Statistics and Probability Letters 239 (2026) 110913

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Nonparametric estimation of splicing points in actuarial loss distributions via data transformation

Benedikt Funke [a,*], Masayuki Hirukawa [b]

[a] Institute for Insurance Studies, TH Köln – University of Applied Sciences, Gustav-Heinemann-Ufer 54, 50968 Köln, Germany
[b] Faculty of Economics, Ryukoku University, 67 Tsukamoto-cho, Fukakusa, Fushimi-ku, Kyoto 612-8577, Japan

ARTICLE INFO

Keywords:
Actuarial loss distributions
Beta kernel
Cross validation
Data transformation
Splicing point

ABSTRACT

We propose a nonparametric method for estimating splicing points in actuarial loss distributions. Our approach transforms non-negative loss data onto the unit interval $[0,1]$ via the transformation based on the cumulative distribution function of the log–logistic (or Fisk) distribution, which magnifies jump discontinuities in the density and facilitates detection of the splicing point even in the sparse tail region. The transformed data are smoothed using the asymmetric beta kernel, and the splicing point estimator is obtained by back-transforming the maximizer of a diagnostic function. It is demonstrated that the proposed splicing point estimator is strongly consistent and asymptotically normal with a convergence rate faster than the parametric one. Monte Carlo simulations and applications to non-life insurance loss data confirm the practical relevance of the approach.

1. Introduction

In non-life insurance, accurate modeling of loss distributions underpins pricing, reserving and solvency capital determination. Because the right tail, where infrequent but severe losses reside, typically follows a different pattern from the bulk of attritional claims, actuaries rely on splicing or composite models that join the bulk and tail distributions at a threshold or splicing point (e.g., Klugman et al., 2019). Misspecifying this point distorts risk measures and reinsurance structures.

Most existing threshold-selection methods, such as heuristic quantile rules, graphical diagnostics and automated procedures built on extreme value theory (EVT) (e.g., Clauset et al., 2009; Bader et al., 2018; Danielsson et al., 2019), either depend on a specific tail model such as the generalized Pareto distribution (GPD) or leave substantial discretion to practitioners. Recently, Funke and Hirukawa (2026) proposed a model-free approach that interprets the splicing point as a density jump and estimates it nonparametrically using the asymmetric gamma kernel on $\mathbb{R}_+$ by Chen (2000). It is strongly consistent and asymptotically normal with a convergence rate faster than the parametric one. However, its finite-sample performance deteriorates when the jump is small or lies deep in the sparse tail.

We address this limitation by transforming the loss data via a known monotone mapping $T:\mathbb{R}_+ \to [0,1]$. The transformation is designed to magnify the jump at the splicing point t_0 and compress tail distances. The splicing point estimator is initially obtained in the transformed scale as the maximizer of the diagnostic function using the asymmetric beta kernel by Chen (1999) and then back-transformed into the original scale. The estimator is strongly consistent and asymptotically normal with a super-consistent convergence rate (i.e., the rate exceeds the parametric one). Monte Carlo simulations indicate that it outperforms the splicing point estimator in the original scale and automated threshold estimators.

* Corresponding author.
E-mail addresses: benedikt.funke@th-koeln.de (B. Funke), hirukawa@econ.ryukoku.ac.jp (M. Hirukawa).

https://doi.org/10.1016/j.spl.2026.110913
Received 21 April 2026; Received in revised form 20 July 2026; Accepted 21 July 2026
Available online 22 July 2026

The remainder of this paper is organized as follows. Section 2 proposes the splicing point estimator via data transformation. In Section 3 asymptotic properties of the estimator are explored. Section 4 conducts Monte Carlo simulations. In Section 5 the proposed estimator is applied to four non-life insurance datasets. Section 6 concludes.

Technical proofs and details of numerical analyses are collected in the online Supplement. Lemma S1 therein delivers a normal approximation to the incomplete beta function ratio. This lemma, in conjunction with a few asymptotic results from Funke and Hirukawa (2026), is a building block for Proposition S2, which demonstrates that a uniform approximation to the diagnostic function has a unique maximum on the prespecified interval. The lemma is also employed for bias and variance approximations in Theorem 2.

