Astin Bulletin (2026), **56**, pp. 1066–1090
doi:10.1017/asb.2026.10109

International Actuarial Association
Association Actuarielle Internationale

RESEARCH ARTICLE

A Normal-Gamma copula for dependence modeling in collective risk models

Jae Youn Ahn[1,†], Himchan Jeong[2,3,†] and Rosy Oh[4]

[1]Department of Statistics, Ewha Womans University, Republic of Korea
[2]Department of Statistics and Actuarial Science, Simon Fraser University, Canada
[3]Department of Next Generation Smart Energy System Convergence, Gachon University, Republic of Korea
[4]Department of Mathematics, Korea Military Academy, Republic of Korea
Corresponding author: Rosy Oh; Email: rosy.oh5@gmail.com

Received: 5 June 2025 UTC; **Revised:** 6 May 2026 UTC; **Accepted:** 20 June 2026 UTC;
First published online: 28 July 2026 UTC

Keywords: Collective risk model; increasing/decreasing conditional variance property; copula; dependence modeling; average severity

Abstract

This paper introduces a new copula family and a corresponding statistical methodology for modeling the dependence between claim frequency and average severity in the collective risk model (CRM). Existing copula-based approaches often struggle to capture the intrinsic dependence structure between these variables. To address this limitation, we propose the Normal-Gamma (NG) copula, which is derived from the NG distribution and naturally exhibits a decreasing conditional variance property. This property aligns well with the empirical characteristics of claim frequency and average severity, making the NG copula particularly suitable for CRM applications. The effectiveness of the proposed copula is demonstrated through numerical studies and real data analysis.

1. Introduction

The collective risk model (CRM) is a statistical tool for modeling the total amount of claims arising from each policyholder. The model accounts for two sources of variability: the number of claims (frequency) and the size of each claim (individual severity). Traditionally, frequency and individual severities were assumed to be independent. However, supported by empirical evidence (Gschlöß l and Czado, 2007) and numerous follow-up research, recent CRM research has relaxed this independence assumption using various statistical tools.

While various methods exist, we are primarily interested in modeling the joint distribution of frequency and average severity, where average severity is defined as the arithmetic mean of the individual severities. In the insurance literature, these statistical methods are mainly classified into one of the following categories: (*i*) In two-part models, the dependence between frequency and average severity is derived by using frequency as an explanatory variable for average severity (Gschlöß l and Czado, 2007; Shi *et al*., 2015; Garrido *et al*., 2016). (*ii*) In random effect models, the dependence between frequency and average severity is derived through common/bivariate mixtures (Belzunce *et al*., 2006; Hernández-Bastida *et al*., 2009; Oh *et al*., 2020; Jeong and Valdez, 2020; Cheung *et al*., 2021; Vernic *et al*., 2022; Chen *et al*., 2024). (*iii*) In copula models, the dependence between frequency and average severity is modeled using a copula (Krämer *et al*., 2013; Czado *et al*., 2012; Frees *et al*., 2016; Lee and Shi, 2019; Oh *et al*., 2021b).

[†] First authors.

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

https://doi.org/10.1017/asb.2026.10109 Published online by Cambridge University Press

This article focuses on copula methods for modeling dependence in CRM. However, as pointed out in Oh *et al*. (2021a), the dependence between frequency and average severity is not trivial because it is induced by the intrinsic construction of the average severity itself. Even under the assumption that the claim frequency and the individual severities are mutually independent, frequency and average severity are not independent, so the independence copula is not appropriate for describing their dependence. More specifically, a numerical simulation study by Oh *et al*. (2021a) shows that the corresponding copula exhibits a decreasing conditional variance property under the same independence assumption. This suggests that standard copula choices, such as the Gaussian copula and commonly used Archimedean copula families, may not be suitable for this setting.

This property is not satisfied by any existing copula family, with the exception of the recent study in Oh and Ahn (2024), making the choice of an appropriate copula family extremely difficult. This limitation makes the selection of an appropriate copula family particularly challenging. To overcome this issue, this article introduces a novel copula family, referred to as the *Normal-Gamma copula* (NG copula). This copula is derived from the well-known NG distribution and naturally exhibits the decreasing conditional variance property. Additionally, we establish an asymptotic result demonstrating that the NG copula accommodates a broad range of dependence structures. Finally, we illustrate the practical application of the NG copula in modeling the joint distribution of frequency and average severity.

Given the mixed nature of frequency and average severity, encompassing both discrete and continuous data, the computation of the likelihood function necessitates numerical integration. To facilitate faster calculation, we provide an analytical expression for the likelihood function and outline a suitable statistical estimation procedure.

The remainder of this paper is structured as follows. Section 2 presents the definitions and discusses the necessity of a new copula approach. Section 3 introduces the new copula family, followed by Section 3.3, which establishes its asymptotic properties. In Section 4, we demonstrate how this new copula family can be applied to model dependence in CRMs. Section 5 reports the results of a simulation study that evaluates the performance of the proposed approach. Section 6 applies the model to real-world insurance data.

2. Motivations

2.1. Notation and definitions

Let $\mathbb{R}$, $\mathbb{R}_{>0}$, and $\mathbb{R}_{\geq 0}$ denote the sets of real numbers, positive real numbers, and nonnegative real numbers, respectively. We also define $\mathbb{N}_{>0}$ and $\mathbb{N}_{\geq 0}$ to be the sets of positive integers and nonnegative integers, respectively. Let $f_1 : \mathbb{R} \to \mathbb{R}$ and $f_2 : \mathbb{R} \to \mathbb{R}_{>0}$ be functions defined in some neighborhood of $a \in \mathbb{R}$. We say that

$$f_1(\theta) = o(f_2(\theta)) \quad \text{as } \theta \to a$$

if, for every $\varepsilon > 0$, there exists a $\delta > 0$ such that for all $\theta \in \mathbb{R}$ satisfying $|\theta - a| < \delta$, it holds that $|f_1(\theta)| < \varepsilon\, |f_2(\theta)|$. For an arbitrary distribution function $F : \mathbb{R} \to [0, 1]$, we define the quantile function of F as

$$F^{-1}(u) = \inf\{x \in \mathbb{R} : u \leq F(x)\}, \quad u \in [0, 1].$$

We denote by $\Phi : \mathbb{R} \to [0, 1]$ the cumulative distribution function and by $\phi : \mathbb{R} \to \mathbb{R}_{>0}$ the probability density function of the standard normal distribution.

Copulas are versatile tools for describing dependence among random variables. For this study, we focus on bivariate copulas defined on the two-dimensional space. A bivariate copula C:$[0, 1]^2 \to [0, 1]$ is a joint distribution function whose marginal distributions are uniform on (0,1). We denote the set of all bivariate copulas by $\mathcal{F}_2$.

https://doi.org/10.1017/asb.2026.10109 Published online by Cambridge University Press

2.2. Modeling dependence in positive collective risk using copulas

We consider the positive integer-valued random variable N^+ and the random variable M^+ defined as

$$M^+ = \frac{1}{N^+} \sum_{j=1}^{N^+} Z_j, \tag{1}$$

where Z_j's are positive random variables. Here, N^+ denotes the positive frequency while M^+ and Z_j's represent to the positive average severity and individual severities, respectively.

This subsection introduces a positive CRM that employs a copula-based approach to capture the dependence between the positive frequency and the average severity. Importantly, this model is designed for scenarios where the frequency is strictly positive. However, the methodology can be readily extended to address the full CRM, including cases with zero frequency. The details of this extension are provided in Section 4.

Model 1. (Positive CRM). Consider a positive integer-valued random variable N^+ with cumulative distribution function $F_1^+ : \mathbb{N}_{>0} \to [0, 1]$ and probability mass function $f_1^+ : \mathbb{N}_{>0} \to \mathbb{R}_{\geq 0}$. The corresponding positive average severity M^+, defined in (1), has cumulative distribution function $F_2^+ : \mathbb{R}_{>0} \to [0, 1]$ and probability density function $f_2^+ : \mathbb{R}_{>0} \to \mathbb{R}_{\geq 0}$.

We model the joint distribution of (N^+, M^+) using a copula $C \in \mathcal{F}_2$,

$$(N^+, M^+) \sim H^+ = C\big(F_1^+, F_2^+\big).$$

For mathematical simplicity, we assume that the copula $C \in \mathcal{F}_2$ is absolutely continuous on $[0, 1]^2$, and

$$\begin{cases} f_2^+(m) > 0, & m > 0; \\ f_2^+(m) = 0, & \text{otherwise.} \end{cases}$$

In Model 1, the corresponding joint density function $h^+ : \mathbb{N}_{>0} \times \mathbb{R}_{>0} \to \mathbb{R}_{\geq 0}$, which is defined as the Radon–Nikodym derivative of the distribution of (N^+, M^+) with respect to the product of the counting and Lebesgue measures, is given by

$$\begin{aligned} h^+(n, m) &:= \frac{\partial}{\partial m}(H^+(n, m) - H^+(n-1, m)) \\ &= f_2^+(m)(\mathbb{P}(N^+ \leq n \mid M^+ = m) - \mathbb{P}(N^+ \leq n-1 \mid M^+ = m)) \\ &= f_2^+(m)\mathbb{P}(N^+ = n \mid M^+ = m) \end{aligned} \tag{2}$$

for $(n, m) \in \mathbb{N}_{>0} \times \mathbb{R}_{>0}$.

A key step is selecting an appropriate copula family, $\{C_\theta \in \mathcal{F}_2 | \, \theta \in \Theta\}$, for some well-defined parameter space Θ. For example, the Gaussian and t-copulas have been used in prior studies such as Czado *et al*. (2012) and Lee and Shi (2019), while Archimedean families such as the Clayton, Frank, Gumbel, and Joe copulas have also been considered in Shi and Zhao (2020). However, as noted in Oh *et al*. (2021a) and Oh *et al*. (2021b), identifying a copula family that effectively captures the dependence between the positive frequency N^+ and the average severity M^+ remains a challenge. The difficulty is intrinsic to the frequency-severity setting. Indeed, even in the simplest benchmark where the frequency and the individual severities are mutually independent, the induced dependence between N^+ and M^+ is already nontrivial: the independence copula is not appropriate for (N^+, M^+) because the frequency itself affects the dispersion of the sample average severity. In particular, as the number of claims increases, the average severity becomes more concentrated around its center, so the frequency contains information about the variability of the average severity. This also suggests that a copula suitable for (N^+, M^+) need not be suitable for (N^+, S^+), where $S^+ := \sum_{j=1}^{N^+} Z_j$ denotes the positive aggregate severity. The following subsection examines this inherent dependence structure in detail.

https://doi.org/10.1017/asb.2026.10109 Published online by Cambridge University Press

2.3. Classical collective risk model

We start with the simplest case where the positive frequency and the individual severities are assumed to be independent. The following classical CRM and subsequent properties demonstrate the challenges of using a copula method to model the dependence between positive frequency and positive average severity.

Model 2. For the joint modeling of the positive frequency and the positive average severity, (N^+, M^+), we consider the following assumptions:

(i) The positive frequency N^+ is a nonnegative integer-valued random variable with finite mean and variance.
(ii) Conditioned on $N^+ = n \in \mathbb{N}_{>0}$, the random variables $Z_1, Z_2, \ldots, Z_n$ are independent and identically distributed (i.i.d.) following the Gamma(a, b) distribution with shape parameter $a > 0$ and rate parameter $b > 0$.

Given this setup, the conditional distribution of M^+ given N^+ is

$$M^+ \mid N^+ \sim \text{Gamma}(aN^+, bN^+)$$

as implied by the properties of the gamma distribution. Furthermore, the conditional mean and variance of M^+ given N^+ are

$$\mathbb{E}[M^+ \mid N^+] = \frac{a}{b} \quad \text{and} \quad \mathrm{V}ar(M^+ \mid N^+) = \frac{a}{b^2 N^+}. \tag{3}$$

This means that, if one attempts to model (N^+, M^+) using a copula, standard choices such as the independence and Gaussian copulas may not be suitable.

