrt{n}}\right) \xrightarrow{\mathbb{P}} 0,$$

which implies, according to Slutsky's Theorem (Kapadia *et al.*, 2017), that:

$$\sqrt{n}\left[\widehat{\gamma_C(\Theta)} - \gamma(\Theta)\right] \xrightarrow{d} \mathcal{N}(0, \sigma_\gamma^2(\Theta)).$$

This completes the proof. □

A.4 The Proof of Theorem 3.2

Proof. According to the exchangeability of variables for $Y_i = (X_i - \overline{X})^2$, $i = 1, 2, \ldots, n$, the minimization problem (3.2) is equivalent to A

$$\min_{a_0, b \in \mathbb{R}} \mathbb{E}\left[\left(\gamma(\Theta) - a_0 - b\tilde{S}_n^2\right)^2\right], \tag{A1}$$

where $b = (n-1)\, a_1 = \cdots = (n-1)\, a_n$. From the standard Lagrange methods, the optimal selection of a_0 and b are the solutions of the following normal equations:

$$\begin{cases} \mathbb{E}\left[\gamma(\Theta)\right] = a_0 + b\mathbb{E}\left[\tilde{S}_n^2\right] \\ \mathrm{Cov}\left(\gamma(\Theta), \tilde{S}_n^2\right) = b\mathrm{Var}\left(\tilde{S}_n^2\right) \end{cases}. \tag{A2}$$

In Appendix 1, we have demonstrated the following conclusions:

$$\mathbb{E}\left(\tilde{S}_n^2 \mid \Theta\right) = \gamma(\Theta) \quad \text{and} \quad \mathrm{Var}\left(\tilde{S}_n^2 \mid \Theta\right) = \delta_n(\Theta),$$

thus we have

$$\mathbb{E}\left[\gamma(\Theta)\right] = \mathbb{E}\left[\tilde{S}_n^2\right] = \gamma_0, \quad \mathrm{Var}\left(\tilde{S}_n^2\right) = \delta_n + \lambda, \quad \mathrm{Cov}\left(\gamma(\Theta), \tilde{S}_n^2\right) = \lambda.$$

By solving the Equation (A2), the solutions for a_0 and b are given by

$$\widehat{a_0} = \left(1 - \widehat{b}\right)\gamma_0, \quad \widehat{b} = \frac{\lambda}{\delta_n + \lambda}.$$

Thus, it follows that

$$\widetilde{\gamma_C(\Theta)} = \widehat{a_0} + \widehat{b}\tilde{S}_n^2 = t_n^* \cdot \tilde{S}_n^2 + \left(1 - t_n^*\right)\gamma_0. \tag{A3}$$

□

A.5 The Proof of Theorem 3.3

Proof. When the mean is unknown, the sample variance can be written as

$$\tilde{S}_n^2 = \frac{n}{n-1}\left[\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}_n^2\right],$$

https://doi.org/10.1017/asb.2026.10094 Published online by Cambridge University Press

where $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ denotes the sample mean. By the strong law of large numbers, conditional on Θ, we almost surely have

$$\frac{1}{n}\sum_{i=1}^{n} X_i \longrightarrow \mu(\Theta), \qquad \frac{1}{n}\sum_{i=1}^{n} X_i^2 \longrightarrow \mathbb{E}\left[X_1^2 \mid \Theta\right],$$

as $n \to \infty$. This immediately yields

$$\tilde{S}_n^2 \longrightarrow \gamma(\Theta) \qquad \text{almost surely.}$$

The credibility estimator is given by

$$\widetilde{\gamma_C(\Theta)} = t_n^* \, \tilde{S}_n^2 + (1 - t_n^*)\, \gamma_0,$$

and the fact that $t_n^* \to 1$ implies

$$\widetilde{\gamma_C(\Theta)} \longrightarrow \gamma(\Theta) \qquad \text{almost surely.}$$

