ROBUST ESTIMATION OF LOSS MODELS FOR LOGNORMAL INSURANCE PAYMENT SEVERITY DATA

BY

CHUDAMANI POUDYAL

ABSTRACT

The primary objective of this scholarly work is to develop two estimation procedures – *maximum likelihood estimator* (MLE) and *method of trimmed moments* (MTM) – for the mean and variance of lognormal insurance payment severity data sets affected by different loss control mechanism, for example, truncation (due to deductibles), censoring (due to policy limits), and scaling (due to coinsurance proportions), in insurance and financial industries. Maximum likelihood estimating equations for both payment-per-payment and payment-per-loss data sets are derived which can be solved readily by any existing iterative numerical methods. The asymptotic distributions of those estimators are established via Fisher information matrices. Further, with a goal of balancing efficiency and robustness and to remove point masses at certain data points, we develop a dynamic MTM estimation procedures for lognormal claim severity models for the above-mentioned transformed data scenarios. The asymptotic distributional properties and the comparison with the corresponding MLEs of those MTM estimators are established along with extensive simulation studies. Purely for illustrative purpose, numerical examples for 1500 US indemnity losses are provided which illustrate the practical performance of the established results in this paper.

KEYWORDS

Dynamic estimation, lognormal insurance severity, maximum likelihood estimators, trimmed moments, insurance payments, loss models, robust estimation, truncated and censored data.

1. INTRODUCTION

The research leading to the results of this work is basically motivated to find and compare two estimation approaches for two types of insurance

Astin Bulletin 51(2), 475-507. doi:10.1017/asb.2021.4 © 2021 by Astin Bulletin. All rights reserved.

https://doi.org/10.1017/asb.2021.4 Published online by Cambridge University Press

benefit payment severity data with the assumption of stand-alone continuous underlying loss distributions. In practice, due to commonly used loss control mechanism in the financial and insurance industries (Ergashev *et al.*, 2016; Klugman *et al.*, 2019), the random variables we observe and wish to model are affected by data truncation (due to deductibles), censoring (due to policy limits), and scaling (due to coinsurance factor). Therefore, the main focus of this paper is to develop *maximum likelihood* and a dynamic *method of trimmed moments* estimators, respectively, called MLE and MTM for short, of location and scale parameters of lognormal insurance benefit payment claims severity.

In the current practice, statistical inference for loss models is almost exclusively likelihood based (see, e.g., Klugman *et al.*, 2019), which typically result in sensitive loss severity models if there is a small perturbation in the underlying assumed model or if the observed sample is coming from a contaminated distribution, Tukey (1960). Further, following the work by Cooray and Ananda (2005), the contemporary actuarial loss modeling literature has been directed toward composite loss modeling (see, e.g., Scollnik, 2007; Scollnik and Sun, 2012; Abu Bakar *et al.*, 2015; Chan *et al.*, 2018) which is also referred to as splicing (see, e.g., Klugman *et al.*, 2019) including mixtures of Erlangs (see, e.g., Verbelen *et al.*, 2015; Reynkens *et al.*, 2017; Gui *et al.*, 2021). The composite models appear to capture the tail behavior of the underlying distribution more accurately (see, e.g., Grün and Miljkovic, 2019, and the references therein). But as mentioned by Punzo *et al.* (2018), often the simple closed-form expressions for the pdf of composite models are not available which could lead to more complexity of deriving analytic sampling distributional results. Therefore, beside many ideas from the mainstream robust statistics literature (see, e.g., Huber, 1964; Hampel, 1974; Huber and Ronchetti, 2009), actuaries have to deal with heavy-tailed and skewed distributions, data truncation and censoring, identification and recycling of outliers, and aggregate loss, etc. Thus, it is appealing to search some estimation procedures which directly work with those mentioned loss control mechanism, are insensitive against small perturbations from the assumed models, are computationally efficient, and still reasonably balance the asymptotic efficiency-robustness trade-offs with respect to the corresponding MLE.

As a member of log-location-scale family, the lognormal distribution has diverse applications in actuarial science, business, and economics (see, e.g., Serfling, 2002, and the references therein) which closely approximates certain types of homogeneous actuarial loss data (Hewitt *et al.*, 1979; Punzo *et al.*, 2018). Further, it has been established that, even for the heterogeneous actuarial losses, lognormal distribution is able to capture the nature of the data set either on the head or on the tail or on both head and tail parts of different composite models, see, for example, Cooray and Ananda (2005), Brazauskas and Kleefeld (2016), Miljkovic and Grün (2016), Punzo *et al.* (2018), Blostein and Miljkovic (2019), and Michael *et al.* (2020). More comprehensive investigation of 256 different composite models have been analyzed by Grün and Miljkovic (2019).

https://doi.org/10.1017/asb.2021.4 Published online by Cambridge University Press

On the other hand, MTM approach can be viewed as a special case of *L*-statistics (Chernoff *et al.*, 1967) and has established itself as sufficiently general and flexible in order to balance between estimator's efficiency and robustness for fitting of continuous parametric models based on complete ground-up loss severity data (see, e.g., Brazauskas *et al.*, 2009; Zhao *et al.*, 2018). But it is yet open to investigate the MTM performance beyond the complete data scenario and this paper will address this issue for lognormal payment-per-payment and payment-per-loss data scenarios. 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). A novel robust estimation procedure called *method of truncated moments* (MTuM) for modeling different actuarial loss data scenarios was initially purposed by Poudyal (2018) and further developed by Poudyal (2021). But the goal and motivation of this research work is different. That is, instead of fitting composite models or designing MTuM, both MLE and MTM estimation procedures will be derived for stand-alone lognormal insurance benefit payment data scenarios with an assumption that a close fit in one tail or both tails of the distribution is not desired but at the same time we want to remove partial point masses accumulated at the truncated and/or censored data points. As the main contribution of this parer, asymptotic distributions, such as normality and consistency, along with asymptotic relative efficiency (ARE) of those estimators (under transformed data) with respect to the corresponding MLEs are established including extensive validating simulation studies. The developed methodologies in this paper are implemented with real 1500 US indemnity loss data set. It will be shown that, when properly redesigned, a dynamic MTM can be a robust and computationally efficient alternative to the MLE-based inference for claim severity models that are affected by deductibles, policy limits, and coinsurance. Some of the technical challenges in implementing the developed estimation procedures in practical data analysis are discussed along with the corresponding simplifying assumptions.

The remainder of the paper is organized as follows. In Section 2, we describe the two types of insurance benefit payment lognormal random variables. In Section 3, the MLE procedures is developed for those two payment type of data sets with their corresponding asymptotic distributional properties. Section 4 is the corresponding section to develop dynamic MTM procedures for the location and scale parameters for those two payment models. Comparison of ARE of those MTM estimators with respect to MLE are also presented. Section 5 is for some special cases which are more common in operational risk modeling with some theoretical results. Section 6 summarizes a detail simulation study. Numerical examples to observe the performance of the developed estimation procedures can be found in Section 7. Finally, summary and concluding remarks are offered in Section 8.

https://doi.org/10.1017/asb.2021.4 Published online by Cambridge University Press

2. INSURANCE PAYMENTS

Consider a ground-up lognormal loss random variable $W \sim LN(w_0, \theta, \sigma)$ with cdf:

$$F_W(w) = \Phi\left(\frac{\log(w - w_0) - \theta}{\sigma}\right), \quad w > w_0,\ -\infty < \theta < \infty,\ \sigma > 0, \tag{2.1}$$

where θ and σ are unknown parameters to be estimated, $w_0 > 0$ is assumed to be known, and Φ is the cdf of the standard normal distribution. Clearly, $X := \log(W - w_0) \sim N(\theta, \sigma^2)$ with cdf and pdf, respectively, given by:

$$F_X(x) = \Phi\left(\frac{x - \theta}{\sigma}\right) \quad \text{and} \quad f_X(x) = \frac{1}{\sigma}\phi\left(\frac{x - \theta}{\sigma}\right), \qquad -\infty < x < \infty, \tag{2.2}$$

where $\phi(x) = \left(1/\sqrt{2\pi}\right) e^{-x^2/2},\ x \in \mathbb{R}$ is the pdf of the standard normal distribution. The corresponding quantile function $F_X^{-1} : (0, 1) \to \mathbb{R}$ is given by $F_X^{-1}(v) = \theta + \sigma\Phi^{-1}(v)$.

Insurance contracts have coverage modifications that need to be taken into account when modeling the underlying loss variable. Usually, the coverage modifications such as deductibles, policy limits, and coinsurance are introduced as loss control strategies, so that unfavorable policyholder behavioral effects (e.g., adverse selection) can be minimized. There are also situations when certain features of the contract emerge naturally (e.g., the value of insured property in general insurance is a natural upper policy limit). Here, we describe two common transformations of the loss variable along with the corresponding cdfs, pdfs, and qfs.

Suppose the lognormal severity random variable W has ordinary deductible d, upper policy limit u, and coinsurance factor c $(0 < c \leq 1)$. These coverage specifications imply that when a loss W is reported, the insurance company is responsible for a proportion c of W exceeding d, but no more than $c(u - d)$. Define

$$\begin{aligned} &t := \log(d - w_0),\ T := \log(u - w_0),\ R := T - t, \\ &\gamma := \frac{t - \theta}{\sigma},\ \text{and } \xi := \frac{T - \theta}{\sigma}. \end{aligned} \tag{2.3}$$

and note that it is possible to have $t < 0$ but $d > 0$. Here, it is important to note that d and u refer for the lognormal left truncation threshold and right-censored point, respectively. On the other hand, t and T are, respectively, the corresponding normal form of left truncation threshold and right-censored point as defined in (2.3). Then, obviously

$$\theta = t - \sigma\gamma \quad \text{and} \quad \xi = \gamma + \frac{R}{\sigma}. \tag{2.4}$$

https://doi.org/10.1017/asb.2021.4 Published online by Cambridge University Press

Also, let $\bar{\Phi}(z) = 1 - \Phi(z)$ be the standard normal survival function at $z \in \mathbb{R}$. Finally, define

$$\Omega_1 := \frac{\phi(\gamma)}{\bar{\Phi}(\gamma) - \bar{\Phi}(\xi)} \quad \text{and} \quad \Omega_2 := \frac{\phi(\xi)}{\bar{\Phi}(\gamma) - \bar{\Phi}(\xi)}. \tag{2.5}$$

Now, if the loss severity W below the deductible d is completely unobservable (even its frequency is unknown), then the *left-truncated*, *right-censored*, and *linearly transformed* (also known as *payment-per-payment* variable) form of W is defined as:

$$Y_w := c\left(\min\left\{W, u\right\} - d\right) \mid W > d = \begin{cases} \text{undefined}, & W \le d; \\ c\,(W - d)\,, & d < W < u; \\ c\,(u - d)\,, & u \le W. \end{cases} \tag{2.6}$$

The corresponding normal form of Y_w is given by

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

The cdf F_Y, pdf f_Y, and qf F_Y^{-1} of the payment-per-payment random variable Y are given by:

$$\begin{aligned} F_Y(y; c, t, T) &= \mathbf{P}\left[c\left(\min\left\{X,\ T\right\} - t\right) \le y \,\middle|\, X > t\right] \\ &= \begin{cases} 0, & y \le 0; \\ \frac{F_X(y/c+t) - F_X(t)}{1 - F_X(t)}, & 0 < y < c(T - t); \\ 1, & y \ge c(T - t), \end{cases} \end{aligned} \tag{2.8}$$

$$f_Y(y; c, t, T) = \begin{cases} \frac{f_X(y/c+t)}{c[1 - F_X(t)]}, & 0 < y < c(T - t); \\ \frac{1 - F_X(T^-)}{1 - F_X(t)}, & y = c(T - t); \\ 0, & \text{elsewhere}, \end{cases} \tag{2.9}$$

and

$$F_Y^{-1}(v; c, t, T) = \begin{cases} c\left[F_X^{-1}\big(v + (1 - v)F_X(t)\big) - t\right], & 0 \le v < \frac{F_X(T) - F_X(t)}{1 - F_X(t)}; \\ c(T - t), & \frac{F_X(T) - F_X(t)}{1 - F_X(t)} \le v \le 1. \end{cases} \tag{2.10}$$

The scenario that no information is available about W below d is likely to occur when modeling is done based on the data acquired from a third party (e.g., data vendor). For payment data collected in house, the information about

https://doi.org/10.1017/asb.2021.4 Published online by Cambridge University Press

the number of policies that did not report claims (equivalently, resulted in a payment of 0) would be available. This minor modification yields different payment variable, say Z_w, which can be treated as *interval-censored* and *linearly transformed W* (also called *payment-per-loss* random variable):

$$Z_w := c\left(\min\left\{W, u\right\} - \min\left\{W, d\right\}\right) = \begin{cases} 0, & W \le d; \\ c\,(W-d)\,, & d < W < u; \\ c\,(u-d)\,, & u \le W. \end{cases} \tag{2.11}$$

The corresponding normal form of Z_w is given by:

$$Z := c \log\left(\frac{Z_w}{c(d-w_0)} + 1\right) = c\left(\min\left\{X, T\right\} - \min\left\{X, t\right\}\right)$$
$$= \begin{cases} 0, & X \le t; \\ c\,(X-t)\,, & t < X < T; \\ c\,(T-t)\,, & T \le X. \end{cases} \tag{2.12}$$

The cdf F_Z, pdf f_Z, and qf F_Z^{-1} are related to F_X, f_X, and F_X^{-1} and given by:

$$F_Z(z; c, t, T) = \mathbf{P}\left[c\big(\min\{X, T\} - \min\{X, t\}\big) \le z\right] \tag{2.13}$$
$$= \begin{cases} 0, & z < 0; \\ F_X(z/c + t), & 0 \le z < c(T-t); \\ 1, & z \ge c(T-t), \end{cases} \tag{2.14}$$

$$f_Z(z; c, t, T) = \begin{cases} F_X(t), & z = 0; \\ f_X(z/c + t)/c, & 0 < z < c(T-t); \\ 1 - F_X(T^-), & z = c(T-t); \\ 0, & \text{elsewhere}, \end{cases} \tag{2.15}$$

and

$$F_Z^{-1}(v; c, t, T) = \begin{cases} 0, & 0 \le v \le F_X(t); \\ c\left(F_X^{-1}(v) - t\right), & F_X(t) < v < F_X(T); \\ c(T-t), & F_X(T) \le v \le 1. \end{cases} \tag{2.16}$$

Note 2.1. *The numbers t and T can be treated as deductible and policy limit, respectively, for the normal random variable* $X \sim N(\theta, \sigma^2)$ *with a possibility of* $t < 0$. □

Note 2.2. *The variable* γ*, defined in (2.3), is considered to be the independent parameter of location,* θ*, for MLE estimation purpose. Then, the mean,* θ*, is a linear function of* γ *given by (2.4).* □

https://doi.org/10.1017/asb.2021.4 Published online by Cambridge University Press

3. MLE

Here, we develop MLE estimation procedures for θ and σ under the transformed data scenarios given by (2.7) and (2.12).