Although the above derivation uses the Gamma distribution, the monotone conditional variance property is not specific to this choice. More generally, suppose that, conditional on $N^+ = n$, the severities $Z_1, \ldots, Z_n$ are i.i.d. from a reproductive exponential dispersion model (Jorgensen, 1997) with mean μ, dispersion parameter τ, and unit variance function V so that $\mathbb{E}\left[Z_j\right] = \mu$ and $\mathrm{V}ar\left(Z_j\right) = \tau V(\mu)$. Then,

$$M^+ \mid N^+ = n$$

belongs to the same exponential dispersion family with mean μ and dispersion parameter τ/n, and therefore,

$$\mathbb{E}[M^+ \mid N^+ = n] = \mu \quad \text{and} \quad \mathrm{V}ar(M^+ \mid N^+ = n) = \frac{\tau V(\mu)}{n}.$$

In summary, although the CRM in Model 2 assumes independence between the positive frequency and individual severities, it induces a nontrivial dependence between the positive frequency N^+ and the positive average severity M^+. This dependence is evident in (3), where the conditional variance of M^+ decreases as N^+ increases – a characteristic known as the *monotone conditional variance property*. For the general case, where dependence exists between the positive frequency and individual severities, a similar argument is provided in Appendix A.

Monotone conditional variance property is also relevant in fields such as economics and climatology (Dette and Pilz, 2009; Sydeman *et al*., 2013). Oh and Ahn (2024) show that the monotone conditional variance property is closely related to the *stochastically monotone wing property*, which is the property of the copula; see Theorem 3.6 below for the definition of the stochastically monotone wing property.

The visualization of the copula density function provided by Oh *et al*. (2021a) offers valuable insight into the dependence structure. Specifically, Figure 1, adapted from Oh *et al*. (2021a), illustrates the jittered version of the copula density corresponding to (N^+, M^+) in Model 2 with $\mathbb{P}(N^+ = n) = 0.2$, $n = 1, 2, 3, 4, 5$.

Although various copula families, including the Gaussian copula and Archimedean copulas, have been used to model dependence between frequency and average severity in the insurance literature, none of the existing choices can fully explain monotone conditional variance property nor stochastically monotone wing property (Erdemir and Meral, 2020; Gao and Li, 2023). To address this limitation, we

https://doi.org/10.1017/asb.2026.10109 Published online by Cambridge University Press

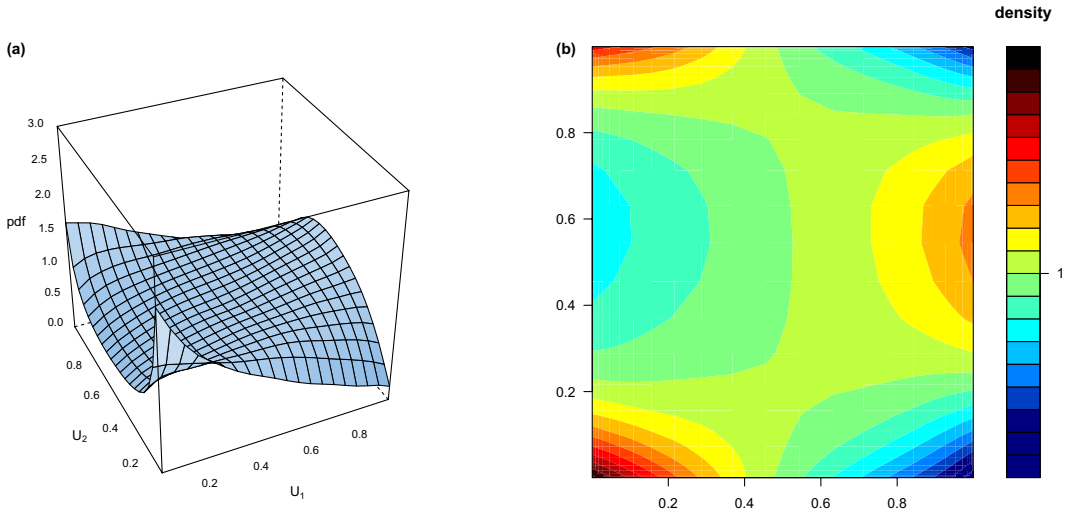


Figure 1. *The jittered version of the copula density corresponding to* (N^+, M^+) *(Oh et al., 2021a).*

introduce a novel copula family derived from the well-known NG distribution (Jewell, 1974; DeGroot, 2005; Bernardo and Smith, 2009). This new copula family possesses both the monotone conditional variance property and the stochastically monotone wing property, making it a promising candidate for modeling the dependence structure between N^+ and M^+.

3. Introduction to a new family of copulas

This section introduces a new copula family derived from the NG distribution. We begin by defining the NG distribution and reviewing its essential properties.

3.1. Normal-Gamma copula

Model 3. [NG Distribution]. A bivariate random variable (X, Y) follows a NG distribution if

$$Y \mid X \sim \mathrm{N}\left(0, \frac{1}{X}\right) \quad \text{and} \quad X \sim \text{Gamma}\,(\alpha, \beta)\,, \tag{4}$$

where $N(\mu, \sigma^2)$ implies a normal distribution with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 > 0$, and Gamma (α, β) denotes a gamma distribution with shape parameter $\alpha > 0$ and rate parameter $\beta > 0$.

The NG distribution defined in Model 3 exhibits a well-known conjugate property, as demonstrated in part *i* of the following corollary. Furthermore, part *ii* shows that the marginal distributions can be expressed analytically.

Corollary 3.1. *For a bivariate random variable (X, Y) as described in Model 3, we have the following properties:*

(i) The conditional distribution of X given $Y = y$ *is*

$$X \mid Y = y \sim \mathit{Gamma}\left(\alpha + \frac{1}{2}, \beta + \frac{1}{2}y^2\right).$$

(ii) The marginal distribution of Y is

$$Y \sim t_{2\alpha}\left(0, \frac{\beta}{\alpha}\right),$$

https://doi.org/10.1017/asb.2026.10109 Published online by Cambridge University Press

where $t_\nu\left(\mu, \tau^2\right)$ *implies the location-scale Student's t distribution with degree of freedom parameter* $\nu > 0$*, location parameter* $\mu \in \mathbb{R}$*, and scale parameter* $\tau^2 > 0$*.*

The following result provides an expression for the copula family corresponding to the NG distributions presented in Model 3.

Proposition 3.2. *For a bivariate random variable (X, Y) as described in Model 3, we obtain the following results:*

(i) The joint distribution function of (X, Y) at $(X, Y) = (x, y)$*, denoted by* $H_{\alpha,\beta} : \mathbb{R}^2 \to [0, 1]$*, is given by*

$$H_{\alpha,\beta}(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} h_{\alpha,\beta}(s, t; \alpha, \beta)\, dt\, ds, \tag{5}$$

where $h_{\alpha,\beta} : \mathbb{R}^2 \to \mathbb{R}_{\geq 0}$ *is the corresponding joint density function defined by*

$$h_{\alpha,\beta}(x, y; \alpha, \beta) = \begin{cases} \dfrac{\beta^\alpha}{\Gamma(\alpha)\sqrt{2\pi}}\, x^{\alpha+\frac{1}{2}-1} \exp\left(-x\left(\beta + \frac{y^2}{2}\right)\right), & x \in \mathbb{R}_{>0},\ y \in \mathbb{R}; \\ 0, & \text{otherwise}. \end{cases} \tag{6}$$

(ii) Let $C_{\alpha,\beta} \in \mathcal{F}_2$ *be a bivariate copula corresponding to* $H_{\alpha,\beta} : \mathbb{R}^2 \to [0, 1]$ *for some* $\alpha, \beta > 0$*. Then, we have*

$$C_{\alpha,\beta}(u, v) = \int_0^u \int_0^v c_{\alpha,\beta}(s, t)\, dt\, ds, \quad u, v \in [0, 1], \tag{7}$$

where $c_{\alpha,\beta} : [0, 1]^2 \to \mathbb{R}_{\geq 0}$ *is the corresponding copula density function defined by*

$$\begin{aligned} & c_{\alpha,\beta}(u, v) \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)\sqrt{2\pi}} \frac{\left(G^{-1}(u; \alpha, \beta)\right)^{\alpha+\frac{1}{2}-1} \exp\left(-G^{-1}(u; \alpha, \beta)\left(\beta + \frac{1}{2}\left(\Phi_{2\alpha}^{-1}\left(v; 0, \frac{\beta}{\alpha}\right)\right)^2\right)\right)}{g\left(G^{-1}(u; \alpha, \beta); \alpha, \beta\right) \phi_{2\alpha}\left(\Phi_{2\alpha}^{-1}\left(v; 0, \frac{\beta}{\alpha}\right); 0, \frac{\beta}{\alpha}\right)} \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)\sqrt{2\pi}} \frac{\left(G^{-1}(u; \alpha, \beta)\right)^{\alpha+\frac{1}{2}-1} \exp\left[-G^{-1}(u; \alpha, \beta)\left\{\beta + \frac{1}{2}\left(\Phi_{2\alpha}^{-1}\left(v; 0, \frac{\beta}{\alpha}\right)\right)^2\right\}\right]}{\frac{\beta^\alpha}{\Gamma(\alpha)}\left(G^{-1}(u; \alpha, \beta)\right)^{\alpha-1} \exp\left[-\beta G^{-1}(u; \alpha, \beta)\right] \phi_{2\alpha}\left(\Phi_{2\alpha}^{-1}\left(v; 0, \frac{\beta}{\alpha}\right); 0, \frac{\beta}{\alpha}\right)} \\ &= \frac{1}{\sqrt{2\pi}} \frac{\left(G^{-1}(u; \alpha, \beta)\right)^{\frac{1}{2}} \exp\left(-\frac{1}{2}G^{-1}(u; \alpha, \beta)\left(\Phi_{2\alpha}^{-1}\left(v; 0, \frac{\beta}{\alpha}\right)\right)^2\right)}{\phi_{2\alpha}\left(\Phi_{2\alpha}^{-1}\left(v; 0, \frac{\beta}{\alpha}\right); 0, \frac{\beta}{\alpha}\right)}, \end{aligned} \tag{8}$$

for $u, v \in (0, 1)$*. Here,* $G(\cdot\,; \alpha, \beta)$ *and* $g(\cdot\,; \alpha, \beta)$ *denote the cumulative distribution function and probability density function of the gamma distribution with shape parameter* $\alpha > 0$ *and rate parameter* $\beta > 0$*, respectively. Similarly,* $\Phi_\nu(\cdot\,; \mu, \tau^2)$ *and* $\phi_\nu(\cdot\,; \mu, \tau^2)$ *denote the cumulative distribution function and probability density function of the location-scale Student's t-distribution with degrees of freedom* $\nu > 0$*, location parameter* $\mu \in \mathbb{R}$*, and scale parameter* $\tau^2 > 0$*, respectively.*

Proof. The proof of part *i* follows directly from Model 3. For part *ii*, Sklar's theorem gives

$$C_{\alpha,\beta}(u, v) = H_{\alpha,\beta}\left(G^{-1}(u; \alpha, \beta), \Phi_{2\alpha}^{-1}\left(v; 0, \frac{\beta}{\alpha}\right)\right).$$

Differentiating this identity with respect to u and v, using Corollary 3.1, gives (8). □

https://doi.org/10.1017/asb.2026.10109 Published online by Cambridge University Press

While it is not immediately obvious from the expression in (8), the following result demonstrates that the parameter $\beta > 0$ is redundant in defining the copula in (7) and hence its corresponding density function in (8).

Lemma 3.3. *The copula in (8) is invariant with respect to the parameter $\beta > 0$.*

Proof. Consider the random variable (X, Y) in (4). The copula corresponding to (X, Y) is identical to the copula corresponding to $(\beta X, Y)$ because the copula is invariant under the strictly monotone transformation of the random variables, as stated in Theorem 2.4.3 of Nelsen (2007). Since $\beta X \sim \text{Gamma}(\alpha, 1)$, the proof is concluded. □

Lemma 3.3 implies that we can set $\beta = 1$ without loss of generality when defining the copula family corresponding to the NG distribution. Consequently, the copula in (8) depends only on the parameter $\alpha > 0$. Furthermore, we found it convenient to use the alternative parameterization $\theta = 1/\alpha$. With $\alpha = 1/\theta$ and $\beta = 1$, the joint density function in (6) and the copula density function in (8) can then be expressed as

$$h_\theta(x, y) := \begin{cases} \dfrac{1}{\Gamma(\frac{1}{\theta})\sqrt{2\pi}}\, x^{\frac{1}{\theta}-\frac{1}{2}} \exp\left(-x\left(1+\frac{y^2}{2}\right)\right), & x, y \in \mathbb{R}_{>0} \times \mathbb{R}; \\ 0, & \text{otherwise}, \end{cases}$$

and

$$c_\theta(u, v) := \frac{1}{\sqrt{2\pi}} \frac{\left(G^{-1}\left(u; \frac{1}{\theta}, 1\right)\right)^{1/2} \exp\left(-\frac{1}{2} G^{-1}\left(u; \frac{1}{\theta}, 1\right)\left(\Phi^{-1}_{2/\theta}\left(v; 0, \theta\right)\right)^2\right)}{\phi_{2/\theta}\left(\Phi^{-1}_{2/\theta}\left(v; 0, \theta\right); 0, \theta\right)}, \tag{9}$$

respectively. This copula family is formally defined in Definition 3.4.

