Method of Winsorized Moments for Robust Fitting of Truncated and Censored Lognormal Distributions

Chudamani Poudyal,1 Qian Zhao,2 and Vytaras Brazauskas3

1Department of Statistics and Data Science, University of Central Florida, Orlando, Florida
2Department of Mathematics, Robert Morris University, Moon Township, Pennsylvania
3Department of Mathematical Sciences, University of Wisconsin–Milwaukee, Milwaukee, Wisconsin

When constructing parametric models to predict the cost of future claims, several important details have to be taken into account: (1) models should be designed to accommodate deductibles, policy limits, and coinsurance factors; (2) parameters should be estimated robustly to control the influence of outliers on model predictions; and (3) all point predictions should be augmented with estimates of their uncertainty. The methodology proposed in this article provides a framework for addressing all of these aspects simultaneously. Using payment per payment and payment per loss variables, we construct the adaptive version of method of winsorized moments (MWM) estimators for the parameters of truncated and censored lognormal distribution. Further, the asymptotic distributional properties of this approach are derived and compared with those of the maximum likelihood estimator (MLE) and method of trimmed moments (MTM) estimators, the latter being a primary competitor to MWM. Moreover, the theoretical results are validated with extensive simulation studies and risk measure robustness analysis. Finally, practical performance of these methods is illustrated using the well-studied dataset of 1500 U.S. indemnity losses.

1. INTRODUCTION

Parametric inference for loss models is commonly based on maximum likelihood estimators (MLEs), which are sensitive to model misspecification and outliers. To address these vulnerabilities, one can employ robust techniques such as method of trimmed moments (MTM) and method of winsorized moments (MWM); see Brazauskas, Jones, and Zitikis (2009) and Zhao, Brazauskas, and Ghora (2018). MTM and MWM, however, were designed for completely observed data, not for insurance payment variables, which often are affected by deductibles, policy limits, or coinsurance factors. MLEs can certainly handle such data transformations, but, as will be shown later, even when data are truncated and censored—and thus within finite range—sensitivity of MLEs to data perturbations at the interval endpoints is inescapable. Therefore, the main motivation for this work is to develop robust estimators for the parameters of lognormal distributions when data are left-truncated and right-censored (these are primary transformations for defining insurance payments).

Applications of the lognormal distribution are diverse and include actuarial science, business, economics, and other areas (see, e.g., Serfling [2002], and the references cited therein). The model has been effectively used with homogeneous loss data (Hewitt and Lefkowitz 1979; Punzo, Bagnato, and Maruotti 2018) as well as with heterogeneous data. In the latter case, it has been shown by Cooray and Ananda (2005), Scollnik (2007), Brazauskas and Kleefeld (2016), Miljkovic and Grun (2016), Punzo, Bagnato, and Maruotti (2018), and Blostein and Miljkovic (2019) that lognormal models can capture the nature of the dataset either in the head or in the tail or in both parts. Also, using approximate Bayesian computation methods and the Bayesian information criterion, Goffard and Laub (2021) demonstrated that a lognormal distribution provides the best fit to claim sizes when compared to gamma and Weibull models.

Further, sensitivity of MLEs to the underlying modeling assumptions has been known for a long time, at least since the early 1960s (Tukey 1960). Moreover, even for finite mixture models, MLEs often produce unstable tail estimates of the underlying parameters (Fung 2022). Though there exist multiple reasons for model misspecification, in the actuarial literature researchers emphasized unobserved heterogeneity of claims, multimodality, and different tail behavior of small and large claim sizes.
amounts (see Delong, Lindholm, and Wüthrich 2021). The important conclusion that can be drawn from these observations is that MLE-based model estimates are usually flawed, resulting in biased risk predictions and/or unreliable assessment of their variability. One of the most popular proposals to deal with such a problem that emerged in the literature is to fit spliced (or mixtures of) loss distributions; see Cooray and Ananda (2005), Scollnik (2007), Brazauskas and Kleefeld (2016), Blosstein and Miljkovic (2019), and Delong, Lindholm, and Wüthrich (2021). In particular, Nadarajah and Bakar (2014) and Tomarchio and Punzo (2020) used the lognormal distribution as one component in mixed composite models to fit the observed claim severity datasets. This approach is intuitively appealing because it provides more flexibility to the data modeler because mixtures of distributions are certainly more flexible than a single loss distribution. However, it still remains to be seen whether such an approach yields stable cost predictions (i.e., stable against outliers and other data perturbations) because fitting of mixture models to data is done using MLEs or equivalent methods.

Furthermore, there are several types of statistics that can be used to construct robust estimation procedures. In the actuarial literature, the first comprehensive paper on robust estimation of lognormal distributions used generalized L-statistics (Serfling 2002). Redesigning such computationally intense methods for insurance payments, however, is not simple. Therefore, we will employ less general (but very effective) techniques based on L-statistics (Chernoff, Gastwirth, and Johns 1967). In this line of research, the robust $t$-score methodology and its variants (Fabián 2008; Stehlik et al., 2010) have been designed for completely observed ground-up heavy-tailed insurance data. For incomplete loss data, Efromovich (2021) developed a nonparametric data-driven approach to robustly estimate the sample mean of the hazard rate and the conditional probability density. On the parametric side, several new proposals have been made by Poudyal (2021a,b), where the first paper developed MTM estimators of the lognormal distribution parameters for payment per payment and payment per loss data scenarios and the second studied robust estimation via data truncation, censoring, and related variants. Both MTM and MWM have been investigated for a single-parameter Pareto model by Poudyal and Brazauskas (2022b). In the current literature, MTM is the most effective parametric approach for robust estimation of truncated and censored lognormal models. But as earlier theoretical and empirical studies have shown (see Zhao, Brazauskas, and Ghorai 2018), for complete data and under lognormal distributional assumptions, MWM outperforms MTM in terms of asymptotically smaller variance (all else being equal). Thus, in this article we will develop MWM estimators for the parameters of left-truncated and right-censored lognormal distributions and compare them with the corresponding MLE and MTM estimators. The asymptotic distributional properties of MWM are derived and compared with those of MLE and MTM. Moreover, the theoretical results are validated with extensive simulation studies and risk measure robustness analysis. They are also illustrated on the well-studied dataset of 1500 U.S. indemnity losses.

The rest of the article is organized as follows. In Section 2, we briefly summarize two types of insurance benefit payments when the underlying distribution is lognormal. Section 3 provides the existing MLE results for the two payment variables along with the estimators of the lognormal distribution parameters for payment per payment and payment per loss data scenarios and the second studied robust estimation via data truncation, censoring, and related variants. Both MTM and MWM have been investigated for a single-parameter Pareto model by Poudyal and Brazauskas (2022b). In the current literature, MTM is the most effective parametric approach for robust estimation of truncated and censored lognormal models. But as earlier theoretical and empirical studies have shown (see Zhao, Brazauskas, and Ghorai 2018), for complete data and under lognormal distributional assumptions, MWM outperforms MTM in terms of asymptotically smaller variance (all else being equal). Thus, in this article we will develop MWM estimators for the parameters of left-truncated and right-censored lognormal distributions and compare them with the corresponding MLE and MTM estimators. The asymptotic distributional properties of MWM are derived and compared with those of MLE and MTM. Moreover, the theoretical results are validated with extensive simulation studies and risk measure robustness analysis. They are also illustrated on the well-studied dataset of 1500 U.S. indemnity losses.

The rest of the article is organized as follows. In Section 2, we briefly summarize two types of insurance benefit payments when the underlying distribution is lognormal. Section 3 provides the existing MLE results for the two payment variables along with the estimators’ asymptotic distributional properties. Section 3 also includes MLE composite models for those two payment data types. Section 4 is focused on the development of MWM procedures for the location and scale parameters under the distributional scenarios of Section 2. Also in this section, comparisons of asymptotic relative efficiencies of MWM estimators with respect to MLE as well as MTM are presented. In Section 5, we conduct a simulation study and risk sensitivity study to complement the theoretical results. Further, the newly designed methodology is implemented on real data and its performance is provided in Section 6. Finally, concluding remarks and future outlook are offered in Section 7.

2. PAYMENT DATA SCENARIOS

Consider a ground-up lognormal loss random variable $W \sim LN(w_0, \theta, \sigma)$ with cumulative distribution function (cdf) $F_W$ and probability density function (pdf) $f_W$, respectively, given by

$$F_W(w) = \Phi\left(\frac{\log(w-w_0) - \theta}{\sigma}\right) \quad \text{and} \quad f_W(w) = \frac{1}{\sigma(w-w_0)} \phi\left(\frac{\log(w-w_0) - \theta}{\sigma}\right),$$

(2.1)

for $w > w_0, \sigma > 0$, and $\theta \in \mathbb{R}$, where $\Phi$ and $\phi(x) = \left(1/\sqrt{2\pi}\right) e^{-x^2/2}$, $x \in \mathbb{R}$ are the cdf and pdf of the standard normal distribution, respectively. Then with policy deductible $d$, policy limit $u$, and coinsurance factor $c$, the two typical insurance payments random variables, payment per payment ($Y_w$) and payment per loss ($Z_w$), are defined as (see Klugman, Panjer, and Willmot 2019, p. 126):

$$Y_w := c(\min\{W, u\} - d) \mid W > d = \begin{cases} \text{undefined,} & W \leq d; \\ c(W - d), & d < W < u; \\ c(u - d), & u \leq W. \end{cases}$$

(2.2)
Clearly, $X := \log (W - w_0) \sim N(\theta, \sigma^2)$ with cdf and pdf, respectively, given by

$$F(x) = \Phi\left(\frac{x - \theta}{\sigma}\right) \quad \text{and} \quad f(x) = \frac{1}{\sigma} \phi\left(\frac{x - \theta}{\sigma}\right), \quad -\infty < x < \infty. \quad (2.4)$$

The corresponding quantile function $F^{-1} : (0, 1) \to \mathbb{R}$ is given by $F^{-1}(v) = \theta + \sigma F^{-1}(v)$. Consider the following notations from Poudyal (2021a):

$$t := \log (d - w_0), \quad T := \log (u - w_0), \quad R := T - t, \quad \gamma := \frac{t - \theta}{\sigma}, \quad \text{and} \quad \xi := \frac{T - \theta}{\sigma}. \quad (2.5)$$

Note that it is possible to have $t < 0$ but $d > 0$. Then, it follows that

$$\theta = t - \sigma \gamma \quad \text{and} \quad \xi = \gamma + \frac{R}{\sigma}. \quad (2.6)$$

With these notations, the corresponding normal form of the random variables $Y_w$ and $Z_w$, respectively, defined by Equations (2.2) and (2.3) are now, respectively, given by

$$Y := c \log \left(\frac{Y_w}{c(d - w_0) + 1}\right) = c(\min\{X, T\} - t) \quad | \quad X > t = \begin{cases} \text{undefined}, & X \leq t; \\ c(X - t), & t < X < T; \\ cR, & T \leq X. \end{cases} \quad (2.7)$$

$$Z := c \log \left(\frac{Z_w}{c(d - w_0) + 1}\right) = c(\min\{X, T\} - \min\{X, t\}) = \begin{cases} 0, & X \leq t; \\ c(X - t), & t < X < T; \\ cR, & T \leq X. \end{cases} \quad (2.8)$$

The cdfs and pdfs of the random variables $Y$ and $Z$ are respectively given by

$$F_Y(y) = \begin{cases} 0, & y \leq 0; \\ \frac{F_X\left(\frac{y}{c} + t\right) - F_X(t)}{1 - F_X(t)}, & 0 < y < cR; \\ 1, & y \geq cR, \end{cases} \quad f_Y(y) = \begin{cases} \frac{F_X\left(\frac{y}{c} + t\right)}{c[1 - F_X(t)]}, & 0 < y < cR; \\ 1 - F_X(t), & y = cR; \\ 0, & \text{elsewhere}, \end{cases} \quad (2.9)$$

$$F_Z(z) = \begin{cases} 0, & z \leq 0; \\ F_X\left(\frac{z}{c} + t\right), & 0 \leq z < cR; \\ 1, & z \geq cR, \end{cases} \quad f_Z(z) = \begin{cases} F_X(t), & z = 0; \\ \frac{F_X\left(\frac{z}{c} + t\right)}{1 - F_X(t)}, & 0 < z < cR; \\ 1 - F_X(T^-), & z = cR; \\ 0, & \text{elsewhere}. \end{cases} \quad (2.10)$$

Note 1. The constants $t$ and $T$ can be treated as transformed deductible and policy limit, respectively, for the normal random variable $X \sim N(\theta, \sigma^2)$ with a possibility of $t < 0$. \hfill \square
3. MLE

If a truncated (both singly and doubly) normal sample data set is available, then the MLE procedures for such data have been developed by Cohen (1950) and the method of moments estimators can be found in Cohen (1951) and Shah and Jaiswal (1966). The corresponding results for payments $Y$ and data types have been established by Poudyal (2021a).