3.1. Payments Y

Let $y_1, \ldots, y_n$ be an *i.i.d.* sample given by pdf (2.9) with policy limit T, deductible t, and coinsurance factor c. Then, with $n_1 := \sum_{i=1}^{n} \mathbb{1}\{0 < y_i < cR\}$ and $n_2 := \sum_{i=1}^{n} \mathbb{1}\{y_i = cR\}$, the corresponding log-likelihood function becomes

$$l_y(\gamma, \sigma) := -\frac{n_1}{2} \log(2\pi) + n_2 \log(\bar{\Phi}(\xi)) - n \log(\bar{\Phi}(\gamma)) - n_1 \log(\sigma) - \frac{1}{2} \sum_{0<y_i<cR} \left(\gamma + \frac{y_i}{c\sigma}\right)^2. \tag{3.1}$$

Thus, setting $\left(\frac{\partial l_y}{\partial \gamma}, \frac{\partial l_y}{\partial \sigma}\right) = (0, 0)$, yields the system of MLE equations:

$$\begin{cases} n\dfrac{\phi(\gamma)}{\bar{\Phi}(\gamma)} - n_2\dfrac{\phi(\xi)}{\bar{\Phi}(\xi)} - \displaystyle\sum_{0<y_i<cR} \left(\gamma + \frac{y_i}{c\sigma}\right) = 0, \\ n_2\dfrac{R\phi(\xi)}{\sigma^2\bar{\Phi}(\xi)} - \dfrac{n_1}{\sigma} + \dfrac{1}{c\sigma^2} \displaystyle\sum_{0<y_i<cR} y_i \left(\gamma + \frac{y_i}{c\sigma}\right) = 0. \end{cases} \tag{3.2}$$

Define

$$\Omega_{y,1} := \frac{n}{n_1}\frac{\phi(\gamma)}{\bar{\Phi}(\gamma)} \quad \text{and} \quad \Omega_{y,2} := \frac{n_2}{n_1}\frac{\phi(\xi)}{\bar{\Phi}(\xi)}, \tag{3.3}$$

then the system of MLE Equations (3.2) takes the form:

$$\begin{cases} \sigma\left(\Omega_{y,1} - \Omega_{y,2} - \gamma\right) - c^{-1}\widehat{\mu}_{y,1} = 0, \\ \sigma^2\left(1 - \gamma(\Omega_{y,1} - \Omega_{y,2} - \gamma) - \frac{\Omega_{y,2}R}{\sigma}\right) - c^{-2}\widehat{\mu}_{y,2} = 0, \end{cases} \tag{3.4}$$

where $\widehat{\mu}_{y,1}$ and $\widehat{\mu}_{y,2}$ are the first and second sample moments, $\widehat{\mu}_{y,j} := n_1^{-1}\sum_{i=1}^{n} \mathbb{1}\{0 < y_i < cR\} y_i^j, j = 1, 2$. To solve the system (3.4) for $\widehat{\gamma}_{\text{y,MLE}}$ and $\widehat{\sigma}_{\text{y,MLE}}$, we initialize the system as below:

$$\sigma_{\text{start}} = \sqrt{c^{-2}\widehat{\mu}_{y,2} - \left(c^{-1}\widehat{\mu}_{y,1}\right)^2} \quad \text{and} \quad \gamma_{\text{start}} = -\frac{c^{-1}\widehat{\mu}_{y,1}}{\sigma_{\text{start}}}. \tag{3.5}$$

https://doi.org/10.1017/asb.2021.4 Published online by Cambridge University Press

Theorem 3.1. *Let* $\boldsymbol{I}_y(\gamma,\sigma)=\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22}\end{bmatrix}$ *be the Fisher information matrix for a sample of size* 1 *for the random variable* Y*, then*

$$a_{11}=-\Lambda r_1(\gamma,\xi),\ a_{12}=a_{21}=-\Lambda\sigma^{-1}r_2(\gamma,\xi),\ \text{and}\ a_{22}=-\Lambda\sigma^{-2}r_3(\gamma,\xi),$$

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

$$\begin{cases} r_1(\gamma,\xi):=-\left[1+\gamma\Omega_1-\xi\Omega_2-\frac{\phi(\gamma)}{\bar{\Phi}(\gamma)}\Omega_1+\frac{\phi(\xi)}{\bar{\Phi}(\xi)}\Omega_2\right], \\ r_2(\gamma,\xi):=\frac{R\Omega_2}{\sigma}\left[\frac{\phi(\xi)}{\bar{\Phi}(\xi)}-\xi\right]+\left[\Omega_1-\Omega_2-\gamma\right], \\ r_3(\gamma,\xi):=\left(\frac{R}{\sigma}\right)^2\Omega_2\left(\xi-\frac{\phi(\xi)}{\bar{\Phi}(\xi)}\right)-\left[2-\gamma(\Omega_1-\Omega_2-\gamma)-\frac{\Omega_2 R}{\sigma}\right]. \end{cases} \tag{3.6}$$

Consequently, the asymptotic normality of $\left(\widehat{\gamma}_{y,MLE},\widehat{\sigma}_{y,MLE}\right)$ *is given by:*

$$\sqrt{n}\left(\widehat{\gamma}_{y,MLE}-\gamma,\widehat{\sigma}_{y,MLE}-\sigma\right)\xrightarrow{D}N\left((0,0),\boldsymbol{I}_y^{-1}(\gamma,\sigma)\right). \tag{3.7}$$

Proof. First note that $\boldsymbol{I}_y(\gamma,\sigma)=n^{-1}\boldsymbol{I}_{y,n}(\gamma,\sigma)$, where $\boldsymbol{I}_{y,n}(\gamma,\sigma)$ is the Fisher information matrix for a sample of size n. Using (3.4), it can be shown that

$$\frac{\partial^2 l_y}{\partial\gamma^2}=n_1R_1(\gamma,\xi),\quad \frac{\partial^2 l_y}{\partial\sigma\,\partial\gamma}=\frac{n_1}{\sigma}R_2(\gamma,\xi),\quad \frac{\partial^2 l_y}{\partial\sigma^2}=\frac{n_1}{\sigma^2}R_3(\gamma,\xi),$$

where

$$\begin{cases} R_1(\gamma,\xi):=-\left[1+\gamma\Omega_{y,1}-\xi\Omega_{y,2}-\frac{n_1}{n}\Omega_{y,1}^2+\frac{n_1}{n_2}\Omega_{y,2}^2\right], \\ R_2(\gamma,\xi):=\frac{R\Omega_{y,2}}{\sigma}\left[\frac{n_1}{n_2}\Omega_{y,2}-\gamma\right]+\left[\Omega_{y,1}-\Omega_{y,2}-\gamma\right], \\ R_3(\gamma,\xi):=\left(\frac{R}{\sigma}\right)^2\Omega_{y,2}\left(\xi-\frac{n_1}{n_2}\frac{\phi(\xi)}{\bar{\Phi}(\xi)}\right)-\left[2-\gamma(\Omega_{y,1}-\Omega_{y,2}-\gamma)-\frac{\Omega_{y,2}R}{\sigma}\right]. \end{cases} \tag{3.8}$$

Since $n_1\sim\text{Binomial}(n,\Lambda)$ and $n_2\sim\text{Binomial}\left(n,\bar{\Phi}(\xi)/\bar{\Phi}(\gamma)\right)$, then it follows that

$$\begin{cases} a_{11}=-\frac{1}{n}E\left[\frac{\partial^2 l_y}{\partial\gamma^2}\right]=-\frac{1}{n}E\left[n_1R_1(\gamma,\xi)\right]=-\Lambda r_1(\gamma,\xi), \\ a_{12}=-\frac{1}{n}E\left[\frac{\partial^2 l_y}{\partial\sigma\,\partial\gamma}\right]=-\frac{1}{n}E\left[\frac{n_1}{\sigma}R_2(\gamma,\xi)\right]=-\Lambda\sigma^{-1}r_2(\gamma,\xi), \\ a_{22}=-\frac{1}{n}E\left[\frac{\partial^2 l_y}{\partial\sigma^2}\right]=-\frac{1}{n}E\left[\frac{n_1}{\sigma^2}R_3(\gamma,\xi)\right]=-\Lambda\sigma^{-2}r_3(\gamma,\xi), \end{cases} \tag{3.9}$$

as desired. □

https://doi.org/10.1017/asb.2021.4 Published online by Cambridge University Press

From Equation (3.7), it follows that

$$(\widehat{\gamma}_{\text{y,MLE}}, \widehat{\sigma}_{\text{y,MLE}}) \sim \mathcal{AN}\left((\gamma, \sigma), \frac{\Lambda^{-1}}{n\left(r_1 r_3 - r_2^2\right)} \begin{bmatrix} -r_3 & \sigma r_2 \\ \sigma r_2 & -\sigma^2 r_1 \end{bmatrix}\right). \tag{3.10}$$

Further, since $(\theta, \sigma) = (t - \sigma\gamma, \sigma)$, then by multivariate delta method (see, e.g., Serfling, 1980, p. 122), we have $(\widehat{\theta}_{\text{y,MLE}}, \widehat{\sigma}_{\text{y,MLE}}) \sim \mathcal{AN}\left((\theta, \sigma), \frac{1}{n}\boldsymbol{S}_{\text{y,MLE}}\right)$, where

$$\boldsymbol{S}_{y,MLE} = \frac{\Lambda^{-1}}{\left(r_1 r_3 - r_2^2\right)} \mathbf{D} \begin{bmatrix} -r_3 & \sigma r_2 \\ \sigma r_2 & -\sigma^2 r_1 \end{bmatrix} \mathbf{D}' \quad \text{and} \quad \mathbf{D} = \begin{bmatrix} -\sigma & -\gamma \\ 0 & 1 \end{bmatrix}. \tag{3.11}$$

3.2. Payments Z

Consider an observed *i.i.d.* sample $z_1, \ldots, z_n$ given by pdf (2.14). Define

$$n_0 := \sum_{i=1}^{n} \mathbb{1}\{z_i = 0\}, \quad n_1 := \sum_{i=1}^{n} \mathbb{1}\{0 < z_i < cR\}, \quad n_2 := \sum_{i=1}^{n} \mathbb{1}\{z_i = cR\}. \tag{3.12}$$

Note that $n = n_0 + n_1 + n_2$. In this case, the log-likelihood function becomes

$$\begin{aligned} l_z(\gamma, \sigma) := &-\frac{n_1}{2}\log(2\pi) + n_0 \log\left(1 - \bar{\Phi}(\xi)\right) - n_1 \log(\sigma) \\ &- \frac{1}{2} \sum_{0<z_i<cR} \left(\gamma + \frac{z_i}{c\sigma}\right)^2 + n_2 \log\left(\bar{\Phi}(\xi)\right). \end{aligned} \tag{3.13}$$

The corresponding likelihood system of equations takes the form:

$$\begin{cases} \frac{\partial l_z}{\partial \gamma} = n_0 \dfrac{\phi(\gamma)}{\Phi(\gamma)} - n_2 \dfrac{\phi(\xi)}{\bar{\Phi}(\xi)} - \displaystyle\sum_{0<z_i<cR} \left(\gamma + \frac{z_i}{c\sigma}\right) = 0, \\ \frac{\partial l_z}{\partial \sigma} = -n_1 \dfrac{1}{\sigma} + \dfrac{1}{\sigma^2} \displaystyle\sum_{0<z_i<cR} \left[\frac{z_i}{c}\left(\gamma + \frac{z_i}{c\sigma}\right)\right] + n_2 \dfrac{\phi(\xi)}{\bar{\Phi}(\xi)} = 0. \end{cases} \tag{3.14}$$

Similarly with (2.5), define

$$\Omega_{z,1} := \frac{n_0}{n_1} \frac{\phi(\gamma)}{\Phi(\gamma)}, \qquad \Omega_{z,2} := \frac{n_2}{n_1} \frac{\phi(\xi)}{\bar{\Phi}(\xi)}, \tag{3.15}$$

then the MLE system of Equations (3.14) becomes

$$\begin{cases} \sigma\left(\Omega_{z,1} - \Omega_{z,2} - \gamma\right) - c^{-1}\widehat{\mu}_{z,1} = 0, \\ \sigma^2\left(1 - \gamma(\Omega_{z,1} - \Omega_{z,2} - \gamma) - \frac{\Omega_{z,2}R}{\sigma}\right) - c^{-2}\widehat{\mu}_{z,2} = 0, \end{cases} \tag{3.16}$$

where $\widehat{\mu}_{z,1}$ and $\widehat{\mu}_{z,2}$ are the first and second sample moments, $\widehat{\mu}_{z,j} := n_1^{-1} \sum_{i=1}^{n} \mathbb{1}\{0 < z_i < cR\} z_i^j$, $j = 1, 2$. With the assumptions of $n_0 > 0$ and $n_2 > 0$,

https://doi.org/10.1017/asb.2021.4 Published online by Cambridge University Press

the system (3.16) with relation (2.4) can be initialized at Cohen (1950):

$$(\gamma_{\text{start}}, \sigma_{\text{start}}) = \left(\Phi^{-1}\left(\frac{n_0}{n}\right), \frac{R}{\xi_{\text{start}} - \gamma_{\text{start}}} \right), \quad \text{where} \quad \xi_{\text{start}} = \Phi^{-1}\left(1 - n_2/n\right).$$

Note that γ_{start} and σ_{start} are simply the empirically estimated values of γ and σ, respectively. And, if $n_2 = 0$ which implies that none of the observations are censored, then (5.9) can provide a satisfactory initialization of the system (3.16).

