British Actuarial Journal (2025), Vol. **30**, e6, pp. 1–31
doi:10.1017/S1357321724000205

Institute and Faculty of Actuaries

CONTRIBUTED PAPER

Cyber-insurance pricing models

James Bardopoulos

Institute and Faculty of Actuaries, London, UK and Science Faculty, University of Cape Town, Western Cape, South Africa
Email: jbardopoulos@gmail.com

Abstract

In the present technological age, where cyber-risk ranks alongside natural and man-made disasters and catastrophes – in terms of global economic loss – businesses and insurers alike are grappling with fundamental risk management issues concerning the quantification of cyber-risk, and the dilemma as to how best to mitigate this risk. To this end, the present research deals with data, analysis, and models with the aim of quantifying and understanding cyber-risk – often described as "holy grail" territory in the realm of cyber-insurance and IT security. Nonparametric severity models associated with cyber-related loss data – identified from several competing sources – and accompanying parametric large-loss components, are determined, and examined. Ultimately, in the context of analogous cyber-coverage, cyber-risk is quantified through various types and levels of risk adjustment for (pure-risk) increased limit factors, based on applications of actuarially founded aggregate loss models in the presence of various forms of correlation. By doing so, insight is gained into the nature and distribution of volatile severity risk, correlated aggregate loss, and associated pure-risk limit factors.

Keywords: Cyber-insurance; increased limit factors; risk adjustment; correlation; severity distributions; aggregate loss models

1. Introduction

Cyber-risk, an umbrella term for risks associated with technology and information (CRO Forum, 2014), is a significant threat with an estimated annual cost to the worldwide economy of over $600bn (McAfee & Center for Strategic and International Studies, 2018). It encompasses a wide host of events caused by inadvertent activities (e.g. loss of data) and criminal threats (e.g. phishing) that can lead to various types of loss (e.g. remediation costs) damage and liability.

Uncertainty in the realm of a nascent insurance market has led to conservative underwriting; premiums are perceived to be large in relation to the level of cover – and thus low product penetration (UK Government and Industry, 2015) and restricted coverage (high deductibles, low policy limits) that fails to protect firms against low frequency events with volatile severity. Many of these obstacles have been attributed to the following characteristics associated with cyber-risk:

1. Lack of reliable (frequency, but mainly severity) data for modelling and quantifying cyber-risk in an "actuarial pricing" context (Cashell et al., 2004; Böhme & Schwartz, 2010)
2. The correlated nature of cyber-risk (Baldwin et al., 2012) and interdependence (i.e. degree of "interconnectedness" between networks and systems) – (Ogut et al., 2005) – precipitated by widespread use of the internet, relatively few Internet Service Providers (ISPs), and reliance upon common IT software (Böhme, 2005; Laszka et al., 2014)
3. Information asymmetry (Bandyopadhyay et al., 2010)

In academic circles, these factors have evidently influenced the development of cyber-risk models in several ways. Due to data related issues, frequency models appear to be more prevalent than severity (i.e. cost) models; aggregate loss models often assume constant severity leading to (possibly mixed) binomial distributions. Overall, the level of empirical support is egregiously low. Correlation and interdependence have led to the consideration of copula (Herath & Herath, 2011), Markov processes (Barracchini & Addessi, 2014), and Bayesian belief nets (Mukhopadhyay et al., 2013). Many of these models, having been developed beyond the framework of economics and computer science, are abstracted from several peculiarities associated with aggregate cyber-risk – especially in the context of cyber-insurance and risk quantification:

- Aggregate loss distributions, risk measures (e.g. variance and value at risk), tail dependence, and the effects of correlation and interdependence in terms of different sections of insurance cover (e.g. business interruption, data breach remediation, etc.) have received little attention
- Loss models are generally underdeveloped in the field of cyber-science – applications concerning (much required) risk theory and aggregate loss modelling techniques have been largely neglected in this domain
- There is very little evidence in academic cyber-related research of Increased limit factors (ILFs: multiples of premiums for different cover limits), which are highly relevant given concerns regarding "low policy limits" and "accurate pricing"

This work contributes to each of these areas. To begin with, Section 2 summarises implemented cyber-risk models (and accompanying data, if utilised), in the context of a model taxonomy by field of study and design. Several sources of data are evaluated, and a primary source is identified for constructing ILFs – this source is described in Section 3. Section 4 derives several models in the context of Individual and Collective Risk frameworks. These reflect different types of correlation and risk adjustments and are used to model ALDs and ILF curves. Section 5 considers various severity and aggregate loss distributions and explores the impact of correlation and risk adjustments at the aggregate level. Section 6 closes with conclusions and recommendations.

2. Review of Models

Taxonomy

A chronological taxonomy that depicts cyber-risk models under the following four broad headings can be found in Appendix A.2:

- **Economic** – models that consider the decisions and behaviours of individuals and organisations in the context of IT security and cyber-insurance. These typically focus on the "demand-side" (Böhme & Schwartz, 2010) of trade-off decisions (e.g. for allocating resources between insurance and IT security) using Utility or Decision theory.
- **Correlation based** – models that include copula and regression techniques, with some models that straddle the economic sphere (Liu et al., 2001; Böhme, 2005).
- **Operational Risk** (OR) – models that stem from OR quantification techniques such as those used to determine regulatory capital requirements, (European Commission, 2017). These encompass Extreme Value Theory (EVT) and risk theory (Section 4.1).
- **Epidemic** (and related) – models that utilise Markov processes and regression techniques, and are analogous to epidemiological compartmental (van Mieghem et al., 2009; Parker & Farkas, 2011) or health insurance (Barracchini & Addessi, 2014) models.

The search strategy underpinning this literary review of models is described in Appendix A.1.

https://doi.org/10.1017/S1357321724000205 Published online by Cambridge University Press

Summary of models

Counting processes and related distributions

As Table 1 shows, a variety of stochastic processes have been considered for count (e.g. number of cyber-related incidents, losses, etc.) and associated interarrival times. The homogeneous Poisson process (i.e. constant rate of arrival; independent, exponentially distributed interarrival times) is one common example (Van Mieghem et al., 2009; Herath & Herath, 2011). Variations (e.g. Pareto) have also been proffered in the context of privacy incidents (Yannacopoulos et al., 2008). The Bernoulli process is another example (Gordon & Loeb, 2002; Böhme, 2005; Böhme & Kataria, 2006). Non-homogeneous processes have also been utilised (Edwards et al., 2016).

Severity and aggregate loss distributions

Constant severity has often been assumed (Böhme, 2005; Böhme & Kataria, 2006; Mukhopadhyay et al., 2013, Section 5.2), which has resulted in several impractical aggregate loss models (characterised by binomial distributions). In the case of Edwards et al. (2016), aggregate loss was estimated using an independent regression model (Jacobs (2014), log skew-normal breach size) and a negative binomial distributed breach count variable. Indeed, few severity models have been based on genuine cyber-related loss data – in the case of (Biener et al., 2015), this entailed an extensive classification exercise in respect of OR data SAS (2015).

Table 1. Extant cyber-risk models. Distributions, models – *green* (recognised or plausible in the context of general insurance), *orange* (data dependent), *red* (unrealistic, misrepresentative), *grey* (out-of-scope, not applicable, unspecified)

	Distribution, model			Output and related exposure		Cyber-risk feature		
Author(s)	Count (N)	Severity ($X_1,\dots,X_N$)	Aggregate ($S=X_1+,\dots,+X_N$)	Output(s)	Exposure-measure	Correlation	Interdependence	Info. asymmetry
Soo Hoo (2000)	Triangular, uniform	Triangular	Not modelled	Expected benefit - security	Customers	×	×	×
Gordon & Loeb (2002)	Bernoulli; probability functions	Given	Not modelled	Optimal investment - security	Not specified	×	×	×
Mukhopadhyay et al. (2005)	Not specified	Not specified	Not modelled (1)	Risk premium - insurance	None	×	×	×
Böhme (2005)	Binomial mixture	Constant (unit cost)	Mixed binomial	Correlation - claims	Risks	✓	×	×
Rachev, Chernobai & Menn (2006)	Poisson (2a)	Variety (2b)	Not modelled (2c)	Parameters - distributions	Loss events	×	×	×
Böhme & Kataria (2006)	Beta-binomial; mixed binomial (EM algorithm) (3)	Constant (unit cost)	Mixed binomial	Correlation, densities	Computers(*)	✓	✓	×
Liu, Tanaka & Matsuura (2007)	Regression - Gordon & Loeb (2002)	Not modelled	Not modelled	Parameters - regression	Companies(*)	✓	×	×
Cope & Antonini (2008)	Not modelled	Not modelled	Empirical	Tail ratios - distributions	Banks(*)	✓	×	×
Kesan, Majuca & Yurcik (2008)	Given	Not modelled	Asset pricing	Premium - insurance Welfare gains - security	Layer of cover	×	×	✓
Yannacopoulos et al. (2008)	Poisson (4a)	Random Utility Model (4b)	Collective risk (4c)	Cost, benefit - insurance	None	×	×	×
Wang & Kim (2009)	Regression	Not modelled	Not modelled	Correlation - residuals	Internet users, GDP	✓	✓	×
van Mieghem, Omic & Kooij (2009)	Poisson	Not modelled	Not modelled	Number, fraction - infected nodes	Not specified	×	×	×
Bandyopadhyay, Mookerjee & Rao (2010)	Not specified	Uniform	Not modelled	Optimal deductible - insurance	Not specified	×	×	✓
Hess (2011)	Poisson	Spliced (exponential, GPD) (5)	Compound Poisson	Parameters - distributions	Companies(*)	×	×	×
Herath et al. (2011)	Poisson	Weibull	Not modelled (6)	Premium - insurance	Computers(*)	✓	×	×
Parker & Farkas (2011)	SEIR (7)	Not modelled	Not modelled	None (descriptive - SEIR)	Systems	×	×	×
Baldwin et al. (2012)	Brownian motion	Not modelled	Not modelled	Correlation - intensity, size	Ports	✓	×	×
Mukhopadhyay et al. (2013)	Gaussian copula; binomial	Constant (unit cost)	Binomial	Risk premium - insurance	Organisation	✓	×	×
Barracchini & Addessi (2014)	Markov	Not modelled	Not modelled	Transition intensity, premium	Computers	×	×	×
Biener, Eling & Wirfs (2015)	Not modelled	Variety (8)	Not modelled	Spliced densities	None	×	×	×
Laube & Böhme (2015)	Gordon & Loeb (2002), probability functions (9a)	Expected cost; parameter-based (9b)	Not modelled	Expected cost, social optima, nash equilibria	None	×	✓	✓
Edwards, Hofmeyr & Forrest (2016)	Negative Binomial, Poisson (10a)	Log-log model; mixed log-skewnormal, lognormal (10b)	Simulation	Records - distributions, Predictions - regression	None	×	×	×

Notes: (1) only moments considered. (2a) Homogeneous; non-homogeneous: lognormal, log-Weibull based functions; (b) exponential, lognormal, Weibull, log-Weibull, Pareto αStable (log, symmetric); (c) compound processes (e.g. Poisson, Cox) described but not applied. (3) *EM* – Expectation Maximisation. (4a–c) Per simulation example, *RUM* (utility – Pareto, random term – Normal). (5) *GPD* – Generalised Pareto Distribution. (6) Per simulation example (single claim per period, with certainty). (7) *SEIR* – Susceptible-Exposed-Infectious-Recovered. (8) Spliced (exponential, *GPD*), Weibull, gamma, lognormal. (9a) With parameters for *interdependence*, *disseminated information*; (b) direct and disclosure costs, security investment. 10a) "Daily" and "large" respectively; (b) log-log (Jacobs, 2014). Outputs: non-exhaustive examples. Exposure (*): conditional (e.g. given breach). Features: Π (considered) ⊆ (otherwise)

3. Data

Data is drawn from Ponemon Institute (2012a–2012i, 2013a–2013j, 2014a–2014k, 2015) global and country-level cost of data breach survey reports (hereafter, 2012–2015 years respectively)

https://doi.org/10.1017/S1357321724000205 Published online by Cambridge University Press

which form part of the PON (2019) data source. These reports feature estimated organisational costs in respect of publicly disclosed data breaches (loss or theft of personally identifiable records such as names and account numbers). Relevant information and basic preparation for subsequent analysis includes:

- Costs are subdivided into four "cost centres" – A: detection and escalation; B: notification; C: ex-post response; and D: lost business (hereafter, classes A–D respectively, with class E being the total)
- Years 2012–2014 (country-reports) – organisation-level costs, by class, are collated and US-dollar converted at prevailing exchange rates
- Year 2015 (global) – class E costs (in US dollars) are depicted in various "one-way" graphs (e.g. by rank of mean time to discover a breach); R-based image-scraping software, Webplotdigitiser (Rohatgi, 2013), is used to obtain this data from Ponemon Institute (2015, Figure 20), before further scrutiny and adjustments (as described shortly)
- Mean and extrema (with respect to costs) are given, by class and year
- In terms of the 2015 year, extracted costs appear to resemble the corresponding data points reasonably well (partly due to the ordering represented, which results in volatile and easily identifiable costs). A graphical comparison reveals 8 discrepancies (<2.5% of the data points)

These are manually corrected; after doing so, the mean cost falls within 0.2% of the given value and extrema are exact. Table 2 summarises classes A–D in terms of underlying activities and reputational damage associated with breaches, alongside examples of first-party coverage (i.e. which protect the insured's assets).