For any $0 \leq s \leq 1$, define $\bar{s} = 1 - s$. Therefore, $\Phi(z) = 1 - \Phi(z)$ is the standard normal survival function at $z \in \mathbb{R}$. Consider

$$\Omega_1 := \frac{\phi(\gamma)}{\Phi(\gamma) - \Phi(\xi)} \quad \text{and} \quad \Omega_2 := \frac{\phi(\xi)}{\Phi(\gamma) - \Phi(\xi)}.$$  

(3.1)

Note 2. For MLE estimation purposes, the variable $\gamma$, defined in (2.5), can be treated the same way as parameter $\theta$. Therefore, the mean $\theta$ is a linear function of $\gamma$ given by (2.6).

3.1. Payments $Y$

Let $y_1, \ldots, y_n$ be an independent and identically distributed (i.i.d.) sample given by pdf $f_Y$ with policy limit $T$, deductible $t$, and coinsurance factor $c$. Define $n_1 := \sum_{i=1}^n 1 \{0 < y_i < cR\}$ and $n_2 := \sum_{i=1}^n 1 \{y_i = cR\}$. Then the MLE system of implicit equations to be solved for $(\hat{\gamma}_{Y,\text{MLE}}, \hat{\sigma}_{Y,\text{MLE}})$ is given by

$$\begin{align*}
\sigma(\Omega_{y,1} - \Omega_{y,2} - \gamma) - c^{-1}\hat{\mu}_{y,1} &= 0, \\
\sigma^2 \left(1 - \gamma(\Omega_{y,1} - \Omega_{y,2} - \gamma) - \frac{\Omega_{y,2}R}{\sigma}\right) - c^{-2}\hat{\mu}_{y,2} &= 0,
\end{align*}$$

(3.2)

where $\hat{\mu}_{y,1}$ and $\hat{\mu}_{y,2}$ are the first and second sample moments, $\hat{\mu}_{y,j} := n_i^{-1} \sum_{i=1}^n 1 \{0 < y_i < cR\}y_i$, $j = 1, 2$, and

$$\Omega_{y,1} := \frac{n \phi(\gamma)}{n \Phi(\gamma)} \quad \text{and} \quad \Omega_{y,2} := \frac{n \phi(\xi)}{n \Phi(\xi)}.$$  

(3.3)

Further, it has been established by Poudyal (2021a) that

$$(\hat{\gamma}_{Y,\text{MLE}}, \hat{\sigma}_{Y,\text{MLE}}) \sim \mathcal{N}\left((\gamma, \sigma), \frac{\Lambda^{-1}}{n(r_1 r_3 - r_2^2)} \begin{bmatrix} -r_3 & \sigma r_2 \\ \sigma r_2 & -\sigma^2 r_1 \end{bmatrix}\right).$$

(3.4)

where $\Lambda := \frac{\phi(\gamma) - \phi(\xi)}{\Phi(\gamma)}$ and

$$\begin{align*}
r_1(\gamma, \xi) &= \left[-1 + \gamma \Omega_1 - \xi \Omega_2 - \frac{\phi(\gamma)}{\Phi(\gamma)} \Omega_1 + \frac{\phi(\xi)}{\Phi(\xi)} \Omega_2\right], \\
r_2(\gamma, \xi) &= \frac{R \Omega_2}{\sigma} \left[\frac{\phi(\xi)}{\Phi(\xi)} - \xi\right] + [\Omega_1 - \Omega_2 - \gamma], \\
r_3(\gamma, \xi) &= \left(\frac{\xi}{\sigma}\right)^2 \Omega_2 \left[\xi - \frac{\phi(\xi)}{\Phi(\xi)}\right] - \left[2 - \gamma(\Omega_1 - \Omega_2 - \gamma) - \frac{\Omega_2 R}{\sigma}\right].
\end{align*}$$

(3.5)

Because $(\theta, \sigma) = (t - \gamma, \sigma)$, then by the multivariate delta method (see, e.g., Serfling [1980], p. 122), we have

$$(\hat{\theta}_{Y,\text{MLE}}, \hat{\sigma}_{Y,\text{MLE}}) \sim \mathcal{N}\left((\theta, \sigma), \frac{1}{n} S_{Y,\text{MLE}}\right),$$

where

$$S_{Y,\text{MLE}} = \frac{\Lambda^{-1}}{(r_1 r_3 - r_2^2)} D \begin{bmatrix} -r_3 & \sigma r_2 \\ \sigma r_2 & -\sigma^2 r_1 \end{bmatrix} D'$$

and

$$D = \begin{bmatrix} -\sigma & -\gamma \\ 0 & 1 \end{bmatrix}. $$

(3.6)
3.2. Payments Z

Consider an observed i.i.d. sample \( z_1, \ldots, z_n \) given by pdf \( f_Z \). Define

\[
n_0 := \sum_{i=1}^{n} 1\{z_i = 0\}, n_1 := \sum_{i=1}^{n} 1\{0 < z_i < cR\}, n_2 := \sum_{i=1}^{n} 1\{z_i = cR\}.
\]

(3.7)

Note that \( n = n_0 + n_1 + n_2 \). Similar to (3.1), define

\[
\Omega_{c,1} := \frac{n_0 \phi(x)}{n_1 \Phi(x)}, \quad \Omega_{c,2} := \frac{n_2 \phi(\xi)}{n_1 \Phi(\xi)}.
\]

(3.8)

Then, the MLE system of equations to be solved for \( \gamma \) and \( \sigma \) becomes

\[
\begin{align*}
\sigma(\Omega_{c,1} - \Omega_{c,2} - \gamma) - c^{-1}\tilde{\mu}_{z,1} &= 0, \\
\sigma^2 \left( 1 - \gamma(\Omega_{c,1} - \Omega_{c,2} - \gamma) - \frac{\Omega_{c,2}R}{\sigma} \right) - c^{-2}\tilde{\mu}_{z,2} &= 0,
\end{align*}
\]

(3.9)

where \( \tilde{\mu}_{z,1} \) and \( \tilde{\mu}_{z,2} \) are the first and second sample moments, \( \tilde{\mu}_{z,j} := n_1^{-1} \sum_{i=1}^{n} 1\{0 < z_i < cR\}z_i, j = 1, 2 \). Further, it has been established by Poudyal (2021a) that

\[
(\tilde{\gamma}_{z,MLE}, \tilde{\sigma}_{z,MLE}) \sim AN\left( (\gamma, \sigma), \Lambda^{-1} \right),
\]

(3.10)

where

\[
\begin{align*}
\psi_1(\gamma, \xi) &:= -\left[ 1 + \gamma \Omega_1 - \xi \Omega_2 + \frac{\phi(x)}{\Phi(x)} \Omega_1 + \frac{\phi(\xi)}{\Phi(\xi)} \Omega_2 \right], \\
\psi_2(\gamma, \xi) &:= \frac{R\Omega_2}{\sigma} \left[ \frac{\phi(\xi)}{\Phi(\xi)} - \xi \right] + [\Omega_1 - \Omega_2 - \gamma], \\
\psi_3(\gamma, \xi) &:= \frac{\Phi(\xi)}{\Phi(x)} \left( \xi - \frac{\phi(x)}{\Phi(x)} \right) - \left[ 2 - \gamma(\Omega_1 - \Omega_2 - \gamma) - \frac{\Omega_2R}{\sigma} \right].
\end{align*}
\]

(3.11)

Then, it follows that \( (\tilde{\gamma}_{z,MLE}, \tilde{\sigma}_{z,MLE}) \sim AN((\theta, \sigma), \frac{1}{n} S_{z,MLE}) \), where

\[
S_{z,MLE} = \frac{\Lambda^{-1}}{\Phi(x)} \left[ \psi_1 \psi_3 - \psi_2^2 \right] D \left[ \begin{array}{cc} -\psi_3 & \sigma \psi_2 \\ \sigma \psi_2 & -\sigma^2 \psi_1 \end{array} \right] D', \quad D \text{ is given by (3.6)}.
\]

(3.12)

4. MWM

MWM estimators are derived by following the standard method-of-moments approach, but instead of standard moments we match sample and population winsorized moments (or their variants). Similar to the MTM estimator, the population winsorized moments also always exist. The following definition lists the formulas of sample and population winsorized moments for the payment per payment and payment per loss data scenarios.

**Definition 1.** Let us denote the sample and population winsorized moments as \( \tilde{W}_j \) and \( W_j(\theta) \), respectively. If \( w_{1:n} \leq \cdots \leq w_{n:n} \) is an ordered realization of variables (2.7) or (2.8) with quantile function (aq) denoted by \( F_V^{-1}(v \mid \theta) \) with \( V \in \{Y, Z\} \), then the sample and population winsorized moments, with the winsorizing proportions \( a \) (lower), \( b \) (upper) and \( b = 1 - b \), have the following expressions:
\[
\widehat{W}_j = \frac{1}{n} \left[ m_n h(w_{mn+1:n}) \right]^j + \sum_{i=mn+1}^{n-m_n} \left[ h(w_{i:n}) \right]^j + m_n^* \left[ h(w_{n-m_n:n}) \right]^j, \quad (4.1)
\]

\[
W_j(\theta) = a \left[ h(F_V^{-1}(a \mid \theta)) \right]^j + \int_a^b \left[ h(F_V^{-1}(v \mid \theta)) \right]^j dv + b \left[ h(F_V^{-1}(b \mid \theta)) \right]^j, \quad (4.2)
\]

where \( j = 1, \ldots, k \), and the winsorizing proportions \( a, b \) and function \( h \) are chosen by the researcher. Also, integers \( m_n \) and \( m_n^* \) (\( 0 \leq m_n < n - m_n^* \leq n \)) are such that \( m_n/n \to a \) and \( m_n^*/n \to b \) when \( n \to \infty \). In finite samples, the integers \( m_n \) and \( m_n^* \) are computed as \( m_n = \lfloor na \rfloor \) and \( m_n^* = \lfloor nb \rfloor \), where \( \lfloor \cdot \rfloor \) denotes the greatest integer part.

Winsorized estimators are found by matching sample winsorized moments (4.1) with population winsorized moments (4.2) for \( j = 1, \ldots, k \), and then solving the system of equations with respect to \( \theta_1, \ldots, \theta_k \). The obtained solutions, which we denote by
\[
\widehat{\theta}_j = g_j(\widehat{W}_1, \ldots, \widehat{W}_k), \quad 1 \leq j \leq k,
\]
are, by definition, the MWM estimators of \( \theta_1, \ldots, \theta_k \). Note that the functions \( g_j \) are such that \( \theta_j = g_j(W_1(\theta), \ldots, W_k(\theta)) \).

The asymptotic theory of MWM estimators as a general class of \( L \)-statistics can be found in Chernoff, Gastwirth, and Johns (1967), and a more computationally efficient expression for completely observed data scenarios has been established by Zhao, Brazauskas, and Ghorai (2018) and is given by Theorem 1.

**Theorem 1.** Suppose an i.i.d. realization of variables (2.7) or (2.8) has been generated by cdf \( F_V(v \mid \theta) \), which, depending upon the data scenario, equals the cdf given by \( F_Y, V \in \{ Y, Z \} \), respectively. Let
\[
\widehat{\theta}_w = (\widehat{\theta}_1, \ldots, \widehat{\theta}_k) = \left( g_1(\widehat{W}_1, \ldots, \widehat{W}_k), \ldots, g_k(\widehat{W}_1, \ldots, \widehat{W}_k) \right)
\]
denote a winsorized estimator of \( \theta \). Then
\[
\widehat{\theta}_w = (\widehat{\theta}_1, \ldots, \widehat{\theta}_k) \quad \text{is} \quad \mathcal{N} \left( (\theta_1, \ldots, \theta_k), \frac{1}{n} \mathbf{D} \Sigma \mathbf{D}^T \right), \quad (4.3)
\]
where \( \mathbf{D} := [d_{ij}]_{i,j=1}^k \) is the Jacobian of the transformations \( g_1, \ldots, g_k \) evaluated at \( (W_1(\theta), \ldots, W_k(\theta)) \) and \( \Sigma := [\sigma_{ij}^{(2:k)}]_{i,j=1}^{2:k} \) is the variance–covariance matrix with the entries
\[
\sigma_{ij}^y = \widehat{\alpha}_{ij}^{(1)} + \widehat{\alpha}_{ij}^{(2)} + \widehat{\alpha}_{ij}^{(3)} + \widehat{\alpha}_{ij}^{(4)}, \quad (4.4)
\]
where the terms \( \widehat{\alpha}_{ij}^{(m)} \), \( m = 1, \ldots, 4 \), are specified in Zhao, Brazauskas, and Ghorai (2018, lemma A.1).