To examine the Bayesian estimator, note that

$$\widetilde{\gamma_B(\Theta)} = \mathbb{E}\left[\gamma(\Theta) \mid X_1, \ldots, X_n\right],$$

which, as a conditional expectation, minimizes the mean squared error among all $\mathcal{F}_n$-measurable estimators. Therefore,

$$0 \le \mathbb{E}\left[\left(\widetilde{\gamma_B(\Theta)} - \gamma(\Theta)\right)^2\right] \le \mathbb{E}\left[\left(\widetilde{\gamma_C(\Theta)} - \gamma(\Theta)\right)^2\right] = \frac{\delta_n \lambda}{\delta_n + \lambda} \longrightarrow 0,$$

which shows that $\widetilde{\gamma_B(\Theta)}$ converges to $\gamma(\Theta)$ in mean square.

Moreover, since $\widetilde{\gamma_B(\Theta)}$ forms a martingale with respect to $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$, the martingale convergence theorem (Schmidt, 1990) guarantees almost sure convergence:

$$\widetilde{\gamma_B(\Theta)} \xrightarrow{a.s.} \gamma(\Theta).$$

To derive the asymptotic normality of the credibility estimator, define $W_i = (X_i - \mu(\Theta))^2$ and

$$\overline{W} = \frac{1}{n}\sum_{i=1}^{n} W_i.$$

Conditional on Θ, we have

$$\mathbb{E}[W_i \mid \Theta] = \gamma(\Theta), \qquad \mathrm{Var}(W_i \mid \Theta) = u_4(\Theta) - \gamma^2(\Theta) = \sigma_v^2(\Theta),$$

and the classical central limit theorem gives

$$\sqrt{n}\left(\overline{W} - \gamma(\Theta)\right) \xrightarrow{d} \mathcal{N}\left(0, \sigma_v^2(\Theta)\right).$$

In addition,

$$\tilde{S}_n^2 = \overline{W} - \left(\overline{X}_n - \mu(\Theta)\right)^2 + O\left(\frac{1}{n}\right),$$

and using $\overline{X}_n - \mu(\Theta) = O(n^{-1/2})$, we obtain

$$\sqrt{n}\left[\tilde{S}_n^2 - \gamma(\Theta)\right] - \sqrt{n}\left[\overline{W} - \gamma(\Theta)\right] \longrightarrow 0 \qquad \text{in probability.}$$

Thus, by Slutsky's theorem,

$$\sqrt{n}\left(\tilde{S}_n^2 - \gamma(\Theta)\right) \xrightarrow{d} \mathcal{N}\left(0, \sigma_v^2(\Theta)\right).$$

Finally,

$$\sqrt{n}\left(\widetilde{\gamma_C(\Theta)} - \gamma(\Theta)\right) = \sqrt{n}\left(\tilde{S}_n^2 - \gamma(\Theta)\right) + \sqrt{n}(1 - t_n^*)\left(\gamma_0 - \tilde{S}_n^2\right),$$

https://doi.org/10.1017/asb.2026.10094 Published online by Cambridge University Press

and the second term is $o_p(1)$ because $(1 - t_n^*) = O(n^{-1})$. Therefore,

$$\sqrt{n}\left(\widetilde{\gamma_C(\Theta)} - \gamma(\Theta)\right) \xrightarrow{d} \mathcal{N}\left(0, \sigma_v^2(\Theta)\right).$$

Combining the above results, we conclude that when the mean is unknown, both the credibility and Bayesian estimators are strongly consistent, and in addition, the credibility estimator is asymptotically normal. □

A.6 Empirical implementation and structural parameter estimation

This appendix provides a comprehensive description of the empirical implementation underlying the practical application in Section 4.2. The presentation follows the chronological order of the empirical procedure, beginning with data preparation, followed by structural parameter estimation, bootstrap approximation of dispersion quantities, and concluding with premium computation.