This paper adopts the following notational conventions: '$a_n = o\left(b_n\right)$' signifies that a_n/b_n converges to 0; '$a_n = O\left(b_n\right)$' means that a_n/b_n is bounded; we say that '$a_n \asymp b_n$' if there exist constants $0 < c_1 < c_2 < \infty$ so that $c_1 a_n \leq b_n \leq c_2 a_n$; and $\mathbf{1}\{\cdot\}$ is an indicator function. For a function $h(x)$ and a point c, $h(c^-) = \lim_{x\uparrow c} h(x)$, $h(c^+) = \lim_{x\downarrow c} h(x)$ and $h^{(m)}(x) = d^m h(x)/dx^m$ denote the left and right limits, and the mth-order derivative, respectively. The abbreviation '*a.s.*' stands for "almost surely". Finally, $B(p,q) = \int_0^1 y^{p-1}(1-y)^{q-1}dy$ for $p, q > 0$ is the beta function.

2. Methodology

2.1. Setup and data transformation

Suppose that the probability density function (pdf) of $X \in \mathbb{R}_+$ has a jump at t_0 strictly inside a prespecified interval $I_0 := [\underline{t}, \bar{t}] \subset (0, \infty)$ lying in the right tail. Practitioners typically have some prior knowledge about the location of t_0 from historical experience, policy limits or empirical quantiles. We model the pdf locally on I_0 as $f_X(x) = g_X(x) + d_0 \mathbf{1}\{x < t_0\}$ for some smooth function $g_X(\cdot)$ and a jump $d_0 := f_X(t_0^-) - f_X(t_0^+) \neq 0$.

When $|d_0|$ is small or the tail is sparse, detecting t_0 directly on $\mathbb{R}_+$ is difficult. For a known smooth monotone transformation $T : \mathbb{R}_+ \to [0,1]$, the pdf of $Y := T(X)$ on $I_T := [T(\underline{t}), T(\bar{t})]$ can be expressed as

$$f_Y(y) = g_Y(y) + d_T \mathbf{1}\{y < t_T\}, \tag{1}$$

where $t_T := T(t_0)$, $d_T := f_Y(t_T^-) - f_Y(t_T^+)$, and $g_Y(\cdot)$ satisfies $g_X(x) = g_Y\{T(x)\}T^{(1)}(x)$. The key relationship is $d_0 = d_T T^{(1)}(t_0)$. If $0 < T^{(1)}(x) < 1$ for $x \in I_0$, then $|d_T| > |d_0|$ or the jump is magnified.

Because it is not straightforward to derive an optimal transformation analytically among all that satisfy Assumption 3 below, we focus on the family of the cumulative distribution function (cdf) of the log–logistic (or Fisk) distribution

$$T_\theta(x) = \frac{(x/t_M)^\theta}{1 + (x/t_M)^\theta} = \frac{1}{1 + (x/t_M)^{-\theta}}, \tag{2}$$

which depends on $t_M := (\underline{t} + \bar{t})/2$, the midpoint of I_0, and the shape parameter $\theta > 0$. It follows from $T_\theta(0) = 0$ and $T_\theta(t_M) = 1/2$ that every family member satisfies Assumption 3(i) below. The implied magnification factor $d_T/d_0 = 1/T^{(1)}(t_0)$ is the inverse of the log–logistic pdf evaluated at $x = t_0$, where the pdf is

$$T_\theta^{(1)}(x) = \frac{\left(\theta/t_M\right)\left(x/t_M\right)^{\theta-1}}{\{1 + \left(x/t_M\right)^\theta\}^2}.$$

2.2. Estimation procedure

Following the jump-location detection strategy by Funke and Hirukawa (2026), we use shifted beta kernels. For a design point $y \in [0,1]$, the smoothing parameter $b > 0$ and the shift $\Delta > 0$, the shifted beta kernels are the densities of $\mathrm{Beta}\{(y \pm \Delta)/b + 1, (1 - y \mp \Delta)/b + 1\}$ and take the forms of

$$K_y^\pm(u) = K_{B(y,b;\pm\Delta)}(u) = \frac{u^{(y\pm\Delta)/b}(1-u)^{(1-y\mp\Delta)/b}}{B\{(y \pm \Delta)/b + 1, (1 - y \mp \Delta)/b + 1\}} \mathbf{1}\{u \in [0,1]\}.$$

These kernels concentrate mass slightly to the left or right of y so that the difference of the two shifted estimators $\hat{f}_Y^\pm(y) := n^{-1}\sum_{i=1}^n K_y^\pm(Y_i)$ can detect discontinuities. Setting $\hat{J}(y) := \hat{f}_Y^-(y) - \hat{f}_Y^+(y)$ and $\hat{t}_T := \arg\max_{y \in I_T} |\hat{J}(y)|$, our splicing point estimator in the original scale is given by $\hat{t}_B := T^{-1}(\hat{t}_T)$.