Costs, by class, are inflation-adjusted to make them comparable for analysis, whilst ensuring associated distributions are not overly distorted as a result. It is worth noting, that, despite having the most desirable characteristics out of 19 other potentially useful sources, the data in hand is arguably of questionable veracity as far as an accurate, representative, experienced-based actuarial pricing exercise is concerned.

Table 2. Costs (classes A–E) and possible coverage. Descriptions for classes A–E are based on "global" cost of data breach reports (Ponemon Institute, 2012i, 2013j, 2014f, 2013j); specimen products are purely illustrative examples of first-party coverage in respect of associated costs: AIG – Illinois (Murphy, 2013); ACE –(Cresenzi & Alibrio, 2016); Federal Insurance – (Daigle & Cresenzi, 2018)

Composition of severity data		Plausible cyber-insurance
Class	**Associated costs**	
A: Detection and escalation	Detect and report breach (e.g. forensics, crisis management, internal communications, audit and assessment)	• *PortfolioSelect* (CyberEdge, Event Management) – AIG, Illinois National (covers A–C)
B: Notification	Notify data subjects (e.g. create contact database, determine regulatory requirements, external experts)	
C: Ex-post response	Assist data subjects in aftermath of privacy event (e.g. help desk, inbound communications, investigations, remediation, legal, product discounts, credit monitoring and identity protection, regulatory fines and penalties)	• *Chubb Cyber Enterprise Risk Management policy* (Cyber Incident Response Fund) – ACE (covers A–C)
D: Lost business	Abnormal churn, reputational damage, and diminished goodwill	• *Forefront portfolio* (CyberSecurity, Business Interruption) – Federal
E: Overall	Sum of *class A–D* costs	

https://doi.org/10.1017/S1357321724000205 Published online by Cambridge University Press

4. Loss Models

This section develops six variants (Models 4.1–4.6) to explore the impact of correlation and risk at the aggregate level. Variants of the theory is well-founded – only key definitions and equations shall be provided here – references are provided for the interested reader. Figure 1 illustrates how this theory relates to the each of the six models.

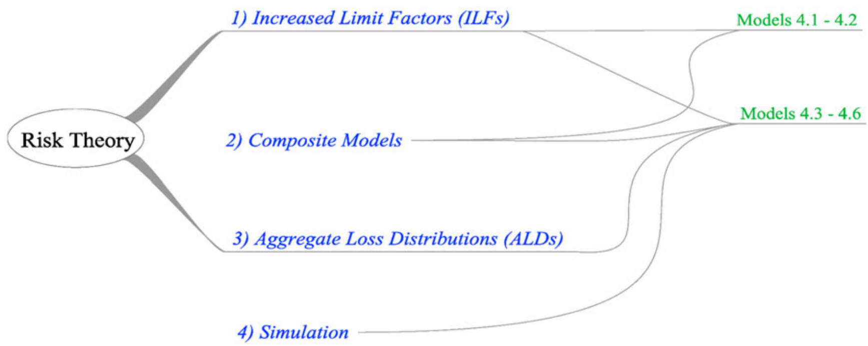


Figure 1. Outline of theory and model links. Theory 1–4 (blue, in addition to risk theory which introduces 1 and 3); Models 4.1–4.6 (green; all models rely upon 1 and 2; 3 and 4 are only utilised in support of Models 4.3–4.6). Generated using *Freemind* (Müller et al., 2004).

4.1. Risk Theory

Aggregate loss, S, represents the total amount for a given period and group of risks,

$$S = X_1 + X_2 +, \ldots, +X_N, \tag{1}$$

where N and X_i can be defined from two perspectives of risk theory, namely:

- Collective Risk (CR): loss count, N, and (non-negative) severities, $X_1, \ldots, X_N$, are random variables with independence assumptions as follows: N does not depend on the severity of loss; given N, X_is are i.i.d., independently with respect to count
- Individual Risk (IR): here, N denotes a fixed number of risks with respective losses, X_is that are independently distributed (as opposed to i.i.d.) random variables with mixed CDFs that may have mass at point zero (i.e. for the probability of no loss)

In terms of (1) – IR, CR models – determining the ALD is one of the classical problems in the realm of risk theory. As there is generally no closed-form solution (Shevchenko, 2010, Section 1) various techniques have been deployed: Fast Fourier Transform (FFT) can be used to reconstruct the density with the aid of the transforms (e.g. characteristic function, CF; moment or probability generating functions, MGF, PGF respectively – provided these exist) – (Kaas et al., 2008, Section 2.1).

4.2. Increased Limit Factors

An ILF is a multiplicative factor that is applied to the premium at a basic limit to determine the premium at an increased limit. Basic limits typically refer to the lowest levels of coverage provided, (Werner & Modlin, 2010). However, in principle, any non-negative limit can be contemplated for this purpose (hereafter, the term base limit is used instead of basic limit).

Limit definitions

A policy limit refers to the maximum amount payable under an insurance policy, either overall, or in respect of a particular section of a policy (Lloyd's, 2019), hereafter, known as "coverage section." This may be expressed on several bases; for demonstrative purposes, the following are assumed applicable for losses associated with classes A–E:

- Per-loss: applies to individual costs (i.e. classes A–D)
- Per-occurrence: applies to total cost (i.e. class E)

The limited random variable $X^{(b)}$ is defined as follows:

$$X^{(b)} = \min(X, b), \tag{2}$$

where X is a random variable and $\{b : b > 0\}$ is some limit. More generally, consider the limited variable $X^{(b)}$ (2), and suppose X has a CDF and PDF denoted by F and f respectively; the limited k$^{\text{th}}$-order moment of X, when limit b applies, can then be expressed in terms of the Riemann-Stieltjes integral:

$$\mathrm{E}\left(X^{(b)k}\right) = \mathrm{E}\left(\min(X, b)^k\right) = \int_0^b x^k dF(x)dx + b^k(1 - F(b)) = \int_0^b kx^{k-1} S_X(x)dx, \tag{3}$$

where $S_X = 1 - F$ and ($k = 1$ yields the LEV). Refer to Lee (1988) for a graphical illustration and Klugman et al. (2004) for a mathematical proof. Now let the aggregate loss in respect of limited severities (hereafter, Limited Aggregate Severity, LAS) be $S(b)$ defined by:

$$S(b) = \sum_{i=1}^{N} X_i^{(b)}, \tag{4}$$

where $b > 0$ is a given limit, and X_is are severities, and N is the loss count, as for the aggregate loss in (2). This gives rise to the following definition: let limit factor, γ, for a given base limit, a, be defined by:

$$\gamma(b) := \gamma(b; a) = \frac{\mathrm{E}(S(b))}{\mathrm{E}(S(a))}, \mathrm{E}(S(a)), \mathrm{E}(S(b)) > 0, \tag{5}$$

where $a, b > 0$. The term limit factor, for the purpose of the present research, refers to both discount factors and ILFs, defined as follows:

- Discount factor: $(a > b > 0) \Rightarrow \gamma(b; a) \in (0, 1)$; in this case, a could represent the highest limit of coverage, or, in the context of coverage without-limits, $a \to \infty$
- ILF: $(0 < a \leq b) \Rightarrow \gamma(b; a) \geq 1$; the conventional definition of an *ILF*, where a and b represent "basic" and increased limits respectively

In terms of (5), CR independence assumptions lead to the following expression for limit factors:

$$\gamma(b) = \frac{\mathrm{E}\left(X^{(b)}\right)}{\mathrm{E}\left(X^{(a)}\right)}. \tag{6}$$

Consistency properties

Limit factors satisfy consistency properties if they are asymptotically constant, have a monotonically decreasing and positive gradient, and are concave down.

https://doi.org/10.1017/S1357321724000205 Published online by Cambridge University Press

4.3. Risk Adjustments

Process risk refers to the inherent variability associated with the stochastic nature of frequency and severity of losses. To allow for this, actuaries may incorporate a risk adjustment into ILF calculations to achieve a certain risk margin (percentage increase in LAS).

Example 4.1. Variance principle risk adjustment

Let $\pi_{\text{var}}(S; w)$ denote the variance-adjusted (pure-risk) premium with respect to the aggregate loss amount, S, and a risk parameter, $w > 0$, be defined by:

$$\pi_{\text{var}}(S; w) = \mathrm{E}(S) + w\mathrm{Var}(S), \tag{7}$$

then the variance-adjusted limit factor, γ_S, can be defined as:

$$\gamma_S(b; a, w) = \frac{\pi_{\text{var}}(S(b); w)}{\pi_{\text{var}}(S(a); w)}, \tag{8}$$

where $S(a)$ and $S(b)$ are LASs (4) with limits $a, b > 0$ respectively. Independence assumptions (1) concerning loss count and i.i.d. severity (i.e. N, Y respectively), with Poisson N, simplify the risk-adjusted limit factor, γ_S (8) to the following:

$$\gamma_Y(b; a, w) = \frac{\pi^*_{\text{var}}(Y^{(b)}; w)}{\pi^*_{\text{var}}(Y^{(a)}; w)}, \pi^*_{\text{var}}(Y^{(b)}; w) = \pi_{\text{var}}(Y^{(b)}; w) + w\big(\mathrm{E}\big(Y^{(b)}\big)\big)^2. \tag{9}$$

Example 4.2. Excess losses with inflation and variance principle risk adjustment

For a compound Poisson "excess" LAS, $S = \sum_{i=1}^{N} Y_i$, where $Y_i = \max(0, vX_i^{(\frac{b}{v})} - d)$, $i = 1, 2, \ldots, N$ (i.e. $N \sim \mathit{Poisson}$), under CR independence assumptions (1), limits $a, b > 0$; deductible d, s.t. $0 \leq d < \min(a, b)$, and constant inflation $v > 1$, the variance-adjusted limit factor, γ_Y (9), with parameter w (as before), becomes:

$$\gamma_Y(b; a, d, w, v) = \frac{\pi^*_{\text{var}}\Big(X^{(\frac{b}{v})}, vw\Big) - \pi^*_{\text{var}}\Big(X^{(\frac{d}{v})}, vw\Big) - 2dw\Big[\mathrm{E}\Big(X^{(\frac{b}{v})}\Big) - \mathrm{E}\Big(X^{(\frac{d}{v})}\Big)\Big]}{\pi^*_{\text{var}}\Big(X^{(\frac{a}{v})}, vw\Big) - \pi^*_{\text{var}}\Big(X^{(\frac{d}{v})}, vw\Big) - 2dw\Big[\mathrm{E}\Big(X^{(\frac{a}{v})}\Big) - \mathrm{E}\Big(X^{(\frac{d}{v})}\Big)\Big]}, \tag{10}$$

where π^*_{var} is defined as previously. This can be shown through substitution $x = yv^{-1}$.