Definition 3.4. *Under the setting in Model 3 with $\beta = 1$ and $\theta = 1/\alpha$, we call the corresponding copula family of the joint distribution $H_{\alpha,\beta}$ in (5)*

$$\mathcal{S} = \{C_\theta \in \mathcal{F}_2 \mid \theta > 0\}$$

as the NG copula family.

3.2. Properties of Normal-Gamma copulas

We first record a symmetry property of the NG copula. Although this property is not the primary modeling motivation in the frequency-severity setting, it plays an important technical role in the sequel and is used repeatedly in the proofs of the main dependence properties. In particular, $C_\theta(u, v)$ is symmetric with respect to the horizontal line

$$\left\{(u, v) \in (0, 1)^2 \;\middle|\; v = \frac{1}{2}\right\}. \tag{10}$$

Lemma 3.5. *A NG copula $C_\theta \in \mathcal{S}$ for some $\theta > 0$ is symmetric along the horizontal line in (10). Specifically, we have*

$$C_\theta(u, v) = u - C_\theta(u, 1 - v), \quad u, v \in (0, 1).$$

Proof. It is enough to show

$$c_\theta\left(u, \frac{1}{2} - v\right) = c_\theta\left(u, \frac{1}{2} + v\right), \quad u \in (0, 1) \quad \text{and} \quad v \in \left(0, \frac{1}{2}\right). \tag{11}$$

First, due to the symmetry of the Student's t distribution, we have

$$\Phi^{-1}_{2/\theta}\left(\frac{1}{2} - v; 0, \theta\right) = -\Phi^{-1}_{2/\theta}\left(\frac{1}{2} + v; 0, \theta\right)$$

https://doi.org/10.1017/asb.2026.10109 Published online by Cambridge University Press

and

$$\phi_{2/\theta}\left(\frac{1}{2}-v;0,\theta\right)=\phi_{2/\theta}\left(\frac{1}{2}+v;0,\theta\right)$$

for $v\in\left(0,\frac{1}{2}\right)$. Applying these results to (9), we have (11). □

The NG copula family also satisfies the stochastically monotone wing property, which was proposed in Oh and Ahn (2024). We note that the stochastically monotone wing property can be viewed as an extension of the stochastically increasing property proposed by Tukey (1958) and Lehmann (1966). It is shown that this property remains invariant under strictly increasing transformations of the marginal distributions, and hence characterizes the copula itself (Oh and Ahn, 2024).

Theorem 3.6. *For a NG copula $C_\theta\in\mathcal{S}$, consider a random vector $(U,V)\sim C_\theta$. Then, we have*

$$\mathbb{P}\left(V>\frac{1}{2}+v\,\middle|\,U=u\right) \tag{12}$$

and

$$\mathbb{P}\left(V<\frac{1}{2}-v\,\middle|\,U=u\right)$$

are decreasing functions of $u\in(0,1)$ for all $v\in(0,1/2)$.

Proof. By the symmetry in Lemma 3.5, it is enough to show (12). First, for $u\in(0,1)$ and $v\in(0,1/2)$, we have

$$\mathbb{P}\left(V>\frac{1}{2}+v\,\middle|\,U=u\right)=1-\Phi\left(\Phi^{-1}_{2/\theta}\left(\frac{1}{2}+v;0,\theta\right)\sqrt{G^{-1}\left(u;\frac{1}{\theta},1\right)}\right) \tag{13}$$

by (4) and Corollary 3.1, where $\Phi(\cdot)$ is a cumulative distribution function of standard normal distribution. Since

$$\Phi^{-1}_{2/\theta}\left(\frac{1}{2}+v;0,\theta\right)>0$$

for any given $v\in(0,1/2)$, it is clear that (13) is a decreasing function of $u\in(0,1)$ for any given $v\in(0,1/2)$. □

The following result, which is a corollary of Theorem 3.6, shows that the NG copula has the monotone conditional variance property, which was proposed in Oh and Ahn (2024).

Theorem 3.7. *Consider a NG copula $C_\theta\in\mathcal{S}$ for some $\theta>0$, and the corresponding bivariate random variable $(U,V)\sim C_\theta$. Then, we have the following properties:*

i. $\mathbb{E}\left[V\mid U=u\right]=\frac{1}{2}$ *for* $u\in(0,1)$.
ii. $\mathrm{V}ar\left(V\mid U=u\right)$ *is a decreasing function of u on the open interval* $(0,1)$.

Proof. The proof of part *i* follows immediately from the symmetry of the copula in Lemma 3.5. Now, we prove part *ii*. For $u\in(0,1)$, we have

$$\begin{aligned}\mathrm{V}ar\left(V\mid U=u\right)&=\mathbb{E}\left[\left(V-\frac{1}{2}\right)^2\,\middle|\,U=u\right]\\&=4\int_0^{\frac{1}{2}}v\mathbb{P}\left(V>\frac{1}{2}+v\,\middle|\,U=u\right)\mathrm{d}v,\end{aligned}$$

where the first equality is from part *i*, and the second equality is from Lemma 3.5 and integration by parts. Then, combining this result with Theorem 3.6, we conclude the proof. □

https://doi.org/10.1017/asb.2026.10109 Published online by Cambridge University Press

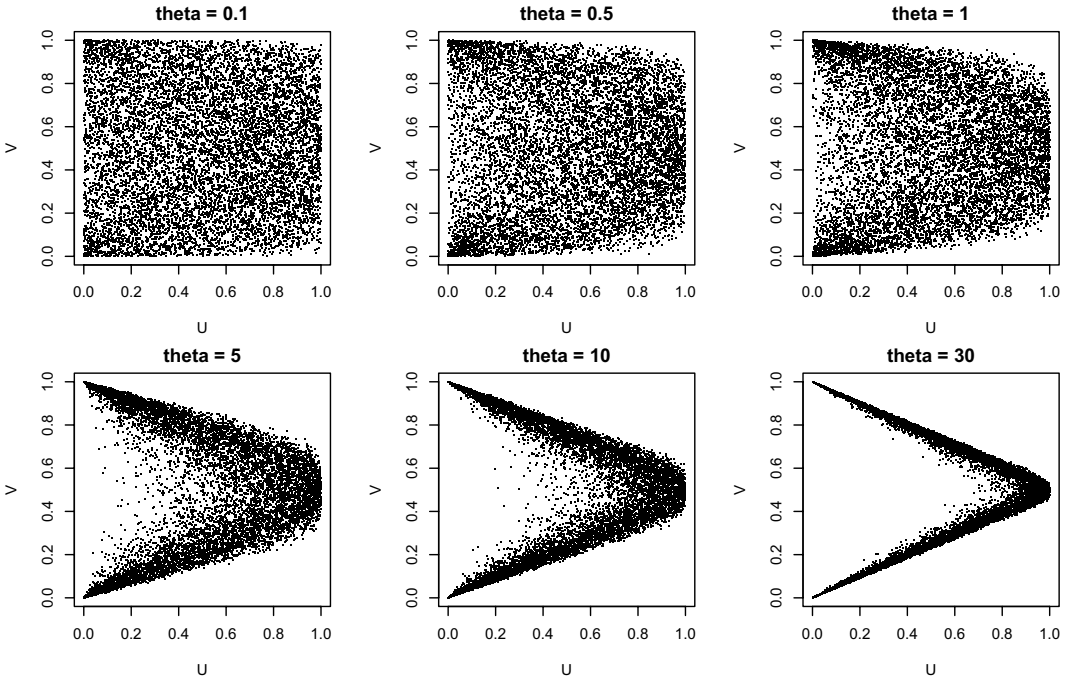


Figure 2. *Scatterplots of 10,000 samples of (U, V) under the NG copula with* $\theta = (0.1, 0.5, 1, 5, 10, 30)$.

Note that the decreasing conditional variance property is also affected by the value of θ. For example, in the case of $\theta \to 0$, the copula $C_\theta(u, v)$ approaches to an independence copula as described in Theorem 3.8 so that $\lim_{\theta\to 0} \mathrm{V}ar\,(V|U = u) = \mathrm{V}ar\,(V) = 1/12$. On the other hand, in the case of $\theta \to \infty$, one can observe that $U = 1 - |2V - 1|$ for $U \in (0, 1)$ from Figure 2 so that $\lim_{\theta\to\infty} \mathrm{V}ar\,(V|U = u) = 0.5 \cdot (1 - 0.5u - 0.5)^2 + 0.5 \cdot (0.5u - 0.5)^2 = 0.25(u - 1)^2$ for $u \in (0, 1)$.

3.3. Limiting behavior of Normal–Gamma copula

In Figure 2, we display scatterplots of random samples generated from the NG copula C_θ for several values of $\theta > 0$. As θ decreases toward zero, the samples exhibit progressively weaker dependence, converging to the pattern of an independence copula. Conversely, as the copula parameter θ tends to infinity, the dependence structure induced by the NG copula converges to a degenerate copula whose probability mass is entirely concentrated on the set

$$\{(u, v) \in (0, 1)^2 : u = 1 - |2v - 1|\}. \tag{14}$$

This section provides the rigorous statements for two asymptotic behaviors of the NG copula; one when $\theta \searrow 0$ and the other $\theta \to \infty$. We start with the asymptotic behavior of the NG copula as θ goes to zero. The proof is given in Appendix B.

Theorem 3.8. *For* $(u, v) \in (0, 1)^2$, *we have*

$$\lim_{\theta\searrow 0} c_\theta(u, v) = 1.$$

Now, as $\theta \to \infty$, we establish that the NG copula converges to a degenerate copula whose probability mass is entirely concentrated on the set in (14). Equivalently, for $(U, V) \sim C_\theta$, this implies that (U, V) are countermonotonic given $V > 1/2$ and comonotonic given $V < 1/2$; see Theorem 3.9 for the formal statement. The proof is given in Appendix C.

https://doi.org/10.1017/asb.2026.10109 Published online by Cambridge University Press

Theorem 3.9. *Let (U_i, V_i), $i = 1, 2$, be independent copies drawn from the NG copula C_θ. Then, as $\theta \to \infty$, we have*

i. Countermonotonicity in the upper half

$$\lim_{\theta\to\infty} \mathbb{P}((V_2 - V_1)(U_2 - U_1) < 0 \mid V_1 > 1/2,\ V_2 > 1/2) = 1.$$

ii. Comonotonicity in the lower half

$$\lim_{\theta\to\infty} \mathbb{P}((V_2 - V_1)(U_2 - U_1) > 0 \mid V_1 < 1/2,\ V_2 < 1/2) = 1.$$

4. Application to the collective risk model

This section presents the CRM, where the NG copula $C_\theta \in \mathcal{S}$ is used to model the dependence between the frequency and the average severity. We start with special case of the positive CRM in Model 1.

Model 4. Consider the setting in Model 1 where a copula $C \in \mathcal{F}_2$ is modeled by an NG copula $C_\theta \in \mathcal{S}$.

Now, we derive the following analytical result on the conditional distribution, which is the key result to calculate the Radon–Nikodym derivative of the distribution of (N^+, M^+) defined in (2). Although conditioning on severity given frequency is more common in collective risk modeling, we consider the reverse conditioning here as a technical step that enables an explicit characterization of the joint density of (N^+, M^+).