The asymptotic performance of the newly designed estimators will be measured via asymptotic relative efficiency (ARE) with respect to MLE, and for the two-parameter case it is defined as (see, e.g., Serfling 1980; van der Vaart 1998):
\[
\text{ARE}(C, \text{MLE}) = \left( \frac{\det(\Sigma_{\text{MLE}})}{\det(\Sigma_C)} \right)^{1/2}, \quad (4.5)
\]
where \( \Sigma_{\text{MLE}} \) and \( \Sigma_C \) are the asymptotic variance–covariance matrices of the MLE and \( C \) estimators, respectively, and \( \det \) stands for the determinant of a square matrix. The main reason why MLE should be used as a benchmark procedure is its optimal performance in terms of asymptotic variance (of course, with the usual caveat of “under certain regularity conditions”); for more details we refer to Serfling (1980, section 4.1).

### 4.1. Payments \( Y \)

Due to the piecewise nature of the qf \( F_Y^{-1} \) with the transition point \( s^* = \frac{F_Y(T) - F_Y(t)}{1 - F_Y(t)} \), there are three possible arrangements among \( s^*, a, \) and \( b: \)
have also been demonstrated. Case II: $0 \leq a < 1 - b \leq 1$ (estimation based on observed data only).

Case III: $0 < s^* \leq a < 1 - b \leq 1$ (estimation based on censored data only).

The empirical estimate $\hat{\tau}_n^*$ of $s^*$ is given by

$$
\hat{\tau}_n^* := \frac{F_n(T) - F_n(t)}{1 - F_n(t)} = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}\{0 < y_i < cR\}, \quad \text{where } F_n \text{ is the empirical cdf.} \tag{4.6}
$$

Case II simply implies the estimation based on observed data only, and this is the most reasonable case to be considered, as mentioned by Poudyal (2021a). Thus, in this article, we will proceed with Case II.

Next, let $y_1, \ldots, y_n$ be an i.i.d. sample of normal payment per payment data defined by (2.7) with qf $F_Y^{-1}$. Then for $k = 1, 2$, we have

$$
\hat{W}_{y,k} = \frac{1}{n} \left[ m_n \left(h_Y(y_{n+1})\right)^k + \sum_{i=m_n+1}^{n-m^*_n} \left(h_Y(y_i)\right)^k + m^*_n \left(h_Y(y_{m^*_n})\right)^k \right]
$$

$$
= \frac{1}{n} \left[ m_n \left(\frac{y_{n+1} - a}{c} + t\right)^k + \sum_{i=m_n+1}^{n-m^*_n} \left(\frac{y_i - a}{c} + t\right)^k + m^*_n \left(\frac{y_{m^*_n} - a}{c} + t\right)^k \right], \tag{4.7}
$$

with $m_n/n \to a$ and $m^*_n/n \to b$. With Case II, choose $m^*_n \geq \sum_{i=1}^{n} \mathbb{1}\{y_i = cR\}$. The corresponding population winsorized moments (4.2) with the qf defined by $F_Y^{-1}$ are given by

$$
W_{y,k}(\theta) = a [h_Y(F_Y^{-1}(a \mid \theta))]^k + \int_a^b [h_Y(F_Y^{-1}(s \mid \theta))]^k ds + b [h_Y(F_Y^{-1}(b \mid \theta))]^k
$$

$$
= a(\theta + \Delta_a)^k + \int_a^b \Delta_s^k \, ds + b(\theta + \Delta_b)^k \tag{4.8}
$$

$$
= \begin{cases} 
\theta + \sigma c_{y,1}, & \text{for } k = 1; \\
\theta^2 + 2\theta\sigma c_{y,1} + \sigma^2 c_{y,2}, & \text{for } k = 2,
\end{cases}
$$

where for $k = 1, 2$,

$$
c_{y,k} \equiv c_{y,k}(\Phi, a, b, \gamma) = a\Delta_a^k + \int_{ab} \Delta_s^k \, ds + b\Delta_b^k. \tag{4.9}
$$

It is important to mention here that $c_{y,k}$ depends on the unknown parameters but does not depend on the parameters to be estimated for completely observed sample data (see, e.g., Zhao, Brazauskas, and Ghorai 2018). Equating $W_{y,k} = \hat{W}_{y,k}$, for $k = 1, 2$ yields the implicit system of equations to be solved for $\theta$ and $\sigma$:

$$
\begin{align*}
\theta &= \hat{W}_{y,1} - c_{y,1} \sigma = g_1(\hat{W}_{y,1}, \hat{W}_{y,2}), \\
\sigma &= \sqrt{\left(\hat{W}_{y,2} - \hat{W}_{y,1}^2\right) / \left(c_{y,2} - c_{y,1}^2\right)} = g_2(\hat{W}_{y,1}, \hat{W}_{y,2}). \tag{4.10}
\end{align*}
$$

The system of equations (4.10) can be solved for $\hat{\theta}_{y,MWM}$ and $\hat{\sigma}_{y,MWM}$ by using an iterative numerical method with the initial and
\[ \sigma_{\text{start}} = \sqrt{\hat{W}_{y,2} - \hat{W}_{y,1}^2} \quad \text{and} \quad \theta_{\text{start}} = \hat{W}_{y,1}. \] (4.11)

From Theorem 1, the entries of the variance–covariance matrix \( \Sigma_y \) calculated using (4.4) are

\[
\begin{align*}
\sigma_{11}^2 &= \sigma^2 c_{y,1}^4, \\
\sigma_{12}^2 &= \sigma_{21}^2 = 2\theta \sigma^2 c_{y,1}^3 + 2\sigma^3 c_{y,2}, \\
\sigma_{22}^2 &= 4\theta^2 \sigma^2 c_{y,1}^2 + 8\theta \sigma^2 c_{y,2}^2 + 4\sigma^4 c_{y,3},
\end{align*}
\]

where the expressions for \( c_{y,k}^r, k = 1, 2, 3 \) are listed in the Appendix. For \( k = 1, 2 \), it follows that

\[
\begin{align*}
\frac{\partial c_{y,k}}{\partial \theta} &= -\frac{k\phi(\gamma)}{\sigma} \left\{ \frac{a\hat{a} \Delta_{k-1}^N}{\phi(\Delta_n)} + \int_0^1 \frac{b^2 \Delta_{k-1}^N}{\phi(\Delta_n)} \, ds + \frac{b^2 \Delta_{k-1}^N}{\phi(\Delta_n)} \right\}, \\
\frac{\partial c_{y,k}}{\partial \sigma} &= -\frac{k(t - \theta)\phi(\gamma)}{\sigma^2} \left\{ \frac{a\hat{a} \Delta_{k-1}^N}{\phi(\Delta_n)} + \int_0^1 \frac{b^2 \Delta_{k-1}^N}{\phi(\Delta_n)} \, ds + \frac{b^2 \Delta_{k-1}^N}{\phi(\Delta_n)} \right\}.
\end{align*}
\] (4.12)

For \( k = 1, 2 \), let us denote

\[
\begin{align*}
\theta_{W,1} := \left. \frac{\partial g_1}{\partial W_{y,1}} \right|_{(W_{y,1}, W_{y,2})} \\
\sigma_{W,1} := \left. \frac{\partial g_2}{\partial W_{y,1}} \right|_{(W_{y,1}, W_{y,2})}
\end{align*}
\]

Consider the following additional notations used by Poudyal (2021a) but for different \( c_{y,k} \) functions:

\[
\begin{align*}
f_{11}(\theta, \sigma) &:= 1 + \sigma \frac{\partial c_{y,1}}{\partial \theta}, \\
f_{12}(\theta, \sigma) &:= c_{y,1} + \sigma \frac{\partial c_{y,1}}{\partial \sigma}, \\
f_{21}(\theta, \sigma) &:= c_{y,1} + \sigma \frac{\partial c_{y,1}}{\partial \sigma}, \\
f_{22}(\theta, \sigma) &:= \frac{\partial c_{y,2}}{\partial \sigma} - 2 c_{y,1} \frac{\partial c_{y,1}}{\partial \sigma}.
\end{align*}
\] (4.13)

By the multivariate chain rule and for \( j, k \geq 1 \), we have

\[
\frac{\partial c_{y,k}}{\partial W_{y,j}} = \frac{\partial c_{y,k}}{\partial \theta} \frac{\partial \theta}{\partial W_{y,j}} + \frac{\partial c_{y,k}}{\partial \sigma} \frac{\partial \sigma}{\partial W_{y,j}} = \frac{\partial c_{y,k}}{\partial \theta} \theta_{W,j} + \frac{\partial c_{y,k}}{\partial \sigma} \sigma_{W,j}. \tag{4.14}
\]

Finally, the entries of the Jacobian matrix \( D_y \), given by Theorem 1, are found by implicitly differentiating the functions \( g_j \) from Equations (4.10) with the help of Equation (4.12), and using Equations (4.13) and (4.14), we have

\[
\begin{align*}
d_{11} &= \theta_{W,1} = \frac{1 - f_{12} \sigma_{W,1}}{f_{11}} = \frac{1 - f_{12} d_{21}}{f_{11}}, \\
d_{12} &= \theta_{W,2} = -\frac{f_{12} \sigma_{W,2}}{f_{11}} = -\frac{f_{12} d_{22}}{f_{11}}, \\
d_{21} &= \sigma_{W,1} = -\frac{K \left[ 2 f_{11} W_{y,1} (c_{y,2} - c_{y,1}^2) + f_{21} (W_{y,2} - W_{y,1}^2) \right]}{f_{11} (c_{y,2} - c_{y,1}^2)^2 + K (W_{y,2} - W_{y,1}^2) (f_{11} f_{22} - f_{12} f_{21})}, \\
d_{22} &= \sigma_{W,2} = \frac{K f_{11} (c_{y,2} - c_{y,1}^2)}{f_{11} (c_{y,2} - c_{y,1}^2)^2 + K (W_{y,2} - W_{y,1}^2) (f_{11} f_{22} - f_{12} f_{21})},
\end{align*}
\]

where \( K := \frac{1}{2} \sqrt{\frac{c_{y,2} - c_{y,1}^2}{W_{y,2} - W_{y,1}}} \). Consequently,

\[
S_{y,WM} := D_y \Sigma_y D_y' = \begin{bmatrix} d_{11} & d_{12} & \sigma_{11}^2 & \sigma_{12}^2 \\
\sigma_{12}^2 & d_{22} & \sigma_{22}^2 \end{bmatrix} \begin{bmatrix} d_{11} & d_{21} \\
\sigma_{11}^2 & d_{22} \end{bmatrix} = \begin{bmatrix} d_{11} & d_{21} \\
\sigma_{12}^2 & \sigma_{22}^2 \end{bmatrix} \begin{bmatrix} d_{11} & d_{21} \\
\sigma_{12}^2 & \sigma_{22}^2 \end{bmatrix}.
\] (4.15)
From (2.8), it follows that payment \( Z \) is a left- and right-censored version of random variable \( X \). Thus, possible permutations between \( a, b, \) and their positioning with respect to \( F(t) \) and \( F(T) \) have to be taken into account because the expressions for \( \sigma_{\epsilon}^2 \) given by (4.4) with \( q_f F_Z^{-1} \) depend on the six possible permutations among \( a, b, F(t), \) and \( F(T) \):

1. \( 0 < a < b < F(t) < F(T) \leq 1. \)
2. \( 0 < a < F(t) < b \leq F(T) \leq 1. \)
3. \( 0 < a < F(t) < F(T) \leq b \leq 1. \)
4. \( 0 \leq F(t) < F(T) \leq a < b \leq 1. \)
5. \( 0 \leq F(t) < a < F(T) \leq b \leq 1. \)
6. \( 0 \leq F(t) < a < b \leq F(T) \leq 1. \)

### 4.2. Payments \( Z \)

From (2.8), it follows that payment \( Z \) is a left- and right-censored version of random variable \( X \). Thus, possible permutations between \( a, b, \) and their positioning with respect to \( F(t) \) and \( F(T) \) have to be taken into account because the expressions for \( \sigma_{\epsilon}^2 \) given by (4.4) with \( q_f F_Z^{-1} \) depend on the six possible permutations among \( a, b, F(t), \) and \( F(T) \):

\[
\begin{align*}
1. & \quad 0 < a < b < F(t) < F(T) \leq 1. \\
2. & \quad 0 < a < F(t) < b \leq F(T) \leq 1. \\
3. & \quad 0 < a < F(t) < F(T) \leq b \leq 1. \\
4. & \quad 0 \leq F(t) < F(T) \leq a < b \leq 1. \\
5. & \quad 0 \leq F(t) < a < F(T) \leq b \leq 1. \\
6. & \quad 0 \leq F(t) < a < b \leq F(T) \leq 1.
\end{align*}
\]
Among the six cases, two of those scenarios (estimation based on censored data only, Cases 1 and 4) have no parameters to be estimated in the formulas of population winsorized moments and three (estimation based on observed and censored data, Cases 2, 3, and 5) are inferior to the estimation scenario based on fully observed data. Thus, from now on we will proceed only with Case 6, which makes the most sense and simplifies the estimation procedure significantly because it uses the available data in the most effective way. Moreover, the MWM estimators based on Case 6 will be resistant to outliers; that is, observations that are inconsistent with the assumed model and most likely appearing at the boundaries \( t \) and \( T \). Case 6 also eliminates heavier point masses given at the censored points \( t \) and \( T \).