3. Large-sample properties

3.1. Regularity conditions

We require the following regularity conditions for convergence results.

Assumption 1. $\{X_i\}_{i=1}^n \in \mathbb{R}_+$ are *i.i.d.* random variables.

Assumption 2. **(i)** $f_Y(y)$ is uniformly bounded on $[0,1]$. **(ii)** The local structure (1) holds, $g_Y^{(2)}$ is uniformly bounded on $[0,1]$, and $g_Y^{(3)}$ is Lipschitz continuous and bounded on I_T.

Assumption 3. **(i)** T is injective with $T(0)=0$ and $T(t_M)=1/2$, where $t_M := (\underline{t}+\bar{t})/2$ is the midpoint of I_0. **(ii)** $T^{(1)}$ is Lipschitz continuous on $\mathbb{R}_+$, and $0 < T^{(1)}(x) < 1$ holds for all $x \in I_0$.

Assumption 4. Tuning parameters b and Δ satisfy, as $n \to \infty$, $b, \Delta \to 0$,

$$\frac{b^{3/4}}{\Delta} + \frac{\Delta}{b^{1/2+\delta_1}} + \frac{b^{1/2-4\delta_1}}{n^{1-\delta_2}\Delta^2} \to 0$$

for some small $\delta_1, \delta_2 > 0$, and $\ln n/(nb^{3/2-\kappa}) = O(1)$ for some $\kappa \in [0,1)$.

Assumptions 1, 2 and 4 are standard for strong uniform consistency of asymmetric kernel estimators (e.g., Hirukawa et al., 2022; Funke and Hirukawa, 2026). Assumption 4 suggests choosing $\Delta \asymp b^\alpha$ for $\alpha \in (1/2, 3/4)$, which makes the effect of Δ asymptotically negligible in the first-order sense.

Assumption 3 refers to the transformation. The log–logistic distribution has the mode at $x=0$ for $0 < \theta \le 1$ and $x_{mo} := t_M\{(\theta-1)/(\theta+1)\}^{1/\theta} (< t_M)$ for $\theta > 1$. Therefore, for the log–logistic cdf family to satisfy Assumption 3, we must additionally impose the following restriction on θ: (i) for $0 < \theta \le 1$, $T_\theta^{(1)}(\underline{t}) = (\theta/t_M)(\underline{t}/t_M)^{\theta-1}/\{1+(\underline{t}/t_M)^\theta\}^2 < 1$; and (ii) for $\theta > 1$, $T_\theta^{(1)}(\underline{t}) < 1$ if $x_{mo} < \underline{t}$ and $T_\theta^{(1)}(x_{mo}) = (\theta+1)^{1+1/\theta}(\theta-1)^{1-1/\theta}/(4t_M\theta) < 1$ if $\underline{t} \le x_{mo}$. The values of θ taken in Sections 4 and 5 fulfill this restriction.

3.2. Consistency and asymptotic normality

The theorems below provide strong consistency and asymptotic normality of the splicing point estimator $\hat{t}_B$. Their proofs, intermediate propositions and lemmata are provided in the Supplement.

Theorem 1 (*Strong Consistency*). *Under Assumptions 1–4,* $|\hat{t}_B - t_0| = O(b^{1/2+\delta_1})$ *a.s. as* $n \to \infty$.

Theorem 2 (*Asymptotic Normality*). *Under Assumptions 1–4,*

$$\sqrt{\frac{n}{b^{1/2}}}\left[\hat{t}_B - t_0 - \left\{\frac{1-T(t_0)/2}{T^{(1)}(t_0)}\right\} b\{1+o_p(1)\}\right] \xrightarrow{d} N(0, V_B)$$

as $n \to \infty$, *where*

$$V_B = V_B(T) := \frac{3\sqrt{\pi}\sqrt{T(t_0)\{1-T(t_0)\}}}{4d_0^2\, T^{(1)}(t_0)} \left\{\frac{f_X(t_0^-) + f_X(t_0^+)}{2}\right\}.$$

Theorem 2 implies super-consistency of $\hat{t}_B$. More specifically, when $b \asymp n^{-\beta}$ for $\beta > 1/2$, both squared leading bias and leading variance terms vanish faster than the $1/n$ rate. Smoothing parameters in Sections 4 and 5 are chosen to satisfy $b \asymp n^{-2/3}$, and thus super-consistency of $\hat{t}_B$ is ensured. The super-consistency also matters in practice: replacing an unknown splicing point t_0 with its estimator $\hat{t}_B$ does not distort subsequent inference or tail goodness-of-fit tests (e.g., Reynkens et al., 2017; Clauset et al., 2009).