Example 4.3. Proportional-Hazard (PH) transform

Let π_{PH} be the mean in respect of the PH transform defined by:

$$\pi_{PH}(Y^{(b)}; b, w) = \int_0^b S_Y(x)^{\frac{1}{w}} dx, \tag{11}$$

where b is a given non-negative limit, and, and $w \geq 1$ (Wang, 1995, 1999a).

Example 4.4. Riebesell curves (power transform)

Let $\gamma(2^k a, a) = (1 + r)^k$, where γ is the ILF in respect of an increased limit and base limit, in this case, $2^k a$ and a respectively, with $a > 0$, $r \in (0, 1)$, and $k > 1$. It follows that:

$$\gamma(b; a, w) = (1 + r)^{\log_2(ba^{-1})} = (ba^{-1})^{\log_2(1+r)} = (ba^{-1})^w, \tag{12}$$

where $w = \log_2(1 + r)$. Refer to Mack and Fackler (2003) for details including origin.

https://doi.org/10.1017/S1357321724000205 Published online by Cambridge University Press

4.4. Correlated ALDs

This section describes CFs for correlated aggregate loss and count, based on pioneering contributions by Wang (1998, 1999a) and conventional techniques for mixture models (Klugman et al., 2004; Mildenhall, 2005).

Definition 4.1. *Covariance coefficient*

For random variables X_i andX_j, with Pearson correlation coefficient ρ_{ij}, means μ_i and standard deviations, σ_i, the covariance coefficient κ_{ij} is given by:

$$\kappa_{ij} = \frac{\text{Cov}(X_i, X_j)}{\mu_i \mu_j} = \frac{\rho_{ij}\sigma_i\sigma_j}{\mu_i\mu_j}. \tag{13}$$

The range of κ_{ij} (13) depends on the shape of marginal distributions for X_i and X_j.

CFs for correlated aggregate loss

Define the joint CF, $C_{\mathbf{S}} := C_{S_1,\dots,S_m}$, for $m \in \mathbf{Z}^+$ random variables, $\mathbf{S} = [S_1, \dots, S_m]$, by:

$$C_{\mathbf{S}}[\mathbf{t}] = \left(1 + \sum\nolimits_{i<j} \kappa_{ij}(1 - C_i[t_i])(1 - C_j[t_j])\right) \prod\nolimits_{i=1}^{m} C_i[t_i], \tag{14}$$

where $S_i, S_j \in \mathbf{S}$ have respective CFs C_i, C_j, and covariance coefficient κ_{ij}, $1 \le i < j \le m$; and $t = [t_1, \dots, t_m]$, Wang (1998, pt. IV). Following (14), let the univariate CF of $S = S_1 +, \dots, S_m$ be C_S, then:

$$C_S[t] = \left(1 + \sum\nolimits_{i<j} \kappa_{ij}(1 - C_i[t])(1 - C_j[t])\right) \prod\nolimits_{k=1}^{m} C_k[t], \tag{15}$$

where κ_{ij}s and C_is are defined as previously. The mean and variance of aggregate loss, S, is:

$$\begin{aligned} \mu &:= \text{E}(S) = \text{E}(S_1 +, \dots, +S_m), \\ \text{Var}(S) &= \sigma^2 + 2\sum\nolimits_{i<j} \kappa_{ij}\text{E}(S_i)\text{E}(S_j) \end{aligned} \tag{16}$$

where $\sigma^2 = \sum_{j=1}^{m} \text{Var}(S_j)$, Wang (1998). The univariate CF (15) is apparently less restrictive, in terms of covariance coefficients (for valid PDF), than is the case for the joint CF (16).

CFs for correlated loss count

Often, there is an exogenous cause for uncertainty regarding the extent or number of losses. This is referred to as parameter risk in the context of stochastic models (Freifelder, 1979, cited by Miccolis (1978). To reflect such uncertainty, a secondary mixture CDF can be incorporated within the model. In this section, a joint PGF for correlated aggregate loss count variables is built up using Poisson mixtures. Refer to Klugman et al. (2004, Section 4.6.10) for examples of various other mixtures with theoretical underpinnings.

Poisson mixture models

Let $\mathbf{N} = [N_1, \dots, N_m]$ be a vector of m discrete random variables with joint PGF given by $P_{\mathbf{N}} := P_{N_1,\dots,N_n}$ and assume there exists a random variable θ with MGF M_θ such that $(N_j|\theta = \omega) \sim Poisson(\lambda_j\omega)$ (B.3) where $\text{E}N_j(\theta = \omega) = \omega\lambda_j$, $j = 1, \dots, m$. The marginal PGF of $N_j|(\theta = \omega)$ is then $P_{N_j|\theta=\omega}[t_j] = e^{w\lambda_j(t_j-1)}$, which leads to the following joint PGF for $\mathbf{N}$:

$$P_{\mathbf{N}}[\mathbf{t}] = \text{E}_\theta\left(\text{E}\left({t_1}^{N_1} \dots {t_m}^{N_m}|\theta\right)\right) = \text{E}_\theta\left(\exp(\theta\boldsymbol{\lambda} \cdot (\mathbf{t}' - \mathbf{1}'_m))\right) = M_\theta[\boldsymbol{\lambda} \cdot (\mathbf{t}' - \mathbf{1}'_m)], \tag{17}$$

where $\boldsymbol{\lambda} = [\lambda_1, \dots, \lambda_m]$, $\mathbf{t} = [t_1, \dots, t_m]$, and $\mathbf{1}_m$ is a (row) vector with m ones.

https://doi.org/10.1017/S1357321724000205 Published online by Cambridge University Press

Example 4.5. Gamma-mixed Poisson model

Suppose $\theta \sim Gamma(\alpha, 1)$, for some $a > 0$, has MGF $M_\theta[t] = (1-t)^{-\alpha}$, then the joint PGF in (17) becomes $P_{\mathbf{N}}[\mathbf{t}] = (1 - \boldsymbol{\lambda} \cdot (\mathbf{t}' - \mathbf{1}'_m))^{-\alpha}$. This specifies a form of multivariate negative binomial CDF

where marginals, $N_j \sim NB(\alpha, \lambda_j)$ (B.4), have respective PGFs, P_{N_j}, $j = 1, \ldots, m$, defined by:

$$P_{N_j}[t_j] = (1 - \lambda_j(t_j - 1))^{-\alpha}, \tag{18}$$

(Wang, 1999b). Refer to Mildenhall (2005) for MGFs with alternative parameterisations, and Reshetar (2008) for practical application in the context of OR.

Example 4.6. Multivariate Negative Binomial (MNB) distribution

From Example 4.5, let $N_j \sim NB(a_j, \lambda_j)$ – the joint PGF, $P_{\mathbf{N}}$, is now:

$$P_{\mathbf{N}}[\mathbf{t}] = (\mathbf{1}_m \cdot \mathbf{k}' - m + 1)^{-\frac{1}{w}}, \tag{19}$$

where $\mathbf{t} = [t_1, \ldots, t_m], \mathbf{1}_m$ is a row vector of m ones; $\mathbf{k} = [k_1, \ldots, k_m]$ with $k_j = (1 - \lambda_j(t_j - 1))^{\alpha_j w}$, $j = 1, \ldots, m$; and $w \neq 0$. This specifies a family of MNB CDFs, with marginals $N_j \sim NB(\alpha_j, \lambda_j)$, in either of the following cases:

1. $0 < w < \min_{j \in [1,m]} \{\alpha_j^{-1}\}$
2. $w < 0$ s.t. $P_{\mathbf{N}}[\mathbf{0}_m] > 0$ and $-\frac{1}{w} \in \mathbf{Z}^+$

where $\mathbf{0}_m$ is a row vector of m zeros, (Wang, 1998).

Here, the random vector $\mathbf{N}$ follows an MNB distribution, denoted by $\mathbf{N} \sim \text{MNB}(\boldsymbol{\alpha}, \boldsymbol{\lambda}, w)$ with vector parameters $\alpha = [\alpha_1, \ldots, \alpha_m]$ and $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_m]$. Suppose $S_1, \ldots, S_m$ represent $m \in \mathbf{Z}^+$ CR loss models (1) that are specified by their severities and loss count variables, (X_i, N_i), $i = 1, \ldots, m$, and only correlated through $\mathbf{N} = [N_1, \ldots, N_m] \sim \text{MNB}\,(\boldsymbol{\alpha}, \boldsymbol{\lambda}, w)$ (Example 4.6). Accordingly, the CF for the overall aggregate loss, $C_S := C_{S_1 +, \ldots, +S_m}$, is defined by:

$$C_S[t] = (\mathbf{1}_m \cdot \mathbf{y}' - m + 1)^{-\frac{1}{w}}, \tag{20}$$

where $\mathbf{1}_m$ is a row vector of m ones, $\mathbf{y} = [y_1, \ldots, y_m]$ with $y_j = (1 - \lambda_j(C_j - 1))^{\alpha_j w}$, C_j is the CF of X_j $j = 1, \ldots, m$ (Meyers & Heckman, 1984; Wang, 1998). As such, FFT reconstructs the CDF of $S = S_1 +, \ldots, +S_m$, from transforms C_S (20). The mean and variance of S can be determined using (16) – substituting κ_{ij} with w, the correlation parameter in (20).

4.5. Severity Model

Define a two component spliced model in terms of n observed severities, ordered as $x_1 < x_2 <, \ldots, < x_n$. Losses in the interval $[0, \tau]$, for a given non-negative threshold, τ (i.e. "splicing point"), are assumed to follow a small loss CDF (in this case, estimated by the empirical CDF, F_n). To cover the interval (τ, ∞), a parametric distribution G is estimated using (observed) losses greater than τ. Now let H be the spliced distribution in question:

$$\begin{aligned} 1 - H(x) &= \begin{cases} 1 - F_n(x) & x \leq \tau \\ (1 - F_n(\tau))\left(1 - \frac{G(x) - G(\tau)}{1 - G(\tau)}\right) & x > \tau \end{cases} \\ &= \begin{cases} 1 - F_n(x) & x \leq \tau \\ (1 - F_n(\tau))\left(\frac{1 - G(x)}{1 - G(\tau)}\right) & x > \tau \end{cases} \end{aligned} \tag{21}$$

https://doi.org/10.1017/S1357321724000205 Published online by Cambridge University Press

where the first component CDF, $F_n(x)/F_n(\tau)$ (for $x \leq \tau$) and second component CDF, $(G(x) - G(\tau))/(1 - G(\tau))$, for $x > \tau$, are spliced with weights $F_n(\tau)$ and $1 - F_n(\tau)$ respectively (Klugman et al., 2004).

Selection (large-loss model)

The following steps are used to select a large-loss CDF, from a set of $k \in \mathbf{Z}^+$ candidate models (e.g. Burr, Weibull, Pareto, etc.) and identify a suitable threshold for application of the spliced model in (21) (i.e. given $\mathbf{x}_n = [x_1, \ldots, x_n]$):

Step 1 Fit $m > 1$ CDFs, $G_{i1}, \ldots, G_{im}$, to the largest $n - i + 1$ severities, for some $i = 2, 3, \ldots, n - k - 1$, where $k \leq n - 2$ is the minimum number parameter estimates for each CDF (e.g. based on Maximum Likelihood Estimation, MLE).

Step 2 Let $G_i^* = \min_j \{c_j\}$, where c_j is the AIC^c for G_{ij}, $j = 1, \ldots, m$.

Step 3 Calculate B_i^*, the KS-ratio (ratio of the Kolmogorov Smirnov test statistic (Glivenko-Cantelli – van der Vaart (1998)) to the critical value at the specified level) for G_i^*.

Steps 1–3 have the following outputs: the large-loss distribution, G_i^*, empirical threshold, x_i, and KS-ratio, B_i^* (valid scores require $i = 2, \ldots, n - k - 1$, as in step 1). In terms of the spliced model, H (21), $G_i^*(\mathrm{x}) = (G(x) - G(\tau))/(1 - G(\tau)), x > \tau$ and $x_i \leq \tau < x_{i+1}$ – if $\tau < x_2$ or $\tau > x_{n-k-1}$, then the unconditional CDFs, G and F_n respectively, might be used. The threshold itself can be expressed in terms of the empirical rank as follows:

$$j = nF_n(\tau), \tag{22}$$

where $j = 1, \ldots, n$; F_n, τ, and x_1 are defined as previously (21).