For practical data analysis purposes, standard empirical estimates of \( F(t) \) and \( F(T) \) provide guidance about the choice of \( a \) and \( b \) and are chosen according to

\[
F_n(t) \leq a < b \leq F_n(T), \quad \text{where} \quad F_n \text{ is the empirical cdf.} \tag{4.18}
\]

Further, define \( h_z(z) := z/c + t \). Now, consider an observed i.i.d. sample \( z_1, \ldots, z_n \) defined by \( F_Z^{-1} \). Let \( z_{1:n}, \ldots, z_{n:n} \) be the corresponding order statistics. Then the sample winsorized moments for \( k = 1, 2 \), are given by (4.1),

\[
\hat{W}_{z,k} = \frac{1}{n} \left[ m_n \left[ h_z(z_{m_n+1:n}) \right]^k + \sum_{i=m_n+1}^{n-m_n} \left[ h_z(z_{i:n}) \right]^k + m_n^* \left[ h_z(z_{n-m_n:n}) \right]^k \right]
\]

\[
= \frac{1}{n} \left[ m_n \left( \frac{z_{m_n+1:n}}{c} + t \right)^k + \sum_{i=m_n+1}^{n-m_n} \left( \frac{z_{i:n}}{c} + t \right)^k + m_n^* \left( \frac{z_{n-m_n:n}}{c} + t \right)^k \right],
\tag{4.19}
\]

with \( m_n/n \to a \) and \( m_n^*/n \to b \). With Case 6, choose \( m_n \geq \sum_{i=1}^{n} 1 \{ z_i = 0 \} \) and \( m_n^* \geq \sum_{i=1}^{n} 1 \{ z_i = cR \} \). By assuming the most general case that \( 0 \leq F(t) \leq a < b \leq F(T) \leq 1 \), the corresponding population winsorized moments (4.2) with the qf \( F_Z^{-1} \) are given by

\[
W_{z,k}(\theta) = a \left[ h_z(F_Z^{-1}(a \mid \theta)) \right]^k + \int_a^b \left[ h_z(F_Z^{-1}(s \mid \theta)) \right]^k \, ds + b \left[ h_z(F_Z^{-1}(b \mid \theta)) \right]^k
\]

\[
= a \left[ F^{-1}(a \mid \theta) \right]^k + \int_a^b \left[ F^{-1}(s \mid \theta) \right]^k \, ds + b \left[ F^{-1}(b \mid \theta) \right]^k
\]

\[
= \begin{cases} 
\theta + \sigma c_1, & \text{for } k = 1; \\
\theta^2 + 2\theta \sigma c_1 + \sigma^2 c_2, & \text{for } k = 2,
\end{cases}
\tag{4.20}
\]

where \( c_k \equiv c_{y,k}, 1 \leq k \leq 4 \), given by (4.9) with \( \gamma = -\infty \) and are listed in the Appendix. Thus, with the assumption \( 0 \leq F(t) \leq a < b \leq F(T) \leq 1 \), this case translates to the complete data case that was fully investigated by Zhao, Brazauskas, and Ghorai (2018), and the MWM estimators of \( \theta \) and \( \sigma \) are

\[
\begin{align*}
\hat{\theta}_{Z,\text{MWM}} &= \frac{\hat{W}_{z,1} - c_1 \hat{\sigma}_{Z,\text{MWM}}}{\sqrt{\hat{W}_{z,2} - \hat{W}_{z,1}^2/(c_2 - c_1^2)}}, \\
\hat{\sigma}_{Z,\text{MWM}} &= \sqrt{\hat{W}_{z,2} - \hat{W}_{z,1}^2/(c_2 - c_1^2)}.
\end{align*}
\tag{4.21}
\]

The corresponding ARE is given by

\[
\text{ARE} \left( \left( \hat{\theta}_{Z,\text{MWM}}, \hat{\sigma}_{Z,\text{MWM}} \right), \left( \hat{\theta}_{Z,\text{MLE}}, \hat{\sigma}_{Z,\text{MLE}} \right) \right) = \left( \text{det}(S_{Z,\text{MLE}})/\text{det}(S_{Z,\text{MWM}}) \right)^{0.5},
\tag{4.22}
\]

where

\[
S_{Z,\text{MWM}} := \frac{\sigma^2}{(c_2 - c_1^2)} \begin{bmatrix}
c_1^2 c_2^2 - 2c_1 c_2 c_3^* + c_3^2 c_3^* & -c_1^* c_1 c_2 + c_2 c_3^* + c_3^2 c_3^* - c_1 c_3^* \\
-c_1^* c_1 c_2 + c_2 c_3^* + c_3^2 c_3^* - c_1 c_3^* & c_1^2 - 2c_1 c_2 + c_3^* 
\end{bmatrix}.
\tag{4.23}
\]
where the expressions for \(c_i^k, k = 1, 2, 3\), as functions of \(a, b, c_1, c_2, c_3\), and \(c_4\) are such that \(c_i^k = c_i^*\) with \(\gamma = -\infty\) and are listed in the Appendix.

Similar to Table 1, several important conclusions can be made based on Table 2.

- Similar to payments \(Y\) (i.e., Table 1), MWM estimators consistently outperform MTM estimators in terms of ARE. The choice of right-censoring point (the upper limit of complete data) significantly affects the level of model efficiency.
- Unlike the payment \(Y\) scheme, the degree of efficiency loss (with complete data) is almost identical from the MWM approach to the MTM method. This can be explained by the fact that payment \(Z\) is an unconditional variable, and the density function used for both MTM and MWM procedures is not restricted by the conditional status of policy deductible and limit. Thus, on such a payment per loss basis, the ARE results only depend on the fraction of \(a\) and \(b\) in this lognormal model.

Clearly, the ARE expressions given by Equations (4.17) and (4.22) are functions of \(d, u\), the parameters to be estimated, and the trimming and winsorizing proportions. It is important to note here that those AREs for completely observed ground-up loss dataset do not depend on the parameters to be estimated; see, for example, Brazauskas, Jones, and Zitikis (2009) and Zhao, Brazauskas, and Ghorai (2018). As functions of policy deductible, \(d\), those AREs are visualized in Figure 1 for some selected

### Table 2

| \(a\) | \(b (u = 5.96 \times 10^3)\) | \(b (u = 1.54 \times 10^3)\) | \(b (u = 7.52 \times 10^2)\) |
|------|-----------------|-----------------|-----------------|
|      | \(0.01\) | \(0.05\) | \(0.10\) | \(0.15\) | \(0.05\) | \(0.10\) | \(0.15\) | \(0.10\) | \(0.15\) |
| MWM  | 0.10  | 0.954 | 0.927 | 0.891 | 0.854 | 0.961 | 0.923 | 0.884 | 0.965 | 0.924 |
|      | 0.15  | 0.896 | 0.867 | 0.830 | 0.791 | 0.899 | 0.860 | 0.819 | 0.899 | 0.857 |
|      | 0.25  | 0.795 | 0.765 | 0.726 | 0.686 | 0.793 | 0.752 | 0.710 | 0.786 | 0.743 |
| MTM  | 0.10  | 0.909 | 0.863 | 0.810 | 0.761 | 0.894 | 0.839 | 0.788 | 0.878 | 0.824 |
|      | 0.15  | 0.855 | 0.812 | 0.761 | 0.712 | 0.841 | 0.788 | 0.738 | 0.824 | 0.771 |
|      | 0.25  | 0.754 | 0.714 | 0.667 | 0.621 | 0.740 | 0.690 | 0.643 | 0.722 | 0.672 |

**FIGURE 1.** ARE\((\hat{\theta}_{\gamma,\text{MWM}}, \hat{\sigma}_{\gamma,\text{MWM}}), (\hat{\theta}_{\gamma,\text{MLE}}, \hat{\sigma}_{\gamma,\text{MLE}})\) Curves as Functions of Policy Deductible \(d\) and with Fixed Policy Limit \(u = 5.96 \times 10^3\). Where (a) \(\nu = Y\) and (b) \(\nu = Z\), Respectively, Represent the Payments \(Y\) and \(Z\) Data Scenarios from an LN\((1, 4, 2)\) Distribution.
winsorizing proportions satisfying Case III as specified in Section 4.1 for payment Y and Case 6 as considered in Section 4.2 for payment Z.

For payment Y (Figure 1, left panel), the ARE curves for the left winsorizing proportions $a \in \{0.00, 0.05\}$ with fixed right winsorizing proportion $b = 0.10$ do not show a significant difference, but when the left winsorizing proportion is increasing, the ARE values are going down. Because the ARE curves do not show a consistent pattern (either increasing or decreasing) as functions of $d$, this makes the practical interpretation of those curves more difficult.

On the other hand, for payment Z, the ARE curves are consistently increasing functions of $d$ (Figure 1, right panel). This is because the higher values of $d$ imply the narrower support for exact loss observations (which are bigger than $d$ and smaller than $u$), which eventually yields the smaller variance of MWM estimators.

5. SIMULATION STUDY

In this section, we use simulations to complement the theoretical results of Section 4. The main goal is to determine the sample size for which the estimators become approximately unbiased (given that they are asymptotically unbiased), justify the asymptotic normality, and assess when the estimators’ finite-sample relative efficiencies (REs) reach the corresponding AREs.

To compute RE of MWM estimators we use MLE as a benchmark. Thus, the definition of ARE given by Equation (4.5) for finite sample performance translates to

\[
RE(\text{MWM, MLE}) = \frac{\text{asymptotic variance of MLE estimator}}{\text{small - sample variance of a competing MWM estimator}}.
\]

where the numerator is as defined in (4.5) and the denominator is given by

\[
\left( \det \begin{bmatrix} \mathbb{E}\left( (\hat{\theta} - \theta)^2 \right) & \mathbb{E}\left( (\hat{\theta} - \theta)(\hat{\sigma} - \sigma) \right) \\ \mathbb{E}\left( (\hat{\theta} - \theta)(\hat{\sigma} - \sigma) \right) & \mathbb{E}\left( (\hat{\sigma} - \sigma)^2 \right) \end{bmatrix} \right)^{1/2}.
\]

5.1. Study Design

From a ground-up lognormal distribution $F_W(w_0 = 1, \theta = 4, \sigma = 2)$, we generate 10,000 samples of a specified length $n$ using Monte Carlo simulations. For each sample we estimate the parameters of $F$ using various MTM and MWM estimators and then compute the average mean and RE of those 10,000 estimates. The standardized ratio $\hat{\theta}/\theta$ that we report is defined as the average of 10,000 estimates divided by the true value of the parameter that we are estimating. The standard error is standardized in a similar manner.