Remark 1. Theorem 2 makes the trade-off behind the magnification explicit. Suppose that t_0 lies in the vicinity of t_M. When $t_0 \approx t_M$, the magnification factor becomes $1/T_\theta^{(1)}(t_0) \approx 4t_M/\theta$. This quantity ranges from 8 at $\theta = 2$ to 32 at $\theta = 0.5$ for $t_M = 4$ (Case (i) of the next section). While a small θ enlarges the magnification factor and makes it easier to detect the jump signal, it inflates both the leading bias and variance coefficients given in Theorem 2. In fact, when $t_0 \approx t_M$,

$$\frac{1-T_\theta(t_0)/2}{T_\theta^{(1)}(t_0)} \approx \frac{3t_M}{\theta} \quad \text{and} \quad \frac{\sqrt{T_\theta(t_0)\{1-T_\theta(t_0)\}}}{T_\theta^{(1)}(t_0)} \approx \frac{2t_M}{\theta},$$

and thus both grow without bound as $\theta \to 0$. Which effect dominates in finite samples is unclear, and we are motivated to find the value of θ that performs best in the simulation study.

4. Finite-sample performance

4.1. Design

We simulate 1000 replications with $n \in \{250, 500\}$ from three distributions, each with a common splicing point $t_0 = 4$ and a small jump size $d_0 \approx 0.05$. Model A uses a log-normal-like density augmented by a quadratic step below t_0. Models B and C splice a Weibull bulk distribution with a GPD tail and a half-normal tail, respectively. Full model specifications and characteristic numbers are given in the Supplement.

Table 1
Monte Carlo RMSE ($n = 500$; 1000 replications).

Estimator	α	Model A			Model B			Model C		
KS	–	0.8886			1.1735			0.9074		
Q-MAD	–	0.4423			0.4936			0.4599		
Q-SUP	–	0.9918			1.5684			1.0779		
AEB	–	0.7115			1.8934			0.6417		
ADST	–	1.4748			1.2689			1.5354		
		(i)	(ii)	(iii)	(i)	(ii)	(iii)	(i)	(ii)	(iii)
SG-ML	0.70	0.5269	0.3122	0.8590	0.6756	0.4317	0.7889	0.6503	0.4183	0.7610
SG-ML-BC	0.70	0.5037	0.2902	0.8322	0.6358	0.3895	0.7496	0.6196	0.3861	0.7305
SB-LS(0.50)	0.55	0.3479	0.3457	0.3299	0.1192	0.1182	0.1227	0.1513	0.1425	0.1627
SB-LS(0.75)	0.55	0.4886	0.4248	0.4484	0.2350	0.2539	0.2120	0.1960	0.2165	0.1734
SB-LS(1.00)	0.55	0.5845	0.4453	0.5257	0.3441	0.3623	0.3081	0.3268	0.3495	0.2828
SB-LS(1.50)	0.55	0.6929	0.4429	0.7438	0.4788	0.4452	0.4137	0.5115	0.4550	0.4382
SB-LS(2.00)	0.55	0.6745	0.4155	0.8823	0.5959	0.4628	0.4882	0.6423	0.4727	0.5457

We compare the shifted beta estimator with five automated methods including the minimum Kolmogorov–Smirnov distance procedure [KS] (Clauset et al., 2009), two versions of minimum quantile-discrepancy criteria [Q-MAD, Q-SUP] and automated Eye-Balling method [AEB] (Danielsson et al., 2019), and Anderson–Darling sequential testing procedure [ADST] (Bader et al., 2018). The shifted gamma kernel estimator and its bias-corrected version implemented by the modified likelihood cross-validation (MLCV) [SG-ML, SG-ML-BC] (Funke and Hirukawa, 2026) serve as kernel benchmarks, where the MLCV criterion can be found in Funke and Hirukawa (2026). The minimizer of the MLCV criterion is chosen among $b = C_G \hat{\sigma}_X n^{-2/3}$ for the sample standard deviation $\hat{\sigma}_X$ over eight grids $C_G \in \{2, 4, 6, \dots, 16\}$. We exclusively consider the exponent $\alpha = 0.70$ for the shift $\Delta = b^\alpha$.