Theorem 3.2. *Let* $\boldsymbol{I}_z(\gamma, \sigma) = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$ *be the Fisher information matrix for a sample of size* 1 *for the random variable Z, then*

$$b_{11} = -\Lambda \bar{\Phi}(\gamma)\, \psi_1(\gamma, \xi),\ b_{12} = b_{21} = -\Lambda \bar{\Phi}(\gamma)\, \sigma^{-1} \psi_2(\gamma, \xi),$$
$$\text{and } b_{22} = -\Lambda \bar{\Phi}(\gamma)\, \sigma^{-2} \psi_3(\gamma, \xi),$$

where Λ *is defined as in Theorem 3.1 and*

$$\begin{cases} \psi_1(\gamma, \xi) := -\left[1 + \gamma\Omega_1 - \xi\Omega_2 + \frac{\phi(\gamma)}{\Phi(\gamma)}\Omega_1 + \frac{\phi(\xi)}{\bar{\Phi}(\xi)}\Omega_2\right], \\ \psi_2(\gamma, \xi) := \frac{R\Omega_2}{\sigma}\left[\frac{\phi(\xi)}{\bar{\Phi}(\xi)} - \xi\right] + \left[\Omega_1 - \Omega_2 - \gamma\right], \\ \psi_3(\gamma, \xi) := \left(\frac{R}{\sigma}\right)^2 \Omega_2 \left(\xi - \frac{\phi(\xi)}{\bar{\Phi}(\xi)}\right) - \left[2 - \gamma(\Omega_1 - \Omega_2 - \gamma) - \frac{\Omega_2 R}{\sigma}\right]. \end{cases} \tag{3.17}$$

Consequently, the asymptotic normality of $(\widehat{\gamma}_{z,MLE}, \widehat{\sigma}_{z,MLE})$ *is given by:*

$$\sqrt{n}\left(\widehat{\gamma}_{z,MLE} - \gamma, \widehat{\sigma}_{z,MLE} - \sigma\right) \xrightarrow{D} N\left((0,0), \boldsymbol{I}_z^{-1}(\gamma, \sigma)\right). \tag{3.18}$$

Proof. Using (3.16), it can be shown that

$$\frac{\partial^2 l_z}{\partial \gamma^2} = n_1 \Psi_1(\gamma, \xi), \quad \frac{\partial^2 l_z}{\partial \sigma\, \partial \gamma} = \frac{n_1}{\sigma} \Psi_2(\gamma, \xi), \quad \frac{\partial^2 l_z}{\partial \sigma^2} = \frac{n_1}{\sigma^2} \Psi_3(\gamma, \xi),$$

where

$$\begin{cases} \Psi_1(\gamma, \xi) := -\left[1 + \gamma\Omega_{z,1} - \xi\Omega_{z,2} + \frac{n_1}{n_0}\Omega_{z,1}^2 + \frac{n_1}{n_2}\Omega_{z,2}^2\right], \\ \Psi_2(\gamma, \xi) := \frac{R\Omega_{z,2}}{\sigma}\left[\frac{n_1}{n_2}\Omega_{z,2} - \xi\right] + \left[\Omega_{z,1} - \Omega_{z,2} - \gamma\right], \\ \Psi_3(\gamma, \xi) := \left(\frac{R}{\sigma}\right)^2 \Omega_{z,2}\left(\xi - \frac{n_1}{n_2}\Omega_{z,2}\right) - \left[2 - \gamma(\Omega_{z,1} - \Omega_{z,2} - \gamma) - \frac{\Omega_{z,2} R}{\sigma}\right]. \end{cases} \tag{3.19}$$

https://doi.org/10.1017/asb.2021.4 Published online by Cambridge University Press

Also, note that $n_0 \sim \text{Binomial}(n, 1-\bar{\Phi}(\gamma))$, $n_1 \sim \text{Binomial}(n, \bar{\Phi}(\gamma)-\bar{\Phi}(\xi))$, and $n_2 \sim \text{Binomial}(n, \bar{\Phi}(\xi))$, then it follows that

$$\begin{cases} b_{11} = -\frac{1}{n}E\left[\dfrac{\partial^2 l_z}{\partial \gamma^2}\right] = -\frac{1}{n}E\left[n_1\Psi_1(\gamma,\xi)\right] = -\Lambda\bar{\Phi}(\gamma)\,\psi_1(\gamma,\xi), \\ b_{12} = -\frac{1}{n}E\left[\dfrac{\partial^2 l_z}{\partial\sigma\,\partial\gamma}\right] = -\frac{1}{n}E\left[\frac{n_1}{\sigma}\Psi_2(\gamma,\xi)\right] = -\Lambda\bar{\Phi}(\gamma)\,\sigma^{-1}\psi_2(\gamma,\xi), \\ b_{22} = -\frac{1}{n}E\left[\dfrac{\partial^2 l_z}{\partial\sigma^2}\right] = -\frac{1}{n}E\left[\frac{n_1}{\sigma^2}\Psi_3(\gamma,\xi)\right] = -\Lambda\bar{\Phi}(\gamma)\,\sigma^{-2}\psi_3(\gamma,\xi), \end{cases} \tag{3.20}$$

as desired. □

From Equation (3.18), it follows that

$$(\widehat{\gamma}_{z,\text{MLE}}, \widehat{\sigma}_{z,\text{MLE}}) \sim \mathcal{AN}\left((\gamma,\sigma), \frac{\Lambda^{-1}}{n\bar{\Phi}(\gamma)\left(\psi_1\psi_3-\psi_2^2\right)}\begin{bmatrix} -\psi_3 & \sigma\psi_2 \\ \sigma\psi_2 & -\sigma^2\psi_1 \end{bmatrix}\right), \tag{3.21}$$

consequently $(\widehat{\theta}_{z,\text{MLE}}, \widehat{\sigma}_{z,\text{MLE}}) \sim \mathcal{AN}\left((\theta,\sigma), \frac{1}{n}\boldsymbol{S}_{z,MLE}\right)$, where

$$\boldsymbol{S}_{z,MLE} = \frac{\Lambda^{-1}}{\bar{\Phi}(\gamma)\left(\psi_1\psi_3-\psi_2^2\right)}\mathbf{D}\begin{bmatrix} -\psi_3 & \sigma\psi_2 \\ \sigma\psi_2 & -\sigma^2\psi_1 \end{bmatrix}\mathbf{D}', \quad \mathbf{D} \text{ is given by (3.11).} \tag{3.22}$$

4. MTM

MTMs are derived by following the standard method-of-moments approach, but instead of standard moments we match sample and population *trimmed* moments (or their variants), see, for example, Brazauskas *et al.* (2009). Here, we develop MTM estimation procedures for θ and σ under the transformed data scenarios given by (2.7) and (2.12).

Definition 4.1 (Method of trimmed moments – MTM). *Let $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)$ be the parameter vector to be estimated. For the random variables defined by (2.7) or (2.12), let us denote the sample and population trimmed moments as $\widehat{\mu}_j$ and $\widehat{\mu}_j(\boldsymbol{\theta})$, respectively. Let $v_{1:n} \le \cdots \le v_{n:n}$ be an ordered realization of variables (2.7) or (2.12) with qf denoted $F_V^{-1}(v \mid \boldsymbol{\theta})$ where $V \in \{Y, Z\}$, then the sample and population trimmed moments, with the trimming proportions a (lower) and b (upper), have the following expressions:*

$$\widehat{\mu}_j = \frac{1}{n-m_n-m_n^*}\sum_{i=m_n+1}^{n-m_n^*}\left[h(v_{i:n})\right]^j, \qquad j=1,\ldots,k, \tag{4.1}$$

https://doi.org/10.1017/asb.2021.4 Published online by Cambridge University Press

$$\mu_j(\boldsymbol{\theta}) = \frac{1}{1-a-b}\int_a^{1-b}\left[h(F_V^{-1}(v\,|\,\boldsymbol{\theta}))\right]^j dv, \qquad j=1,\ldots,k. \tag{4.2}$$

The trimming proportions a and b and function h are chosen by the researcher. The function h is specially chosen for "mathematical convenience" depending upon the nature of the underlying distribution function, for more details, see, for example, Brazauskas et al. (2009). *Also, integers* m_n *and* m_n^* $(0 \le m_n < n - m_n^* \le n)$ *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* $m_n = [na]$ *and* $m_n^* = [nb]$, *where* $[\cdot]$ *denotes the greatest integer part.*

The MTM estimators are then found by matching sample trimmed moments (4.1) with population trimmed 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* $\widehat{\theta}_j = g_j(\widehat{\mu}_1,\ldots,\widehat{\mu}_k)$, $1 \le j \le k$, *are, by definition, the MTM estimators of* $\theta_1,\ldots,\theta_k$. *Note that the functions* s_j *are such that* $\theta_j = g_j(\mu_1(\boldsymbol{\theta}),\ldots,\mu_k(\boldsymbol{\theta}))$.

MTM estimators belong to the class of *L*-statistics whose general asymptotic properties have been established by Chernoff *et al.* (1967). A computationally more efficient formulation has been derived by Brazauskas *et al.* (2009) and is given by Theorem 4.1.

Theorem 4.1. *Suppose an i.i.d. realization of variables (2.6) or (2.11) has been generated by cdf* $F_V(v\,|\,\boldsymbol{\theta})$ *which depending upon the data scenario equals to cdf (2.8) or (2.13), respectively. Let* $\widehat{\boldsymbol{\theta}}_{MTM} = (\widehat{\theta}_1,\ldots,\widehat{\theta}_k) = (g_1(\widehat{\mu}_1,\ldots,\widehat{\mu}_k),\ldots,g_k(\widehat{\mu}_1,\ldots,\widehat{\mu}_k))$ *denote an MTM estimator of* $\boldsymbol{\theta}$. *Then,*

$$\widehat{\boldsymbol{\theta}}_{MTM} = (\widehat{\theta}_1,\ldots,\widehat{\theta}_k) \;\; \textit{is} \;\; \mathcal{AN}\left((\theta_1,\ldots,\theta_k),\, \frac{1}{n}\mathbf{D}\boldsymbol{\Sigma}\mathbf{D}'\right), \tag{4.3}$$

where $\mathbf{D} := \left[d_{ij}\right]_{i,j=1}^k$ *is the Jacobian of the transformations* $g_1,\ldots,g_k$ *evaluated at* $\left(\mu_1(\boldsymbol{\theta}),\ldots,\mu_k(\boldsymbol{\theta})\right)$ *and* $\Sigma := \left[\sigma_{ij}^2\right]_{i,j=1}^k$ *is the variance–covariance matrix with the entries:*

$$\sigma_{ij}^2 = \frac{1}{(1-a-b)^2}\int_a^{1-b}\int_a^{1-b}\left(\min\{v,w\} - vw\right) d\left[h\left(F_V^{-1}(v)\right)\right]^j \, d\left[h\left(F_V^{-1}(w)\right)\right]^i. \tag{4.4}$$

The asymptotic performance of the newly designed estimators will be measured via ARE with respect to MLE and for two parameter case it is defined as (see, e.g., Serfling, 1980; van der Vaart, 1998):

$$ARE(\mathcal{C}, MLE) = \left(\frac{\det\left(\boldsymbol{\Sigma}_{\text{MLE}}\right)}{\det\left(\boldsymbol{\Sigma}_{\mathcal{C}}\right)}\right)^{1/2}, \tag{4.5}$$

https://doi.org/10.1017/asb.2021.4 Published online by Cambridge University Press

where $\mathbf{\Sigma}_{\text{MLE}}$ and $\mathbf{\Sigma}_{\mathcal{C}}$ are the asymptotic variance–covariance matrices of the MLE and $\mathcal{C}$ estimators, respectively, and det stands for the determinant of a square matrix. The main reason why MLE should be used as a benchmark is its optimal asymptotic performance in terms of variability (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

For payment-per-payment data, there are three different cases to be considered. From the quantile function (2.10), define $s^* = \frac{F_X(T)-F_X(t)}{1-F_X(t)}$, then we get the following arrangements:

Case 1: $0 \le a < s^* \le 1-b \le 1$ (estimation based on observed and censored data).
Case 2: $0 \le a < 1-b \le s^* \le 1$ (estimation based on observed data only).
Case 3: $0 < s^* \le a < 1-b \le 1$ (estimation based on censored data only).

In all these cases, the sample trimmed moments (4.1) can be easily computed by first empirically estimating the probability:

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

then selecting a, b, and finally choosing $h_Y(y) := y/c + t$. Note that c, t, and T are known constants.

More specifically, if both uncensored and censored sample observations participate in $\widehat{\mu}_j$, then we end up with the first case. If the censored observations, that is, $y_i = cR$, $1 \le i \le n$ are not involved in computing $\widehat{\mu}_j$, then we end up with the second case. And, finally if $\widehat{\mu}_j$ is computed only with censored observations, then we are in the third case, but in this case the population trimmed moment μ_j is no longer a function of the parameter to be estimated. We may also rule out the third case by choosing $a_j = 0$, that is, no trimming on the left which is reasonable since the sample data set is already left-truncated at t. Further, due to space limitations and to stay with the observed data only, from now on we only focus on Case 2.