Threshold determination

A score-based approach (Klugman et al., 2004, Section 13.5.3) is adopted using a similar set-up adopted for Maximum Likelihood by Ralucavernic (2009), but with greater emphasis being placed on tail fit and limit-factor consistency. Differentiability and continuity requirements (Cerchiara & Acri, 2016) are not explicitly allowed for, however, model selection incorporates the corrected Akaike Information Criteria (AIC^C); refer to Akaike (1998) and Burnham and Anderson (2002) for details. This provides a practical and simplified means to identify both parametric CDF and threshold – additional considerations pertain to limit-factor consistency and mean excess (ME) plots.

Criteria 4.1 Splicing point

The following criteria are contemplated for determining threshold, τ, in terms of output from steps 1–3:

1. τ, with the greatest rank, i.
2. τ, with the lowest KS-ratio, B_i^*.

In this way, larger thresholds are favoured through the first criterion, whilst the second attempts to optimise tail fit. Upper bounds are established subjectively by considering ME plots.

Normalising scores

According to the set of Criteria 4.1, preference is given to higher and lower values of x_is and B_i^*s respectively (i.e. steps 1–3). Equivalently, higher values of α_i and β_i, defined as follows, are favoured over lower values:

https://doi.org/10.1017/S1357321724000205 Published online by Cambridge University Press

$$\alpha_i = \frac{x_i}{x_{n-k-1}}, \beta_i = \frac{\min\limits_{i \in [1,n]}\{B_i^*\}}{B_i^*} \forall B_i^* > 0 \tag{23}$$

The weighted average score, z_i, with respect to measures α_i and β_i (23), is determined by:

$$z_i = w_i \alpha_i + (1 - w_i)\beta_i, \tag{24}$$

With i defined as previously in step 1; thus $\alpha_i, \beta_i \in (0, 1]$ are on the same scale. $w_i^{(2)}$, which reduces as τ increases, can be defined as follows:

$$w_i^{(2)} = \frac{n - i}{n} \tag{25}$$

Algorithm 1. Optimal threshold and large-loss CDF
For a given group (i.e. class) of n ordered, homogeneous, and independent severities, $x_1, \ldots, x_n$, with empirical CDF $F_n = \frac{1}{n}\sum\limits_{i=1}^{n} 1_{\{x_i \le x\}}$; steps 1-3 (p. 14) are run for each $i \in [2, n]$ to produce the following input vectors for this algorithm:

- $\mathbf{G} = [G_2^*, G_3^*, \ldots, G_{n-k-1}^*]$ (i.e. selected large-loss distributions from step 3).
- $\mathbf{x} = [x_2^*, \ldots, x_{n-k-1}^*]$ (i.e. vector of "thresholds").
- $\mathbf{B} = [B_2^*, B_3^*, \ldots, B_{n-k-1}^*]$ (i.e. associated vector of KS-ratios).

Next, (23) is applied to $\mathbf{x}$ and $\mathbf{B}$ (element by element) to obtain the vector of scores $\boldsymbol{\alpha} = [\alpha_2, \ldots, \alpha_n]$ and $\boldsymbol{\beta} = [\beta_2, \ldots, \beta_n]$ respectively. For a given vector of weights $\mathbf{w} = [w_2, \ldots, w_n]$, where $w_i \in (0, 1) \forall i = 2, 3, \ldots, n$, the vector of (calculated) weighted scores, $\mathbf{z} = [z_2, \ldots, z_n]$, is determined using (24). The optimal threshold, τ^*, is x_{i*}, where $i^* \in \{2, 3, \ldots, n\}$ is the optimal index value that yields the solution to the following:

$$z_{i*} = \max\{z_i : i = 2, 3, \ldots, n\}. \tag{26}$$

The corresponding (parameterised) optimal distribution is then $G^*_{nF_n(\tau*)} = G^*_{i*}$ (which follows from (22) with $j := i^*$). Thus, the outputs of this algorithm are the optimal threshold, optimal index value, and optimal distribution (i.e. τ^*, i^*, and G^*_{i*} respectively).

Algorithm 2. Model confidence sets – Kullback-Leibler
This algorithm follows the bootstrap approach of Burnham & Anderson 2002 (Section 4.5), which is based on essential Kullback & Leibler (1951) theory associated with AIC and other such information criteria. For each candidate CDF (i.e. parametric family), G_i, and bootstrap sample indexed $i = 1, \ldots, m$ and$j = 1, \ldots, M$ respectively, $m, M > 2$, determine Akaike differences, δ_{ij}, in relation to the minimum AIC^C, $A_j^* = \min\limits_{i=1,..,m} \{A_{ij}\}$, and associated Akaike weights, w_{ij} (that sum to one for each sample) as follows:

$$\delta_{ij} = A_{ij} - A_j^* \quad w_{ij} = \frac{\exp(-0.5\delta_{ij})}{\sum_{u=1}^{m} \exp(-0.5\delta_{ij})} \tag{27}$$

where A_{ij} is the AIC^C score for CDF G_i, parameterised (e.g. using MLE) in respect of data for sample $j \in \{1, \ldots, M\}$. Differences and weights accompanying the M samples can provide insight into model (in this case, CDF) selection uncertainty. For instance, in terms of the following "model confidence set" and selection probability estimates:

https://doi.org/10.1017/S1357321724000205 Published online by Cambridge University Press

- The 100 α% "Kullback-Leibler" (KB) confidence set, for specified CDF with (common) index $s \in \{1, \ldots, m\}$,comprises the set of candidate CDFs with corresponding Akaike differences below the 100α% empirical quantile, $q^{(\alpha)}$,of Akaike differences for the specified CDF; the probability that CDF indexed $i = 1, \ldots, m$ is in such a confidence set, $c_i^{(\alpha)}$, can be estimated from the samples as follows: $\hat{c}_i^{(\alpha)} = M^{-1} \sum_{j=1}^{M} 1_{\{A_{ij} - A_j^* \le q^{(\alpha)}\}}$(where, in general, indicator $1_{\{A\}} = 1$ if a given event A occurs – failing which, $1_{\{A\}} = 0$).
- Correspondence between the average weight, $\hat{w}_i = M^{-1} \sum_{j=1}^{M} w_{ij}$, for a given CDF with index $i = 1, \ldots, m$, and the proportion of (M) minimum Akaike scores that correspond to the CDF in question, $\hat{\pi}_i = M^{-1} \sum_{j=1}^{M} 1_{\{\delta_{ij}=0\}}$, attests to the veracity of the (aforementioned) KL confidence set, and associated model inference uncertainty.

4.6. Limit Factor and Aggregate Loss Models

This section describes and formulates various models, which are grouped in Figure 2 according to whether correlation is recognised, and how loss count, N, is modelled:

- IR framework: $N = n$ is given.
- CR framework: N is a random variable with a given PDF.

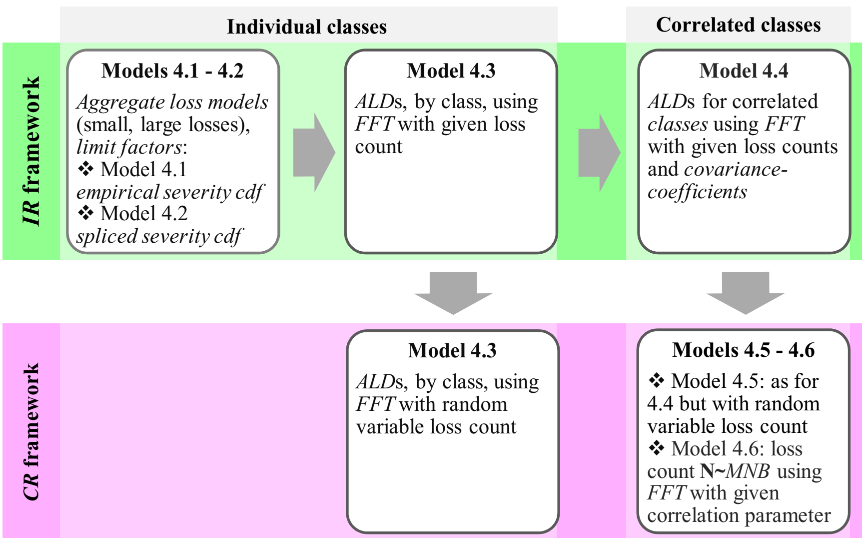


Figure 2. Flow chart for Models 4.1–4.6. Models 4.4–4.5 and Model 4.6 assume correlated aggregate loss amounts and counts (*classes A–D*) respectively. Adjustments (e.g. inflation, risk) may apply to limit factors based on any of these models.

In this way, IR represents a special type of CR, where N has a degenerate distribution such that $Pr(N = n) = 1$, as contemplated by Klugman, Panjer & Willmot (2004, Section 6.1).

Models 4.1 and 4.2 Limit factors for independent, individual classes (IR model)

The following is an overview of Models 4.1–4.6, as depicted in Figure 2:

https://doi.org/10.1017/S1357321724000205 Published online by Cambridge University Press

- Models 4.1–4.2 model aggregate losses in respect of small and large severities, using empirical CDFs and the spliced-severity model (Section 4.5); relevant limited moments (3) are used to determine the risk-adjusted LAS (7) and limit factors (8) in an IR framework with consideration for possible application in a CR framework.
- Model 4.3 (IR and CR) derives ALDs in respect of classes A–D (subject to per-loss limits), and class E (subject to a per-occurrence limit) from which limit factors are determined in respect of ground-up or excess losses; inflation and risk adjustments (10).
- Models 4.4 and 4.5 rely on given covariance coefficients between aggregate losses in classes A–D (15).
- Model 4.6 applies (20) with relevant parameters for the (correlated) marginal loss count CDFs (NB, Table B.2: B.4).

Models 4.1 and 4.2 are formerly defined in this section; Models 4.3–4.6 are more descriptive in nature and are framed in the context of tailored FFT steps, with compound Poisson and negative binomial applications for Models 4.5–4.6.

Assumptions for ILFs

The "top-slicing" method is used to determine ILFs; in respect of the risk premium. This assumes that severities, by class, are homogenous, independent, and independent of loss count; non-risk elements (e.g. expenses) are negligible (or proportional); and there is no anti-selection (e.g. by size of limit).

Variables and definitions

Define the following for a given class with n observed severities:

- F_n and τ: empirical CDF and splicing point respectively.
- $x_1 \leq, \ldots, \leq x_u \leq \tau$: the smallest, ordered, u (i.i.d.) severities with LAS, LEV, and "limited" variance denoted by $Z_S(b) = \sum_{i=1}^{u} x_i^{(b)}, \mu_{S;b} = \mathrm{E}\left(X_S^{(b)}\right) = \frac{1}{u}\sum_{i=1}^{u} x_i^{(b)}$, and $\sigma^2_{S;b} = \mathrm{Var}\left(X_S^{(b)}\right) = \frac{1}{u}\sum_{i=1}^{u} (x_i^{(b)} - \mu_{S;b})^2$ respectively, where $b > 0$ is a single limit that applies to severity (maximum payable in respect of individual claims); $u = nF_n(\tau) \in \{0, 1, \ldots, n\}$; $X_S \in \{x_1, \ldots, x_u\}$ is the small severity random variable: $x_i \sim^d X_S, i = 1, \ldots, u$, $X_S \sim F_n$.
- $X_1, \ldots, X_{n-u}$: $n-u$ random variable "large" severities with LAS, LEV, and limited variance $Z_L(b) = \sum_{i=1}^{u} X_i^{(b)}$, $\mu_{L;b} = \mathrm{E}\left(X_L^{(b)}\right)$, and $\sigma^2_{L;b} = \mathrm{Var}\left(X_L^{(b)}\right)$ respectively, where X_i are i.i.d. such that $X_i \sim^d X_L, i = 1, \ldots, n-u$; $X_L \sim G$, where X_L and G are large severity and its CDF (unconditional with respect to τ), respectively; $X_L \perp X_S$; b is the limit as before.