Lemma 4.1. *Consider the setting in Model 1. Then, we have*

$$\begin{aligned}
&\mathbb{P}(N^+ = n \mid M^+ = m) \\
&\quad = G\left(G^{-1}\left(F_1^+(n); \frac{1}{\theta}, 1\right); \frac{1}{\theta} + \frac{1}{2},\ 1 + \frac{1}{2}\left(\Phi_{2/\theta}^{-1}\left(F_2^+(m); 0, \theta\right)\right)^2\right) \\
&\qquad - G\left(G^{-1}\left(F_1^+(n-1); \frac{1}{\theta}, 1\right); \frac{1}{\theta} + \frac{1}{2},\ 1 + \frac{1}{2}\left(\Phi_{2/\theta}^{-1}\left(F_2^+(m); 0, \theta\right)\right)^2\right)
\end{aligned}$$

for $m \in \mathbb{R}_{>0}$ and $n \in \mathbb{N}_{>0}$.

Proof. For $N^+ \sim F_1^+$, consider

$$N^* := N^+ - U,$$

where $U \sim \text{Unif}(0, 1)$, and N^+ and U are independent. Then, if we let F_1^* be a cumulative distribution function of N^*, we have

$$F_1^*(n) = F_1^+(n), \quad n \in \mathbb{N}_{>0}$$

and

$$F_1^*(N^+ - U) \sim \text{Unif}(0, 1).$$

Then, for $C_\theta \in \mathcal{S}$ with some $\theta > 0$, consider

$$(N^*, M^+) \sim C_\theta\left(F_1^*, F_2^+\right).$$

Since both F_1^* and F_2^+ are absolutely continuous cumulative distribution functions, we have

$$\left(F_1^*(N^*), F_2^+(M^+)\right) \sim C_\theta$$

and

$$\left(G^{-1}\left(F_1^*(N^*); \frac{1}{\theta}, 1\right), \Phi_{2/\theta}^{-1}\left(F_2^+(M^+); 0, \theta\right)\right) \sim H_\theta,$$

https://doi.org/10.1017/asb.2026.10109 Published online by Cambridge University Press

where H_θ is the joint distribution of the NG distribution defined in (5), and C_θ is the corresponding copula function defined in (7). Then, Corollary 3.1 yields

$$\begin{aligned}
&\mathbb{P}\left(N^+ \le n \mid M^+ = m\right)\\
&= \mathbb{P}\left(N^* \le n \mid M^+ = m\right)\\
&= \mathbb{P}\left(G^{-1}\left(F_1^*(N^*); \frac{1}{\theta}, 1\right) \le G^{-1}\left(F_1^*(n); \frac{1}{\theta}, 1\right) \,\middle|\, \Phi_{2/\theta}^{-1}\left(F_2^+(M^+); 0, \theta\right) = \Phi_{2/\theta}^{-1}\left(F_2^+(m); 0, \theta\right)\right)\\
&= G\left(G^{-1}\left(F_1^*(n); \frac{1}{\theta}, 1\right); \frac{1}{\theta} + \frac{1}{2},\ 1 + \frac{1}{2}\left(\Phi_{2/\theta}^{-1}\left(F_2^+(m); 0, \theta\right)\right)^2\right)\\
&= G\left(G^{-1}\left(F_1^+(n); \frac{1}{\theta}, 1\right); \frac{1}{\theta} + \frac{1}{2},\ 1 + \frac{1}{2}\left(\Phi_{2/\theta}^{-1}\left(F_2^+(m); 0, \theta\right)\right)^2\right)
\end{aligned}$$

for $m \in \mathbb{R}_{>0}$ and $n \in \mathbb{N}_{>0}$. This concludes the proof. □

Since Model 1 assumes positive frequency, it cannot be directly applied to real-world data that includes zero frequencies. In the following, we present a modified version of Model 1, which accommodates zero frequency.

Model 5. (General CRM). Consider the setting described in Model 4. Given a constant $p \in (0, 1)$ and a random variable $R \sim \text{Ber}(p)$, where $\text{Ber}(p)$ denotes the Bernoulli distribution with parameter p, and R is independent of (N^+, M^+), we define

$$N := R\,N^+ \quad \text{and} \quad M := R\,M^+.$$

We refer to N as the frequency and M the average severity. The corresponding joint density function $h : \mathbb{N}_{\ge 0} \times \mathbb{R}_{\ge 0} \to \mathbb{R}_{\ge 0}$ is defined as the Radon–Nikodym derivative of the distribution of (N, M) with respect to the product of the counting measure on $\mathbb{N}_{\ge 0}$ and Lebesgue measures on $\mathbb{R}_{\ge 0}$:

$$h(n, m) := \frac{\partial}{\partial\, m}\left(\mathbb{P}\left(N \le n,\ M \le m\right) - \mathbb{P}\left(N \le n - 1,\ M \le m\right)\right) \tag{15}$$

for $n \in \mathbb{N}_{\ge 0},\ m \in \mathbb{R}_{\ge 0}$.

The following corollary is a direct result of Model 5.

Corollary 4.2. *Under the settings in Model 15, the Radon–Nikodym derivative of the distribution of (N, M) in (15) is obtained by*

$$h(n, m) = \begin{cases} 1 - p, & n = 0 \quad \textit{and} \quad m = 0;\\ p\,h^+(n, m), & n \in \mathbb{N}_{>0} \quad \textit{and} \quad m > 0;\\ 0, & \textit{otherwise.} \end{cases}$$

Here, $h^+(n, m)$ can be calculated using (2) and Lemma 4.1.

Remark 1. Although a copula can in principle be applied directly to the mixed pair (N, M), we separate the zero-claim and positive-claim regimes. The reason is that $(0, 0)$ is a structural point: when $N = 0$, the average severity is not observed and is set equal to zero by convention. Hence, if $p = \mathbb{P}(N > 0)$, then a copula fitted directly to (N, M) would have to allocate probability mass $\mathbb{P}(N = 0, M = 0) = 1 - p$ to the lower-left structural zero component. Thus, when zero-claim observations form a large portion of the portfolio, as is common in auto insurance portfolios, a substantial part of the copula probability mass on $[0, 1] \times [0, 1]$ is used to represent this deterministic zero-zero association. Only the remaining probability mass p is available to describe the actual dependence between frequency and average severity among policies with positive claims. Our Bernoulli construction separates the claim-occurrence mechanism from the positive-claim dependence structure and applies the NG copula fully devoted to (N^+, M^+), where average severity is genuinely observed and the dependence of interest is modeled directly by the copula.

https://doi.org/10.1017/asb.2026.10109 Published online by Cambridge University Press

5. Numerical study

5.1. Finite sample property

We conduct a simulation study to investigate the finite sample properties of the parameter estimation and the dependence between frequency and the average severity based on Model 5. To accommodate the realistic setting, we consider $\mathbf{x}$ as a row vector of the explanatory variables, and the column vectors β_1 and β_2 are corresponding sets of regression coefficients for the frequency and severity parts, respectively. We also consider the exposure w of each policy. Then, for given $\mathbf{x}$ and w, we consider $R \mid \mathbf{x}, w \sim \text{Ber}(p)$ and $N^+ \mid \mathbf{x}, w \sim F_1^+$ such that we have

$$N \mid \mathbf{x}, w \sim \text{NB}(\lambda w, rw) \quad \text{with} \quad \lambda = \exp(\mathbf{x}\beta_1), \tag{16}$$

where $N \sim \text{NB}(\mu, s)$ implies the negative binomial distribution with $\mathbb{E}[N] = \mu$ and $\text{Var}(N) = \mu + \mu^2/s$. Specifically, we set $p = 1 - \mathbb{P}(N = 0 \mid \mathbf{x}, w)$ and

$$\mathbb{P}\left(N^+ = n \mid \mathbf{x}, w\right) = \frac{\mathbb{P}(N = n \mid \mathbf{x}, w)}{p}, \quad n \in \mathbb{N}_{>0}.$$

For the marginal distribution of the positive average severity, we assume generalized beta of the second kind (GB2) distribution,

$$M^+|\mathbf{x}, w \sim \text{GB2}(1, \xi, w\psi, wk + 1) \quad \text{with} \quad \xi = \exp(\mathbf{x}\beta_2), \tag{17}$$

with

$$\mathbb{E}\left[M^+ \mid \mathbf{x}, w\right] = \xi \frac{\psi}{k},$$

and, whenever the second moment exists, $wk > 1$,

$$\text{Var}\left(M^+ \mid \mathbf{x}, w\right) = \xi^2 \frac{\psi(\psi + k)}{k^2(wk - 1)}.$$

where $\text{GB2}(a, b, p, q)$ indicates that GB2 distribution with scale parameter $b > 0$ and shape parameters $a, p, q > 0$ and its probability density function is given by

$$g(s; a, b, p, q) = \frac{a(s/b)^{ap-1}}{bB(p, q)(1 + (s/b)^a)^{p+q}}.$$

Note that the motivation of the using the exposure in (17) is in Remark 2.

Remark 2. For the frequency component, the exposure specification in (16) is standard, and in this case, both

$$\mathbb{E}[N \mid \mathbf{x}, w] = \lambda w \quad \text{and} \quad \text{Var}(N \mid \mathbf{x}, w) = \lambda w + \frac{\lambda^2}{r} w,$$

are of order w.

By contrast, there is no equally canonical way to incorporate exposure into the positive average severity. To motivate our specification, suppose that the distribution of the individual severity does not depend on w and define

$$S^+ := \sum_{j=1}^{N^+} Z_j \quad \text{and} \quad M^+ := \frac{S^+}{N^+},$$

as in Model 2. Under the benchmark assumption that N^+ is independent of the individual severities and that the Z_j's are i.i.d., we have

$$\mathbb{E}\left[M^+ \mid \mathbf{x}, w\right] = \mathbb{E}\left[Z_j \mid \mathbf{x}\right],$$

and

$$\text{Var}\left(M^+ \mid \mathbf{x}, w\right) = \text{Var}\left(Z_j \mid \mathbf{x}\right) \mathbb{E}\left[\frac{1}{N^+} \,\middle|\, \mathbf{x}, w\right].$$

https://doi.org/10.1017/asb.2026.10109 Published online by Cambridge University Press

Table 1. *Parameter settings for the 12 scenarios.*

	Parameter									
Scenario	β_0	β_1	β_2	γ_0	γ_1	γ_2	r	ψ	k	θ
1	−1.5	0.5	1.0	6.5	0.1	0.5	0.2	1.6	0.2	0.1
2	−1.5	0.5	1.0	6.5	0.1	0.5	0.2	1.6	0.2	0.5
3	−1.5	0.5	1.0	6.5	0.1	0.5	0.2	1.6	0.2	1
4	−1.5	0.5	1.0	6.5	0.1	0.5	0.2	1.6	0.2	5
5	−1.5	0.5	1.0	6.5	0.1	0.5	0.2	1.6	0.2	10
6	−1.5	0.5	1.0	6.5	0.1	0.5	0.2	1.6	0.2	30
7	−1.5	0.5	2.0	6.5	0.1	0.5	0.2	1.6	0.2	0.1
8	−1.5	0.5	2.0	6.5	0.1	0.5	0.2	1.6	0.2	0.5
9	−1.5	0.5	2.0	6.5	0.1	0.5	0.2	1.6	0.2	1
10	−1.5	0.5	2.0	6.5	0.1	0.5	0.2	1.6	0.2	5
11	−1.5	0.5	2.0	6.5	0.1	0.5	0.2	1.6	0.2	10
12	−1.5	0.5	2.0	6.5	0.1	0.5	0.2	1.6	0.2	30

Hence, even in this simplest benchmark model, the mean of M^+ is invariant to the exposure, while its variance decreases with w through the term

$$\mathbb{E}\left[\frac{1}{N^+}\,\middle|\,\mathbf{x}, w\right].$$

Moreover, outside the independence setting, the exact form of this exposure effect depends on the dependence structure between frequency and severity. For this reason, instead of imposing an offset structure analogous to that of the frequency model, we adopt the GB2 specification in (17), which preserves an exposure-invariant mean while allowing the variance to decrease flexibly with w through the additional parameter $k > 0$.

A portfolio of 5000 policyholders ($I = 5000$) is generated from the proposed model under 12 scenarios motivated by the real data analysis in Section 6. Table 1 provides the parameter settings. Each simulation uses two categorical variables, x_1 and x_2, generated independently from Ber(0.5); hence, we have $\mathbf{x} = [1,\ x_1,\ x_2]$.