The empirical analysis is based on the dataset `fremotor2freq9907u` from the `CASdatasets` package in `R`. For each policy, we retain the policy identifier `IDpol` and the annual claim-count variables

$$\texttt{NbClaim1999}, \texttt{NbClaim2000}, \ldots, \texttt{NbClaim2006}.$$

The available annual claim counts for each policy i are arranged as a longitudinal claim history $\{X_{ij}\}_{j=1}^{n_i}$, where n_i denotes the number of non-missing annual observations. Policies with fewer than two observed years are excluded, ensuring that $(n_i - 1) \neq 0$ for the estimation of within-policy variance. Each policy is treated as a single credibility unit associated with an unobserved risk parameter Θ_i. The analysis is conducted directly on annual claim counts without constructing exposure-adjusted frequencies.

Based on the sample, the structural parameters are computed as follows. Let K denote the number of policies in the sample and $N = \sum_{i=1}^{K} n_i$ the total number of observations. The structural mean and within-policy variance are estimated by the unbiased formulas

$$\widehat{\mu}_0 = \frac{1}{N} \sum_{i=1}^{K} n_i \overline{X}_{n_i}, \qquad \widehat{\gamma}_0 = \frac{1}{\sum_{i=1}^{K} (n_i - 1)} \sum_{i=1}^{K} \sum_{j=1}^{n_i} (X_{ij} - \overline{X}_{n_i})^2,$$

where $\overline{X}_{n_i} = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ij}$ is the mean claim count for policy i. The between-policy variance is estimated as

$$\widehat{\tau}^2 = \frac{N}{N^2 - \sum_{i=1}^{K} n_i^2} \left[\sum_{i=1}^{K} n_i (\overline{X}_{n_i} - \overline{\overline{X}})^2 - (K-1)\widehat{\gamma}_0 \right],$$

with $\overline{\overline{X}} = \frac{1}{N} \sum_{i=1}^{K} n_i \overline{X}_{n_i}$ being the overall mean.

The variance-based credibility estimator additionally requires the dispersion quantities

$$\delta_{n_i} = \mathbb{E}[\mathrm{Var}(\tilde{S}_{n_i}^2 \mid \Theta_i)], \qquad \lambda = \mathrm{Var}[\gamma(\Theta_i)],$$

where $\tilde{S}_{n_i}^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (X_{ij} - \overline{X}_{n_i})^2$ and $\gamma(\Theta_i) = \mathrm{Var}(X_{ij} \mid \Theta_i)$. Since closed-form estimators for δ and λ are generally unavailable, these quantities are approximated via a nonparametric bootstrap applied independently within each policy. Specifically, for $m = 10{,}000$ bootstrap replications, bootstrap samples $\{X_{ij}^{(t)} : i = 1, \ldots, K;\ j = 1, \ldots, n_i\},\ t = 1, \ldots, m$ are generated. For each replication, the within-policy variance is computed as

$$S_{i(t)}^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (X_{ij}^{(t)} - \overline{X}_{n_i}^{(t)})^2, \qquad \overline{X}_{n_i}^{(t)} = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ij}^{(t)}.$$

The overall estimates of δ and λ are then obtained by

$$\widehat{\delta} = \frac{1}{K(m-1)} \sum_{i=1}^{K} \sum_{t=1}^{m} n_i (S_{i(t)}^2 - \overline{S_i^2})^2, \qquad \widehat{\lambda} = \frac{1}{m(K-1)} \sum_{i=1}^{K} \sum_{t=1}^{m} (S_{i(t)}^2 - \overline{S_t^2})^2,$$

https://doi.org/10.1017/asb.2026.10094 Published online by Cambridge University Press

where $\overline{S_i^2} = \frac{1}{m}\sum_{t=1}^{m} S_{i(t)}^2$ and $\overline{S_t^2} = \frac{1}{K}\sum_{i=1}^{K} S_{i(t)}^2$. The policy-specific dispersion parameters are then defined as $\widehat{\delta}_{n_i} = \widehat{\delta}/n_i$ for $i = 1, \ldots, K$.

Finally, premium estimates are computed using the estimated structural parameters and corresponding credibility factors. For illustration, Table 7 reports premiums for seven policies selected from the sample. Each reported policy corresponds to a single credibility unit, and the premiums incorporate the calibrated loading coefficients ρ_i as described in Section 4.2.

https://doi.org/10.1017/asb.2026.10094 Published online by Cambridge University Press