Thus, $\mu_{L;b} = \int_0^b S_X(x)dx$ and $\sigma^2_{L;b} = 2\int_0^b xS_X(x)dx - \mu_{L;b}{}^2$, which follows from Equation (3) with $k = 1, 2$ respectively, and $S_X = 1 - F_X$ where $F_X(\mathrm{x}) = (G(x) - G(\tau))/(1 - G(\tau))$, $x > \tau$ ($F_X(x) = 0, x \leq \tau$). The overall aggregate loss, Z, its mean, μ_Z, variance, σ_Z^2, and associated (variance principle) risk-adjusted LAS, $\pi_Z := \pi_{\mathrm{var}}$ ((7), $S = Z$), and limit factor, $\gamma_Z := \gamma_S$ ((8), $S = Z$), are defined by Models 4.1–4.2, in an IR framework, as follows:

$$Z(b) = Z_S(b) + Z_L(b) = \sum_{i=1}^{u} x_i^{(b)} + \sum_{i=1}^{n-u} X_i^{(b)}$$
$$\mathrm{E}(Z(b)) = \mu_{Z;b} = u\mu_{S;b} + (n-u)\mu_{L;b}; \mathrm{Var}(Z(b)) = \sigma^2_{Z;b} = u\sigma^2_{S;b} + (n-u)\sigma^2_{L;b} \tag{28}$$
$$\pi_{Z;b} = \mu_{Z;b} + w\sigma^2_{Z;b}; \gamma_{Z;a,b} = \frac{\pi_{Z;b}}{\pi_{Z;a}}$$

https://doi.org/10.1017/S1357321724000205 Published online by Cambridge University Press

where $a, b > 0$; $(u, n, b, Z_S, Z_L, \mu_{S;b}, \mu_{L;b}, \sigma^2_{L;b})$ as before; and $\mathrm{Cov}(X_S X)_L = 0$. Models 4.1–4.2 can now be distinguished from one another as follows:

- Model 4.1 – by setting $u = n$ (or equivalently, $\tau \geq x_n$, the maximum observed severity), X_L and associated terms in Equation (28) become redundant and Z, μ_Z, $\sigma^2_{Z},.$ $\pi_{Z},.$ and γ_Z are expressed solely in terms of x_i, $i = 1, \ldots, n$) and calculated numerically.
- Model 4.2 – this relies on the spliced-severity model (and associated algorithms) developed in Section 4.5, by setting τ, u, and G to the optimal outputs (i.e. threshold τ^*, index i^*, and large-loss CDF, G^*_{i*} respectively); analytical solutions, are checked using Model Risk by Vose (2019), risk analysis software and simulation (e.g. Appendix C)

ILFs and associated measures for Models 4.1–4.2 can then be determined for a range of different splicing points and associated (small and large) severity CDFs. Model 4.2 can easily be amended to cater for the CR framework.

Attention is now turned to Models 4.3–4.6, which utilise FFT as summarised in Table 3.

Model 4.3 ALD for independent classes (IR, CR models)

Of Models 4.3–4.6, this model represents the most straightforward application of FFT. In terms of steps 1–4 (Table 3), consider a class with n observed severities. Model 4.3 (IR) proceeds with step 1 by discretising the spliced-severity distribution (of limited severities) using the rounding method. The corresponding vector of CFs (determined in step 2) are raised to the power of n (element by element) to obtain CFs in respect of ALDs (step 3a), which are yielded using the inverse Fourier transform (step 4). Model 4.3 (CR) is very similar except, instead of raising severity CFs to the power of n (step 3a), the PGF of an assumed loss count CDF (in this case, Poisson) is incorporated (steps 1, 2, and 4 remain otherwise unchanged.

Table 3. FFT steps for ALDs (Models 4.3–4.6) (✓) if step is relevant, (x) otherwise

Model:	**Steps 1 - 2** 1) *Discretise* (limited, *spliced*) severity *cdf*s in respect of *classes A-E*; 2) Apply *FFT* (element by element) to obtain their *cfs*	**Step 3a** *Cf*s (*step 2*): raise to power of *n* (i.e. given loss count), or apply within *pgf* of *N* (i.e. random variable *loss count*) to obtain aggregate loss *cf*s	**Step 3b** Combine *cf*s (*step 3a*) to obtain overall aggregate loss *cf* using given *covariance coefficients* or *Multi-NegBin model*	**Step 4** Reconstruct *ALD*(s) from *cf*(s) in penultimate step (i.e. *step 3a* or *step 3b*), using *inverse FFT*
Model 4.3 (*IR*)	✓	Raise *cf*s to power of *n*	x	*Inverse FFT* (*cf*s: *step 3a*)
Model 4.3 (*CR*)	✓	Apply *cfs* in *pgf* of *N*	x	*Inverse FFT* (*cf*s: *step 3a*)
Model 4.4	✓	Raise *cf*s to power of *n*	Combine using *cov coeff*	*Inverse FFT* (*cf*: *step 3b*)
Model 4.5	✓	Apply *cfs* in *pgf* of *N*	Combine using *cov coeff*	*Inverse FFT* (*cf*: *step 3b*)
Model 4.6	✓	Apply *cfs* in NegBin *pgf*s	Combine with *MNB*	*Inverse FFT* (*cf*: *step 3b*)

Model 4.4 ALD for correlated aggregate losses (IR model)

Step 3a is relevant for Model 4.4 as this is based on the IR framework which assumes each class has a (deterministic) loss count, n. The CF for each of the classes A–D is thus raised to the power of n (element by element) to obtain corresponding (class-level) CFs in respect of their marginal ALDs. Step 3b combines these using (15) (with $m = 4$, and assumed covariance coefficients,

https://doi.org/10.1017/S1357321724000205 Published online by Cambridge University Press

$\kappa_{ij} = \kappa$), before taking the inverse Fourier transform in step 4 to yield the aggregate loss CDF (i.e. joint CDF for correlated marginal ALDs with respect to classes A–D).

Model 4.5 ALD for correlated aggregate losses (CR model)

Model 4.5 is the CR analogue to Model 4.4. Instead of raising the CFs in each of the classes A–D to the power of a deterministic count parameter, n, as is the case for Model 4.4 in step 3a, the PGF of an assumed loss count variable is incorporated within the CF (element by element). This yields the CFs in respect of the (marginal) ALDs for each of the classes A–D. Step 3b (i.e. application of (15) with given marginals and covariance coefficients) and step 4 (i.e. inverse Fourier transform) used in this model are otherwise identical to those used for Model 4.4. For variance principle adjustments regarding limit factors, (16) is utilised.

Model 4.6 ALD for correlated loss count (CR model)

Model 4.6 utilises a (multivariate) mixture model, as considered for Example 4.6.

In particular, step 3a assumes that the class has random variable loss count, N_j, with $NB(a_j, \lambda_j)$ CDF and specified parameters a_j, λ_j, $j = 1, 2, 3, 4$ (Table B.2: B.4). The associated PGF is thus incorporated (element by element) within CFs in step 2 to produce (class-level) vectors of CFs (step 3a) for respective ALDs. These are then combined using (15) (with $m = 4$, and assumed correlation parameter, w) in step 3b, before using the inverse Fourier transform to yield the aggregate loss CDF in step 4 (i.e. joint CDF in respect of classes with correlated aggregate NB loss count variables).

5. Results

Severity distributions are first identified; these are used to devise ALDs from which ILFs are determined to study the impact of risk adjustments in the presence of correlation.

5.1. Severity CDFs

Final selections (CDFs, thresholds), based on Algorithm 1 (Table B.1) **are in** Table 4.

Table 4. Selected large-loss CDFs and splicing points. Threshold: dollar value of splicing point; Burr represents inverse Burr (i.e. Dagum CDF); CDFs fit using MLE to severities from Ponemon Institute (2012a–2012i, 2013a–2013j, 2014a–2014k, 2015, inflated to end of 2016

	A	*B*	*C*	*D*	*E*
Threshold (\$m)	1.40	0.29	1.67	4.14	6.50
Percentile	87.3%	77.5%	81.0%	92.1%	83.9%
Distribution	Weibull	Burr	Burr	Weibull	Weibull
Shape	0.76	2.12, 0.53	2.13, 0.57	1.56	1.11
Scale (\$m)	0.82	0.38	1.34	3.37	3.81
Location (\$m)	1.40	0.29	1.67	4.14	6.50

Turning to model confidence sets (Algorithm 2), the first column of Table 5 show the top four models according to how frequently they were selected on the basis of AIC^{C}. Key observations include:

- Selected %($\hat{\pi}$, following (27)), AIC weight($\hat{w}$), KS and Anderson Darling (AD) ratios (i.e. test statistic to critical value) are in agreement; $\hat{\pi}$ and $\hat{w}$ are highest for selected CDFs, except for

https://doi.org/10.1017/S1357321724000205 Published online by Cambridge University Press

C (Weibull, the highest, fails the AD-test, 5% critical; also, the selected Burr CDF has a similar 90% confidence set success rate, $\hat{c}^{(90\%)}$).

- Light-tailed CDF selections are confirmed for D, E (with average shape $\alpha > 1$).
- Lowest and highest $\hat{c}^{(90\%)}$ can be seen for D (due to high, 92.5% truncation, Table 4) and E (due to additional 350 observations, year 2015) respectively.

Table 5. Bootstrap results. 10k samples; selected % achieving minimum AIC^{C}; 90% confidence sets based on Kullback-Leibler distance estimate for selected CDF (colour coded font, A–E – average shape parameter for Weibull CDF selections). Tail-fit ratios (KS, AD – 5% critical); consistent ILFs (rate per 100). Underlying costs based on Ponemon Institute (2012a–2012i, 2013a–2013j, 2014a–2014k, 2015 inflated to 2016

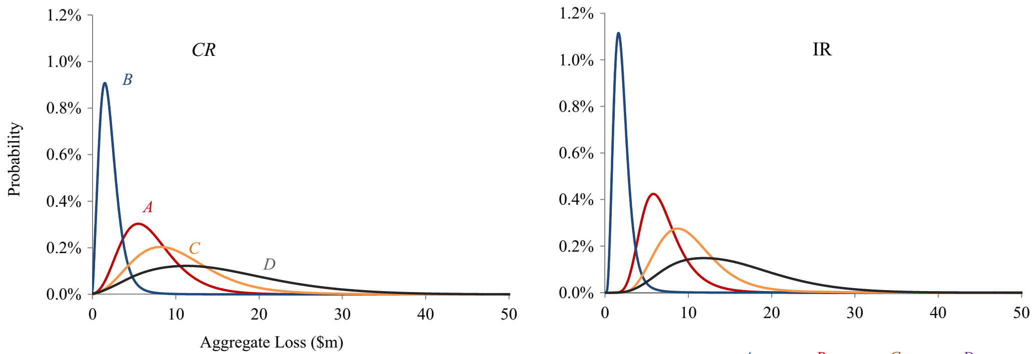
	Model	*Selected %*	*AIC weight*	*Confidence set %*	*KS-ratio*	*AD-ratio*	*Consistent ILFs*
A	Weibull3 (α=0.76)	79%	77%	68%	0.5	0.6	99.8
	Burr	17%	18%	25%	1.1	1.3	54.4
	Fatigue	3%	3%	2%	2.5	17.8	24.5
	LogLaplace	1%	1%	2%	3.4	67.8	99.5
B	Burr	74%	67%	48%	0.7	0.5	99.9
	Weibull3	22%	22%	23%	0.7	1.1	99.9
	LogGamma	3%	7%	17%	0.4	0.2	100
	GEV	1%	2%	8%	0.4	0.2	100
C	Weibull3	55%	52%	26%	0.7	1.0	99.9
	Burr	29%	29%	25%	0.8	0.6	98.0
	LogGamma	14%	12%	17%	0.4	0.2	100
	LogLaplace	1%	2%	8%	1.0	0.3	99.9
D	Weibull3 (α=1.20)	49%	34%	11%	0.6	0.6	99.8
	Fatigue	21%	21%	4%	2.2	14.0	41.9
	Burr	9%	9%	10%	1.5	1.5	52.1
	Pearson5	8%	6%	8%	0.5	0.2	100
E	Weibull3 (α=1.04)	85%	80%	74%	0.5	0.6	100
	Burr	8%	10%	13%	0.9	0.3	95.6
	LogGamma	7%	8%	12%	0.4	0.2	100
	LogLaplace	0%	1%	1%	1.7	0.7	99.3

5.2. ALDs

ALDs in Figures 3 and 4 are now considered in terms of underlying costs (Table 2):

- B: this has the lowest mean (Figure 3) and largest kurtosis – in keeping with the fact that these costs are not significant drivers of overall loss (e.g. data recreation, expert engagement, possibly customer notification); and the element of "determining regulatory requirements," suggesting a heavier tail than otherwise (i.e. in support of Burr, Table 4)
- A, C: most similar in terms of ALDs and moments – this agrees with underlying cost types which appear to be overlapping in some respects (e.g. forensic, investigative, communication, assessment costs); however, the nature of other costs in C (legal, regulatory fines and penalties, product discounts, and credit monitoring) would explain its relatively larger moments and heavier tail.
- D: the largest mean and, as implied by the lowest kurtosis and skewness (relative to mean), lightest tail (severity CDF and ALD) – this appears to reflect the nature of the underlying extrapolated cost estimate that has been derived from some other distribution. Further investigation shows that positive correlation with D is a key driver for the bimodal feature that can be seen in Figure 4 for scenario 3 (more prominent in Model 4.4 than 4.5).