We observe the performance of different estimators for the parameters of the specified lognormal distribution. The simulation study is designed using the following parameters:

(i) Sample size: $n = 100, 500, \infty$.
(ii) Coinsurance rate: $c = 1$.
(iii) Truncation and censoring thresholds for both variables Y and Z:
(iv) $d = 3$ (corresponding to about 5% left truncation);
(v) $u_1 = 5.96 \times 10^7$ (corresponding to about 1% right-censoring);
(vi) $u_2 = 1.54 \times 10^7$ (corresponding to about 5% right-censoring).
(vii) Selection of trimming and winsorizing proportions:
(viii) As mentioned in Section 4.1, for payments Y, the trimmed and winsorized estimators are derived under the condition that $0 \leq a < 1 - b \leq s^*$ (i.e., Case II). In simulations, however, the proportion $s^*$ is random, which can easily result in violations of the specified condition. If $s^*$ is estimated empirically—that is, by replacing $F$ with its empirical estimator $F_n$—then the variance of such an estimator is equal to $\sigma^2_n = \frac{s^*(1-s^*)}{n}$. Thus, to minimize the number of possible violations, the right trimming proportion $b$ is chosen to satisfy the following inequality:
the confidence level or risk appetite. The specific expressions of these measures are given by

\[ \text{VaR}_p[F_W] = w_0 + \frac{1}{p} \left[ \Phi^{-1}(p) \right], \]

\[ \text{TVaR}_p[F_W] = w_0 + \frac{1}{p} \left[ \Phi\left( \frac{\ln \left( \frac{w_0}{\sigma} \right) - \theta}{\sigma} \right) \right], \]

and

\[ \text{PHDRM}_p[F_W] = w_0 + \int_{w_0}^{\infty} \left[ 1 - \Phi\left( \frac{\ln \left( \frac{w_0}{\sigma} \right) - \theta}{\sigma} \right) \right]^p \, dw. \]

and, to demonstrate the difference between sufficiently and insufficiently robust trimmed and winsorized estimators, some values of \( b \) are chosen to violate the condition \( 1 - b \leq s^* - 2\sigma^* \).

- Again, as mentioned in Section 4.2 and for payments \( Z \), the equivalent condition is \( F(d) \leq a < 1 - b \leq F(u) \). Similar arguments as those for payments \( Y \) lead to \( \sigma^2 F(d) = \frac{F(d)[1-F(d)]}{n} \) and \( \sigma^2 = \frac{F(u)[1-F(u)]}{n} \). In this case, the left and right trimming/winsorizing proportions \( a \) and \( b \), respectively, are chosen such that

\[ F(d) + 2\sigma^*_{F,d(a)} \leq a \quad \text{and} \quad 1 - b \leq F(u) - 2\sigma^*_{F,u(b)}. \]

- Estimators of \( \theta \) and \( \sigma \): MLE, MTM, and MWM estimators with the trimming and winsorizing proportions as specified in Table 3.

### 5.2. Convergence and Relative Bias

Simulation results of both MTM and MWM are recorded in Table 4 (for payment per payment variable \( Y \)) and Table 5 (for payment per payment variable \( Z \)). The entries of the last column \( (n \to \infty) \) represent analytic results that were established in Section 4. This helps us compare large- and small-sample properties when the deductible and policy limit are present.

The relative bias is defined as the ratio of the expectation of parameter estimate to its true value, thus making the value of one the target. As is seen in Table 4, all MWM and MTM estimators successfully estimate both the log-location parameter \( \mu \) and log-scale parameter \( \sigma \). Indeed, they practically become unbiased for samples of size \( n = 500 \). The situation is similar in Table 5. Furthermore, we notice that the simulated REs of these estimators for \( n = 500 \) are almost identical to the corresponding AREs for large policy limit \( u = 5.96 \times 10^3 \). When \( u \) moves to \( 1.54 \times 10^3 \), however, the REs of MWM in finite-size samples are far from corresponding ARE for \( b = 0.05 \) because the condition (5.2) is violated, which leads to poor parameter estimation.

### 5.3. Robustness Studies

In this section, we investigate robustness properties of all of the estimators presented in this article. The study builds upon the ideas of Poudyal and Brazauskas (2022a), where a similar sensitivity analysis study was performed for a single-parameter Pareto model.

Using the ground-up variable \( W \sim LN(w_0, \theta, \sigma) \) (see Section 2), we consider estimation of three risk measures: Value-at-Risk \( \text{VaR}_p \), tail Value-at-Risk \( \text{TVaR}_p \), and proportional hazard distortion risk measure \( \text{PHDRM}_p \), where \( p \) represents the confidence level or risk appetite. The specific expressions of these measures are given by

\[ \text{VaR}_p[F_W] = w_0 + e^{\theta + \sigma \Phi^{-1}(p)}, \]

\[ \text{TVaR}_p[F_W] = w_0 + \frac{1}{p} \left[ \Phi\left( \sigma - \Phi^{-1}(p) \right) \right], \]

and

\[ \text{PHDRM}_p[F_W] = w_0 + \int_{w_0}^{\infty} \left[ 1 - \Phi\left( \frac{\ln \left( \frac{w_0}{\sigma} \right) - \theta}{\sigma} \right) \right]^p \, dw. \]
For both payments $Y$ and $Z$, $m = 10,000$ samples of $n = 100$ data points each are generated from the following contamination model:

$$F_c(x) = (1 - \epsilon)F_0(x) + \epsilon F_c(x),$$

(5.3)

where $\epsilon$ represents the probability that a sample observation comes from the contaminating distribution $F_c$ instead of the central (assumed) model $F_0$. That is, $\epsilon = 0$ simply corresponds to $F_0$, which is our assumed model $LN(w_0 = 1, \theta = 4, \sigma = 2)$. The main objective is to see how risk measure estimates (under $F_0$) behave when the actual data slightly deviate from the assumed model.

### TABLE 4

Lognormal Payment per Payment Actuarial Loss Scenario, $LN(w_0 = 1, \theta = 4, \sigma = 2)$ with $d = 3$ and Two Selected Values of Right-Censoring Point $u$.

| Proportion | $n = 100$ | $n = 500$ | $n \to \infty$ |
|------------|-----------|-----------|----------------|
| $a$        | $b$       | MWM       | MTM           | MWM           | MTM           | MWM           | MTM           |
| $\hat{\theta}/\theta$ | $\hat{\alpha}/\alpha$ | $\hat{\theta}/\theta$ | $\hat{\alpha}/\alpha$ | $\hat{\theta}/\theta$ | $\hat{\alpha}/\alpha$ | $\hat{\theta}/\theta$ | $\hat{\alpha}/\alpha$ |

**Mean values of $\hat{\theta}/\theta$ and $\hat{\alpha}/\alpha$**

- **MLE**
  - $1.00$ $1.00$ $1.00$ $1.00$ $1.00$ $1.00$ $1.00$ $1.00$ $1.00$ $1.00$ $1.00$ $1.00$ $1.00$
  - $0.00$ $0.05$ $1.00$ $1.00$ $1.00$ $1.00$ $1.00$ $1.00$ $1.00$ $1.00$ $1.00$ $1.00$ $1.00$
  - $0.00$ $0.10$ $1.00$ $1.00$ $1.00$ $1.01$ $1.00$ $1.00$ $1.00$ $1.00$ $1.00$ $1.00$ $1.00$
  - $0.00$ $0.15$ $1.00$ $1.00$ $0.99$ $1.01$ $1.00$ $1.00$ $1.00$ $1.00$ $1.00$ $1.00$ $1.00$
  - $0.05$ $0.05$ $1.00$ $0.99$ $0.99$ $1.01$ $1.00$ $1.00$ $1.00$ $1.00$ $1.00$ $1.00$ $1.00$
  - $0.10$ $0.10$ $1.00$ $0.99$ $0.99$ $1.01$ $1.00$ $1.00$ $1.00$ $1.00$ $1.00$ $1.00$ $1.00$
  - $0.15$ $0.15$ $1.00$ $0.99$ $0.99$ $1.01$ $1.00$ $1.00$ $1.00$ $1.00$ $1.00$ $1.00$ $1.00$

**Finite-sample efficiencies (RE) of MWMs and MTMs relative to MLEs**

- **MLE**
  - $0.95$ $0.96$ $0.98$ $0.98$ $0.98$ $1.00$ $1.00$
  - $0.00$ $0.05$ $1.11$ $0.97$ $1.11$ $0.97$ $1.00$
  - $0.00$ $0.10$ $0.91$ $0.84$ $0.92$ $0.86$ $0.93$
  - $0.00$ $0.15$ $0.83$ $0.73$ $0.87$ $0.79$ $0.87$
  - $0.05$ $0.05$ $1.10$ $0.97$ $1.10$ $0.96$ $0.99$
  - $0.10$ $0.10$ $0.89$ $0.81$ $0.90$ $0.84$ $0.91$
  - $0.15$ $0.15$ $0.80$ $0.71$ $0.82$ $0.75$ $0.83$

**Note:** The entries are mean values based on 10,000 samples. The standard errors for the entire entries in this table are reported to be $\leq 0.0015$. For both payments $Y$ and $Z$, $m = 10,000$ samples of $n = 100$ data points each are generated from the following contamination model:
TABLE 5
Lognormal Payment per Loss Actuarial Loss Scenario, $LN(\omega_0 = 1, \theta = 4, \sigma = 2)$ with $d = 3$ and Two Selected Values of Right-Censoring Point $u$.

\begin{tabular}{cccccccccc}
\hline
Proportion & \multicolumn{2}{c}{n = 100} & \multicolumn{2}{c}{n = 500} & \multicolumn{2}{c}{n $\to \infty$} \\
\hline
\hline
\multirow{2}{*}{$u = 5.96 \times 10^3$} & $a$ & $b$ & MWM & MTM & MWM & MTM & MWM & MTM & MWM & MTM \\
\hline
MLE & 1.00 & 0.99 & 1.00 & 0.99 & 1.00 & 1.00 & 1.00 & 1.00 & 1.00 & 1.00 \\
0.10 & 0.05 & 1.00 & 0.99 & 1.00 & 1.00 & 1.00 & 1.00 & 1.00 & 1.00 & 1.00 \\
0.15 & 0.10 & 1.00 & 0.99 & 1.00 & 1.00 & 1.00 & 1.00 & 1.00 & 1.00 & 1.00 \\
0.10 & 0.15 & 1.00 & 0.99 & 1.00 & 1.00 & 1.00 & 1.00 & 1.00 & 1.00 & 1.00 \\
0.15 & 0.10 & 1.00 & 0.99 & 1.00 & 1.00 & 1.00 & 1.00 & 1.00 & 1.00 & 1.00 \\
0.10 & 0.15 & 1.00 & 0.99 & 1.00 & 1.00 & 1.00 & 1.00 & 1.00 & 1.00 & 1.00 \\
0.25 & 0.25 & 1.00 & 0.98 & 1.00 & 1.01 & 1.00 & 1.00 & 1.00 & 1.00 & 1.00 \\
\hline
\end{tabular}

Mean values of $\hat{\theta}/\hat{\sigma}$ and $\hat{\sigma}/\hat{\sigma}$

\begin{tabular}{cccccccccc}
\hline
\multirow{2}{*}{$u = 1.54 \times 10^3$} & $a$ & $b$ & MWM & MTM & MWM & MTM & MWM & MTM & MWM & MTM \\
\hline
MLE & 0.99 & 1.00 & 1.00 & 1.00 & 1.00 & 1.00 & 1.00 & 1.00 & 1.00 & 1.00 \\
0.10 & 0.05 & 0.92 & 0.87 & 0.93 & 0.87 & 0.927 & 0.863 & 0.927 & 0.863 \\
0.15 & 0.10 & 0.82 & 0.79 & 0.84 & 0.77 & 0.830 & 0.761 & 0.830 & 0.761 \\
0.10 & 0.15 & 0.83 & 0.76 & 0.84 & 0.77 & 0.854 & 0.761 & 0.854 & 0.761 \\
0.10 & 0.10 & 0.87 & 0.81 & 0.88 & 0.82 & 0.891 & 0.810 & 0.891 & 0.810 \\
0.15 & 0.15 & 0.78 & 0.71 & 0.80 & 0.72 & 0.791 & 0.712 & 0.791 & 0.712 \\
0.25 & 0.25 & 0.60 & 0.53 & 0.61 & 0.54 & 0.603 & 0.534 & 0.603 & 0.534 \\
\hline
\end{tabular}

Finite-sample efficiencies (RE) of MWMs and MTMs relative to MLEs

Note: The entries are mean values based on 10,000 samples. The standard errors for the entire entries in this table are reported to be $\leq 0.0013$.