Now, let $y_1, \ldots, y_n$ be an *i.i.d.* sample of normal payment-per-payment data defined by (2.7) with qf (2.10). Then,

$$\begin{cases} \widehat{\mu}_{y,1} = \dfrac{1}{n-m_n-m_n^*}\displaystyle\sum_{i=m_n+1}^{n-m_n^*} h_Y(y_{i:n}) = \dfrac{1}{n-m_n-m_n^*}\displaystyle\sum_{i=m_n+1}^{n-m_n^*}\left(\frac{y_{i:n}}{c}+t\right), \\ \widehat{\mu}_{y,2} = \dfrac{1}{n-m_n-m_n^*}\displaystyle\sum_{i=m_n+1}^{n-m_n^*} (h_Y(y_{i:n}))^2 = \dfrac{1}{n-m_n-m_n^*}\displaystyle\sum_{i=m_n+1}^{n-m_n^*}\left(\frac{y_{i:n}}{c}+t\right)^2, \end{cases} \tag{4.7}$$

https://doi.org/10.1017/asb.2021.4 Published online by Cambridge University Press

with $m_n/n \to a$ and $m_n^*/n \to b$. With Case 2, choose $m_n^* \geq \sum_{i=1}^n \mathbb{1}\{y_i = cR\}$. The corresponding population trimmed moments (4.2) with the qf defined by (2.10) are given by:

$$\begin{aligned}
\mu_{y,1} &= \frac{1}{1-a-b}\int_a^{1-b} h_Y\left(F_Y^{-1}(s)\right)\, ds \\
&= \frac{1}{1-a-b}\int_a^{1-b} \left[F_X^{-1}\big(s+(1-s)F_X(t\,|\,\boldsymbol{\theta})|\boldsymbol{\theta}\big)\right]\, ds, \\
&= \theta + \sigma c_{y,1}, \qquad (4.8)\\
\mu_{y,2} &= \frac{1}{1-a-b}\int_a^{1-b} \left[h_Y\left(F_Y^{-1}(s)\right)\right]^2\, ds \\
&= \frac{1}{1-a-b}\int_a^{1-b} \left[F_X^{-1}\big(s+(1-s)F_X(t\,|\,\boldsymbol{\theta})|\boldsymbol{\theta}\big)\right]^2\, ds \\
&= \theta^2 + 2\theta\sigma c_{y,1} + \sigma^2 c_{y,2}, \qquad (4.9)
\end{aligned}$$

where for $k = 1, 2$;

$$c_{y,k} \equiv c_{y,k}(\Phi, a, b, t) := \frac{1}{1-a-b}\int_a^{1-b} \left[\Phi^{-1}\left(s+(1-s)\Phi(\gamma)\right)\right]^k\, ds. \qquad (4.10)$$

Note 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., Brazauskas *et al.*, 2009). Equating $\mu_{y,1} = \widehat{\mu}_{y,1}$ and $\mu_{y,2} = \widehat{\mu}_{y,2}$ yield the implicit system of equations to be solved for θ and σ:

$$\begin{cases} \theta &= \widehat{\mu}_{y,1} - c_{y,1}\sigma =: g_1(\widehat{\mu}_{y,1}, \widehat{\mu}_{y,2}), \\ \sigma &= \sqrt{(\widehat{\mu}_{y,2} - \widehat{\mu}_{y,1}^2)/(c_{y,2} - c_{y,1}^2)} =: g_2(\widehat{\mu}_{y,1}, \widehat{\mu}_{y,2}). \end{cases} \qquad (4.11)$$

The system of Equations (4.11) can be solved for $\widehat{\theta}_{y,\text{MTM}}$ and $\widehat{\sigma}_{y,\text{MTM}}$ using an iterative numerical method with the initializing values:

$$\sigma_{\text{start}} = \sqrt{\widehat{\mu}_{y,2} - \widehat{\mu}_{y,1}^2} \quad \text{and} \quad \theta_{\text{start}} = \widehat{\mu}_{y,1}. \qquad (4.12)$$

From Theorem 4.1, the entries of the variance–covariance matrix $\boldsymbol{\Sigma}_y$ calculated using (4.4) are

$$\sigma_{11}^2 = \sigma^2 c_{y,1}^*,\ \ \sigma_{12}^2 = 2\theta\sigma^2 c_{y,1}^* + 2\sigma^3 c_{y,2}^* \text{ and } \sigma_{22}^2 = 4\theta^2\sigma^2 c_{y,1}^* + 8\theta\sigma^3 c_{y,2}^* + 4\sigma^4 c_{y,3}^*,$$

https://doi.org/10.1017/asb.2021.4 Published online by Cambridge University Press

where the expressions for $c^*_{y,k}$, $k = 1, 2, 3$ are listed in Appendix A. For $k = 1, 2$; it follows evidently that

$$\begin{cases} \frac{\partial c_{y,k}}{\partial \theta} = -\frac{2^{k-1}\phi(\gamma)}{\sigma(1-a-b)} \int_a^{1-b} \frac{(1-s)\left[\Phi^{-1}(s+(1-s)\Phi(\gamma))\right]^{k-1}}{\phi\left[\Phi^{-1}(s+(1-s)\Phi(\gamma))\right]} ds, \\ \frac{\partial c_{y,k}}{\partial \sigma} = -\frac{2^{k-1}(t-\theta)\phi(\gamma)}{\sigma^2(1-a-b)} \int_a^{1-b} \frac{(1-s)\left[\Phi^{-1}(s+(1-s)\Phi(\gamma))\right]^{k-1}}{\phi\left[\Phi^{-1}(s+(1-s)\Phi(\gamma))\right]} ds. \end{cases} \tag{4.13}$$

For $k = 1, 2$, let us denote

$$\theta_{\mu_{y,k}} := \left.\frac{\partial g_1}{\partial \widehat{\mu}_{y,k}}\right|_{(\mu_{y,1};\mu_{y,2})} \quad \text{and} \quad \sigma_{\mu_{y,k}} := \left.\frac{\partial g_2}{\partial \widehat{\mu}_{y,k}}\right|_{(\mu_{y,1};\mu_{y,2})}.$$

Consider the following more notations:

$$f_{11}(\theta,\sigma) := 1 + \sigma\frac{\partial c_{y,1}}{\partial \theta}, \qquad f_{12}(\theta,\sigma) := c_{y,1} + \sigma\frac{\partial c_{y,1}}{\partial \sigma},$$
$$f_{21}(\theta,\sigma) := \frac{\partial c_{y,2}}{\partial \theta} - 2c_{y,1}\frac{\partial c_{y,1}}{\partial \theta}, \qquad f_{22}(\theta,\sigma) := \frac{\partial c_{y,2}}{\partial \sigma} - 2c_{y,1}\frac{\partial c_{y,1}}{\partial \sigma}.$$

The entries of the matrix $\boldsymbol{D}_y$, given by Theorem 4.1, are found by implicitly differentiating the functions g_j (with multivariate chain rule) from Equations (4.11) with the help of Equations (4.13):

$$d_{11} = \theta_{\mu_{y,1}} = \frac{1 - f_{12}\sigma_{\mu_{y,1}}}{f_{11}} = \frac{1 - f_{12}d_{21}}{f_{11}},$$

$$d_{12} = \theta_{\mu_{y,2}} = -\frac{f_{12}\sigma_{\mu_{y,2}}}{f_{11}} = -\frac{f_{12}d_{22}}{f_{11}},$$

$$d_{21} = \sigma_{\mu_{y,1}} = -\frac{K\left[2f_{11}\mu_{y,1}(c_{y,2} - c^2_{y,1}) + f_{21}(\mu_{y,2} - \mu^2_{y,1})\right]}{f_{11}(c_{y,2} - c^2_{y,1})^2 + K(\mu_{y,2} - \mu^2_{y,1})(f_{11}f_{22} - f_{12}f_{21})},$$

$$d_{22} = \sigma_{\mu_{y,2}} = \frac{Kf_{11}(c_{y,2} - c^2_{y,1})}{f_{11}(c_{y,2} - c^2_{y,1})^2 + K(\mu_{y,2} - \mu^2_{y,1})(f_{11}f_{22} - f_{12}f_{21})},$$

where $K := \frac{1}{2}\sqrt{\frac{c_{y,2} - c^2_{y,1}}{\mu_{y,2} - \mu^2_{y,1}}}$. Hence, the asymptotic result (4.3) becomes

$$(\widehat{\theta}_{\text{y,MTM}}, \widehat{\sigma}_{\text{y,MTM}}) \sim \mathcal{AN}\left((\theta,\sigma), n^{-1}\boldsymbol{S}_{\text{y,MTM}}\right), \ \boldsymbol{S}_{\text{y,MTM}} := \boldsymbol{D}_y\boldsymbol{\Sigma}_y\boldsymbol{D}_y'. \tag{4.14}$$

From (3.11) and (4.14), it follows that

$$\text{ARE}\left(\left(\widehat{\theta}_{\text{y,MTM}}, \widehat{\sigma}_{\text{y,MTM}}\right), \left(\widehat{\theta}_{\text{y,MLE}}, \widehat{\sigma}_{\text{y,MLE}}\right)\right) = \left(\det\left(\boldsymbol{S}_{\text{y,MLE}}\right)/\det\left(\boldsymbol{S}_{\text{y,MTM}}\right)\right)^{0.5}. \tag{4.15}$$

https://doi.org/10.1017/asb.2021.4 Published online by Cambridge University Press

TABLE 1

ARE $\left(\left(\widehat{\theta}_{Y,MTM}, \widehat{\sigma}_{Y,MTM}\right), \left(\widehat{\theta}_{Y,MLE}, \widehat{\sigma}_{Y,MLE}\right)\right)$ FOR FIXED $d = 4$ AND SELECTED a AND b AND VARIOUS CHOICES OF RIGHT-CENSORING POINT u FROM LN(1, 5, 3).

	b (when $u = 2 \times 10^5$)					b (when $u = 2.4 \times 10^4$)				b (when $u = 8.5 \times 10^3$)		
a	0.01	0.05	0.10	0.15	0.25	0.05	0.10	0.15	0.25	0.10	0.15	0.25
0	0.987	0.904	0.821	0.747	0.616	0.960	0.871	0.793	0.654	0.934	0.850	0.701
0.05	0.984	0.904	0.821	0.749	0.620	0.959	0.872	0.795	0.658	0.935	0.852	0.705
0.10	0.971	0.893	0.813	0.742	0.615	0.948	0.863	0.788	0.653	0.925	0.844	0.700
0.15	0.948	0.874	0.796	0.726	0.602	0.927	0.845	0.771	0.639	0.906	0.827	0.685
0.25	0.885	0.816	0.742	0.676	0.556	0.867	0.788	0.718	0.590	0.845	0.769	0.633

For some selected trimming proportions a and b, numerical values of the AREs given by Equation (4.15) are provided in Table 1.

Note 4.1. *If we empirically estimate $\Phi(\gamma)$ in Equation (4.10) using σ_{start} and θ_{start} given by (4.12), then $c_{y,k}$ given by (4.10) is no longer a function of the parameters θ and σ to be estimated and hence the explicit estimated values of θ and σ are given by the system (4.11) which is equivalent to the complete data scenario. This technique could be implemented to eliminate all the computational complexity we discussed in this section.* □

4.2. Payments Z

From (2.12), it follows that payment Z is left- and right-censored form of complete random variable X. Thus, possible permutations between a, b, and their positioning with respect to $F_X(t)$ and $F_X(T)$ have to be taken into account, since the expressions for σ_{ij}^2 given by (4.4) with qf (2.16) depend on the six possible permutations among a, b, $F_X(t)$, and $F_X(T)$:

1. $0 \le a < 1 - b \le F_X(t) < F_X(T) \le 1$.
2. $0 \le a \le F_X(t) < 1 - b \le F_X(T) \le 1$.
3. $0 \le a \le F_X(t) < F_X(T) \le 1 - b \le 1$.
4. $0 \le F_X(t) < F_X(T) \le a < 1 - b \le 1$.
5. $0 \le F_X(t) \le a < F_X(T) \le 1 - b \le 1$.
6. $0 \le F_X(t) \le a < 1 - b \le F_X(T) \le 1$.

Among these 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 trimmed 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 most sense and simplifies the estimation procedure significantly because it uses the available data in the most effective way. Moreover, the MTM 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.

https://doi.org/10.1017/asb.2021.4 Published online by Cambridge University Press

Note 4.2. *The MTM estimators with* $a > 0$ *and* $b > 0$ $(0 \le F_X(t) \le a < 1 - b \le F_X(T) \le 1)$ *are globally robust with the* lower *and* upper *breakdown points given by* LBP $= a$ *and* UBP $= b$*, respectively. The robustness of such estimators against small or large outliers comes from the fact that in the computation of estimates the influence of the order statistics with the index less than* $n \times$ LBP *or higher than* $n \times (1 -$ UBP$)$ *is limited. For more details on* LBP *and* UBP*, see Brazauskas and Serfling* (2000) *and Serfling* (2002). □

For practical data analysis purpose, standard empirical estimates of $F_X(t)$ and $F_X(T)$ provide guidance about the choice of a and $1 - b$ and are chosen according to

$$F_n(t) \le a < 1 - b \le F_n(T), \quad \text{where} \quad F_n \text{ is the empirical cdf.} \tag{4.16}$$