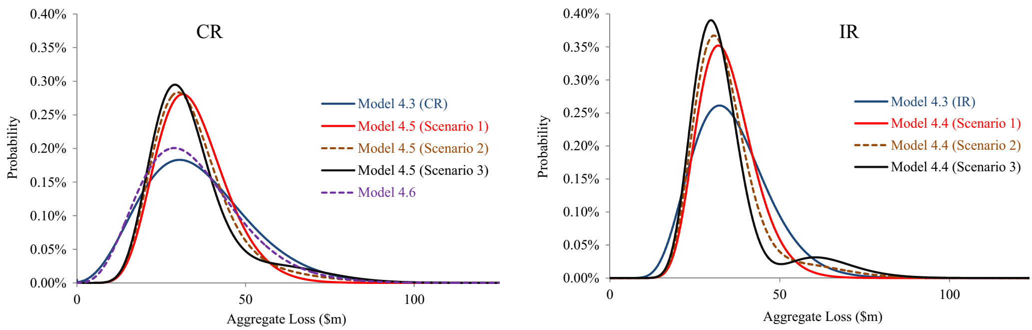
https://doi.org/10.1017/S1357321724000205 Published online by Cambridge University Press



CR	A	B	C	D
Mean	7.05	2.12	10.35	14.99
Min	0	0	0	0
Std dev	3.48	1.39	5.24	8.08
Skew	1.00	3.03	1.10	0.84
Kurt	4.67	28.60	5.25	3.87
VaR $_{1\%}$	17.58	6.56	26.87	38.38

IR	A	B	C	D
Mean	7.05	2.12	10.35	14.99
Min	0.54	0	0.61	0.09
Std dev	2.67	1.22	4.10	6.54
Skew	1.23	3.93	1.39	0.77
Kurt	5.74	43.29	6.91	3.68
VaR $_{1\%}$	15.55	6.08	24.36	33.61

Figure 3. ALDs: Model 4.3 Loss count: *CR* – *Poisson*(10); *IR* – 10 (deterministic). Ponemon Institute (2012a–2012i, 2013a–2013j, 2014a–2014k, 2015) costs inflated to end of 2016. Per-loss limit ($20m, *A-D*).



	Model 4.3 (*CR*)	Scenario 1	Scenario 2	Scenario 3	Model 4.6
Mean	35.84	34.46	34.46	34.46	34.46
Min	0	1.75	1.68	1.61	0
Std dev	15.83	10.33	12.12	13.67	14.75
Skew	0.66	0.59	1.27	1.52	0.75
Kurt	3.56	3.51	5.85	6.08	3.79
VaR $_{1\%}$	79.85	62.88	75.40	81.58	76.44

	Model 4.3 (*IR*)	Scenario 1	Scenario 2	Scenario 3
Mean	35.84	34.46	34.46	34.46
Min	3.50	7.22	7.15	7.08
Std dev	11.06	8.26	10.41	12.18
Skew	0.62	0.61	1.63	1.84
Kurt	3.50	3.55	7.08	6.79
VaR $_{1\%}$	66.28	57.23	72.28	77.31

Figure 4. ALDs: Models 4.3–4.6 $m; Scenarios 1–3: constant covariance coefficients of 0%, 5%, 10% resp., for Models 4.4 (IR) and 4.5 (CR). Loss count: Poisson(10) (Models 4.3–4.5, CR); MNB(10,1,0.09) for Model 4.6; IR: 10 (deterministic). Ponemon Institute (2012a–2012i, 2013a–2013j, 2014a–2014k, 2015) costs inflated to end of 2016. Per-occurrence limit (class *E*).

It is worth considering the point regarding class D and the bimodal formation in greater detail. The implications of a correlated class D (business interruption) may be exacerbated in the presence of interdependent organisational structures (processes, activities). For instance, vertically integrated businesses that suffer losses due to a common cause may find there is inadequate coverage (e.g. if such losses erode a common aggregate limit). Organising structures that avoid upstream (or downstream) dependences should assist in preserving coverage limits. However, this may not always be feasible (logistics, costs, etc.).

https://doi.org/10.1017/S1357321724000205 Published online by Cambridge University Press

5.3. Risk-Adjusted ILFs

Table 6 compares ILFs for several major league insurers to those based on Models 4.5–4.6 (low–high risk), filed by Cresenzi and Alibrio (2016) on behalf of ACE (Chubb, 2017). See Figure B.2 for a description of low-high-risk environments referenced in this table.

Table 6. Insurer ILF comparison (per-loss limits). Insurer comparison: 2016 ACE *SERFF* filing – *Chubb Enterprise Risk Management Cyber* and *Digitech* products (Cresenzi & Alibrio, 2016), with reference to (2015 year) *SERFF* filings by: *AIG* (*Speciality Risk Protector*) [AGNY-130104025], *Travellers* (*Cyber-Essentials*) [TRVD-130748646], *Philadelphia* (*Cyber-Security Liability*) [PHLX-G128091742], and *ACE* (*MPL Advantage*) [ACEH-125807939]. *$100m: *ILF*s estimated with *Riebesell* curve (implied at $10m limit). *Base limit*: $1m; retention: $10k. Shading: model range within insurer range (*A*:*B*):=(min, max); partial if ranges overlap. "Median": model ILF range. Ponemon Institute (2012a–2012i, 2013a–2013j, 2014a–2014k, 2015, inflated to end of 2016 (ILFs: adjusted to 2015)

	Limit						
Insurer	*$1m*	*$2m*	*$3m*	*$4m*	*$5m*	*$10m*	*$100m**
Chubb	1	1.29 - 1.50	1.49 - 1.89	1.65 - 2.21	1.77 - 2.50	2.20 - 3.60	4.84 - 12.96
AIG National	1	1.50	1.88	2.14	2.35	3.04	9.24
Travelers	1	1.42	1.62	1.83	1.99	2.73	7.44
Philadelphia	1	1.58	1.98	2.27	2.47	3.15	9.92 - 9.92
ACE	1	1.30 - 1.50	1.50 - 1.89	1.65 - 2.22	1.78 - 2.51	2.21 - 3.62	4.88 - 13.10
Overall range (A , B)	1	1.29 - 1.58	1.49 - 1.98	1.65 - 2.27	1.77 - 2.51	2.20 - 3.62	4.84 - 13.10
Models 4.5 - 4.6							
Low risk	1	1.37	1.56 - 1.57	1.66 - 1.68	1.72 - 1.75	1.84 - 1.89	3.39 - 3.57
(Median - A) / (B - A)	-	*27%*	*15%*	*3%*	*-5%*	*-24%*	*-16%*
Medium risk	1	1.43 - 1.45	1.67 - 1.73	1.81 - 1.90	1.90 - 2.02	2.09 - 2.31	4.37 - 5.32
(Median - A) / (B - A)	-	*52%*	*43%*	*33%*	*26%*	*0%*	*0%*
High risk	1	1.49 - 1.54	1.78 - 1.90	1.96 - 2.13	2.07 - 2.30	2.33 - 2.73	5.44 - 7.48
(Median - A) / (B - A)	-	*77%*	*71%*	*64%*	*57%*	*23%*	*20%*

Costs covered by the insurance products underlying Table 6 correspond with A–D categories (Table 2). Since insurer ILFs incorporate a base retention and base limit of $10k and $1m respectively, model ILFs are derived using Equation (10) with $v = 1.025$ (based on inflation used for E, year 2015), $d = \$10\,\text{k}$, and $a = \$1\,\text{m}$. Riebesell estimation ($100m) is based upon Equation (12) with $a = \$1\,\text{m}, b = \$100\,\text{m}$, and an insurer-specific parameter w that ranges from (0.34, 0.56). Overall, there appears to be reasonable correspondence between modelled and insurer ILFs. However, as might be expected, there is greater alignment with high-risk parameter as the limit increases.

6. Conclusions, recommendations

The model review (Section 2, Appendix A) found cyber-pricing models to be in need of further development and empirical support – particularly derelict aspects included severity and aggregate loss; there was no evidence of ILF related models. Empirical support, based on statistically viable severity data, featured only once (Biener et al. (2015) and included almost 1,000 cases. Key contributions made by the present research include:

1. Model confidence sets for various severity CDFs, derived in relation to key forms of first-party data breach coverage.
2. New insight into aspects associated with correlated ALDs and risk-adjusted ILFs.

https://doi.org/10.1017/S1357321724000205 Published online by Cambridge University Press

This was done in terms of nonparametric models based on empirical data, extracted from data breach survey reports (4×800 : A–D; 1 150 : E). There was no evidence of such applications or findings in the model review (or, to the best knowledge of the author, elsewhere in cyber-related research).

6.1. Conclusions

Conclusions, some of which are data or model dependent (i.e. not necessarily applicable in every situation) include:

- Severity distributions, based on data breach costs, were heavy tailed in the main, although D, representing business interruption, often affiliated with issues such as interdependence in the realm of insurance, was found to be light-tailed.
- Correlation between D and other classes (i.e. A–C) was found to have the greatest impact on the ALD in its tail (in the case where the aggregate loss model was used, the peak of the second mode of a bimodal distribution was intensified). The Value at Risk, however, was less affected by this compared to other risk measures (e.g. standard deviation).
- Empirical evidence suggests insurers are indeed avoiding volatile severity risk associated with increased cover limits, not only through low upper limits, but through increasing implied risk margins. Reducing Riebesell parameters support this view; in some (isolated) cases, this led to ILF consistency not being observed.

Enriched empirical data, as a basis for actuarial experience rating, may represent a source of value, despite the notion that it "quickly goes stale" due to the dynamic nature of the technological environment. This is demonstrated by reconciling modelled (i.e. "experience-based") and insurer (exposure-based) ILFs, and introduces the recommendations made in Section 6.2.

6.2 Recommendations

Wider audience

Onus should be placed on all stakeholders concerned to establish a unified approach to deal with common cyber-risk management issues – whilst industry groups and international initiatives are reportedly underway; actions to "better" address basic data issues are still highly anticipated.

Developing an anonymised "community-wide" data base (with key elements for quantifying cyber-risk) may be fraught with wider issues concerning cooperation, funding, administration, and governance. However, there would appear to be some incentive to collaborate more effectively, given the \$600bn (and growing) cyber-cost estimate previously mentioned (Section 1).

This would align with academic interests in support of such an initiative – although a unified approach may also be required here – possibly through a multidisciplinary academic interest group. Such cross-pollination would accelerate the development of cyber-risk and associated pricing models.

Specific directions – academia

There were only two "actuarial" contributions (according to title) that featured in the model review, neither of which appeared to have emerged from that domain. Given this, it is worth emphasising that further actuarial contribution to this specialised field of academia is essential.

Specific areas that warrant greater input include the following:

- Correlation and interdependence: risks within a class were assumed to be independent – simulation (e.g. common shock model) would be useful for understanding interdependence with respect to business interruption.

https://doi.org/10.1017/S1357321724000205 Published online by Cambridge University Press

- Information asymmetry: anti-selection (e.g. different limits attracting different types, levels of risk) could be explored using SERFF ILFs (e.g. Hanover, 2015) which differentiate by turnover; or considering class D divided by customer churn); empirical insight into the notion of secondary loss (Bandyopadhyay et al., 2010) and associated asymmetries (e.g. insureds' claiming strategy) could be investigated in terms of "retention factors" (for pricing different deductibles).