The loss control parameters, policy deductible $d$ and limit $u$, are chosen as follows:

\[ d = 5 \text{(equivalently, } F_0(d) \approx 0.10) \text{ and } u = 2000 \text{(equivalently, } F_0(u) \approx 0.04). \]

For the contaminating distribution $F_\alpha$, we choose a single-parameter Pareto with the shape parameter $\alpha = 0.5$ and the scale parameter $x_0 = 1000$. This model has a much higher likelihood than $F_\omega$ to generate extremes in the upper (right) tail. Further, the probability of contamination is set to be $\epsilon = (\epsilon_1, \epsilon_2, \epsilon_3) = (0.00, 0.03, 0.06)$. Finally, the left- and right-trimming/winsorizing proportion vectors are chosen as follows:
For payment per payment variable $Y$:
- Condition required: \( 0 \leq a < 1 - b \leq s^* = 0.96 \leq 1 \).
- Estimators: MLE, \( T_i \), and \( W_i \) for \( i = 1, 2, 3 \) as specified in Table 3.

For payment per loss variable $Z$:
- Condition required: \( 0 \leq F(d) = 0.10 < a < \bar{b} \leq F(u) = 0.96 \leq 1 \).
- Estimators: MLE, \( T_i \), and \( W_i \) for \( i = 1, 2, 3 \) as specified in Table 3.

The simulated risk measure estimates are summarized using box plots. The resulting box plots for payment $Y$ and $Z$ variables are presented in Figures 2 and 3, respectively.

Note 3 (Discussion of Figures 2 and 3 and Tables 6 and 7). We first notice a very strong performance (i.e., shorter boxplots) of $T$- and $W$-estimators compared to MLE as the contamination proportion $\epsilon$ increases, especially in the second and third columns of Figures 2 and 3. This fact can also be seen in Tables 6 and 7, where the mean square errors of $T$- and $W$-estimated risk measures are smaller than those of the corresponding MLEs (when $\epsilon > 0$). Further, the estimators with $b \geq \epsilon$ perform better than those with $b < \epsilon$. To see this, let us focus on the “box” part of the box plots (i.e., the middle 50\% of estimates) and consider the case of $\epsilon = 0.06$. We clearly see that the estimates based on $W_3$ are closer to the target value (the horizontal green line, $y = 0$) than those based on $W_1$ or $W_2$. Similar conclusions can be drawn among $T_1$, $T_2$, and $T_3$.

Further, focusing on specific numbers, we see that the bias and variability of estimates get more inflated for nonrobust estimates (MLEs) than those for the robust ones. For example, as reported in Table 6, when $\epsilon$ changes from 0 to 0.06, the...
percentage change of $Q_2$ in TVaR is 211% for MLE, whereas the corresponding change for $W_3$ is 118%. The variability metrics change even more: $\sqrt{\text{MSE}}$ in TVaR change 3.613% for MLE and 7.83% for $W_3$. Or, if we consider $Q_3 - Q_1$ in TVaR, then it changes 42.8% for MLE and 25.3% for $W_3$.

Finally, the estimates for payment per loss variable $(Z)$ exhibit similar patterns to those for payment per payment variable $(Y)$. However, there is one crucial difference here: in general, the box plots presented in Figure 3 are shorter than those in Figure 2. This is because variable $Z$ is left-censored, whereas $Y$ is left-truncated.

6. REAL DATA ILLUSTRATIONS

In this section, we apply the MWM, MTM, and MLE estimators to analyze 1500 indemnity losses in the United States provided by Insurance Service Office, Inc., which has been widely studied in the actuarial literature (see, e.g., Frees and Valdez 1998; Beirlant et al. 2004; Michael, Miljkovic, and Melnykov 2020). The dataset is available in the R package copula with the data name loss (Kojadinovic and Yan 2010). Our goal is to investigate what effect initial assumptions and parameter estimation methods have on model fit and corresponding insurance contract pricing.

A preliminary diagnostics, which we based on the histogram and normal Q-Q-plot of the log-transformed losses, shows that a lognormal distribution may provide an adequate fit to indemnity losses (see Figure 4). Thus, the effectiveness of the proposed methodology will be illustrated using this dataset. The impact of trimming/winsorizing proportions on the lognormal model performance will be examined as well.
| $\varepsilon$ | Risk Measure | Estimators |
|-------------|--------------|------------|
|             | MLE $T_1$    | $T_2$ | $T_3$ | $W_1$ | $W_2$ | $W_3$ |
| 0.00 VaR    | 716          | 720    | 727   | 731   | 701   | 711   | 712   |
| (709)       |              |        |       |       |       |       |       |
| Q1          | 575          | 577    | 571   | 565   | 570   | 565   | 560   |
| Q2          | 690          | 694    | 695   | 696   | 682   | 681   | 681   |
| Q3          | 829          | 836    | 846   | 859   | 814   | 828   | 829   |
| $\sqrt{MSE}$ | 197          | 198    | 216   | 233   | 180   | 203   | 212   |
| TVaR Mean   | 3,443        | 3,493  | 3,700 | 3,909 | 3,206 | 3,438 | 3,531 |
| (3,082)     |              |        |       |       |       |       |       |
| Q1          | 2,242        | 2,258  | 2,196 | 2,130 | 2,196 | 2,145 | 2,114 |
| Q2          | 2,989        | 3,071  | 3,095 | 3,114 | 2,943 | 2,942 | 2,962 |
| Q3          | 4,119        | 4,262  | 4,448 | 4,626 | 4,189 | 4,428 | 4,248 |
| $\sqrt{MSE}$ | 1,928        | 1,851  | 2,16  | 2,38  | 1,390 | 1,967 | 2,385 |
| PHDRM Mean  | 644          | 653    | 694   | 738   | 600   | 644   | 663   |
| (574)       |              |        |       |       |       |       |       |
| Q1          | 428          | 431    | 421   | 409   | 421   | 412   | 407   |
| Q2          | 559          | 573    | 578   | 580   | 550   | 5510  | 5540  |
| Q3          | 759          | 787    | 818   | 850   | 731   | 722   | 781   |
| $\sqrt{MSE}$ | 363          | 341    | 467   | 657   | 250   | 365   | 433   |
| 0.03 VaR    | 958          | 911    | 903   | 884   | 893   | 911   | 887   |
| (709)       |              |        |       |       |       |       |       |
| Q1          | 735          | 707    | 681   | 663   | 712   | 690   | 667   |
| Q2          | 909          | 875    | 854   | 834   | 863   | 866   | 834   |
| Q3          | 1,123        | 1,064  | 1,058 | 1,043 | 1,025 | 1,081 | 1,041 |
| $\sqrt{MSE}$ | 398          | 343    | 358   | 355   | 319   | 359   | 353   |
| TVaR Mean   | 6,821        | 5,529  | 5,749 | 5,720 | 5,167 | 5,717 | 5,597 |
| (3,082)     |              |        |       |       |       |       |       |
| Q1          | 3,568        | 3,257  | 2,974 | 2,766 | 3,299 | 3,055 | 2,825 |
| Q2          | 5,109        | 4,620  | 4,390 | 4,169 | 4,397 | 4,522 | 4,159 |
| Q3          | 7,780        | 6,676  | 6,777 | 6,521 | 7,951 | 7,041 | 6,466 |
| $\sqrt{MSE}$ | 9,464        | 4,356  | 5,647 | 6,693 | 4,896 | 4,286 | 5,260 |
| PHDRM Mean  | 1,314        | 1,033  | 1,088 | 1,095 | 963   | 1,074 | 1,065 |
| (574)       |              |        |       |       |       |       |       |
| Q1          | 662          | 609    | 559   | 522   | 615   | 572   | 532   |
| Q2          | 939          | 850    | 809   | 771   | 811   | 833   | 766   |
| Q3          | 1,437        | 1,227  | 1,243 | 1,195 | 1,090 | 1,293 | 1,186 |
| $\sqrt{MSE}$ | 2,736        | 848    | 1,185 | 1,546 | 768   | 971   | 1,418 |
| 0.06 VaR    | 1,286        | 1,220  | 1,160 | 1,106 | 1,247 | 1,200 | 1,155 |
| (709)       |              |        |       |       |       |       |       |
| Q1          | 947          | 904    | 845   | 797   | 909   | 886   | 822   |
| Q2          | 1,197        | 1,137  | 1,085 | 1,023 | 1,154 | 1,143 | 1,058 |
| Q3          | 1,528        | 1,451  | 1,384 | 1,318 | 1,507 | 1,439 | 1,379 |
| $\sqrt{MSE}$ | 754          | 680    | 629   | 591   | 707   | 662   | 647   |
| TVaR Mean   | 76,384       | 11,594 | 10,367| 9,788 | 12,173| 11,676| 10,958|
| (3,082)     |              |        |       |       |       |       |       |
| Q1          | 5,890        | 5,080  | 4,336 | 3,762 | 5,071 | 4,886 | 4,023 |
| Q2          | 9,303        | 7,885  | 6,878 | 5,954 | 8,613 | 8,004 | 6,454 |
| Q3          | 15,801       | 13,135 | 11,550| 10,259| 15,656| 12,500| 11,554|
| $\sqrt{MSE}$ | 71,587       | 15,477 | 15,455| 19,549| 14,146| 22,670| 20,177|
| PHDRM Mean  | 71,501       | 2,287  | 2,056 | 1,993 | 2,382 | 2,382 | 2,236 |
| (574)       |              |        |       |       |       |       |       |
| Q1          | 1,081        | 933    | 802   | 700   | 931   | 901   | 744   |
| Q2          | 1,712        | 1,450  | 1,262 | 1,092 | 1,583 | 1,469 | 1,186 |
| Q3          | 2,998        | 2,466  | 2,142 | 1,892 | 2,957 | 2,328 | 2,145 |
| $\sqrt{MSE}$ | 35,315       | 3,583  | 3,861 | 5,872 | 3,037 | 6,602 | 5,819 |

Note: $Q_1$, $Q_2$, $Q_3$ denote the first, second, and third quartiles, respectively. The boxed values in the table represent the quantities to be compared along that row and the corresponding metrics for the different contamination proportions given by column $\varepsilon$. 
### TABLE 7
Summary Statistics for Different Estimated Risk Measures; VaR \(0.00[F_0]\), TVaR \(0.00[F_0]\), and PHDRM \(0.00[F_0]\) for Payment per Loss Data Scenario

| \(\epsilon\) | Risk Measure | Estimators |
|--------------|---------------|------------|
|              |               | MLE | \(T_1\) | \(T_2\) | \(T_3\) | \(W_1\) | \(W_2\) | \(W_3\) |
| 0.00 \(\text{(709)}\) | VaR \(Q_1\)  | 730 | 736 | 748 | 747 | 704 | 722 | 711 |
|          | \(Q_2\)      | 580 | 581 | 579 | 569 | 565 | 563 | 551 |
|          | \(Q_3\)      | 698 | 705 | 711 | 708 | 682 | 688 | 678 |
|          | \(\sqrt{MSE}\) | 848 | 861 | 879 | 883 | 824 | 847 | 831 |
|          | TVaR \(Q_1\) | 3,345 | 3,391 | 3,556 | 3,522 | 3,067 | 3,289 | 3,147 |
|          | \(Q_2\)      | 2,276 | 2,282 | 2,255 | 2,200 | 2,161 | 2,143 | 2,070 |
|          | \(Q_3\)      | 2,993 | 3,062 | 3,121 | 3,075 | 2,842 | 2,926 | 2,829 |
|          | \(\sqrt{MSE}\) | 4,027 | 4,170 | 4,357 | 4,322 | 4,051 | 4,051 | 3,851 |
|          | PHDRM \(Q_1\) | 622 | 630 | 660 | 654 | 572 | 612 | 587 |
|          | \(Q_2\)      | 431 | 432 | 428 | 419 | 411 | 409 | 395 |
|          | \(Q_3\)      | 559 | 571 | 580 | 572 | 532 | 546 | 530 |
|          | \(\sqrt{MSE}\) | 743 | 767 | 802 | 796 | 696 | 747 | 711 |
| 0.03 \(\text{(709)}\) | VaR \(Q_1\)  | 1,012 | 966 | 958 | 915 | 729 | 705 | 665 |
|          | \(Q_2\)      | 765 | 735 | 706 | 674 | 729 | 705 | 665 |
|          | \(Q_3\)      | 951 | 920 | 898 | 858 | 894 | 896 | 842 |
|          | \(\sqrt{MSE}\) | 1,189 | 1,141 | 1,137 | 1,086 | 1,076 | 1,140 | 1,065 |
| 0.06 \(\text{(709)}\) | VaR \(Q_1\)  | 1,411 | 1,378 | 1,290 | 1,166 | 1,401 | 1,332 | 1,204 |
|          | \(Q_2\)      | 1,025 | 982 | 902 | 822 | 978 | 939 | 827 |
|          | \(Q_3\)      | 1,303 | 1,262 | 1,180 | 1,066 | 1,291 | 1,248 | 1,087 |
|          | \(\sqrt{MSE}\) | 1,673 | 1,653 | 1,558 | 1,399 | 1,736 | 1,609 | 1,442 |
|          | TVaR \(Q_1\) | 5,647 | 5,217 | 5,306 | 4,808 | 4,748 | 5,120 | 4,519 |
|          | \(Q_2\)      | 3,452 | 3,278 | 3,046 | 2,830 | 3,180 | 2,990 | 2,721 |
|          | \(Q_3\)      | 4,808 | 4,581 | 4,397 | 4,054 | 4,244 | 4,326 | 3,820 |
|          | \(\sqrt{MSE}\) | 6,869 | 6,396 | 6,540 | 5,901 | 5,601 | 6,346 | 5,525 |
|          | PHDRM \(Q_1\) | 1,040 | 961 | 979 | 888 | 876 | 944 | 835 |
|          | \(Q_2\)      | 641 | 610 | 568 | 530 | 593 | 560 | 511 |
|          | \(Q_3\)      | 882 | 842 | 809 | 749 | 783 | 798 | 707 |
|          | \(\sqrt{MSE}\) | 1,255 | 1,170 | 1,196 | 1,080 | 1,026 | 1,160 | 1,012 |