For each scenario, Tables 2 and 3 summarize the simulation results from 100 independent Monte Carlo sample generations. The tables report the relative bias and mean squared error (MSE) of the parameter estimates. Table 2 demonstrates that the estimates in all scenarios closely align with the true model parameters, exhibiting small relative bias and MSE. While Table 3 indicates relatively larger MSE values for $\theta = 30$ in Scenarios 6 and 12, Table 2 confirms that the corresponding relative biases remain well-controlled.

5.2. Variance and thickness of tail

We now investigate how the variance and right-tail behavior of the predictive distribution of total aggregate loss, $\sum_{i=1}^{I} S_i$, are affected by the NG copula parameter θ. Recall that a larger θ indicates a stronger decreasing conditional variance property. This could lead to less variation in M given a larger N, potentially resulting in reduced variance and a lighter right tail for the distribution of S as θ increases.

Figure 3 clearly illustrates this phenomenon. It displays the right tail of the predictive distribution of total aggregate loss based on 5000 simulations for each scenario in Table 1. In the left panel, Scenarios 1–6 have identical parameters for the marginal distributions of frequency and average severity. We observe that as θ increases, the right tail of the total aggregate loss distribution becomes lighter. The same trend is evident in the right panel of Figure 3 for Scenarios 7–12.

https://doi.org/10.1017/asb.2026.10109 Published online by Cambridge University Press

Table 2. *Relative bias in % for all the parameters from the 12 scenarios.*

	Relative bias (%)									
Scenario	β_0	β_1	β_2	r	γ_0	γ_1	γ_2	ψ	k	θ
1	−0.09	−2.34	0.89	−0.13	−2.55	−1.57	0.81	2.30	3.80	3.64
2	−0.01	−2.06	1.00	−0.06	−3.19	−1.00	0.81	1.74	3.79	1.41
3	0.07	−1.75	1.08	0.00	−2.59	−0.46	0.79	1.21	3.80	1.26
4	0.24	−1.17	1.28	0.05	−1.53	0.11	0.76	0.36	2.72	0.99
5	0.17	−0.95	1.08	0.07	−1.68	−0.06	0.68	0.24	2.56	1.07
6	0.00	−0.30	0.50	0.23	−0.46	−0.05	0.43	−0.27	4.82	1.63
7	0.45	1.50	0.36	−0.13	−4.05	0.75	−0.19	0.93	1.37	1.16
8	0.44	1.44	0.39	−0.07	−2.46	1.13	−0.23	0.59	1.36	1.16
9	0.41	1.42	0.36	−0.12	−2.30	1.25	−0.25	0.57	−0.02	1.23
10	0.13	0.68	0.22	−0.03	−1.59	0.23	−0.42	0.63	1.39	0.86
11	−0.09	0.10	0.20	−0.02	−1.70	0.18	−0.58	0.72	2.49	0.91
12	−0.06	−0.15	0.08	−0.05	−0.27	−0.01	−0.42	0.47	0.81	1.29

Table 3. *Mean squared error for all the parameters from the 12 scenarios.*

	MSE									
Scenario	β_0	β_1	β_2	r	γ_0	γ_1	γ_2	ψ	k	θ
1	0.0044	0.0051	0.0056	0.0306	0.0084	0.0091	0.0001	0.0227	0.0062	0.0009
2	0.0042	0.0047	0.0052	0.0272	0.0063	0.0070	0.0001	0.0167	0.0054	0.0039
3	0.0039	0.0044	0.0049	0.0248	0.0042	0.0048	0.0001	0.0138	0.0047	0.0074
4	0.0028	0.0026	0.0035	0.0158	0.0008	0.0008	0.0001	0.0101	0.0033	0.0512
5	0.0021	0.0014	0.0023	0.0155	0.0003	0.0003	0.0001	0.0099	0.0031	0.1865
6	0.0013	0.0005	0.0008	0.0087	0.0001	0.0001	0.0001	0.0053	0.0021	2.1321
7	0.0038	0.0046	0.0042	0.0260	0.0047	0.0060	0.0000	0.0148	0.0047	0.0006
8	0.0038	0.0044	0.0040	0.0200	0.0033	0.0047	0.0000	0.0130	0.0033	0.0026
9	0.0036	0.0041	0.0038	0.0206	0.0023	0.0034	0.0000	0.0131	0.0034	0.0052
10	0.0028	0.0027	0.0023	0.0150	0.0004	0.0007	0.0000	0.0098	0.0025	0.0348
11	0.0021	0.0015	0.0015	0.0089	0.0001	0.0003	0.0000	0.0062	0.0016	0.1079
12	0.0011	0.0004	0.0006	0.0050	0.0000	0.0001	0.0000	0.0027	0.0011	1.3275

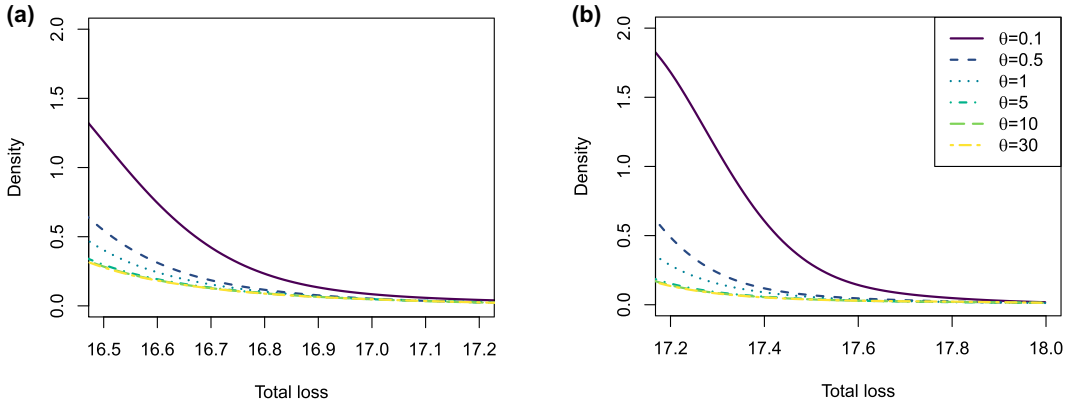


Figure 3. *Right tail of the predictive distribution for the logarithm of the total aggregated losses under (a) Scenarios 1–6 and (b) Scenarios 7–12.*

https://doi.org/10.1017/asb.2026.10109 Published online by Cambridge University Press

Table 4. *Observable policy characteristics and claims information.*

Covariates	Description			Proportions (%)
TypeCity	Indicator for city entity:			11.51
TypeCounty	Indicator for county entity:			13.34
TypeMisc	Indicator for miscellaneous entity:			2.94
TypeSchool	Indicator for school entity:			39.41
TypeTown	Indicator for town entity:			12.97
TypeVillage	Indicator for village entity:			19.83
ClassO	Indicator for an old vehicle			49.94
Claims		Minimum	Mean	Maximum
N	Number of observed claims over five years	0	1.67	47
M	Average annual claim amounts over five years	0	3066	244,364

Table 5. *LRT for influence of N^+ to the dispersion of M^+.*

	−2logLik	*df*	$\hat{p}$	*p*-value
Under H_0	608.1587	8	0	
Under H_1	566.6817	9	−0.5353	
Difference	41.4770			< 0.0001

6. Real data analysis

We applied our model to the LGPIF dataset (Frees *et al*., 2016), a common benchmark in the actuarial literature (Shi and Lee, 2022; Jeong *et al*., 2023). We focused on instances with positive exposure in CO and CN claims, combining five years (2006–2010) of training data using the policy number for merging. We treated the categorical variables O and N as binary covariates. After preprocessing, the dataset contained 817 observations, 329 of which had positive claims. Table 4 summarizes the policy characteristics and claims. The dataset includes a categorical variable representing entity location (City, County, Miscellaneous, School, Town, and Village). We encoded this as five binary indicators (City, County, School, Town, and Village), using Miscellaneous as the reference category. In the real data analysis, we adopt a simplifying assumption that the copula dependence structure is covariate-invariant. This assumption does not affect the MCV property established above, as the result follows from the structural properties of the proposed copula and the marginal specifications, irrespective of whether the dependence parameter is constant or covariate-dependent. Allowing covariate-dependent dependence may increase model flexibility but does not alter the theoretical validity of the MCV property.

Before the dataset is analyzed based on the proposed framework, here we provide another preliminary analysis to detect possible influence of the positive claim frequency (N^+) on the conditional variance of M^+. To be precise, we want to test the following hypothesis:

$$H_0: \text{ Dispersion of } M^+ \text{ conditioned on } N^+ \text{ is not affected by } N^+;$$
$$H_1: \text{ Dispersion of } M^+ \text{ conditioned on } N^+ \text{ is affected by } N^+.$$

To formalize this idea, we adopt an assumption that $M^+|N^+$ follows a gamma distribution where $\mathbb{E}\left[M^+|N^+\right] = \mu = \exp(\mathbf{x}\beta)$ and $\mathrm{V}ar\,(M^+|N^+) = \mu^2 \cdot \psi \cdot (N^+)^p$. Here, we stick to a gamma distribution only for inference purpose to capture the effect of N^+ on the dispersion of M^+. For example, if we use a GB2 distribution (which has indeed better fit for the severity part of the dataset), then the dispersion is already determined by two shape parameters so it is not intuitive to incorporate the impact of N^+ in that case. By doing so, the above hypothesis testing can simply be written as

$$H_0: p = 0 \quad \text{vs} \quad H_1: p \neq 0.$$

https://doi.org/10.1017/asb.2026.10109 Published online by Cambridge University Press

Table 6. *Estimated coefficients and goodness-of-fit of the bivariate copula models.*

	Independent model		Gaussian copula model		Joe copula model		NG copula model	
	Estimate	Std.err	Estimate	Std.err	Estimate	Std.err	Estimate	Std.err
Frequency part								
(Intercept)	−2.0726	0.4127	−2.0823	0.4136	−2.0725	0.4127	−2.0236	0.4066
TypeCity	1.3915	0.4328	1.4023	0.4338	1.3915	0.4328	1.3062	0.4268
TypeCounty	2.6276	0.4240	2.6384	0.4250	2.6276	0.4240	2.5270	0.4178
TypeSchool	0.4172	0.4210	0.4286	0.4220	0.4172	0.4210	0.3755	0.4148
TypeTown	−0.9157	0.4969	−0.9081	0.4977	−0.9157	0.4969	−0.8361	0.4859
TypeVillage	0.3695	0.4325	0.3771	0.4334	0.3695	0.4325	0.3567	0.4263
ClassO	−0.0150	0.1153	−0.0147	0.1154	−0.0150	0.1153	−0.0077	0.1134
r	0.1852	0.0198	0.1848	0.0197	0.1852	0.0198	0.1837	0.0196
Severity part								
(Intercept)	6.2283	0.5800	6.1617	0.5835	6.2283	0.5800	6.3028	0.5536
TypeCity	0.2458	0.3752	0.2573	0.3732	0.2458	0.3752	0.1422	0.3424
TypeCounty	0.7747	0.3665	0.7756	0.3645	0.7747	0.3665	0.6159	0.3349
TypeSchool	0.1896	0.3671	0.1912	0.3651	0.1896	0.3671	−0.0553	0.3374
TypeTown	0.3405	0.4312	0.3409	0.4290	0.3405	0.4311	0.1286	0.3917
TypeVillage	0.3736	0.3771	0.3689	0.3751	0.3736	0.3771	0.1979	0.3435
ClassO	0.1059	0.0989	0.1010	0.0985	0.1059	0.0989	0.0865	0.0914
ψ	1.6676	0.5788	1.7600	0.6281	1.6676	0.5788	1.8671	0.6404
k	0.1771	0.0414	0.1738	0.0404	0.1771	0.0414	0.1757	0.0385
Copula part								
θ	–	–	0.0738	0.0575	1.0000	0.0009	0.1573	0.0617
Goodness-of-fit								
Loglik	−4237.886		−4237.083		−4237.888		−4227.274	
AIC	8509.772		8510.165		8511.777		8490.549	
BIC	8578.005		8582.412		8584.023		8562.795	

Table 5 summarizes the goodness-of-fits for the reduced dispersion model (fitted under H_0) and the full dispersion model (fitted under H_1). Note that the degree of freedom for the full dispersion model is 9 as it has seven covariates including the intercept for the mean part, and two covariates (intercept and N^+) for the dispersion part. From the table, it is shown that the observed dispersion of M^+ from the dataset is severely impacted by N^+, and it is decreasing as N^+ increases (from the sign of $\hat{p}$, −0.5353).