Acknowledgements. The author would like to express gratitude to Professor Thiart for her dedication, guidance, and patience throughout this research. The support provided by Stephen Wheldon and associates of the Institute and Faculty of Actuaries, as well as the external and independent examiner(s) and Chris Olsen, the External Supervisor, was instrumental in making this research possible. Special thanks are extended to the author's family members for their unwavering support and encouragement.

References

Akaike, H. (1998). Information Theory and an Extension of the Maximum Likelihood Principle, in *Selected Papers of Hirotugu Akaike* (E. Parzen, K. Tanabe, & G. Kitagawa) New York, NY, Springer, 199–213. https://doi.org/10.1007/978-1-4612-1694-0_15.

Baldwin, A., Gheyas, I., Ioannidis, C., Pym, D. & Williams, J. (2012). *Contagion in Cyber Security Attacks*. Berlin, DE, available at https://web.archive.org/web/20150813192331/http://infosecon.net/workshop/bibliography.php (internet archive) (accessed 23 July 2016).

Bandyopadhyay, T., Mookerjee, V. & Rao, R. (2010). A model to analyze the unfulfilled promise of cyber insurance: the impact of secondary loss. University of Texas at Dallas (working paper), available at ~https://web.archive.org/web/20190331231242/http://www.utdallas.edu/~rrao/CyberBMR[1].PDF (internet archive); (accessed 24 July 2015).

Bank for International Settlements (2003) *The 2002 loss data collection exercise for operational risk: summary of the data collected*, available at https://www.bis.org/bcbs/qis/ldce2002.htm (accessed 25 June 2015).

Barracchini, C. & Addessi, M.E. (2014). Cyber risk and insurance coverage: An actuarial multistate approach. *Review of Economics & Finance*, **4**(1), 57–69.

Biener, C., Eling, M. & Wirfs, J.H. (2015). Insurability of cyber risk: An empirical analysis. *Geneva Papers on Risk and Insurance – Issues and Practice*, **40**(1), 131–158. https://doi.org/10.1057/gpp.2014.19.

Böhme, R. (2005). Cyber-insurance revisited, in *4th Workshop on the Economics of Information Security (WEIS 2005)*. Cambridge, MA, available at https://web.archive.org/web/20150813192331/http://infosecon.net/workshop/bibliography.php (internet archive); (accessed 05 August 2015).

Böhme, R. & Kataria, G. (2006). Models and measures for correlation in cyber-insurance, in *5th Workshop on the Economics of Information Security (WEIS 2007 Working Paper)*. Cambridge, MA, available at https://web.archive.org/web/20150813192331/http://infosecon.net/workshop/bibliography.php (internet archive); (accessed 21 July 2015).

Böhme, R. & Schwartz, G. (2010). Modeling cyber-insurance: Towards a unifying framework, in *9th Workshop on the Economics of Information Security (WEIS 2010)*. USA, Harvard University, available at https://web.archive.org/web/20150324183146/http://infosecon.net/workshop/bibliography.php (internet archive); (accessed 22 March 2015).

Box, G.E.P. (1979). Robustness in the strategy of scientific model building, in *Robustness in Statistics* (ed. R.L. Launer & G.N. Wilkinson). USA, Academic Press, 201–236. https://doi.org/10.1016/B978-0-12-438150-6.50018-2

Bridwell, L. (2004). *ICSA Labs 9th Annual Computer Virus Prevalence Survey*, available at https://www.icsalabs.com/ (accessed 22 March 2015).

Burnham, K.P. & Anderson, D.R. (2002). *Model Selection and Multimodel Inference: a Practical Information-theoretic Approach*, 2nd ed. Springer.

Cashell, B., Jackson, W.D., Jickling, M. & Webel, B. (2004). *Economic impact of cyber-attacks*. Washington DC, available at https://fas.org/sgp/crs/misc/RL32331.PDF (accessed 30 January 2018).

Cerchiara, R.R. & Acri, F. (2016). Aggregate loss distribution and dependence: composite models, copula functions and fast Fourier transform for the Danish fire insurance data (working paper), available at https://web.archive.org/web/20190331232859/https://www.ivass.it/pubblicazioni-e-statistiche/pubblicazioni/att-sem-conv/2017/conf-131407/CerchiaraAcri_Paper.PDF (internet archive); (accessed 16 February 2018).

Chubb. (2017). Quarter century club, available at https://www.chubb.com/us-en/about-chubb/quarter-century-club.aspx (accessed 16 September 2017).

Cresenzi, C. & Alibrio, C. (2016). *ACE Filling – Cyber and Digitech (SERFF Tracking: ACEH-130778328)*. Washington State, available at https://www.insurance.wa.gov/insurers-regulated-entities (accessed 16 September 2017).

CRO Forum (2014), available at https://www.thecroforum.org/wp-content/uploads/2015/01/Cyber-Risk-Paper-version-24-1.PDF (accessed 20 July 2018).

Daigle, D. & Cresenzi, C. (2018). *Federal Filing – Forefront Portfolio 3.0 (SERFF tracking: ACEH-131628857)*. District of Columbia, available at http://serff.disb.dc.gov/.

Doyle, A.C. (1901). *A Study in Scarlet*. London, Ward, Lock & Co., Ltd., available at https://hdl.handle.net/2027/mdp.39015083421969 (permanent link); (accessed 19 August 2019).

Edwards, B., Hofmeyr, S. & Forrest, S. (2016). Hype and heavy tails: a closer look at data breaches. *Journal of Cybersecurity*, **2**(1), 3–14. https://doi.org/10.1093/cybsec/tyw003.

European Commission (2017). *Banking and Finance*, available at https://ec.europa.eu/ (accessed 08 February 2016).

Freifelder, L.R. (1979). Exponential utility theory ratemaking: an alternative ratemaking approach. *The Journal of Risk and Insurance.* **46**(3):515–530. https://doi.org/10.2307/252462.

Gordon, L.A. & Loeb, M.P. (2002). The economics of information security investment. *ACM Transactions on Information and System Security*, **5**(4), 438–457. https://doi.org/10.1145/581271.581274.

Hanover, (2015). Hanover Insurance 2015. Hanover religious institutions filling (SERFF tracking: HNVX-G129911675). District of Columbia, available at http://serff.disb.dc.gov/ (accessed 08 June 2015).

Herath, H. & Herath, T. (2011). Copula-based actuarial model for pricing cyber-insurance policies. *Insurance Markets and Companies: Analyses and Actuarial Computations*, **2**(1), 7–20, available at https://businessperspectives.org/author/hemantha-s-b-herath (accessed 21 March 2015).

Homer, D.L. & Rosengarten, R.A. (2011). Method for efficient simulation of the collective risk model. In Casualty Actuarial Society E-Forum (Spring). 1–41, available athttps://www.casact.org/pubs/forum/11spforum/ (accessed 23 November 2016).

Institute and Faculty of Actuaries (2019). Heritage online, available at https://actuaries.cirqahosting.com (accessed 25 February 2019).

Jacobs, J. (2014). *Analyzing Ponemon Cost of Data Breach*, available at https://datadrivensecurity.info/blog/posts/2014/Dec/ponemon/ (accessed 11 December 2016).

Jobs, S. (2010), available at https://news.stanford.edu/news/2005/june15/jobs-061505.html (accessed 23 June 2019).

Kaas, R., Goovaerts, M., Dhaene, J. & Denuit, M. (2008). *Modern Actuarial Risk Theory*. Berlin, Heidelberg, Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-70998-5.

Klugman, S.A., Panjer, H.H. & Willmot, G.E. (2004). *Loss Models: From Data to Decisions*, 2nd ed. Wiley Series in Probability and Statistics. Hoboken, NJ, Wiley Interscience.

Kullback, S. & Leibler, R.A. (1951). On information and sufficiency. *Annals of Mathematical Statistics*, **22**(1), 79–86. https://doi.org/10.1214/aoms/1177729694.

Laszka, A., Felegyhazi, M. & Buttyan, L. (2014). A survey of interdependent information security games. *ACM Computing Surveys*, **47**(2), 1–38. https://doi.org/10.1145/2635673.

Lee, Y. (1988). The mathematics of excess of loss coverages and retrospective rating – a graphical approach. *PCAS LXXV*, 49–77, available at http://casact.net/pubs/proceed/proceed88/88049.PDF.

Liu, S., Semenciw, R., Ugnat, A.M. & Mao, Y. (2001). Increasing thyroid cancer incidence in Canada, 1970–1996: Time trends and age-period-cohort effects. *British Journal of Cancer*, **85**(9), S139.

Liu, W., Tanaka, H. & Matsuura, K. (2007). Empirical-Analysis methodology for information-security investment and its application to reliable survey of Japanese firms. *IPSJ Digital Courier.* **3**:585–599. https://doi.org/10.2197/ipsjdc.3.585.

Lloyd's (2019). *Glossary and Acronyms*, available at https://www.lloyds.com/common/help/glossary (accessed 12 October 2017).

Mack, T. & Fackler, M. (2003). Exposure-rating in liability reinsurance. *Blätter der DGVFM*, **26**(2), 229–238. https://doi.org/10.1007/BF02808374.

McAfee & Center for Strategic and International Studies (2018). *Economic Impact of Cybercrime — No Slowing Down*, available at https://www.mcafee.com/enterprise/en-us/assets/reports/restricted/economic-impact-cybercrime.PDF (accessed 16 July 2018).

Meyers, G. & Heckman, P. (1984). The calculation of aggregate loss distributions from claim severity and claim count distributions, in *Casualty Actuarial Society (163 – May 1993).* **V. LXX**, 22–61, available at https://www.casact.org/pubs/proceed/proceed83/ (accessed 11 June 2016).

Miccolis, R.S. (1978). On the theory of increased limits excess of loss pricing, in *Casualty Actuarial Society (121 – May 1977).* **V. LXIV**, 27–59, available at https://www.casact.org/pubs/proceed/proceed77/.

Mildenhall, S.J. (2005). *Correlation and Aggregate Loss Distributions with An Emphasis on the Iman-Conover Method*, available at https://www.casact.org/pubs/forum/06wforum/06w107.PDF.

Ministry of Economy Trade Industry (2004). *Information Processing Survey Report*, available at https://web.archive.org/web/20071130225516/http://www.meti.go.jp/statistics/zyo/zyouhou/result-2/h15jyojitsu.html (internet archive); (accessed 28 February 2016).

Moher, D., Liberati, A., Tetzlaff, J. & Altman, D.G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLOS Medicine.* **6**(7):1–6. https://doi.org/10.1371/journal.pmed.1000097.

Mukhopadhyay, A., Chatterjee, S., Saha, D., Mahanti, A. & Sadhukhan, S.K. (2013). Cyber-risk decision models: To insure IT or not? *Decision Support Systems*, **56**(1), 11–26. https://doi.org/10.1016/j.dss.2013.04.004.