The transformed observations $\{Y_i = T_\theta(X_i)\}_{i=1}^n$ for our shifted beta estimator are computed using the log–logistic cdf family (2) over the grid $\theta \in \{0.50, 0.75, 1.00, 1.50, 2.00\}$, and the smoothing parameter b is determined by the least squares cross-validation (LSCV) [SB-LS(θ)]. Let $\hat{f}^{\pm}_{Y,b,-i}(y;\alpha) := \sum_{j=1, j\neq i}^n K_y^{\pm}(Y_j)/(n-1)$ be density estimates using the sample with the ith observation eliminated and the exponent $\alpha \in \{0.55, 0.60, 0.65, 0.70\}$ for the shift $\Delta = b^\alpha$. Also denote the number of observations falling into I_T as $n_0 := \sum_{i=1}^n \mathbf{1}\{Y_i \in I_T\}$. Then, the LSCV criterion is defined as $CV_{LS}(b;\alpha) = CV^-_{LS}(b;\alpha) + CV^+_{LS}(b;\alpha)$, where

$$CV^{\pm}_{LS}(b;\alpha) := \int_{I_T} \left\{\hat{f}^{\pm}_{Y,b}(y;\alpha)\right\}^2 dx - \frac{2}{n_0} \sum_{i:Y_i \in I_T} \hat{f}^{\pm}_{Y,b,-i}(Y_i;\alpha).$$

The minimizer of $CV_{LS}(b;\alpha)$ is chosen among $b = C_B IDR_Y n^{-2/3}$ for the interdecile range IDR_Y over eight grids $C_B \in \{2, 4, 6, \dots, 16\}$. The reason for adopting the interdecile range as a measure of dispersion is to avoid selecting extremely small values of b. Notice that setting $b \asymp n^{-2/3}$ ensures super-consistency of gamma and beta kernel estimators. Furthermore, three choices of I_0, namely, (i) $[3, 5]$ ($t_0 = t_M$), (ii) $[3.5, 5.5]$ ($t_0 < t_M$) and (iii) $[2.5, 4.5]$ ($t_0 > t_M$), are investigated for the kernel estimators.

4.2. Results

To save space, Table 1 reports the root mean squared error (RMSE) of each estimator over 1000 replications for $n = 500$. For SB-LS(θ), we present the results with $\alpha = 0.55$ as this case dominates other choices of α. The Supplement provides comprehensive simulation results for two sample sizes, including other performance measures.

The results indicate that the RMSE of SB-LS(θ) tends to increase with θ. To put it another way, stronger magnification by a small θ generally sharpens the detection of a jump in the transformed scale. In particular, the SB-LS estimator with $\theta = 0.50$ dominates the original-scale SG benchmarks and all automated methods, with the exception of Model A - Case (ii), confirming that transformation is more effective than direct estimation in the original scale when the jump is small. We further examine SB-LS(0.50) with $\alpha = 0.55$ using real insurance datasets in the next section. Moreover, Case (ii) gives the best results for most of kernel estimators, suggesting that I_0 be set longer on the right-tail side.

5. Empirical applications

In this section our shifted beta estimator is applied to four non-life insurance datasets, including (A) Danish fire losses, which have been analyzed since McNeil (1997), (B) Norwegian fire, (C) Belgian motor, and (D) French motor losses. Their descriptions and summary statistics are available in the Supplement. Our focus is on SB-LS(0.50) with $\alpha = 0.55$, the configuration favored by the simulation. We compare it with the competing estimators considered in Section 4. Kernel estimators are implemented as therein.

Table 2 presents estimates of splicing points $\hat{t}_0$. The SB-LS estimate is of the same order as KS and ADST for (A) and (B) and as KS and Q-MAD for (C) and (D). ADST rejects all candidates for (C) and (D), whereas SG-ML fails to find a maximizer strictly inside I_0 for (A)–(C).

In contrast, SB-LS with $\theta = 0.50$ and $\alpha = 0.55$ preferred by the simulation continues to work well for real data. Indeed, the transformation-based estimates are stable: the diagnostic function attains an interior maximum in the transformed scale and the back-transformed splicing point estimate stays well inside I_0 for every dataset. It is confirmed that our proposed method serves as a viable complement to existing approaches, especially where automated and/or original-scale methods break down.