Further, define $h_Z(z) := z/c + t$. Now, consider an observed *i.i.d.* sample $z_1, \ldots, z_n$ defined by (2.16). Let $z_{1:n}, \ldots, z_{n:n}$ be the corresponding ordered statistics. Then the sample trimmed moments given by (4.1) are given by:

$$\begin{cases} \widehat{\mu}_{z,1} = \dfrac{1}{n - m_n - m_n^*} \displaystyle\sum_{i=m_n+1}^{n-m_n^*} h_Z(z_{i:n}) = \dfrac{1}{n - m_n - m_n^*} \displaystyle\sum_{i=m_n+1}^{n-m_n^*} \left(\dfrac{z_{i:n}}{c} + t \right), \\ \widehat{\mu}_{z,2} = \dfrac{1}{n - m_n - m_n^*} \displaystyle\sum_{i=m_n+1}^{n-m_n^*} (h_Z(z_{i:n}))^2 = \dfrac{1}{n - m_n - m_n^*} \displaystyle\sum_{i=m_n+1}^{n-m_n^*} \left(\dfrac{z_{i:n}}{c} + t \right)^2, \end{cases} \tag{4.17}$$

with $m_n/n \to a$ and $m_n^*/n \to b$. With Case 6, choose $m_n \ge \sum_{i=1}^n \mathbb{1}\{z_i = 0\}$ and $m_n^* \ge \sum_{i=1}^n \mathbb{1}\{z_i = cR\}$. By assuming the most general case that $0 \le F_X(t) \le a < 1 - b \le F_X(T) \le 1$, the corresponding population trimmed moments (4.2) with the qf (2.16) are given by:

$$\begin{aligned} \mu_{z,1} &:= \frac{1}{1 - a - b} \int_a^{1-b} h_Z\left(F_Z^{-1}(s)\right) ds \\ &= \frac{1}{1 - a - b} \int_a^{1-b} F_X^{-1}(s \mid \boldsymbol{\theta}) \, ds \\ &= \theta + \sigma c_1, \\ \mu_{z,2} &:= \frac{1}{1 - a - b} \int_a^{1-b} \left[h_Z\left(F_Z^{-1}(s)\right)\right]^2 ds \\ &= \frac{1}{1 - a - b} \int_a^{1-b} \left[F_X^{-1}(s \mid \boldsymbol{\theta})\right]^2 ds \\ &= \theta^2 + 2\theta\sigma c_1 + \sigma^2 c_2, \end{aligned}$$

where $c_k \equiv c_{y,k}$, $1 \le k \le 4$, given by (4.10) with $\gamma = -\infty$ and are listed in Appendix A. Thus, with the assumption $0 \le F_X(t) \le a < 1 - b \le F_X(T) \le 1$,

https://doi.org/10.1017/asb.2021.4 Published online by Cambridge University Press

TABLE 2

ARE $\left(\left(\widehat{\theta}_{\text{z,MTM}}, \widehat{\sigma}_{\text{z,MTM}}\right), \left(\widehat{\theta}_{\text{z,MLE}}, \widehat{\sigma}_{\text{z,MLE}}\right)\right)$ FOR FIXED $d = 4$ AND SELECTED a AND b AND VARIOUS CHOICES OF RIGHT-CENSORING POINT u FROM LN(1, 5, 3).

	b (when $u = 2 \times 10^5$)					b (when $u = 2.4 \times 10^4$)				b (when $u = 8.5 \times 10^3$)		
a	0.01	0.05	0.10	0.15	0.25	0.05	0.10	0.15	0.25	0.10	0.15	0.25
0.10	0.948	0.900	0.844	0.793	0.695	0.933	0.876	0.822	0.720	0.914	0.858	0.752
0.15	0.891	0.846	0.793	0.742	0.647	0.877	0.822	0.770	0.671	0.858	0.804	0.701
0.25	0.786	0.745	0.695	0.647	0.556	0.772	0.720	0.671	0.577	0.752	0.701	0.602
0.49	0.550	0.516	0.471	0.428	0.343	0.535	0.489	0.444	0.355	0.510	0.464	0.371

this case translates to the complete data case which is fully investigated by Brazauskas *et al.* (2009) and the the MTM estimators of θ and σ are

$$\begin{cases} \widehat{\theta}_{\text{z,MTM}} = \widehat{\mu}_{z,1} - c_1\, \widehat{\sigma}_{\text{z,MTM}} \\ \widehat{\sigma}_{\text{z,MTM}} = \sqrt{\left(\widehat{\mu}_{z,2} - \widehat{\mu}_{z,1}^2\right) / \left(c_2 - c_1^2\right)} \end{cases} \tag{4.18}$$

And, the corresponding ARE is given by:

$$\text{ARE}\left(\left(\widehat{\theta}_{\text{z,MTM}}, \widehat{\sigma}_{\text{z,MTM}}\right), \left(\widehat{\theta}_{\text{z,MLE}}, \widehat{\sigma}_{\text{z,MLE}}\right)\right) = \left(\det\left(\boldsymbol{S}_{\text{z,MLE}}\right) / \det\left(\boldsymbol{S}_{\text{z,MTM}}\right)\right)^{0.5}, \tag{4.19}$$

where

$$\boldsymbol{S}_{\text{z,MTM}} := \frac{\sigma^2}{\left(c_2 - c_1^2\right)^2} \begin{bmatrix} c_1^* c_2^2 - 2c_1 c_2 c_2^* + c_1^2 c_3^* & -c_1^* c_1 c_2 + c_2 c_2^* + c_1^2 c_2^* - c_1 c_3^* \\ -c_1^* c_1 c_2 + c_2 c_2^* + c_1^2 c_2^* - c_1 c_3^* & c_1^* c_1^2 - 2c_1 c_2^* + c_3^* \end{bmatrix},$$

and the expressions for $c_k^*,\ k = 1, 2, 3$, as functions of a, b, c_1, c_2, c_3, and c_4 are such that $c_k^* \equiv c_{y,k}^*$ with $\gamma = -\infty$ and are listed in Appendix A.

Numerical values of the AREs given by Equation (4.19) are summarized in Table 2 for some selected values of a and b.

5. SPECIAL CASES

Singly left-truncated, that is, $u \to \infty$, which is equivalent to $T \to \infty$, sample data set is very common in insurance industries as well as in operational risk modeling (see, e.g., Ergashev *et al.*, 2016). So, in this section, we summarize the analogous results from Sections 3 and 4 for singly left-truncated lognormal sample data. Clearly, for $T \to \infty$ implies $\Omega_2 = 0,\ \Omega_{y,2} = 0,$ and $\Omega_{z,2} = 0$. Thus, the system of Equations (3.4) becomes

$$\begin{cases} \sigma\left(\Omega_{y,1} - \gamma\right) - c^{-1}\widehat{\mu}_{y,1} = 0, \\ \sigma^2\left(1 - \gamma(\Omega_{y,1} - \gamma)\right) - c^{-2}\widehat{\mu}_{y,2} = 0, \end{cases} \tag{5.1}$$

https://doi.org/10.1017/asb.2021.4 Published online by Cambridge University Press

The system (5.1) can be rearranged as a nonlinear equation of γ only as:

$$G(\gamma) = \delta_y \quad \text{where} \quad G(\gamma) := \frac{1}{\Omega_{y,1} - \gamma}\left[\frac{1}{\Omega_{y,1} - \gamma} - \gamma\right],$$
$$\Omega_{y,1} = \frac{\phi(\gamma)}{\bar{\Phi}(\gamma)}, \quad \text{and } \delta_y := \frac{\widehat{\mu}_{y,2}}{\widehat{\mu}_{y,1}^2} \tag{5.2}$$

It is important to note that Equation (5.2) is a nonlinear equation to be solved for γ, but the system (3.4) should be solved simultaneously both for γ and σ. It was shown by Barrow and Cohen (1954) that the function $G(\cdot)$ is monotonically increasing on the real line $\mathbb{R}$ with $\lim_{\gamma \to -\infty} G(\gamma) = 1$ and $\lim_{\gamma \to \infty} G(\gamma) = 2$. Thus, from Equation (5.2), we have

Theorem 5.1. *The equation $G(\gamma) = \delta_y$ has a unique solution if and only if $1 < \delta_y < 2$ and the corresponding unique solution $(\widehat{\gamma}_{y,MLE}, \widehat{\sigma}_{y,MLE})$ of the system (5.1) is given by:*

$$(\widehat{\gamma}_{y,MLE}, \widehat{\sigma}_{y,MLE}) = \left(G^{-1}(\delta_y), \frac{c^{-1}\widehat{\mu}_{y,1}}{\frac{\phi(\widehat{\gamma}_{y,MLE})}{\bar{\Phi}(\widehat{\gamma}_{y,MLE})} - \widehat{\gamma}_{y,MLE}}\right), \tag{5.3}$$

and consequently $\widehat{\theta}_{y,MLE} = t - \widehat{\sigma}_{y,MLE}\,\widehat{\gamma}_{y,MLE}$.

Further, it has been established by Ergashev *et al.* (2016) that the unique solution $(\widehat{\gamma}_{y,\mathrm{MLE}}, \widehat{\sigma}_{y,\mathrm{MLE}})$ given by Theorem 5.1 is, in fact, the point of global maximum of the log-likelihood surface given by Equation (3.1) with the adjustment of $T \to \infty$. Therefore, it follows that $1 < \delta_y < 2$ is the necessary and sufficient condition for the system (5.1) to have the unique global MLE solution $(\widehat{\gamma}_{y,\mathrm{MLE}}, \widehat{\sigma}_{y,\mathrm{MLE}})$.

The asymptotic result (3.10) takes the form:

$$(\widehat{\gamma}_{y,\mathrm{MLE}}, \widehat{\sigma}_{y,\mathrm{MLE}}) \sim \mathcal{AN}\left((\gamma, \sigma), \frac{1}{n\left(r_1 r_3 - r_2^2\right)}\begin{bmatrix} -r_3 & \sigma r_2 \\ \sigma r_2 & -\sigma^2 r_1 \end{bmatrix}\right), \tag{5.4}$$

where

$$r_1(\gamma, \xi) = -\left[1 + \gamma\Omega_1 - \frac{\phi(\gamma)}{\bar{\Phi}(\gamma)}\Omega_1\right], \quad r_2(\gamma, \xi) = \Omega_1 - \gamma,$$
$$r_3(\gamma, \xi) = -\left[2 - \gamma(\Omega_1 - \gamma)\right]. \tag{5.5}$$

Similarly, for payment-per-loss data scenario, the system (3.16) takes the form:

$$\begin{cases} \sigma\left(\Omega_{z,1} - \gamma\right) - c^{-1}\widehat{\mu}_{z,1} = 0, \\ \sigma^2\left(1 - \gamma(\Omega_{z,1} - \gamma)\right) - c^{-2}\widehat{\mu}_{z,2} = 0, \end{cases} \tag{5.6}$$

https://doi.org/10.1017/asb.2021.4 Published online by Cambridge University Press

and a nonlinear equation to be solved for γ is given by:

$$\frac{1}{\Omega_{z,1}-\gamma}\left[\frac{1}{\Omega_{z,1}-\gamma}-\gamma\right]=\frac{\widehat{\mu}_{z,2}}{\widehat{\mu}_{z,1}^2},\quad \text{where}\quad \Omega_{z,1}=\frac{n_0}{n_1}\,\frac{\phi(\gamma)}{\Phi(\gamma)}. \tag{5.7}$$

Clearly, the empirical estimates of $\Phi(\gamma)$ and $\bar{\Phi}(\gamma)$ are n_0/n and n_1/n, respectively. Thus, if we replace the ratios n_0/n and n_1/n by $\Phi(\gamma)$ and $\bar{\Phi}(\gamma)$, respectively, the nonlinear Equation (5.7) takes the form:

$$G(\gamma)=\delta_z,\quad \text{where}\quad \delta_z:=\frac{\widehat{\mu}_{z,2}}{\widehat{\mu}_{z,1}^2}. \tag{5.8}$$

Then with the condition $1<\delta_z<2$, Theorem 5.1 is applicable again to get the unique global MLE solution of the system (5.6) with Equation (5.8) as:

$$(\widehat{\gamma}_{\text{z,MLE}},\widehat{\sigma}_{\text{z,MLE}})=\left(G^{-1}(\delta_z),\frac{c^{-1}\widehat{\mu}_{z,1}}{\frac{\phi(\widehat{\gamma}_{\text{z,MLE}})}{\bar{\Phi}(\widehat{\gamma}_{\text{z,MLE}})}-\widehat{\gamma}_{\text{z,MLE}}}\right), \tag{5.9}$$

and consequently $\widehat{\theta}_{\text{z,MLE}}=t-\widehat{\sigma}_{\text{z,MLE}}\,\widehat{\gamma}_{\text{z,MLE}}$.

Finally, the asymptotic result (3.21) takes the form:

$$(\widehat{\gamma}_{\text{z,MLE}},\widehat{\sigma}_{\text{z,MLE}})\sim\mathcal{AN}\left((\gamma,\sigma),\frac{1}{n\bar{\Phi}(\gamma)\left(\psi_1\psi_3-\psi_2^2\right)}\begin{bmatrix}-\psi_3 & \sigma\psi_2\\ \sigma\psi_2 & -\sigma^2\psi_1\end{bmatrix}\right), \tag{5.10}$$

where

$$\begin{aligned}&\psi_1(\gamma,\xi)=-\left[1+\gamma\Omega_1+\frac{\phi(\gamma)}{\Phi(\gamma)}\Omega_1\right],\quad \psi_2(\gamma,\xi)=\Omega_1-\gamma,\\ &\psi_3(\gamma,\xi)=-\left[2-\gamma(\Omega_1-\gamma)\right].\end{aligned} \tag{5.11}$$

Now, we discuss the analogous results for MTM approach for singly left truncated lognormal data scenarios. For payments Y, since $T\to\infty$, then the value of s^* as defined in Section 4.1 is simply 1 which is Case 2 as discussed in Section 4.1 and the entire theory developed in Section 4.1 is for Case 2.

Similarly, for payments Z with $T\to\infty$, all the six possible permutations mentioned in Section 4.2 shrank to only three Cases 1, 2, and 6. Further, with a goal of finding the most resistant estimators, we discussed only Case 6 in Section 4.2 and the same results also hold for $T\to\infty$.