We refer to the proposed model (Model 5), incorporating the marginal models from Section 5, as the NG copula model. For comparison, we also consider various alternative models that utilize different copula components while retaining the same marginals from Section 5, negative binomial for N^+ as in (16) and GB2 for M^+ as in (17). Note that appropriateness of GB2 distribution to model the claim severity with the LGPIF data is supported by existing actuarial literature. For a comprehensive model selection procedure, we refer the reader to Frees *et al*. (2016). We compare the following alternative specifications for the copula component of Model 5, assuming identical marginal models. **Independent model**: This model assumes independence between N^+ and M^+; **Gaussian copula model**: This model captures dependence using a bivariate Gaussian copula ($\theta \in (-1, 1)$). Note that $\theta = 0$ corresponds to independence; **Joe copula model**: This model captures dependence using a bivariate Joe copula given as

$$C(u, v) = 1 - \left[(1-u)^\theta + (1-v)^\theta - (1-u)^\theta(1-v)^\theta\right]^{1/\theta}, \quad \theta \in [1, \infty).$$

Note that $\theta = 1$ corresponds to independence.

https://doi.org/10.1017/asb.2026.10109 Published online by Cambridge University Press

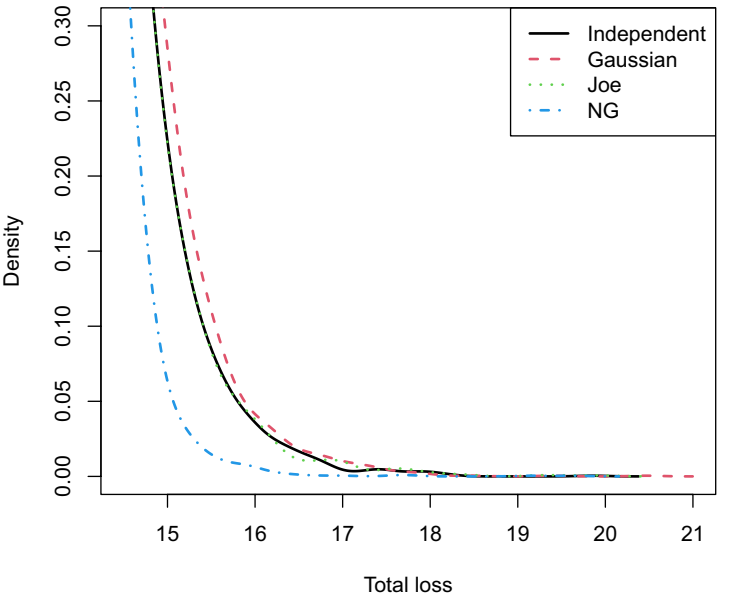


Figure 4. *Right tail of the predictive distribution for the logarithm of total aggregated losses (hold-out sample).*

Table 6 summarizes the estimation results. The estimated copula parameters for the Gaussian and Joe models (0.0738 and 1.0000, respectively) indicate weak dependencies, suggesting they offer minimal improvement over the independent copula model. In contrast, the estimated NG copula parameter (0.1573) implies a moderate level of the monotone decreasing conditional variance property, which the Gaussian and Joe copulas cannot adequately model. The NG copula model demonstrates the best goodness-of-fit (in terms of AIC and BIC), highlighting the importance of capturing the dataset's monotone conditional variance characteristic.

To validate these findings, we compared the four models' total loss predictions for the 716 policyholders in the hold-out sample (year 2011). Figure 4 shows the right tails of the predictive distributions for the logarithm of total aggregated losses. The NG copula model exhibits a significantly lighter right tail, attributed to its decreasing conditional variance property. This aligns with the observations from the numerical study in Figure 3. Figure 5 shows predictive distributions for the logarithm of total aggregated losses from the four models (based on 5000 simulations using each model's estimated parameters). The solid vertical line marks the actual hold-out sample loss (14.02), and the dashed line represents the mean estimated total loss.

Lastly, Table 7 summarizes the predictive distributions of the total aggregated losses from the hold-out sample (minimum, quartiles, maximum, mean, and standard deviation). It is shown that the NG copula model's predictive distribution exhibits the least variation, and its mean is closest to the actual hold-out sample loss. The Gaussian and Joe copula models, with their estimated θ values, essentially behave like the independent copula. They fail to capture the dataset's dependence structure. Conversely, the NG copula model effectively models this dependence through its decreasing conditional variance property. In other words, it naturally reflects the situation where the dispersion of average severity decreases as we have more claim frequencies. This leads to its outperformance on both hold-out sample prediction and in-sample goodness-of-fit.

https://doi.org/10.1017/asb.2026.10109 Published online by Cambridge University Press

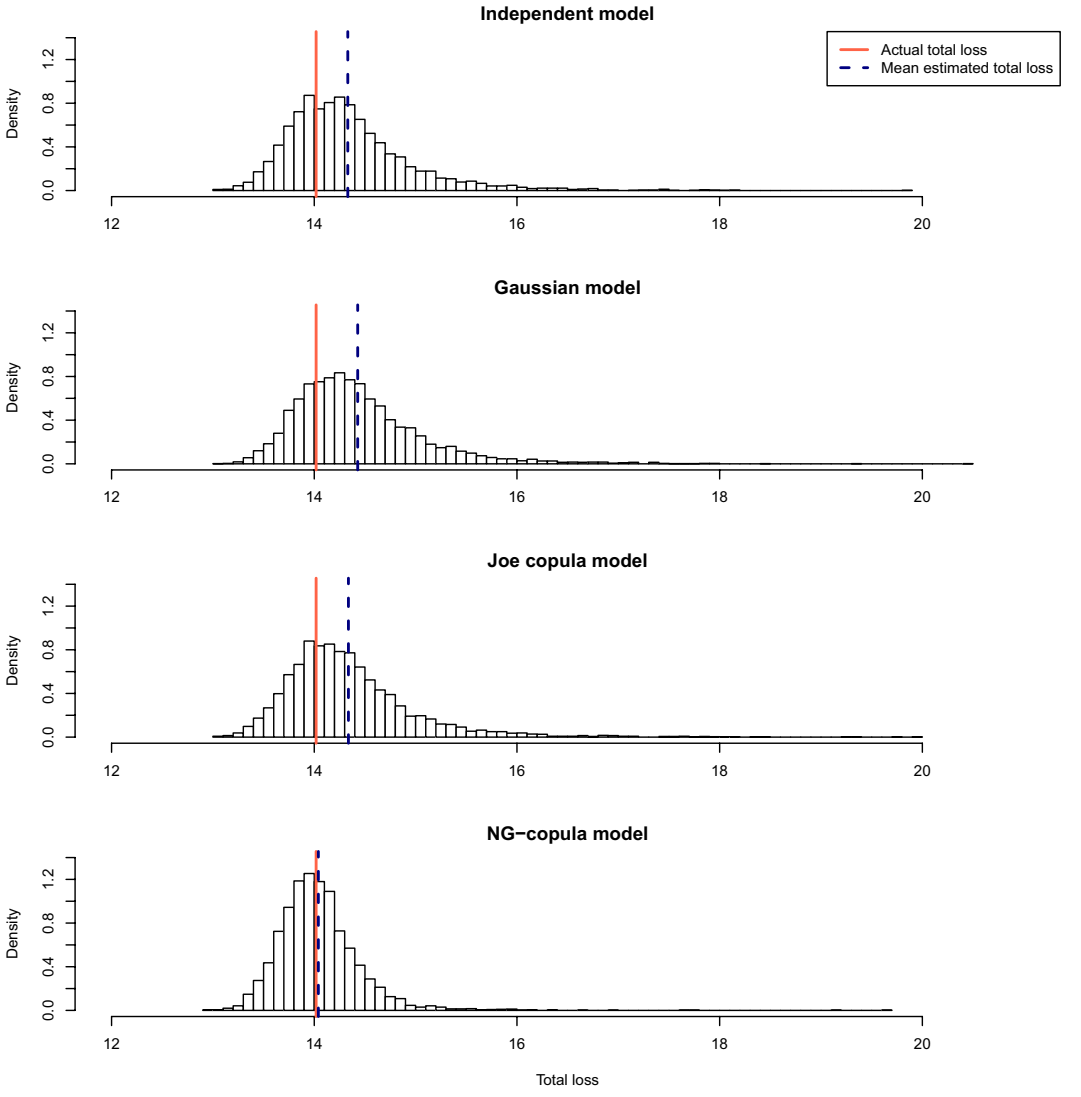


Figure 5. *Predictive distribution for the logarithm of total aggregated losses (hold-out sample).*

7. Conclusion

This paper studied copula-based dependence modeling between claim frequency and average severity in CRMs. We showed that this dependence has a nonstandard structure, especially through the stochastically monotone wing property/the monotone conditional variance property, which are not well captured by standard copula families. To address this issue, we introduced the NG copula and demonstrated, through numerical and real data analyses, that it can effectively model this type of dependence.

One limitation of the current framework is that it is mainly developed for a one-period setting. In particular, the treatment of zero frequency through an auxiliary variable is convenient here, but makes multi-year extensions less straightforward. Extending the proposed framework to longitudinal settings and developing other copula families with similar dependence properties are natural directions for future research.

https://doi.org/10.1017/asb.2026.10109 Published online by Cambridge University Press

Table 7. *Summary statistics for the total aggregated losses (log scale) from the hold-out sample.*

Model	Min.	Q_1	Q_2	Q_3	Max.	Mean	SD
Independent	13.03	13.93	14.23	14.59	19.86	14.33	0.62
Gaussian	13.09	14.00	14.32	14.70	20.47	14.43	0.64
Joe	13.02	13.93	14.23	14.59	19.99	14.34	0.64
NG copula	12.96	13.79	14.00	14.22	19.63	14.04	0.40

Acknowledgments. The authors thank an editor and anonymous referees for their insightful comments, which significantly improved the paper.

Funding statement. Jae Youn Ahn was supported by the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2021R1A6A1A10039823) and by the Ministry of Science and ICT (RS-2025-23524530). Himchan Jeong was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Fund (RGPIN-2021-02622).

Data availability statement. The dataset used in this paper is publicly available at https://sites.google.com/a/wisc.edu/jed-frees/.

Competing interests. The authors have no potential competing or conflicting interests to declare.

References

Belzunce F., Ortega E.-M., Pellerey F. and Ruiz J.M. (2006) Variability of total claim amounts under dependence between claims severity and number of events. *Insurance: Mathematics and Economics* **38**(3), 460–468.

Bernardo J.M. and Smith A.F. (2009) *Bayesian Theory*, vol. 405. New York: John Wiley & Sons.

Chen Z., Dassios A. and Tzougas G. (2024) Em estimation for bivariate mixed poisson inar (1) claim count regression models with correlated random effects. *European Actuarial Journal* **14**(1), 225–255.

Cheung E.C., Ni W., Oh R. and Woo J.-K. (2021) Bayesian credibility under a bivariate prior on the frequency and the severity of claims. *Insurance: Mathematics and Economics* **100**, 274–295.

Czado C., Kastenmeier R., Brechmann E.C. and Min A. (2012) A mixed copula model for insurance claims and claim sizes. *Scandinavian Actuarial Journal* **2012**(4), 278–305.

DeGroot M.H. (2005) *Optimal Statistical Decisions*. Hoboken, New Jersey: John Wiley & Sons.

Dette H. and Pilz K. (2009) On the estimation of a monotone conditional variance in nonparametric regression. *Annals of the Institute of Statistical Mathematics* **61**, 111–141.

Erdemir Ö.K. and Meral S. (2020) A comparative study on modeling of dependence between claim severity and frequency with Archimedean copulas. *İstatistikçiler Dergisi: İstatistik ve Aktüerya* **13**(1), 18–29.

Frees E.W., Lee G. and Yang L. (2016) Multivariate frequency-severity regression models in insurance. *Risks* **4**(1), 4.

Gao G. and Li J. (2023) Dependence modeling of frequency-severity of insurance claims using waiting time. *Insurance: Mathematics and Economics* **109**, 29–51.