Table 2
Estimates of splicing points.

Data	Estimator	$\hat{t}_0$	Estimator	α	I_0	$\hat{t}_0$
(A)	KS	1.375	SG-ML	0.70	[1, 30]	30.000[b]
($n = 2492$)	Q-MAD	29.037	SG-ML-BC	–	–	30.087
	Q-SUP	11.123				
	AEB	25.288	SB-LS(0.50)	0.55	[1, 30]	1.994
	ADST	1.406				
(B)	KS	2.221	SG-ML	0.70	[1, 55]	55.000[b]
($n = 3064$)	Q-MAD	35.794	SG-ML-BC	–	–	55.106
	Q-SUP	123.274				
	AEB	47.528	SB-LS(0.50)	0.55	[1, 55]	3.322
	ADST	2.121				
(C)	KS	3.233	SG-ML	0.70	[1, 40]	40.000[b]
($n = 3151$)	Q-MAD	3.022	SG-ML-BC	–	–	40.036
	Q-SUP	40.035				
	AEB	18.592	SB-LS(0.50)	0.55	[1, 40]	2.112
	ADST	29.451[a]				
(D)	KS	1.318	SG-ML	0.70	[1, 20]	1.624
($n = 2638$)	Q-MAD	3.204	SG-ML-BC	–	–	1.908
	Q-SUP	7.000				
	AEB	18.900	SB-LS(0.50)	0.55	[1, 20]	1.690
	ADST	37.149[a]				

[a] All candidates are rejected in ADST, in which case the 99.5% empirical percentile is reported.
[b] An estimation failure occurs in SG-ML.

6. Conclusion

We have proposed a splicing point estimator using the beta kernel and the transformed data that magnify the density jump. The estimator is strongly consistent and asymptotically normal at a super-consistent rate. Simulation results confirm that the largest magnification case of SB-LS with $\theta = 0.50$ and $\alpha = 0.55$ outperforms the original-scale gamma and automated threshold estimators. The same configuration yields stable, sensible estimates on the four insurance datasets, including cases where competing methods fail.

Funding

This work was supported by the Japan Society for the Promotion of Science under the grant number 23K01340 (M. Hirukawa).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The authors would like to thank an anonymous referee and participants in CFE-CMStatistics 2024 for their constructive comments and suggestions. Open Access funding enabled and organized by Projekt DEAL and the Cologne Research Centre for Reinsurance at TH Köln.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.spl.2026.110913.

Data availability

Data will be made available on request.

References

Bader, B., Yan, J., Zhang, X., 2018. Automated threshold selection for extreme value analysis via ordered goodness-of-fit tests with adjustment for false discovery rate. Ann. Appl. Stat. 12, 310–329.

Chen, S.X., 1999. Beta kernel estimators for density functions. Comput. Statist. Data Anal. 31, 131–145.

Chen, S.X., 2000. Probability density function estimation using gamma kernels. Ann. Inst. Statist. Math. 52, 471–480.

Clauset, A., Shalizi, C.R., Newman, M.E.J., 2009. Power-law distributions in empirical data. SIAM Rev. 51, 661–703.

Danielsson, J., Ergun, L.M., de Haan, L., de Vries, C.G., 2019. Tail index estimation: Quantile-driven threshold selection. Bank of Canada Staff Working Paper 2019-28.

Funke, B., Hirukawa, M., 2026. Nonparametric estimation of splicing points in skewed cost distributions: A kernel-based approach. J. Nonparametr. Stat. 38, 546–585.

Hirukawa, M., Murtazashvili, I., Prokhorov, A., 2022. Uniform convergence rates for nonparametric estimators smoothed by the beta kernel. Scand. J. Stat. 49, 1353–1382.

Klugman, S.A., Panjer, H.H., Willmot, G.E., 2019. Loss Models: From Data to Decisions, fifth ed. Wiley, Hoboken, NJ.

McNeil, A.J., 1997. Estimating the tails of loss severity distributions using extreme value theory. ASTIN Bull. 27, 117–137.

Reynkens, T., Verbelen, R., Beirlant, J., Antonio, K., 2017. Modelling censored losses using splicing: A global fit strategy with mixed erlang and extreme value distributions. Insurance Math. Econom. 77, 65–77.