6. Simulation Study

This section supplements the theoretical results we developed in Section 4 via simulation. The main goal is to access the size of the sample such that the estimators are free from bias (given that the estimators are asymptotically

https://doi.org/10.1017/asb.2021.4 Published online by Cambridge University Press

unbiased), justify the asymptotic normality, and their finite sample relative efficiencies (REs) to reach the corresponding AREs. To compute RE of MTM estimators, we use MLE as a benchmark. Thus, the definition of ARE given by Equation (4.5) for finite sample performance translates to:

$$RE(\text{MTM, MLE}) = \frac{\text{asymptotic variance of MLE estimator}}{\text{small-sample variance of a competing MTM estimator}}, \tag{6.1}$$

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

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

The design of the simulation is as below:

1. *Ground-up distribution*, that is, a model for W given by (2.1): LN $(w_0 = 1, \theta = 5, \sigma = 3)$.
2. *Coinsurance rate*: $c = 1$.
3. *Truncation and censoring thresholds for both variables Y and Z*:
   $d = 4$ ($\approx 10\%$ left truncation under LN $(w_0 = 1, \theta = 5, \sigma = 3)$);
   $u = 2 \times 10^5$ ($\approx 1\%$ right-censoring under LN $(w_0 = 1, \theta = 5, \sigma = 3)$);
   $u = 2.4 \times 10^4$ ($\approx 4.5\%$ right-censoring under LN $(w_0 = 1, \theta = 5, \sigma = 3)$).
4. *Estimators of θ and σ*:
   - MLE (no trimming on either tail)
   - MTM with different left and right trimming proportions a and b, respectively, satisfying
     (a) $0 \le a < 1 - b \le s^* \le 1$, for payment-per-payment variable Y;
     (b) $0 \le F_X(t) \le a < 1 - b \le F_X(T) \le 1$, for payment-per-loss variable Z.
5. *Sample size*: $n = 100,\ 250,\ 500,\ 1000$.

For a selected random variable (Y or Z), we generate 10^4 samples of a specified size n using the corresponding quantile function ((2.10) for Y and (2.16) for Z). For each sample, depending on the underlying scenario, we estimate the parameters θ and σ using MLE and MTM from both loss variables Y and Z.

Simulation results are recorded in Tables 3 (for payment-per-payment variable Y) and 4 (for payment-per-loss variable Z). The entries are mean values based on 10^4 samples. The columns corresponding to $n \to \infty$ represent analytic results and are found in Section 4, not from simulations.

As can be seen from Table 3 that for $n = 100,\ 250$, both estimators $\widehat{\theta}$ and $\widehat{\sigma}$ successfully estimate their corresponding parameter θ and σ, respectively, with less than $\pm 2\%$ of the relative bias. The more important observation is that they practically become unbiased as soon as $n \ge 500$. The situation is even better in Table 4. That is, both estimators $\widehat{\theta}$ and $\widehat{\sigma}$ are practically unbiased as soon as $n \ge 100$. On the other hand, as seen in Tables 3 and 4, the finite REs

https://doi.org/10.1017/asb.2021.4 Published online by Cambridge University Press

TABLE 3

LOGNORMAL PAYMENT-PER-PAYMENT ACTUARIAL LOSS SCENARIO, $LN(w_0 = 1, \theta = 5, \sigma = 3)$ WITH $d = 4$ AND TWO SELECTED VALUES OF RIGHT-CENSORING POINT u. THE ENTRIES ARE MEAN VALUES BASED ON 10,000 SAMPLES.

	Proportion		$n = 100$		$n = 250$		$n = 500$		$n = 1000$		$n \to \infty$	
	a	b	$\widehat{\theta}/\theta$	$\widehat{\sigma}/\sigma$	$\widehat{\theta}/\theta$	$\widehat{\sigma}/\sigma$	$\widehat{\theta}/\theta$	$\widehat{\sigma}/\sigma$	$\widehat{\theta}/\theta$	$\widehat{\sigma}/\sigma$	$\widehat{\theta}/\theta$	$\widehat{\sigma}/\sigma$
			Mean values of $\widehat{\theta}/\theta$ and $\widehat{\sigma}/\sigma$									
	MLE		0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1	1
	0.05	0.05	0.99	1.01	0.99	1.02	1.00	1.00	1.00	1.00	1	1
	0.10	0.10	0.98	1.02	0.99	1.01	1.00	1.00	1.00	1.00	1	1
	0.15	0.15	0.98	1.02	0.99	1.02	1.00	1.00	1.00	1.00	1	1
	0.00	0.05	0.99	1.01	1.00	1.00	1.00	1.00	1.00	1.00	1	1
	0.00	0.10	0.99	1.01	0.99	1.00	1.00	1.00	1.00	1.00	1	1
$u = 2 \times 10^5$	0.00	0.25	0.98	1.02	0.99	1.01	1.00	1.00	1.00	1.00	1	1
			Finite sample efficiencies of MTMs relative to MLEs									
	MLE		0.93		0.99		0.98		1.00		1	
	0.05	0.05	0.80		0.85		0.89		0.90		0.904	
	0.10	0.10	0.69		0.78		0.79		0.80		0.813	
	0.15	0.15	0.57		0.66		0.70		0.71		0.726	
	0.00	0.05	0.80		0.87		0.89		0.90		0.904	
	0.00	0.10	0.70		0.79		0.81		0.81		0.821	
	0.00	0.25	0.41		0.56		0.59		0.60		0.616	
			Mean values of $\widehat{\theta}/\theta$ and $\widehat{\sigma}/\sigma$									
	MLE		0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1	1
	0.10	0.10	0.98	1.02	0.99	1.01	1.00	1.00	1.00	1.00	1	1
	0.15	0.15	0.98	1.02	0.99	1.02	1.00	1.00	1.00	1.00	1	1
	0.00	0.10	0.99	1.01	0.99	1.00	1.00	1.00	1.00	1.00	1	1
	0.05	0.10	0.98	1.01	0.99	1.01	1.00	1.00	1.00	1.00	1	1
$u = 2.4 \times 10^4$	0.00	0.25	0.98	1.02	0.99	1.01	1.00	1.00	1.00	1.00	1	1
			Finite sample efficiencies of MTMs relative to MLEs.									
	MLE		0.93		0.96		0.98		1.00		1	
	0.10	0.10	0.76		0.81		0.84		0.86		0.863	
	0.15	0.15	0.65		0.69		0.74		0.77		0.771	
	0.00	0.10	0.77		0.83		0.85		0.87		0.871	
	0.05	0.10	0.77		0.80		0.85		0.87		0.872	
	0.00	0.25	0.51		0.61		0.62		0.64		0.654	

Note: The standard errors for the entire entries in this table are reported to be ≤ 0.002.

https://doi.org/10.1017/asb.2021.4 Published online by Cambridge University Press

TABLE 4

LOGNORMAL PAYMENT-PER-LOSS ACTUARIAL LOSS SCENARIO, $LN(w_0 = 1, \theta = 5, \sigma = 3)$ WITH $d = 4$ AND TWO SELECTED VALUES OF RIGHT-CENSORING POINT u. THE ENTRIES ARE MEAN VALUES BASED ON 10,000 SAMPLES.

	Proportion		$n = 100$		$n = 250$		$n = 500$		$n = 1000$		$n \to \infty$	
	a	b	$\widehat{\theta}/\theta$	$\widehat{\sigma}/\sigma$	$\widehat{\theta}/\theta$	$\widehat{\sigma}/\sigma$	$\widehat{\theta}/\theta$	$\widehat{\sigma}/\sigma$	$\widehat{\theta}/\theta$	$\widehat{\sigma}/\sigma$	$\widehat{\theta}/\theta$	$\widehat{\sigma}/\sigma$
			Mean values of $\widehat{\theta}/\theta$ and $\widehat{\sigma}/\sigma$									
$u = 2 \times 10^5$	MLE		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1	1
	0.10	0.10	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1	1
	0.15	0.15	1.00	1.00	1.00	1.01	1.00	1.00	1.00	1.00	1	1
	0.25	0.25	1.00	1.01	1.00	1.01	1.00	1.00	1.00	1.00	1	1
	0.10	0.05	1.00	1.01	1.00	1.00	1.00	1.00	1.00	1.00	1	1
	0.10	0.15	1.00	1.01	1.00	1.00	1.00	1.00	1.00	1.00	1	1
	0.10	0.25	1.00	1.01	1.00	1.01	1.00	1.00	1.00	1.00	1	1
			Finite sample efficiencies of MTMs relative to MLEs									
	MLE		0.97		1.01		1.01		0.99		1	
	0.10	0.10	0.89		0.86		0.88		0.85		0.844	
	0.15	0.15	0.75		0.73		0.76		0.74		0.742	
	0.25	0.25	0.55		0.55		0.57		0.55		0.556	
	0.10	0.05	0.94		0.91		0.94		0.90		0.933	
	0.10	0.15	0.84		0.81		0.83		0.80		0.822	
	0.10	0.25	0.74		0.71		0.72		0.68		0.695	
			Mean values of $\widehat{\theta}/\theta$ and $\widehat{\sigma}/\sigma$									
$u = 2.4 \times 10^4$	MLE		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1	1
	0.10	0.10	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1	1
	0.15	0.15	1.00	1.00	1.00	1.01	1.00	1.00	1.00	1.00	1	1
	0.25	0.25	1.00	1.00	1.00	1.01	1.00	1.00	1.00	1.00	1	1
	0.10	0.15	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1	1
	0.15	0.10	1.00	1.00	1.00	1.01	1.00	1.00	1.00	1.00	1	1
	0.20	0.25	1.00	1.00	1.00	1.01	1.00	1.00	1.00	1.00	1	1
			Finite-sample efficiencies of MTMs relative to MLEs.									
	MLE		0.98		0.99		0.99		1.00		1	
	0.10	0.10	0.92		0.90		0.89		0.88		0.876	
	0.15	0.15	0.78		0.76		0.77		0.76		0.770	
	0.25	0.25	0.58		0.57		0.58		0.57		0.577	
	0.10	0.15	0.87		0.85		0.84		0.83		0.822	
	0.15	0.25	0.68		0.67		0.67		0.67		0.671	
	0.20	0.25	0.63		0.62		0.62		0.62		0.623	

Note: The standard errors for the entire entries in this table are reported to be ≤ 0.001.

https://doi.org/10.1017/asb.2021.4 Published online by Cambridge University Press

are obviously biased (as expected) for small sample sizes and converge slower to their corresponding ARE levels. But what is clear is that all those REs are asymptotically unbiased.

7. Numerical examples

This section is to observe the performance of the estimation methods developed in the previous sections. For the illustration purpose only, we consider the 1500 US indemnity losses which is widely studied in actuarial literature, see, for example, Frees and Valdez (1998), Punzo *et al.* (2018), Michael *et al.* (2020). We then transform the data to fit the scenarios of payment-per-payment random variable Y and payment-per-loss random variable Z. That is, from an insurer's prospective, we consider an insurance benefit equal to the amount by which a loss exceeds US$500.00 (deductible, $d = 500$) but with a maximum benefit of US$100,000.00 (policy limit, $u = 10^5$). Further, without loss of generality, since the asymptotic variances of all the estimators investigated in this paper do not depend on the coinsurance factor c, we simply assume that $c \equiv 1$. We consider this example purely for illustrative purpose and a preliminary diagnostics, see Figure 1, shows that the lognormal distribution provides a satisfactory, thought not perfect, fit to the indemnity data. Moreover, the complete ground-up data (though the data set contains some right-censored observations) set is identified to fit $LN(w_0 = 0, \theta, \sigma)$, see, for example, Punzo *et al.* (2018) and Michael *et al.* (2020), distribution. The Kolmogorov–Smirnov (KS) test statistics (see, e.g., Klugman *et al.*, 2019, Section 15.4.1, p. 360) is computed to be 0.0266 and with the significance level of 5% the corresponding critical value is 0.0351. Therefore, the lognormal model is a plausible model for the indemnity loss data set at 5% significance level. The specified insurance contract data transformation is summarized in Table 5.

A couple of crucial points are in queue to be mentioned here before starting real data analysis. First, for payment Y, the MTM estimators as well as their asymptotic results are based on the assumption of $0 \le a < 1 - b \le s^* \le 1$. Initially, we set

$$\hat{s}^*_{\mathrm{E}} := \frac{F_n(T) - F_n(t)}{1 - F_n(t)} = n_1/n, \quad \text{where } F_n \text{ is the empirical cdf,}$$

and choose b such that $1 - b \le \hat{s}^*_{\mathrm{E}}$. But after estimating $\widehat{\theta}_{\mathrm{y,MTM}}$ and $\widehat{\sigma}_{\mathrm{y,MTM}}$, it could be the case that

$$\hat{s}^*_{\mathrm{P}} := \frac{\widehat{F}_X(T) - \widehat{F}_X(t)}{1 - \widehat{F}_X(t)} < 1 - b, \quad \text{where } \widehat{F}_X \text{ is the estimated parametric cdf,}$$

which is unsatisfactory for statistical inferences as it violates the assumption (Case 2 from Section 4.1). Therefore, the value of b should be chosen

https://doi.org/10.1017/asb.2021.4 Published online by Cambridge University Press

TABLE 5

PAYMENT-PER-PAYMENT, Y AND PAYMENT-PER-LOSS, Z TRANSFORMATION SCENARIOS FOR THE US INDEMNITY LOSSES.

Scenario	Left truncation		Right censorship		n_0	n_1	n_2	n	R
	LN – d	Normal – t	LN – u	Normal – T					
Payment Y	500	6.2146	10^5	11.5129	–	1299	152	1451	5.2983
Payment Z	500	6.2146	10^5	11.5129	49	1299	152	1451	5.2983

Note: LN stands for Lognormal, $t = \log(d - w_0)$, and $T = \log(u - w_0)$.