Garrido J., Genest C. and Schulz J. (2016) Generalized linear models for dependent frequency and severity of insurance claims. *Insurance: Mathematics and Economics* **70**, 205–215.

Gschlößl S. and Czado C. (2007) Spatial modelling of claim frequency and claim size in non-life insurance. *Scandinavian Actuarial Journal* **2007**(3), 202–225.

Hernández-Bastida A., Fernández-Sánchez M. and Gómez-Déniz E. (2009) The net Bayes premium with dependence between the risk profiles. *Insurance: Mathematics and Economics* **45**(2), 247–254.

Jeong H., Tzougas G. and Fung T.C. (2023) Multivariate claim count regression model with varying dispersion and dependence parameters. *Journal of the Royal Statistical Society Series A: Statistics in Society* **186**(1), 61–83.

Jeong H. and Valdez E.A. (2020) Predictive compound risk models with dependence. *Insurance: Mathematics and Economics* **94**, 182–195.

Jeong H., Valdez E.A., Ahn J.Y. and Park S. (2021) Generalized linear mixed models for dependent compound risk models. *Variance* **14**(1). https://variancejournal.org/article/28245-generalized-linear-mixed-models-for-dependent-compound-risk-models.

Jewell W.S. (1974) Credible means are exact Bayesian for exponential families. *ASTIN Bulletin: The Journal of the IAA* **8**(1), 77–90.

Johnson N.L., Kotz S. and Balakrishnan N. (1995) *Continuous Univariate Distributions, Volume 2*, vol. 289. New York: John Wiley & Sons.

Jorgensen B. (1997) *The Theory of Dispersion Models*. London, UK: CRC Press.

Krämer N., Brechmann E.C., Silvestrini D. and Czado C. (2013) Total loss estimation using copula-based regression models. *Insurance: Mathematics and Economics* **53**(3), 829–839.

https://doi.org/10.1017/asb.2026.10109 Published online by Cambridge University Press

Lee G.Y. and Shi P. (2019) A dependent frequency–severity approach to modeling longitudinal insurance claims. *Insurance: Mathematics and Economics* **87**, 115–129.
Lehmann E.L. (1966) Some concepts of dependence. *The Annals of Mathematical Statistics* **37**(5), 1137–1153.
Nelsen R.B. (2007) *An Introduction to Copulas.* New York: Springer Science & Business Media.
Oh R. and Ahn J.Y. (2024) Stochastic monotone wing property: A new dependence structure for copulas. *Fuzzy Sets and Systems* **484**, 108932.
Oh R., Ahn J.Y. and Lee W. (2021a) On copula-based collective risk models: From elliptical copulas to vine copulas. *Scandinavian Actuarial Journal* **2021**(1), 1–33.
Oh R., Jeong H., Ahn J.Y. and Valdez E.A. (2021b) A multi-year microlevel collective risk model. *Insurance: Mathematics and Economics* **100**, 309–328.
Oh R., Shi P. and Ahn J.Y. (2020) Bonus-malus premiums under the dependent frequency-severity modeling. *Scandinavian Actuarial Journal* **2020**(3), 172–195.
Park S.C., Kim J.H. and Ahn J.Y. (2018) Does hunger for bonuses drive the dependence between claim frequency and severity? *Insurance: Mathematics and Economics* **83**, 32–46.
Pedersen H.L. (2026) On the asymptotic behavior of the quantiles in the gamma distribution. *Experimental Mathematics* **35**(1), 64–78.
Shi P., Feng X. and Ivantsova A. (2015) Dependent frequency–severity modeling of insurance claims. *Insurance: Mathematics and Economics* **64**, 417–428.
Shi P. and Lee G.Y. (2022) Copula regression for compound distributions with endogenous covariates with applications in insurance deductible pricing. *Journal of the American Statistical Association* **117**(539), 1094–1109.
Shi P. and Zhao Z. (2020) Regression for copula-linked compound distributions with applications in modeling aggregate insurance claims. *Annals of Applied Statistics* **14**(1), 357–380.
Sydeman W.J., Santora J.A., Thompson S.A., Marinovic B. and Lorenzo E.D. (2013) Increasing variance in North Pacific climate relates to unprecedented ecosystem variability off California. *Global Change Biology* **19**(6), 1662–1675.
Tukey J.W. (1958) A problem of Berkson, and minimum variance orderly estimators. *The Annals of Mathematical Statistics* **29**(2), 588–592.
Vernic R., Bolancé C. and Alemany R. (2022) Sarmanov distribution for modeling dependence between the frequency and the average severity of insurance claims. *Insurance: Mathematics and Economics* **102**, 111–125.

Appendix A. Example of motivation

The collective risk model specified in Model 2 assumes independence between the positive frequency and the individual severities. However, this assumption seldom holds in real-world data analysis. To address this limitation, Garrido *et al*. (2016) proposed a generalization that models the dependence between the positive frequency and the individual severities, where the independence constraint is relaxed by allowing the conditional mean of the severity to depend on the frequency.

Specifically, for the joint modeling of the positive frequency and the positive average severity, (N^+, M^+), consider the setting in Model 2 with part *ii* replaced with

ii. Conditioned on $N^+ = n \in \mathbb{N}_{>0}$, the random variables $Z_1, Z_2, \ldots, Z_n$ are independent and identically distributed (i.i.d.) following the Gamma $\left(a, \frac{b}{\exp(cn)}\right)$ distribution for some constant $c \in \mathbb{R}$.

Under this modification, we have

$$M^+ \mid N^+ \sim \text{Gamma}\left(aN^+, \frac{bN^+}{\exp\left(cN^+\right)}\right).$$

Furthermore, the conditional mean and the conditional variance are given by

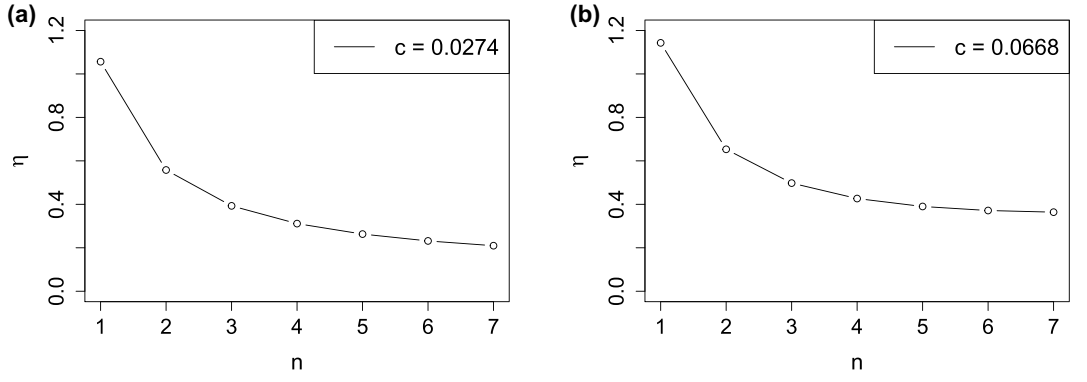
$$\mathbb{E}\left[M^+ \mid N^+\right] = \frac{a}{b}\exp\left(cN^+\right) \quad \text{and} \quad \text{Var}\left(M^+ \mid N^+\right) = \frac{\left(\mathbb{E}\left[M^+ \mid N^+\right]\right)^2}{aN^+}, \tag{A1}$$

respectively. We emphasize that the conditional variance in (A1) decreases with N^+ if and only if the function $\eta{:}\mathbb{N}_{>0} \to \mathbb{R}_{>0}$, defined as

$$\eta(n) = \frac{\exp\left(2cn\right)}{n}, \qquad n \in \mathbb{N}_{>0},$$

is a decreasing function.

https://doi.org/10.1017/asb.2026.10109 Published online by Cambridge University Press



Park et al. (2018) Jeong et al. (2021)

Figure A1. *Decreasing patterns of the η function for various c values.*

When there is a negative dependence between the frequency and the severity, equivalently $c < 0$, then η is a decreasing function. Consequently, the conditional variance in (A1) decreases as the frequency increases. In contrast, when there is a positive dependence, equivalently $c > 0$, the conditional variance of M^+ is generally not guaranteed to decrease. However, when $c > 0$ is small – a situation commonly observed in real data analysis within the insurance literature (Shi *et al.*, 2015; Jeong *et al.*, 2021; Park *et al.*, 2018) – the conditional variance in (A1) is typically observed to decrease with the frequency $N^+ = n$ for moderate values of $n \in \mathbb{N}_{>0}$, as illustrated in Figure A1.

Appendix B. Proof of Theorem 3.8

Before we give the proof of Theorem 3.8, we explore some auxiliary results.

Lemma B.1. *For a fixed $u \in (0, 1)$, let*

$$x^* = G^{-1}(u; 1/\theta, 1). \tag{B1}$$

Then, as $\theta \to 0$,

$$\lim_{\theta \searrow 0} \theta\, x^* = 1.$$

Proof. Let $n = \frac{1}{\theta}$. Define $W_i \overset{\text{iid}}{\sim} G(1, 1)$, we have

$$X_n := W_1 + W_2 + \cdots + W_n \sim G(n, 1).$$

Define

$$F_n(z) = \mathbb{P}\left(\frac{X_n - n}{\sqrt{n}} \le z\right).$$

By the central limit theorem,

$$F_n(z) = \Phi(z) + o(1) \quad \text{as} \quad n \to \infty. \tag{B2}$$

Next, from (B1) we have

$$u = \mathbb{P}\big(X_n \le x^*\big) = \Phi\left(\frac{x^* - n}{\sqrt{n}}\right) + o(1),$$

https://doi.org/10.1017/asb.2026.10109 Published online by Cambridge University Press

where the last step follows from (B2). Since $\Phi(\cdot)$ is strictly increasing and continuously differentiable, the mean-value theorem applied to the above asymptotic implies

$$\frac{x^* - n}{\sqrt{n}} = \Phi^{-1}(u) + o(1) \quad \text{as} \quad n \to \infty,$$

which completes the proof. □

Lemma B.2. *For a fixed $v \in (0, 1)$, let $y^* = \Phi^{-1}_{2/\theta}(v; 0, \theta)$. Then*

$$\lim_{\theta \searrow 0} \frac{y^*}{\sqrt{\theta}} = \Phi^{-1}(v).$$

Proof. Define

$$\widetilde{\Phi}_\theta(z) = \Phi_{2/\theta}\big(z\sqrt{\theta}; 0, \theta\big), \quad z \in \mathbb{R}.$$

Note that the t-distribution is a location-scale family, and hence we have

$$\widetilde{\Phi}_\theta(z) = \Phi_{2/\theta}(z\sqrt{\theta}; 0, \theta) = \Phi_{2/\theta}(z; 0, 1).$$

Since $y^* = \Phi^{-1}_{2/\theta}(v; 0, \theta)$, setting $z^* = y^*/\sqrt{\theta}$ implies that

$$\widetilde{\Phi}_\theta(z^*) = \Phi_{2/\theta}(y^*; 0, \theta) = v.$$

□

Next, by the convergence of the t-distribution to the normal distribution as the degrees of freedom tend to infinity, we have

$$\sup_{z \in \mathbb{R}} \big|\widetilde{\Phi}_\theta(z) - \Phi(z)\big| = \sup_{z \in \mathbb{R}} \big|\Phi_{2/\theta}(z; 0, 1) - \Phi(z)\big| \longrightarrow 0 \quad \text{as} \quad \theta \searrow 0.$$

Hence, for our specific z^*, $v = \widetilde{\Phi}_\theta(z^*) = \Phi(z^*) + o(1)$. Since $\Phi(\cdot)$ is strictly increasing and continuous, applying its inverse function to both sides yields

$$z^* = \Phi^{-1}(\Phi(z^*)) = \Phi^{-1}(v - o(1)).$$

A Taylor expansion of Φ^{-1} around v then implies that $z^* = \Phi^{-1}(v) + o(1)$. Consequently,

$$y^* = \sqrt{\theta}\, z^* = \sqrt{\theta}\, \Phi^{-1}(v) + o(\sqrt{\theta}),$$

which completes the proof.