Note: \(Q_1, Q_2, Q_3\) denote the first, second, and third quartiles, respectively.
We consider estimation of the loss severity component of the pure premium for an insurance benefit that equals the amount by which an indemnity loss \( W \) exceeds 500 (deductible, \( d = 500 \)) dollars with a maximum benefit of 100,000 dollars (policy limit, \( u = 105 \)). Without loss of generality, we assume that \( c = 1 \) because the asymptotic variances of all of the estimators investigated in this article do not depend on the coinsurance factor \( c \). The distribution and density functions of the lognormal payment per payment, \( FY \) and \( fY \), are given by (2.9), and those for payment per loss variables, \( FZ \) and \( fZ \), by (2.10). The specific transformations of \( Y \) and \( Z \) and the number of loss observations within each range of contract are summarized in Table 8.

Next, we discuss the robustness properties and fit of the MTM and MWM estimators when policy deductible and policy limit are considered in the risk pricing. Note that in the subsequent analysis the trimming/winsorizing proportions \( a \) and \( b \) are chosen according to the guidelines of Section 5.1 (see item iv). Specifically, we start with the fact that the dataset has 49 losses among 1500 claims below the deductible \( d \) and 152 losses above the policy limit \( u \). Because the quality of the fit can be evaluated only on observed data (i.e., noncensored data), the proportions \( a \) and \( b \) are chosen to produce data trimming/winsorizing inside the range between \( d \) and \( u \). In addition, to see the impact of \( b \) on the policy limit \( u \), we fix \( a = 0 \) and set different values of \( b \) from small (150/1451) to large (750/1451). The point and 95% interval estimates of log-location parameter \( \theta \) and log-scale parameter \( \sigma \) are reported in Tables 9 and 10. In the tables, we also provide the actuarial premiums (LEV or limited expected value; see, e.g., Klugman, Panjer, and Willmot [2019], p. 137):

\[
\text{LEV} = \begin{cases} 
\mathbb{E}[Y] = \frac{\mathbb{E}[W \wedge u] - \mathbb{E}[W \wedge d]}{1 - F_W(d)}, & \text{for payment per payment;} \\
\mathbb{E}[Z] = \mathbb{E}[W \wedge u] - \mathbb{E}[W \wedge d], & \text{for payment per loss,}
\end{cases}
\]

where
TABLE 9
MLE, MTM, and MWM Estimators of $\theta$ and $\sigma$ with Their Corresponding Asymptotic Confidence Intervals and Estimated Values of AREs for Payment Y Data Scenario

| Estimators | $\hat{\theta}$ | $\hat{\sigma}$ | $\tilde{\hat{\sigma}}$ | 95% CI for $\theta$ | $\theta$ | $\sigma$ | 95% CI for $\sigma$ | LEV ($\times 10^4$) | ARE |
|------------|---------------|----------------|-----------------|----------------|--------|--------|----------------|----------------|------|
| MLE        | 9.43          | 1.59           | 0.90 0.90       | (9.34, 9.52)   | (1.51, 1.67) | 2.675  1.00 |
| MTM        | 9.43          | 1.59           | 0.90 0.90       | (9.34, 9.52)   | (1.51, 1.67) | 2.675  1.00 |
| MWM        | 9.43          | 1.59           | 0.90 0.90       | (9.34, 9.52)   | (1.51, 1.67) | 2.675  1.00 |

Note: For standard method of moments (i.e., MTM/MWM with $a = b = 0$), $\hat{\theta} = 9.43$, $\hat{\sigma} = 1.35$, LEV = $2.669 \times 10^4$.

TABLE 10
MLE, MTM, and MWM Estimators of $\theta$ and $\sigma$ with Their Corresponding Asymptotic Confidence Intervals and Estimated Values of AREs for Payment Z Data Scenario

| Estimators | $\hat{\theta}$ | $\hat{\sigma}$ | $\tilde{\hat{\sigma}}$ | $\hat{\sigma}$ | 95% CI for $\theta$ | $\theta$ | $\sigma$ | 95% CI for $\sigma$ | LEV ($\times 10^4$) | ARE |
|------------|---------------|----------------|-----------------|--------|----------------|--------|--------|----------------|----------------|------|
| MLE        | 9.39          | 1.64           | 0.90 0.90       | (9.30, 9.47)   | (1.50, 1.67) | 2.600  1.00 |
| MTM        | 9.39          | 1.64           | 0.90 0.90       | (9.30, 9.47)   | (1.50, 1.67) | 2.600  1.00 |
| MWM        | 9.39          | 1.64           | 0.90 0.90       | (9.30, 9.47)   | (1.50, 1.67) | 2.600  1.00 |

Note: For standard method of moments (i.e., MTM/MWM with $a = b = 0$), $\hat{\theta} = 9.33$, $\hat{\sigma} = 1.45$, LEV = $2.582 \times 10^4$. 

Moreover, the estimated AREs for MWM, MTM, and MLE are presented so their estimation accuracy can be compared. As Tables 9 and 10 suggest, the MWM and MTM estimators with appropriate proportions \(a\) and \(b\) (i.e., \(a = 0, b = 150/1451\) in Table 9 and \(a = 75/1500, b = 150/1500\) in Table 10) lead to parameter and premium estimates that are fairly close to those of nonrobust MLE. This is because the lognormal model fits the given data set well (see Figure 4). If the fits were of lesser quality, which happens when data are contaminated, we would see significant differences among the estimators (this effect was illustrated in Section 5). On the other hand, the MWM and MTM estimates that differ substantially from MLEs are those that are most robust but least efficient (i.e., \(b = 700/1451\) or \(b = 650/1451\)). This suggests that for a particular dataset the actuary should verify model fits based on multiple estimators and aim at striking a balance between method’s robustness and efficiency. Finally, the standard method-of-moments estimators (i.e., MWM/MTM with \(a = b = 0\)) are reported for comparison. They yield estimates that moderately deviate from MLE, mostly for parameter \(\sigma\).

7. CONCLUSION

In this article, we introduced and developed a new estimation procedure—method of winsorized moments (MWM)—for robust fitting of truncated and censored lognormal distributions, which are appropriate for insurance payment per payment and payment per loss data. Large-sample properties of the MWM estimators were established, and their small-sample performance was investigated through simulations and compared to that of the main competitor (method of trimmed moments, MTM) and the benchmark MLE. Among the three methods, MLE is not robust, whereas MTM and MWM are both robust and exhibit similar robustness properties. In terms of efficiency, MLE is most efficient (as expected) and MWM consistently outperforms MTM under the lognormal model with policy deductible and limit. Moreover, as the sample size increases, finite-sample measures of the estimators’ performance approach their asymptotic counterparts.

Further, an extensive robustness study using contaminated models was conducted to assess the performance of MWM, MTM, and MLE when estimating the Value-at-Risk (\(\text{VaR}_{0.99}\)), tail Value-at-Risk (\(\text{VTaR}_{0.99}\)), and proportional hazard distortion risk measure (\(\text{PHDRM}_{0.99}\)). It was found that as the level of contamination increased, MLE-based risk measure estimates drifted further away from the corresponding targets and their variability increased more than for the corresponding characteristics of the robust and efficient estimates.

Finally, the numerical examples based on 1500 U.S. indemnity losses revealed that evaluation of the net premiums using robust and efficient MWM and MTM estimators leads to reasonable results.

ACKNOWLEDGMENTS

The authors are very appreciative of valuable comments and constructive criticism provided by the Editor Yijia Lin and two anonymous referees, which led to many improvements in the article.

FUNDING

The first author acknowledges research startup funding from the Department of Statistics and Data Science, University of Central Florida, Orlando.

ORCID

Chudamani Poudyal http://orcid.org/0000-0003-4528-867X
Vytaaras Brazauskas http://orcid.org/0000-0003-2161-053X

REFERENCES

Beirlant, J., Y. Goegebeur, J. Teugels, and J. Segers. 2004. Statistics of extremes: Theory and applications. Wiley Series in Probability and Statistics. Chichester, UK: John Wiley & Sons.
Blostein, M., and T. Miljkovic. 2019. On modeling left-truncated loss data using mixtures of distributions. Insurance: Mathematics & Economics 85:35–46. doi: 10.1016/j.insmatheco.2018.12.001.
Brazauskas, V., B. L. Jones, and R. Zitikis. 2009. Robust fitting of claim severity distributions and the method of trimmed moments. *Journal of Statistical Planning and Inference* 139 (6):2028–43. doi: 10.1016/j.jspi.2008.09.012.

Brazauskas, V., and A. Kleefeld. 2016. Modeling severity and measuring tail risk of Norwegian fire claims. *North American Actuarial Journal* 20 (1):1–16. doi: 10.1080/10920277.2015.1062784.

Chernoff, H., J. L. Gastwirth, and M. V. Johns, Jr. 1967. Asymptotic distribution of linear combinations of functions of order statistics with applications to estimation. *Annals of Mathematical Statistics* 38 (1):52–72. doi: 10.1214/aoms/1177699058.

Cohen, A. C., Jr. 1950. Estimating the mean and variance of normal populations from singly truncated and doubly truncated samples. *Annals of Mathematical Statistics* 21 (4):557–69. doi: 10.1214/aoms/1177729751.

Cohen, A. C., Jr. 1951. On estimating the mean and variance of singly truncated normal frequency distributions from the first three sample moments. *Annals of the Institute of Statistical Mathematics* 3:37–44. doi: 10.1007/BF02949774.

Cooray, K., and M. M. A. Ananda. 2005. Modeling actuarial data with a composite lognormal-Pareto model. *Scandinavian Actuarial Journal* 2005 (5):321–34. doi: 10.1080/03461230510009763.

Delong, L., M. Lindholm, and M. V. Wüthrich. 2021. Gamma mixture density networks and their application to modelling insurance claim amounts. *Insuranc* *e: Mathematics & Economics* 101:240–61. doi: 10.1016/j.insmatheco.2021.08.003.

Efroymovich, S. 2021. Nonparametric curve estimation for truncated and censored data without product-limit. *Variance* 14 (2):16.

Fabián, Z. 2008. New measures of central tendency and variability of continuous distributions. *Communications in Statistics: Theory and Methods* 37 (1-2): 159–74. doi: 10.1080/03610920701648987.

Frees, E. W., and E. A. Valdez. 1998. Understanding relationships using copulas. *North American Actuarial Journal* 2 (1):1–25. doi: 10.1080/1092077.1998.10595667.

Fung, T. C. 2022. Maximum weighted likelihood estimator for robust heavy-tail modelling of finite mixture models. *Insurance: Mathematics & Economics* 107:180–98. doi: 10.1016/j.insmatheco.2022.08.008.