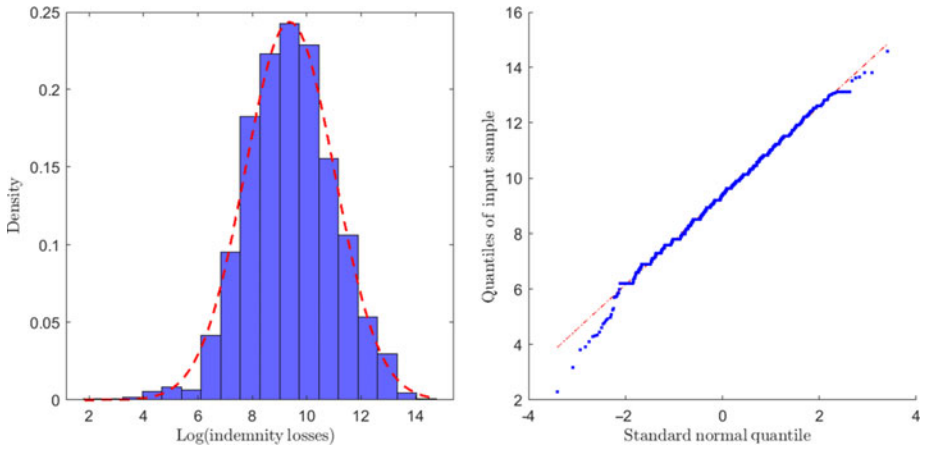


FIGURE 1: Histogram of log-transformed indemnity losses with the fitted normal density (left panel) and the corresponding quantile–quantile plot (right panel).

dynamically such that $1 - b \le \min\{\hat{s}_{\mathrm{E}}^*, \hat{s}_{\mathrm{P}}^*\}$. Similarly, for payment Z, the left and right trimming proportions a and b, respectively, should be chosen dynamically as:

$$\max\{F_n(t), \widehat{F}_X(t)\} \le a \quad \text{and} \quad 1 - b \le \min\{F_n(T), \widehat{F}_X(T)\}.$$

The trimming proportions in Tables 6 and 7 are chosen accordingly.

For payment-per-payment data scenario, the MLE system of Equations (3.4) takes the form:

$$\begin{cases} \sigma\left(\Omega_{y,1} - \Omega_{y,2} - \gamma\right) - 2.9762 = 0, \\ \sigma^2\left(1 - \gamma(\Omega_{y,1} - \Omega_{y,2} - \gamma) - \frac{5.2983\Omega_{y,2}}{\sigma}\right) - 10.3394 = 0, \end{cases} \tag{7.1}$$

https://doi.org/10.1017/asb.2021.4 Published online by Cambridge University Press

https://doi.org/10.1017/asb.2021.4 Published online by Cambridge University Press

TABLE 6

MLE AND MTM ESTIMATORS OF θ AND σ WITH THEIR CORRESPONDING ASYMPTOTIC CONFIDENCE INTERVALS AND ESTIMATED VALUES OF AREs FOR PAYMENT Y DATA SCENARIO.

Estimators			$\widehat{\theta}$	$\widehat{\sigma}$	$\hat{s}^*$		95% CI for		$\widehat{\text{ARE}}$	KS test	
					$\hat{s}^*_{\text{E}}$	$\hat{s}^*_{\text{P}}$	θ	σ		D_y	h_y
MLE			9.43	1.59	–	–	(9.34, 9.52)	(1.52, 1.67)	1.00	0.032	0
	a	b	Condition required: $1 - b \leq \min\{\hat{s}^*_{\text{E}}, \hat{s}^*_{\text{P}}\}$.								
MTM	0.00	150/1451	9.42	1.56	0.90	0.91	(9.34, 9.51)	(1.49, 1.65)	0.94	0.034	0
	0.00	200/1451	9.42	1.55	0.90	0.91	(9.33, 9.51)	(1.47, 1.64)	0.89	0.034	0
	0.00	300/1451	9.42	1.54	0.90	0.91	(9.33, 9.50)	(1.45, 1.63)	0.80	0.034	0
	0.00	700/1451	9.37	1.47	0.90	0.93	(9.27, 9.47)	(1.35, 1.59)	0.48	0.043	1
	10/1451	150/1451	9.42	1.57	0.90	0.91	(9.33, 9.51)	(1.49, 1.65)	0.94	0.033	0
	50/1451	200/1451	9.41	1.59	0.90	0.91	(9.32, 9.50)	(1.50, 1.67)	0.89	0.030	0
	100/1451	300/1451	9.40	1.59	0.90	0.90	(9.31, 9.50)	(1.50, 1.69)	0.79	0.028	0
	650/1451	650/1451	9.26	2.09	0.90	0.85	(8.96, 9.56)	(1.56, 2.81)	0.24	0.064	1

Note: For the KS test column, D_y stands for Kolmogorov–Smirnov test statistics and $h_y \in \{0, 1\}$ stands for the corresponding decision. $h_y = 0$ means that the assumed model is plausible and $h_y = 1$ means that the model is rejected.

https://doi.org/10.1017/asb.2021.4 Published online by Cambridge University Press

TABLE 7

MLE AND MTM ESTIMATORS OF θ AND σ WITH THEIR CORRESPONDING ASYMPTOTIC CONFIDENCE INTERVALS AND ESTIMATED VALUES OF AREs FOR PAYMENT Z DATA SCENARIO.

	Estimators		$\widehat{\theta}$	$\widehat{\sigma}$	$\widehat{F}_X(t)$	$\widehat{F}_X(T)$	95% CI for		$\widehat{\text{ARE}}$	KS test	
							θ	σ		D_z	h_z
	MLE		9.39	1.64	–	–	(9.30, 9.47)	(1.58, 1.71)	1.00	0.027	0
	a	b	Condition required: $\max\{F_n(t), \widehat{F}_X(t)\} \le a$ and $1 - b \le \min\{F_n(T), \widehat{F}_X(T)\}$.								
MTM	75/1500	150/1500	9.38	1.62	0.03	0.91	(9.30, 9.47)	(1.55, 1.69)	0.92	0.027	0
	75/1500	225/1500	9.38	1.61	0.02	0.91	(9.30, 9.47)	(1.54, 1.69)	0.86	0.027	0
	75/1500	375/1500	9.38	1.60	0.02	0.91	(9.29, 9.46)	(1.53, 1.69)	0.76	0.027	0
	75/1500	750/1500	9.36	1.59	0.02	0.91	(9.26, 9.47)	(1.49, 1.70)	0.52	0.028	0
	150/1500	150/1500	9.38	1.63	0.03	0.90	(9.30, 9.47)	(1.55, 1.70)	0.86	0.026	0
	225/1500	225/1500	9.38	1.63	0.03	0.90	(9.29, 9.46)	(1.55, 1.72)	0.76	0.026	0
	375/1500	375/1500	9.38	1.61	0.02	0.91	(9.29, 9.47)	(1.50, 1.71)	0.57	0.027	0
	700/1500	700/1500	9.38	2.36	0.09	0.82	(9.23, 9.52)	(1.92, 2.91)	0.16	0.107	1

Note: $(F_n(t), F_n(T)) = (49/1500, 1348/1500) \approx (0.03, 0.90)$. For the KS Test column, D_z stands for Kolmogorov–Smirnov test statistics and $h_z \in \{0, 1\}$ stands for the corresponding decision. $h_z = 0$ means that the assumed model is plausible and $h_z = 1$ means that the model is rejected.

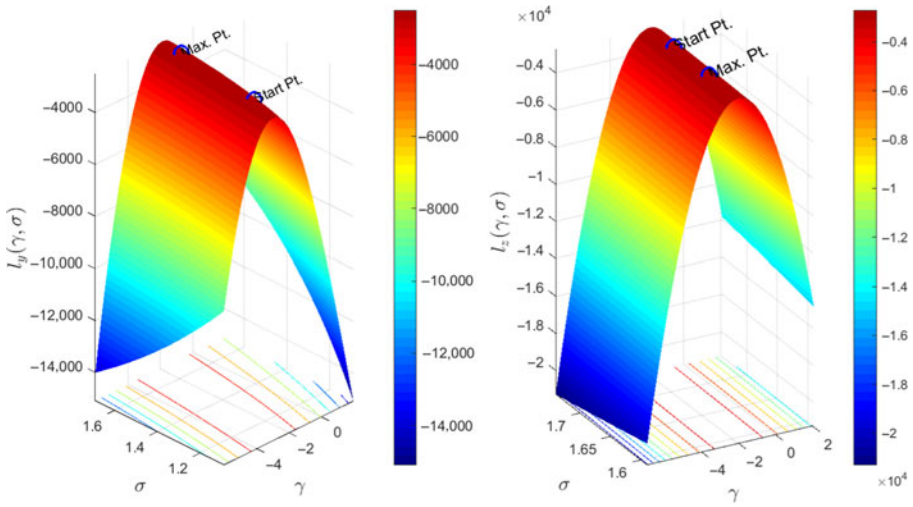


FIGURE 2: Log-likelihood surfaces for payment-per-payment (left panel) and payment-per-loss (right panel) data scenarios given by Equations (3.1) and (3.13), respectively.

with the first approximations: $(\gamma_{\text{start}}, \sigma_{\text{start}}) = (-2.4453, 1.2171)$, denoted by "Start Pt." on the corresponding log-likelihood surface given by Figure 2 (left panel), but note that $\Omega_{y,1}$ and $\Omega_{y,2}$ also depend on γ and σ. The final MLE solution $\left(\widehat{\gamma}_{y,\text{MLE}}, \widehat{\sigma}_{y,\text{MLE}}\right)$ of the system (7.1) is identified with "Max. Pt." in Figure 2 (left panel). Similarly, with different left and right trimming proportions, the system of MTM equations given by (4.11) can be solved with their respective first approximations given by (4.12). The estimated values with their corresponding asymptotic 95% confidence intervals are displayed in Table 6. The confidence intervals for σ provided are log-transformed confidence intervals (see, e.g., Klugman *et al.*, 2019, pp. 315–316), since some of the linear confidence intervals for σ include 0 which are unsatisfactory as $\sigma > 0$. For completely observed ground-up lognormal severity data, ARE $\left(\left(\widehat{\theta}_{\text{MTM}}, \widehat{\sigma}_{\text{MTM}}\right), \left(\widehat{\theta}_{\text{MLE}}, \widehat{\sigma}_{\text{MLE}}\right)\right)$ does not depend on any unknown parameters to be estimated (Brazauskas *et al.*, 2009). But for payment Y and payment Z data scenarios, the corresponding AREs, respectively, given by (4.15) and (4.19) depend on the unknown parameters θ and σ. Therefore, for comparison purpose, in Tables 6 and 7, the $\widehat{\text{ARE}}$ column is computed as a bivariate function of MLE estimated values $\widehat{\theta}_{\text{MLE}}$ and $\widehat{\sigma}_{\text{MLE}}$. As seen in Table 6, the estimated values of θ and σ via MTM are seem to underestimate the MLE estimated values $\left(\widehat{\theta}_{\text{MLE}}, \widehat{\sigma}_{\text{MLE}}\right) = (9.43, 1.59)$, but it is important to note that all the MTM estimated values of θ and σ are close enough to the corresponding MLE estimator even after removing some of the point mass at the right-censored point. It is also noticeable that with trimming pair $(a, b) = (100/1451, 300/1451)$, the estimated values of θ and σ are almost equal to the corresponding MLE estimators but still maintaining the efficiency about of 79%. These all demonstrate the robustness of MTM estimators.

https://doi.org/10.1017/asb.2021.4 Published online by Cambridge University Press

In order to conduct hypothesis test for the various fitted models, we use the KS test statistic. The KS test statistic for payments Y data set is defined as (see, e.g., Klugman *et al.*, 2019, Section 15.4.1, p. 360):

$$D_y := \max_{0<y\leq c(u-d)} \left|F_n(y) - \widehat{F}_Y(y)\right|, \tag{7.2}$$

where $\widehat{F}_Y$ is the estimated cdf given by (2.8). For each fitted model, the corresponding KS test statistics (D_y) and the decision (h_y) of the hypothesis test are given in the last two columns of Table 6 where $h_y = 0$ indicates that the lognormal model is plausible and $h_y = 1$ means that the the lognormal model is rejected at the significance level of 5%. Only one model fitted with $(a, b) = (650/1451, 650/1451) \approx (0.45, 0.45)$ is rejected which is close to median-type estimator.

Similarly, for payment-per-loss data scenario, the MLE system of Equations (3.16) takes the form:

$$\begin{cases} \sigma\left(\Omega_{z,1} - \Omega_{z,2} - \gamma\right) - 2.9762 = 0, \\ \sigma^2\left(1 - \gamma(\Omega_{z,1} - \Omega_{z,2} - \gamma) - \frac{5.2983\Omega_{z,2}}{\sigma}\right) - 10.3394 = 0, \end{cases} \tag{7.3}$$

with the first approximations: $(\gamma_{\text{start}}, \sigma_{\text{start}}) = (-1.8430, 1.6998)$. The starting point and the final MLE solution, respectively, are represented by "Start Pt." and "Max. Pt." on the log-likelihood surface $l_z(\gamma, \sigma)$ given by Figure 2 (right panel). Including the corresponding MTM calculation, a summary statistics is displayed in Table 7. In this case, as expected all the MTM estimators of θ are same which clearly demonstrate the stability of the MTM methodology. On the other hand, highly robust MTM estimated value of $\widehat{\sigma} = 2.36$ with $(a, b) = (700/1500, 700/1500) \approx (0.47, 0.47)$ seems to be more volatile and loosing about 84% of asymptotic efficiency.

Further, the KS test statistic for payments Z data set is defined by:

$$D_z := \max_{0\leq z\leq c(u-d)} \left|F_n(z) - \widehat{F}_Z(z)\right|, \tag{7.4}$$

where $\widehat{F}_Z$ is the estimated parametric cdf given by (2.14). As can be seen in Table 7, similar to 7, only one model fitted for payment-per-loss data scenario with $(a, b) = (700/1500, 700/1500) \approx (0.47, 0.47)$ is rejected with the significance level of 5%, but this model is loosing about 84% of the ARE compared to the corresponding MLE model.