Lemma B.3. *For fixed $v \in (0, 1)$,*

$$\lim_{\theta \searrow 0} \left[\phi_{2/\theta}(y^*; 0, \theta) - \frac{1}{\sqrt{2\pi\theta}} \exp\left(-\frac{1}{2} \left[\Phi^{-1}(v)\right]^2 \right) \right] = 0 \tag{B3}$$

and

$$\lim_{\theta \searrow 0} \frac{1}{\theta} \left[\frac{1}{\phi_{2/\theta}(y^*; 0, \theta)} - \sqrt{2\pi\theta} \exp\left(\frac{1}{2} \left[\Phi^{-1}(v)\right]^2 \right) \right] = 0 \tag{B4}$$

as $\theta \searrow 0$, where $y^* = \Phi^{-1}_{2/\theta}(v; 0, \theta)$.

Proof. The following is well-known

$$\phi_n(z; 0, 1) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2} \right) + o(1) \quad \text{as } n \to \infty \tag{B5}$$

for $z \in \mathbb{R}$; see, for example Johnson *et al*. (1995). Next, recall from Lemma 9.2 that for fixed $v \in (0, 1)$,

$$y^* = \sqrt{\theta}\, \Phi^{-1}(v) + o(\sqrt{\theta}), \quad \text{as } \theta \searrow 0, \tag{B6}$$

where $y^* = \Phi^{-1}_{2/\theta}(v; 0, \theta)$. Define $z^* = y^*/\sqrt{\theta}$, so that (B6) implies $z^* = \Phi^{-1}(v) + o(1)$ as $\theta \searrow 0$. We apply (B5) to get

$$\phi_{2/\theta}(z^*; 0, 1) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\big(\Phi^{-1}(v) + o(1)\big)^2}{2} \right) \quad \text{as } \theta \searrow 0.$$

https://doi.org/10.1017/asb.2026.10109 Published online by Cambridge University Press

By the standard location-scale relation for the t-density,

$$\phi_{2/\theta}(y^*; 0, \theta) = \frac{1}{\sqrt{\theta}}\, \phi_{2/\theta}\Big(\frac{y^*}{\sqrt{\theta}}; 0, 1\Big).$$

Hence,

$$\begin{aligned}\phi_{2/\theta}(y^*; 0, \theta) &= \frac{1}{\sqrt{\theta}}\ \phi_{2/\theta}(z^*; 0, 1)\\ &= \frac{1}{\sqrt{\theta}}\ \frac{1}{\sqrt{2\pi}}\ \exp\left(-\frac{[\Phi^{-1}(v) + o(1)]^2}{2}\right)\\ &= \frac{1}{\sqrt{2\pi\,\theta}}\ \exp\left(-\frac{[\Phi^{-1}(v)]^2}{2}\right) + o(1),\end{aligned}$$

where the final equality uses a Taylor expansion in the $o(1)$ term. This proves (B3). □

We now prove (B4). From (B3) we see that

$$\phi_{2/\theta}(y^*; 0, \theta) = \frac{A}{\sqrt{\theta}} + o(1), \tag{B7}$$

where

$$A := \frac{1}{\sqrt{2\pi}}\ \exp\left(-\frac{1}{2}\left[\Phi^{-1}(v)\right]^2\right).$$

Inverting (B7) and using the Taylor expansion, we obtain

$$\frac{1}{\phi_{2/\theta}(y^*; 0, \theta)} = \frac{\sqrt{\theta}}{A + o(\sqrt{\theta})} = \frac{\sqrt{\theta}}{A} + o(\theta).$$

Since $\frac{1}{A} = \sqrt{2\pi}\ \exp\big(\frac{1}{2}\,(\Phi^{-1}(v))^2\big)$, we have

$$\frac{1}{\phi_{2/\theta}(y^*; 0, \theta)} = \sqrt{2\pi\,\theta}\ \exp\left(\frac{1}{2}\left[\Phi^{-1}(v)\right]^2\right) + o(\theta),$$

which establishes (B4) and completes the proof.

Proof of Theorem 3.8. First, note that the copula density function is given by

$$c_\theta(u, v) = \frac{1}{\sqrt{2\pi}}\ (x^*)^{1/2} \exp\left(-\frac{1}{2}\, x^*\, (y^*)^2\right) \frac{1}{\phi_{2/\theta}(y^*; 0, \theta)},$$

where $x^* = G^{-1}(u; 1/\theta, 1)$ and $y^* = \Phi^{-1}_{2/\theta}(v; 0, \theta)$.

Lemma B.1 implies

$$(x^*)^{1/2} = \Big(\tfrac{1}{\theta} + o\big(\tfrac{1}{\theta}\big)\Big)^{1/2} = \frac{1}{\sqrt{\theta}}\left[1 + o(1)\right] \quad \text{as} \quad \theta \searrow 0. \tag{B8}$$

From Lemma B.2,

$$y^* = \sqrt{\theta}\ \Phi^{-1}(v) + o\big(\sqrt{\theta}\big) \quad \text{as} \quad \theta \searrow 0,$$

we also have

$$(y^*)^2 = \theta\left[\Phi^{-1}(v)\right]^2 + o(\theta) \quad \Longrightarrow \quad x^*\,(y^*)^2 = \left[\Phi^{-1}(v)\right]^2 + o(1)$$

as $\theta \searrow 0$. Therefore, using a Taylor expansion and Lemma 9.1, we have

$$\exp\left(-\frac{1}{2}\, x^*\, (y^*)^2\right) = \exp\left(-\frac{1}{2}\left[\Phi^{-1}(v)\right]^2\right)\left[\,1 + o(1)\right] \quad \text{as} \quad \theta \searrow 0. \tag{B9}$$

https://doi.org/10.1017/asb.2026.10109 Published online by Cambridge University Press

Finally, using (B8), (B9), and Lemma B.3, we get

$$\begin{aligned} c_\theta(u,v) &= \frac{1}{\sqrt{2\pi}}\,(x^*)^{1/2}\exp\left(-\frac{1}{2}\,x^*\,(y^*)^2\right)\frac{1}{\phi_{2/\theta}(y^*;0,\theta)} \\ &= \frac{1}{\sqrt{2\pi}}\left(\frac{1}{\sqrt{\theta}}[1+o(1)]\right)\left(\exp\left(-\frac{1}{2}[\Phi^{-1}(v)]^2\right)[1+o(1)]\right)\left(\sqrt{2\pi\,\theta}\,\exp\left(\frac{1}{2}[\Phi^{-1}(v)]^2\right)+o(\theta)\right) \\ &= (1+o(1))\,\exp\left(-\frac{1}{2}[\Phi^{-1}(v)]^2\right)\left(\left(\exp\left(\frac{1}{2}[\Phi^{-1}(v)]^2\right)\right)+o(\sqrt{\theta})\right) \\ &= 1+o(1) \end{aligned}$$

as $\theta \searrow 0$. □

Appendix C. Proof of Theorem 3.9

First, define the constant

$$s_0 = \int_0^{e^{-\gamma}} \eta_1(s)\log s\,ds + \int_{e^{-\gamma}}^{\infty} \eta_2(s)\log s\,ds,$$

where $\eta_1(s) = (e^{-s}-1)/s$, $\eta_2(s) = e^{-s}/s$, and γ denotes the Euler–Mascheroni constant defined as

$$\gamma = \lim_{n\to\infty}\left(\sum_{k=1}^{n}\frac{1}{k} - \ln n\right),$$

which satisfies $\gamma > 0$. Numerical evaluation gives $0 < s_0 \simeq 0.4617819158$ (Pedersen, 2026). The following lemma describes the asymptotic behavior of the quantiles of the gamma distribution as the shape parameter approaches zero with the rate parameter fixed at 1.

Lemma C.1 *For $u \in (0,1)$ and $\theta > 0$, let $q_\theta(u)$ denote the u-th quantile of the* Gamma$(1/\theta, 1)$ *distribution. Then,*

i. As $\theta \to \infty$,

$$q_\theta(u) = u^\theta(c + O(1/\theta)),$$

where $c = e^{-\gamma}s_0 > 0$.

ii. For any $u_1, u_2 \in (0,1)$ with $u_1 > u_2$,

$$\lim_{\theta\to\infty}\frac{q_\theta(u_2)}{q_\theta(u_1)} = 0.$$

iii. Let $U_1, U_2 \sim$ Unif $(0,1)$ and $Z_1, Z_2 \sim N(0,1)$ be mutually independent, and define

$$Y_i = \frac{|Z_i|}{q_\theta(U_i)}, \quad i = 1,2.$$

Then,

$$\lim_{\theta\to\infty}\mathbb{P}((Y_2 - Y_1)(q_\theta(U_2) - q_\theta(U_1)) < 0) = 1.$$

Proof. Part *i* follows from the main expansion in Pedersen (2026). The proof of part *ii* immediately follows from part *i*. □

For the proof of part *iii*, since $\mathbb{P}(|Z_1| > 0,\ U_1 > 0) = 1$ for given $\theta > 0$, it is enough to prove

$$\lim_{\theta\to\infty}\mathbb{P}\left((Y_2 - Y_1)\,(q_\theta\,(U_2) - q_\theta\,(U_1)) < 0 \mid\ |Z_1| > 0, U_1 > 0\right) = 1. \tag{C1}$$

https://doi.org/10.1017/asb.2026.10109 Published online by Cambridge University Press

By the symmetricity, for the proof of (C1), it is enough to show

$$\begin{aligned}
&\lim_{\theta\to\infty} \mathbb{P}\left(Y_2 > Y_1 \mid U_2 < U_1, |Z_1| > 0, U_1 > 0\right) \\
&\quad = \lim_{\theta\to\infty} \mathbb{P}\left(\frac{|Z_2|}{|Z_1|} > \frac{q_\theta(U_2)}{q_\theta(U_1)} \,\middle|\, U_2 < U_1, |Z_1| > 0, U_1 > 0 \right) \\
&\quad = \mathbb{P}\left(\frac{|Z_2|}{|Z_1|} > 0 \,\middle|\, U_2 < U_1, |Z_1| > 0, U_1 > 0 \right) \\
&\quad = \mathbb{P}\left(\frac{|Z_2|}{|Z_1|} > 0 \,\middle|\, |Z_1| > 0 \right) \\
&\quad = 1,
\end{aligned}$$

where the first equality is from the monotonicity of the quantile function, and the second equality comes from the dominant convergence theorem as well as the result in part *ii*, and the third equality is from the independence assumption.

Proof of Theorem 3.9. We provide the proof of part *i*; the proof of part ii follows similarly.

Showing part *i* is equivalent to show that as $\theta \to \infty$ the NG copula exhibits asymptotically countermonotonic behavior in the upper tail (i.e. when $V > 1/2$). By invariance of the copula under strictly increasing transformations, it suffices to work with the representation of the NG distribution. Using the representation in Model 3, we represent the random vector from the NG distribution $H_{1/\theta,1}$, via the quantile transformation. Let

$$X_i = q_\theta(U_i), \quad Y_i = \frac{Z_i}{q_\theta(U_i)}, \quad i = 1, 2,$$

where $q_\theta(u)$ denotes the u-th quantile of the Gamma $(1/\theta, 1)$ distribution, $U_i \sim \text{Unif}\,(0, 1)$, and $Z_i \sim N(0, 1)$ are mutually independent. We also define

$$Y_i^* = \frac{|Z_i|}{q_\theta(U_i)}.$$

Then, we have that (X_1, X_2) and (Y_1, Y_2) are independent copies drawn from $H_{1/\theta,1}$.

By the symmetry of the Student's t-distribution, the countermonotonic behavior of the NG copula in the upper tail $V > 1/2$ is equivalent to the countermonotonic behavior of the NG distribution in the upper tail $Y > 0$. Thus, it suffices to show

$$\lim_{\theta\to\infty} \mathbb{P}((Y_2 - Y_1)(X_2 - X_1) < 0 \mid Y_1 > 0,\ Y_2 > 0) = 1.$$

This can be obtained from

$$\begin{aligned}
\lim_{\theta\to\infty} \mathbb{P}\left((Y_2 - Y_1)(X_2 - X_1) < 0 \mid Y_1 > 0, Y_2 > 0\right) &= \mathbb{P}\left((Y_2^* - Y_1^*)(X_2 - X_1) < 0\right) \\
&= 1,
\end{aligned}$$

where the first equality is from the symmetric property of the standard normal distribution, and the second equation is from part *iii* of Lemma C.1. □

https://doi.org/10.1017/asb.2026.10109 Published online by Cambridge University Press