Goffard, P. O., and P. J. Laub. 2021. Approximate Bayesian computations to fit and compare insurance loss models. *Insurance: Mathematics & Economics* 100:350–71. doi: 10.1016/j.insmatheco.2021.06.002.

Hewitt, C. C., Jr., and B. Lefkowitz. 1979. Methods for fitting distributions to insurance loss data. In *Proceedings of the Casualty Actuarial Society*, vol. LXVI, 139–60. Arlington County, VA: Casualty Actuarial Society. https://www.casact.org/sites/default/files/database/procceed_proceed79_1979.pdf.

Klugman, S. A., H. H. Panjer, and G. E. Willmot. 2019. *Loss Models: From Data to Decisions*. New York: John Wiley & Sons.

Kojadinovic, I., and J. Yan. 2010. Modeling multivariate distributions with continuous margins using the copula R package. *Journal of Statistical Software* 34 (9):1–20. doi: 10.18637/jss.v034.i09.

Michael, S., T. Miljkovic, and V. Melnykov. 2020. Mixture modeling of data with multiple partial right-censoring levels. *Advances in Data Analysis and Classification* 14 (2):355–78. doi: 10.1007/s11634-020-00391-x.

Miljkovic, T., and B. Grün. 2016. Modeling loss data using mixtures of distributions. *Insurance: Mathematics & Economics* 70:387–96. doi: 10.1016/j.insmatheco.2016.06.019.

Nadaraja, S., and S. A. A. Bakar. 2014. New composite models for the Danish fire insurance data. *Scandinavian Actuarial Journal* 2014 (2):180–87. doi: 10.1080/03461238.2012.695748.

Poudyal, C. 2021a. Robust estimation of loss models for lognormal insurance payment severity data. *ASTIN Bulletin* 51 (2):475–507. doi: 10.1017/ash.2021.4.

Poudyal, C. 2021b. Truncated, censored, and actuarial payment-type moments for robust fitting of a single-parameter Pareto distribution. *Journal of Computational and Applied Mathematics* 388:113310. 18. doi: 10.1016/j.cam.2020.113310.

Poudyal, C., and V. Brazauskas. 2022a. Finite-sample performance of the T- and W-estimators for the Pareto tail index under data truncation and censoring. *Journal of Statistical Computation and Simulation* 1–21. doi: 10.1080/00949655.2022.2146114.

Poudyal, C., and V. Brazauskas. 2022b. Robust estimation of loss models for truncated and censored severity data. *Variance* 15 (2):1–20.

Punzo, A., L. Bagnato, and A. Maruotti. 2018. Compound unimodal distributions for insurance losses. *Insurance: Mathematics & Economics* 81:95–107. doi: 10.1016/j.insmatheco.2017.10.007.

Scollnik, D. P. M. 2007. On composite lognormal-Pareto models. *Scandinavian Actuarial Journal* 2007 (1):20–33. doi: 10.1080/03461230601110447.

Serfling, R. J. 1980. *Approximation theorems of mathematical statistics*. New York: John Wiley & Sons.

Shah, S. M., and M. C. Jaiswal. 1966. Estimation of parameters of doubly truncated normal distribution from first four sample moments. *Annals of the Institute of Statistical Mathematics* 18:107–11. doi: 10.1007/BF02869520.

Stehlik, M., R. Potocky, H. Waldl, and Z. Fabian. 2010. On the favorable estimation for fitting heavy tailed data. *Computational Statistics* 25 (3):485–503. doi: 10.1007/s00180-010-0189-1.

Tomarchio, S. D., and A. Punzo. 2020. Dichotomous unimodal compound models: Application to the distribution of insurance losses. *Journal of Applied Statistics* 47 (13–15):2328–53. doi: 10.1080/02664763.2020.1789076.

Tukey, J. W. 1960. A survey of sampling from contaminated distributions. In *Contributions to probability and statistics*, ed. Ingram Olkin, 448–85. Stanford, CA: Stanford University Press.

van der Vaart, A. W. 1998. *Asymptotic statistics*. Cambridge: Cambridge University Press.

Zhao, Q., V. Brazauskas, and J. Ghorai. 2018. Robust and efficient fitting of severity models and the method of winsorized moments. *ASTIN Bulletin* 48 (1): 275–309. doi: 10.1017/ash.2017.30.
APPENDIX. AUXILIARY RESULTS

Recall that, for \( 0 \leq s \leq 1 \), we define \( \tilde{s} = 1 - s \) and \( \Delta_s := \Phi^{-1}(s + \tilde{s}\Phi(\gamma)) \).

I. The expressions for \( c_{y,k} \), \( 1 \leq k \leq 4 \) used in Section 4.1 are given by:

\[
c_{y,k} \equiv c_{y,k}(\Phi, a, b, \gamma) = a\Delta_a^k + \int_a^b \Delta_s^k \ ds + b\Delta_b^k,
\]

and consequently the corresponding expressions for \( c_k \) used in Section 4.2 are computed as follows:

\[
c_k \equiv c_k(\Phi, a, b) = \lim_{\gamma \to -\infty} c_{y,k} = a[\Phi^{-1}(a)]^k + \int_a^b [\Phi^{-1}(s)]^k \ ds + b[\Phi^{-1}(b)]^k.
\]

I. Following the appendix of Zhao, Brazaunas, and Ghorai (2018), the expressions presented in the entries of the variance-covariance matrix \( \Sigma_y \) from Section 4.1 are as follows:

\[
c_{y,1}^* = c_{y,2}^* = c_{y,1} = c_{y,2} - a - b \left( \frac{\partial(c_{y,2} - c_{y,1})}{\partial a} \right) - b \left( \frac{\partial(c_{y,2} - c_{y,1})}{\partial b} \right) + a\left( \frac{\partial c_{y,1}}{\partial a} \right)^2 + b\left( \frac{\partial c_{y,1}}{\partial b} \right)^2 - 2ab \frac{\partial c_{y,1}}{\partial a} \frac{\partial c_{y,1}}{\partial b},
\]

\[
2c_{y,2}^* = c_{y,3} - c_{y,1} \cdot c_{y,2} - a \left( \frac{\partial(c_{y,3} - c_{y,1}c_{y,2})}{\partial a} \right) - b \left( \frac{\partial(c_{y,3} - c_{y,1}c_{y,2})}{\partial b} \right) + a\left( \frac{\partial c_{y,1}}{\partial a} \right)^2 + b\left( \frac{\partial c_{y,1}}{\partial b} \right)^2 - ab \left( \frac{\partial c_{y,1}}{\partial a} \right)^2 + ab \left( \frac{\partial c_{y,1}}{\partial b} \right)^2 - 2ab \frac{\partial c_{y,1}}{\partial a} \frac{\partial c_{y,1}}{\partial b},
\]

\[
4c_{y,3}^* = c_{y,4} - c_{y,2}^2 - a \left( \frac{\partial(c_{y,4} - c_{y,2})}{\partial a} \right) - b \left( \frac{\partial(c_{y,4} - c_{y,2})}{\partial b} \right) + a\left( \frac{\partial c_{y,2}}{\partial a} \right)^2 + b\left( \frac{\partial c_{y,2}}{\partial b} \right)^2 - 2ab \frac{\partial c_{y,2}}{\partial a} \frac{\partial c_{y,2}}{\partial b},
\]

where the involved derivatives are listed below:

\[
\frac{\partial c_{y,k}}{\partial a} = \Delta_a^k + a \frac{k\Delta_a^{k-1}}{\Phi(\Delta_a)} \Phi(\gamma) - \Delta_a^k = a\Phi(\gamma) \frac{k\Delta_a^{k-1}}{\phi(\Delta_a)}, \quad \frac{\partial c_{y,k}}{\partial b} = -b\Phi(\gamma) \frac{k\Delta_b^{k-1}}{\phi(\Delta_b)},
\]

\[
c_{y,k} \cdot c_{y,j} = a^2 \Delta_a^{k+j} + ab(\Delta_a^k \Delta_b^j + \Delta_a^j \Delta_b^k) + b^2 \Delta_b^{k+j} + (a\Delta_a^k + b\Delta_b^j) \int_a^b \Delta_s^k \ ds + (a\Delta_a^j + b\Delta_b^j) \int_a^b \Delta_s^j \ ds + \left( \int_a^b \Delta_s^k \ ds \right) \left( \int_a^b \Delta_s^j \ ds \right),
\]

\[
\frac{\partial(c_{y,k} \cdot c_{y,j})}{\partial a} = c_{y,j} \frac{\partial c_{y,k}}{\partial a} + c_{y,k} \frac{\partial c_{y,j}}{\partial a} = a\Phi(\gamma) \frac{k\Delta_a^{k-1}}{\phi(\Delta_a)} c_{y,j} + a\Phi(\gamma) \frac{j\Delta_a^{j-1}}{\phi(\Delta_a)} c_{y,k}
\]

\[
= \frac{a\Phi(\gamma)}{\phi(\Delta_a)} \left[ k\Delta_a^{k-1} c_{y,j} + j\Delta_a^{j-1} c_{y,k} \right],
\]

\[
\frac{\partial(c_{y,k} \cdot c_{y,j})}{\partial b} = c_{y,j} \frac{\partial c_{y,k}}{\partial b} + c_{y,k} \frac{\partial c_{y,j}}{\partial b} = b\Phi(\gamma) \frac{k\Delta_b^{k-1}}{\phi(\Delta_b)} c_{y,j} - b\Phi(\gamma) \frac{j\Delta_b^{j-1}}{\phi(\Delta_b)} c_{y,k}
\]

\[
= -\frac{b\Phi(\gamma)}{\phi(\Delta_b)} \left[ k\Delta_b^{k-1} c_{y,j} + j\Delta_b^{j-1} c_{y,k} \right].
\]
Similarly, the expressions used in Equation (4.23) are such that \( c_k^* = \lim_{\gamma \to -\infty} c_{\gamma,k}^* \) and are given by

\[
c_1' = c_2 - c_1 - a \frac{\partial(c_2 - c_1^2)}{\partial a} - b \frac{\partial(c_2 - c_1^2)}{\partial b} + a \frac{\partial c_1}{\partial a} \left( \frac{\partial c_1}{\partial a} \right)^2 + b \frac{\partial c_1}{\partial b} \left( \frac{\partial c_1}{\partial b} \right)^2 - 2ab \frac{\partial c_1}{\partial a} \frac{\partial c_1}{\partial b}.
\]

\[
2c_2' = c_3 - c_1 \cdot c_2 - a \frac{\partial(c_3 - c_1 c_2)}{\partial a} - b \frac{\partial(c_3 - c_1 c_2)}{\partial b} + a \frac{\partial c_1}{\partial a} \frac{\partial c_2}{\partial a} + b \frac{\partial c_1}{\partial b} \frac{\partial c_2}{\partial b} - ab \left( \frac{\partial c_1}{\partial a} \frac{\partial c_2}{\partial b} + \frac{\partial c_1}{\partial b} \frac{\partial c_2}{\partial a} \right),
\]

\[
4c_3' = c_4 - c_2^2 - a \frac{\partial(c_4 - c_2^2)}{\partial a} - b \frac{\partial(c_4 - c_2^2)}{\partial b} + a \frac{\partial c_2}{\partial a} \left( \frac{\partial c_2}{\partial a} \right)^2 + b \frac{\partial c_2}{\partial b} \left( \frac{\partial c_2}{\partial b} \right)^2 - 2ab \frac{\partial c_2}{\partial a} \frac{\partial c_2}{\partial b},
\]

where the involved derivatives are listed below:

\[
\frac{\partial c_k}{\partial a} = \frac{ka \left[ \Phi^{-1}(c) \right]^{k-1}}{\phi(\Phi^{-1}(c))}, \quad \frac{\partial c_k}{\partial b} = -\frac{kb \left[ \Phi^{-1}(c) \right]^{k-1}}{\phi(\Phi^{-1}(c))},
\]

\[
\frac{\partial(c_k c_j)}{\partial a} = \frac{a}{\phi(\Phi^{-1}(c))} \left[ k \left[ \Phi^{-1}(c) \right]^{k-1} c_j + j \left[ \Phi^{-1}(c) \right]^{j-1} c_k \right],
\]

\[
\frac{\partial(c_k c_j)}{\partial b} = -\frac{b}{\phi(\Phi^{-1}(c))} \left[ k \left[ \Phi^{-1}(c) \right]^{k-1} c_j + j \left[ \Phi^{-1}(c) \right]^{j-1} c_k \right].
\]