In conclusion, a take-home knowledge from Tables 6 and 7 is that since the sample data set is already either truncated and/or censored, then a robust estimation procedure still maintaining high relative efficiency with respect to the corresponding MLE is simply to eliminate the point masses at the truncated and censored values. For example, $MTM(a = 0, b = 150/1451)$ and $MTM(a = 10/1451, b = 150/1451)$ from Table 6 and $MTM(a = 75/1500, b = 150/1500)$ from Table 7.

https://doi.org/10.1017/asb.2021.4 Published online by Cambridge University Press

8. Summary and future direction

In this paper, we have developed two estimation procedures – *maximum likelihood* and a dynamic MTM for lognormal insurance payment-per-payment and payment-per-loss loss severity data. A series of theoretical results about estimators' existence and asymptotic normality are established. As seen in Tables 1 and 2, the dynamic MTM estimators sacrifice little asymptotic efficiency with respect to MLE but are more robust as well as computationally more efficient than MLE if properly implemented, see, for example, (4.18) and Note 4.1, and these properties are highly desirable in practice. Finally, the developed estimators are implemented on a real-life data set to analyze the 1500 US indemnity losses which is widely studied in the actuarial literature (see, e.g., Frees and Valdez, 1998; Michael *et al.*, 2020), and it is found that most of the fitted models are plausible for this data set.

The results of this paper motivate open problems and generate several ideas for further research. First, this paper is specifically focused on lognormal insurance payment severity models, but they could be extended to more general (log) location-scale and exponential dispersion families. Second, several contaminated loss severity models are proposed in the literature as mentioned in Section 1 Introduction, so it could even produce a better model still maintaining a reasonable balance between efficiency and robustness if one implements the dynamic MTM procedures for those mixture models, but it is recommended to design the MTM for the complete mixture models first. Third, in this paper, we only discussed about model-fitting procedures and the corresponding finite sample performance, but the ultimate goal of model fitting is to apply the fitted models in decision-making mechanism. Therefore, it is yet to measure how the estimators developed in this paper act in calculating actuarial pure premium as well as in different risk analysis in practice.

Finally, it is important to emphasize again that the motivation to implement the MTM approach rather than MLE even for truncated and/or censored loss data sets is to simultaneously remove partial point masses assigned by MLE at the truncated and/or censored data points (i.e., achieving a desired degree of robustness) and maintaining a high degree of ARE, too. Further, the dynamic MTM procedures designed in this paper can only be implemented if a close fit in one or both tails of the assumed underlying distribution is not desired, otherwise extreme value modeling is recommended.

Acknowledgements

The author is very appreciative of valuable comments and constructive criticism provided by an Editor and two anonymous Referees. These have led to many improvements in the paper.

https://doi.org/10.1017/asb.2021.4 Published online by Cambridge University Press

REFERENCES

ABU BAKAR, S.A., HAMZAH, N.A., MAGHSOUDI, M. and NADARAJAH, S. (2015) Modeling loss data using composite models. *Insurance: Mathematics & Economics*, **61**, 146–154.

BARROW, D.F. and COHEN, JR., A.C. (1954) On some functions involving Mill's ratio. *Annals of Mathematical Statistics*, **25**, 405–408.

BLOSTEIN, M. and MILJKOVIC, T. (2019) On modeling left-truncated loss data using mixtures of distributions. *Insurance: Mathematics & Economics*, **85**, 35–46.

BRAZAUSKAS, V., JONES, B.L. and ZITIKIS, R. (2009). Robust fitting of claim severity distributions and the method of trimmed moments. *Journal of Statistical Planning and Inference*, **139**(6), 2028–2043.

BRAZAUSKAS, V. and KLEEFELD, A. (2016) Modeling severity and measuring tail risk of Norwegian fire claims. *North American Actuarial Journal*, **20**(1), 1–16.

BRAZAUSKAS, V. and SERFLING, R. (2000) Robust and efficient estimation of the tail index of a single-parameter Pareto distribution. *North American Actuarial Journal*, **4**(4), 12–27.

CHAN, J.S.K., CHOY, S.T.B., MAKOV, U.E. and LANDSMAN, Z. (2018). Modelling insurance losses using contaminated generalized beta type-II distribution. *ASTIN Bulletin*, **48**(2), 871–904.

CHERNOFF, H., GASTWIRTH, J.L. and JOHNS, JR., M.V. (1967) Asymptotic distribution of linear combinations of functions of order statistics with applications to estimation. *Annals of Mathematical Statistics*, **38**(1), 52–72.

COHEN, JR., A.C. (1950) Estimating the mean and variance of normal populations from singly truncated and doubly truncated samples. *Annals of Mathematical Statistics*, **21**(4), 557–569.

COHEN, JR., A.C. (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.

COORAY, K. and ANANDA, M.M.A. (2005) Modeling actuarial data with a composite lognormal-Pareto model. *Scandinavian Actuarial Journal*, **2005**(5), 321–334.

ERGASHEV, B., PAVLIKOV, K., URYASEV, S. and SEKERIS, E. (2016) Estimation of truncated data samples in operational risk modeling. *The Journal of Risk and Insurance*, **83**(3), 613–640.

FREES, E.W. and VALDEZ, E.A. (1998) Understanding relationships using copulas. *North American Actuarial Journal*, **2**(1), 1–25.

GRÜN, B. and MILJKOVIC, T. (2019) Extending composite loss models using a general framework of advanced computational tools. *Scandinavian Actuarial Journal*, **2019**(8), 642–660.

GUI, W., HUANG, R. and Lin, X.S. (2021) Fitting multivariate Erlang mixtures to data: A roughness penalty approach. *Journal of Computational and Applied Mathematics*, **386**, 113216, 17.

HAMPEL, F.R. (1974) The influence curve and its role in robust estimation. *Journal of the American Statistical Association*, **69**, 383–393.

HEWITT, C.C., JR. and LEFKOWITZ, B. (1979) Methods for fitting distributions to insurance loss data. In *Proceedings of the Casualty Actuarial Society*, vol. LXVI, pp. 139–160. VA: Casualty Actuarial Society.

HUBER, P.J. (1964) Robust estimation of a location parameter. *Annals of Mathematical Statistics*, **35**(1), 73–101.

HUBER, P.J. and RONCHETTI, E.M. (2009) *Robust Statistics*. 2nd edn. Hoboken, NJ: John Wiley & Sons, Inc.

KLUGMAN, S.A., PANJER, H.H. and WILLMOT, G.E. (2019) *Loss Models: From Data to Decisions*. 5th edn. Hoboken, NJ: John Wiley & Sons.

MICHAEL, S., MILJKOVIC, T. and MELNYKOV, V. (2020) Mixture modeling of insurance loss data with multiple level partial censoring. *Advances in Data Analysis and Classification*, **14**, 355–378.

MILJKOVIC, T. and GRÜN, B. (2016) Modeling loss data using mixtures of distributions. *Insurance: Mathematics & Economics*, **70**, 387–396.

POUDYAL, C. (2018) *Robust estimation of parametric models for insurance loss data*. ProQuest LLC, Ann Arbor, MI. Thesis (Ph.D.)–The University of Wisconsin - Milwaukee.

https://doi.org/10.1017/asb.2021.4 Published online by Cambridge University Press

POUDYAL, C. (2021) 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.
PUNZO, A., BAGNATO, L. and MARUOTTI, A. (2018) Compound unimodal distributions for insurance losses. *Insurance: Mathematics & Economics*, **81**, 95–107.
REYNKENS, T., VERBELEN, R., BEIRLANT, J. and ANTONIO, K. (2017) Modelling censored losses using splicing: A global fit strategy with mixed Erlang and extreme value distributions. *Insurance: Mathematics & Economics*, **77**, 65–77.
SCOLLNIK, D.P.M. (2007) On composite lognormal-Pareto models. *Scandinavian Actuarial Journal*, **2007**(1), 20–33.
SCOLLNIK, D.P.M. and SUN, C. (2012) Modeling with Weibull-Pareto models. *North American Actuarial Journal*, **16**(2), 260–272.
SERFLING, R. (2002) Efficient and robust fitting of lognormal distributions. *North American Actuarial Journal*, **6**(4), 95–109.
SERFLING, R.J. (1980) *Approximation Theorems of Mathematical Statistics*. New York: John Wiley & Sons.
SHAH, S.M. and JAISWAL, M.C. (1966) Estimation of parameters of doubly truncated normal distribution from first four sample moments. *Annals of the Institute of Statistical Mathematics*, **18**, 107–111.
TUKEY, J.W. (1960) A survey of sampling from contaminated distributions. In *Contributions to Probability and Statistics*, pp. 448–485. Stanford, CA: Stanford University Press.
VAN DER VAART, A.W. (1998) *Asymptotic Statistics*. Cambridge: Cambridge University Press.
VERBELEN, R., GONG, L., ANTONIO, K., BADESCU, A. and LIN, S. (2015) Fitting mixtures of Erlangs to censored and truncated data using the EM algorithm. *ASTIN Bulletin*, **45**(3), 729–758.
ZHAO, Q., BRAZAUSKAS, V. and GHORAI, J. (2018) Robust and efficient fitting of severity models and the method of Winsorized moments. *ASTIN Bulletin*, **48**(1), 275–309.

CHUDAMANI POUDYAL (Corresponding author)
Department of Mathematics
Tennessee Technological University
Cookeville, TN 38505, USA
E-Mail: cpoudyal@tntech.edu

APPENDIX A. AUXILIARY RESULTS

Define $\bar{a} := 1 - a$, $\bar{b} := 1 - b$, and $\tau := 1 - a - b$.

1. The expressions for c_k, $1 \le k \le 4$ mentioned in Section 4.2 are given by:

$$c_k \equiv c_k(\Phi, a, b) = \frac{1}{\tau} \int_a^{\bar{b}} \left[\Phi^{-1}(s)\right]^k \, ds.$$

2. The expressions for c_k^*, $1 \le k \le 3$ used in Section 4.2 are given as below:

$$\begin{aligned} c_1^* &= \left(\frac{1}{\tau}\right)^2 \int_a^{\bar{b}} \int_a^{\bar{b}} \frac{\min\{u, v\} - uv}{\phi\left[\Phi^{-1}(v)\right] \phi\left[\Phi^{-1}(u)\right]} \, dv \, du \\ &= \left(\frac{1}{\tau}\right)^2 \left\{ a\bar{a}\left[\Phi^{-1}(a)\right]^2 + b\bar{b}\left[\Phi^{-1}(\bar{b})\right]^2 - 2ab\Phi^{-1}(a)\Phi^{-1}(\bar{b}) \right. \\ &\quad \left. -2\tau c_1 \left[a\Phi^{-1}(a) + b\Phi^{-1}(\bar{b})\right] - \tau^2 c_1^2 + \tau c_2 \right\}. \end{aligned}$$

https://doi.org/10.1017/asb.2021.4 Published online by Cambridge University Press

$$c_2^* = \left(\frac{1}{\tau}\right)^2 \int_a^{\bar{b}} \int_a^{\bar{b}} \frac{(\min\{u,v\} - uv)\ \Phi^{-1}(u)}{\phi\left[\Phi^{-1}(v)\right]\phi\left[\Phi^{-1}(u)\right]}\, dv\, du$$

$$= \frac{1}{2\tau^2}\left\{a\bar{a}\left[\Phi^{-1}(a)\right]^3 + b\bar{b}\left[\Phi^{-1}(\bar{b})\right]^3 - ab\Phi^{-1}(a)\Phi^{-1}(\bar{b})\left[\Phi^{-1}(a) + \Phi^{-1}(\bar{b})\right]\right.$$

$$\left. -\tau c_1\left[a\left[\Phi^{-1}(a)\right]^2 + b\left[\Phi^{-1}(\bar{b})\right]^2\right] - \tau c_2\left[a\Phi^{-1}(a) + b\Phi^{-1}(\bar{b})\right] - \tau^2 c_1 c_2 + \tau c_3\right\}.$$

$$c_3^* = \left(\frac{1}{\tau}\right)^2 \int_a^{\bar{b}} \int_a^{\bar{b}} \frac{(\min\{u,v\} - uv)\ \Phi^{-1}(v)\Phi^{-1}(u)}{\phi\left[\Phi^{-1}(v)\right]\phi\left[\Phi^{-1}(u)\right]}\, dv\, du$$

$$= \frac{1}{4\tau^2}\left\{a\bar{a}\left[\Phi^{-1}(a)\right]^4 + b\bar{b}\left[\Phi^{-1}(\bar{b})\right]^4 - 2ab\left[\Phi^{-1}(a)\right]^2\left[\Phi^{-1}(\bar{b})\right]^2\right.$$

$$\left. -2\tau c_2\left[a\left[\Phi^{-1}(a)\right]^2 + b\left[\Phi^{-1}(\bar{b})\right]^2\right] - \tau^2 c_2^2 + \tau c_4\right\}.$$

3. The entries of the variance–covariance matrix $\mathbf{\Sigma}_y$ from Section 4.1 are as follows:

$$c_{y,1}^* = \left(\frac{\bar{\Phi}(\gamma)}{\tau}\right)^2 \int_a^{\bar{b}} \int_a^{\bar{b}} \frac{\min\{u,v\} - uv}{\phi\left[\Phi^{-1}(v + (1-v)\Phi(\gamma))\right]\phi\left[\Phi^{-1}(u + (1-u)\Phi(\gamma))\right]}\, dv\, du.$$

$$c_{y,2}^* = \left(\frac{\bar{\Phi}(\gamma)}{\tau}\right)^2 \int_a^{\bar{b}} \int_a^{\bar{b}} \frac{[\min\{u,v\} - uv]\ \Phi^{-1}(u + (1-u)\Phi(\gamma))}{\phi\left[\Phi^{-1}(v + (1-v)\Phi(\gamma))\right]\phi\left[\Phi^{-1}(u + (1-u)\Phi(\gamma))\right]}\, dv\, du.$$

$$c_{y,3}^* = \left(\frac{\bar{\Phi}(\gamma)}{\tau}\right)^2 \int_a^{\bar{b}} \int_a^{\bar{b}} \left[\frac{[\min\{u,v\} - uv]\ \Phi^{-1}(v + (1-v)\Phi(\gamma))}{\phi\left[\Phi^{-1}(v + (1-v)\Phi(\gamma))\right]}\right.$$

$$\left. \times \frac{\Phi^{-1}(u + (1-u)\Phi(\gamma))}{\phi\left[\Phi^{-1}(u + (1-u)\Phi(\gamma))\right]}\right] dv\, du.$$

https://doi.org/10.1017/asb.2021.4 Published online by Cambridge University Press