Müller, J., Polansky, D., Novak, P., Polivaev, D., Foltin, C. & Lavarde, E. (2004), available at http://freemind.sourceforge.net/wiki/index.php/Main_Page (accessed 10 February 2019).
Murphy, W. (2013). *Illinois National filing – PortfolioSelect (SERFF Tracking: AGNY-129243435)*. South Dakota, available at https://apps.sd.gov/cc57serffportal/ (accessed 04 September 2016).
Obama, B. (2009). *Remarks by the President on Securing Our Nation's Cyber Infrastructure*, available at https://web.archive.org/web/20170126105102/https://obamawhitehouse.archives.gov/the-press-office/remarks-president-securing-our-nations-cyber-infrastructure [internet archive]; [accessed 01 February 2017].
Ogut, H., Raghunathan, S. & Menon, N. (2005). *Information Security Risk Management through Self-Protection and Insurance*, available at https://www.researchgate.net/publication/229014862_Information_security_risk_management_through_self-protection_and_insurance (accessed 28 June 2018).
Operational Riskdata eXchange Association (2017). *ORX Loss Data*, available at https://www.orx.org.
Parker, D. & Farkas, C. (2011). Modeling estimated risk for cyber attacks: Merging public health and cyber security. *Journal of Information Assurance and Security (JIAS)*, **2**, 32–36, available at ~http://www.cse.sc.edu/~farkas/publications/j18.PDF (accessed 22 March 2015).
Ponemon Institute (2012a). *2011 Cost of Data Breach Study: Australia (sponsored by Symantec).*
Ponemon Institute (2012b). *2011 Cost of Data Breach Study: France (sponsored by Symantec).*
Ponemon Institute (2012c). *2011 Cost of Data Breach Study: Germany (sponsored by Symantec).*
Ponemon Institute (2012d). *2011 Cost of Data Breach Study: India (sponsored by Symantec).*
Ponemon Institute (2012e). *2011 Cost of Data Breach Study: Italy (sponsored by Symantec).*
Ponemon Institute (2012f). *2011 Cost of Data Breach Study: Japan (sponsored by Symantec).*
Ponemon Institute (2012g). *2011 Cost of Data Breach Study: United Kingdom (sponsored by Symantec).*
Ponemon Institute (2012h). *2011 Cost of Data Breach Study: United States (sponsored by Symantec).*
Ponemon Institute (2012i). *2011 Cost of Data Breach Study: Global Analysis (sponsored by Symantec)*, available at http://www.ponemon.org/library/.
Ponemon Institute (2013a). *2013 Cost of Data Breach Study: Australia (sponsored by Symantec).*
Ponemon Institute (2013b). *2013 Cost of Data Breach Study: Brazil (sponsored by Symantec).*
Ponemon Institute (2013c). *2013 Cost of Data Breach Study: France (sponsored by Symantec).*
Ponemon Institute (2013d). *2013 Cost of Data Breach Study: Germany (sponsored by Symantec).*
Ponemon Institute (2013e). *2013 Cost of Data Breach Study: India (sponsored by Symantec).*
Ponemon Institute (2013f). *2013 Cost of Data Breach Study: Italy (sponsored by Symantec).*
Ponemon Institute (2013g). *2013 Cost of Data Breach Study: Japan (sponsored by Symantec).*
Ponemon Institute (2013h). *2013 Cost of Data Breach Study: United Kingdom (sponsored by Symantec).*
Ponemon Institute (2013i). *2013 Cost of Data Breach Study: United States (sponsored by Symantec).*
Ponemon Institute (2013j). *2013 Cost of Data Breach Study: Global Analysis (sponsored by Symantec).*
Ponemon Institute (2014a). *2014 Cost of Data Breach Study: Arabian region (sponsored by IBM).*
Ponemon Institute (2014b). *2014 Cost of Data Breach Study: Australia (sponsored by IBM).*
Ponemon Institute (2014c). *2014 Cost of Data Breach Study: Brazil (sponsored by IBM).*
Ponemon Institute (2014d). *2014 Cost of Data Breach Study: France (sponsored by IBM).*
Ponemon Institute (2014e). *2014 Cost of Data Breach Study: Germany (sponsored by IBM).*
Ponemon Institute (2014g). *2014 Cost of Data Breach Study: Italy (sponsored by IBM).*
Ponemon Institute (2014h). *2014 Cost of Data Breach Study: Japan (sponsored by IBM).*
Ponemon Institute (2014i). *2014 Cost of Data Breach Study: United Kingdom (sponsored by IBM).*
Ponemon Institute (2014j). *2014 Cost of Data Breach Study: United States (sponsored by IBM).*
Ponemon Institute (2014k). *2014 Cost of Data Breach Study: India (sponsored by IBM).*
Ponemon Institute (2014f). *2014 Cost of Data Breach Study: Global Analysis (sponsored by IBM).*
Ponemon Institute (2015). *2015 Cost of Data Breach Study: Global Analysis (sponsored by IBM).*
Ponemon Institute (2019). *Measuring Trust in Privacy and Security.*
Pouget, F., Dacier, M. & Pham, V.H. (2005). Leurre.com: On the advantages of deploying a large scale distributed honeypot platform, in *E-Crime and Computer Conference (ECCE)*. Monaco, available at http://www.eurecom.fr/publication/1558 (accessed 28 February 2016).
Ralucavernic, S.T. (2009). Some composite exponential-pareto models for actuarial prediction. *Romanian Journal of Economic Forecasting*, **4**(12), 82–100, available at http://www.ipe.ro/rjef/rjef4_09/rjef4_09_5.PDF (accessed 16 February 2018).
Reshetar, G. (2008). Dependence of operational losses and the capital at risk. *SSRN Electronic Journal* (January, 7). https://doi.org/10.2139/ssrn.1081256.
Rohatgi, A. (2013), available at https://automeris.io/WebPlotDigitizer (accessed 01 May 2015).
Sagan, C. (1983). *Cosmos*. 1st pbk, ed. (ed. A. Freedgood). New York & Canada, Random House, available at http://www.worldcat.org/oclc/12736033 (permalink); (accessed 18 August 2019).

SAS (2015), available at https://web.archive.org/web/20150807100828/https://www.sas.com/resources/product-brief/sas-oprisk-globaldata-brief.PDF (internet archive) (accessed 11 July 2015).

Shevchenko, P.V. (2010). Calculation of aggregate loss distributions. *Journal of Operational Risk*, **5**(2), 3–40. https://doi.org/10.21314/JOP.2010.077.

Soo Hoo, K.J.K.J. (2000). How much is enough? A risk management approach to computer security (working paper). California, Stanford University: Citeseer, available at https://doc.uments.com/h-how-much-is-enough-a-risk-management-approach-to-computer-security.pdf (accessed 05 September 2015).

SysAdmin Audit Admin and Security Technology Institute (2019). *Internet Storm Center – DShield*, available at https://web.archive.org/web/20151112153715/https://isc.sans.edu/ (internet archive); (accessed 12 November 2015).

UK Government and Industry (2015). Cyber Risk Report (March).

University of Cape Town. (2019). UCT home, available at https://www.uct.ac.za/ (accessed 06 March 2016).

van der Vaart, A.W. (1998). *Asymptotic Statistics*. Cambridge, Cambridge University Press. https://doi.org/10.1017/CBO9780511802256.

van Mieghem, P., Omic, J. & Kooij, R. (2009). Virus spread in networks. *IEEE/ACM Transactions on Networking*, **17**(1), 1–14. https://doi.org/10.1109/TNET.2008.925623.

Vose (2019). *Risk Analysis Software for Excel*, available at https://www.vosesoftware.com/products/modelrisk/.

Wang, S. (1995). Insurance pricing and increased limits ratemaking by proportional hazards transforms. *Insurance: Mathematics and Economics*, **17**(1), 43–54. https://doi.org/10.1016/0167-6687(95)00010-P.

Wang, S.S. (1998). *Aggregation of Correlated Risk Portfolios: Models & Algorithms (research contract with COTOR)*, available at https://www.casact.org/research/cotor/ (accessed 16 May 2015).

Wang, S.S. (1999a). Implementation of proportional hazard transforms in ratemaking, in *Casualty Actuarial Society (163 – Nov 1998, congress catalog – HG9956.C3)*. V. **LXXXV** (ed. E.C. Connell, D.A. Crifo, W.F. Dove, D.R. Edlefson, E.M. Gardiner, J.F. Golz, G.R. Josephson, K.E. Kufera, M. Lewis, R.A. Moody, D. Schwab, & T.A. Turnacioglu). United Book Press, 940–979, available at https://www.casact.org/pubs/proceed/proceed98/ (accessed 28 December 2017).

Wang, S.S. (1999b). *Aggregation of correlated risk portfolios: Models and algorithms*, in *Casualty Actuarial Society (163 – Nov 1998, congress catalog – HG9956.C3)*. **V. LXXXV** (E.C. Connell, D.A. Crifo, W.F. Dove, D.R. Edlefson, E.M. Gardiner, J.F. Golz, G.R. Josephson, K.E. Kufera, M. Lewis, R.A. Moody, D. Schwab, & T.A. Turnacioglu) Toronto, 848–939, available at https://www.casact.org/pubs/proceed/proceed98/ (accessed 13 October 2015).

Werner, G. & Modlin, C. (2010). *Basic Ratemaking*. 4th ed, available at https://web.archive.org/web/20150616074102/https://www.casact.org/library/studynotes/Werner_Modlin_Ratemaking.PDF (internet archive); (accessed 05 July 2015).

Workshop on the Economics of Information Security. (2019). WEIS, available at: https://weis2019.econinfosec.org/organization/previous-workshops/ (accessed 04 July 2019).

World Bank (2019). *Open Data*, available at https://web.archive.org/web/20151231235345/https://data.worldbank.org/ (internet archive); (accessed 27 December 2015).

WorldCat. (2019). World's largest library catalog, available at http://www.worldcat.org/ (accessed 02 March 2019).

Yannacopoulos, A.N., Lambrinoudakis, C., Gritzalis, S., Xanthopoulos, S.Z. & Katsikas, S.N. (2008). Modeling privacy insurance contracts and their utilization in risk management for ICT firms, in *Computer Security – ESORICS 2008, 13th European Symposium on Research in Computer Security* (ed. S. Jajodia & J. López). Lecture Notes in Computer Science, vol. **5283**. Málaga, Spain, Springer Berlin Heidelberg, 207–222. https://doi.org/10.1007/978-3-540-88313-5_14.

Appendix A. Literature review

A.1. Search Strategy

The search strategy used to identify studies in the model review (Section 2) is illustrated in Figure A.1. This incorporates various filters (e.g. language, content, etc.) and utilises the University of Cape Town [UCT] (2019) online search engine. Titles and keywords are searched using strings that are made up of one word from each of the following groups:

- Group 1: "cyber," "information," and "interdependent"
- Group 2: "risk management," "insurance" (and derivatives, such as insurability), and "*security*"

The UCT (2019) online search, used to generate these results, accesses databases such as WorldCat (2019), which is self-proclaimed as "world's largest network of libraries." Incorporated in it are supplementary sources to compliment the search, such as Workshop on the Economics of Information Security [WEIS] (2019) – (archives of papers on information security and privacy), and Association for Computing Machinery [ACM] (2019) – (an international society for learned computing). The library catalogue of the Institute and Faculty of Actuaries (2019) was also considered.

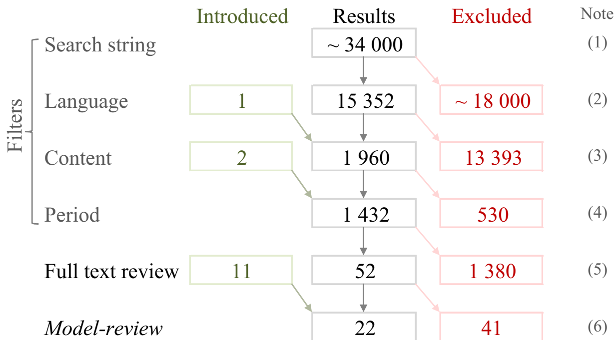


Figure A.1. Identification of studies. Notes: (1) Search string: "*ti:((cyber | information | interdependent) + (risk management | insur* | security)) kw: (model | empirical)*" – which applies to titles (i.e. "*ti*") and keywords (i.e. "*kw*"), through the UCT (n.d.) search engine; (2) English-only; identified Barracchini & Addessi (2014) from a similar (but excluded) Italian manuscript, ; (3) Full-text, peer-reviewed (re-included Soo Hoo (2000), Liu et al. (2007) – not peer-reviewed); (4) Period: 2000 – mid 2016; (5) 52 studies identified for full-text review by scanning titles, then abstracts, and introduced 11 new studies from online searches; references; and archived libraries (e.g. WEIS (2019); (6) eliminated 41 studies based on full-text review, leaving 22 for the *model review*. Motivated by *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA) – (Moher *et al.*, 2009), and Biener et al. (2015) search strategy for cyber-related losses.

The 22 studies that are identified in Figure A.1 constitute studies in the taxonomy (Figure A.2) – this excludes the study Edwards et al. (2016), which fell outside the review period (2000–mid-2006).

https://doi.org/10.1017/S1357321724000205 Published online by Cambridge University Press

A.2. Taxonomy

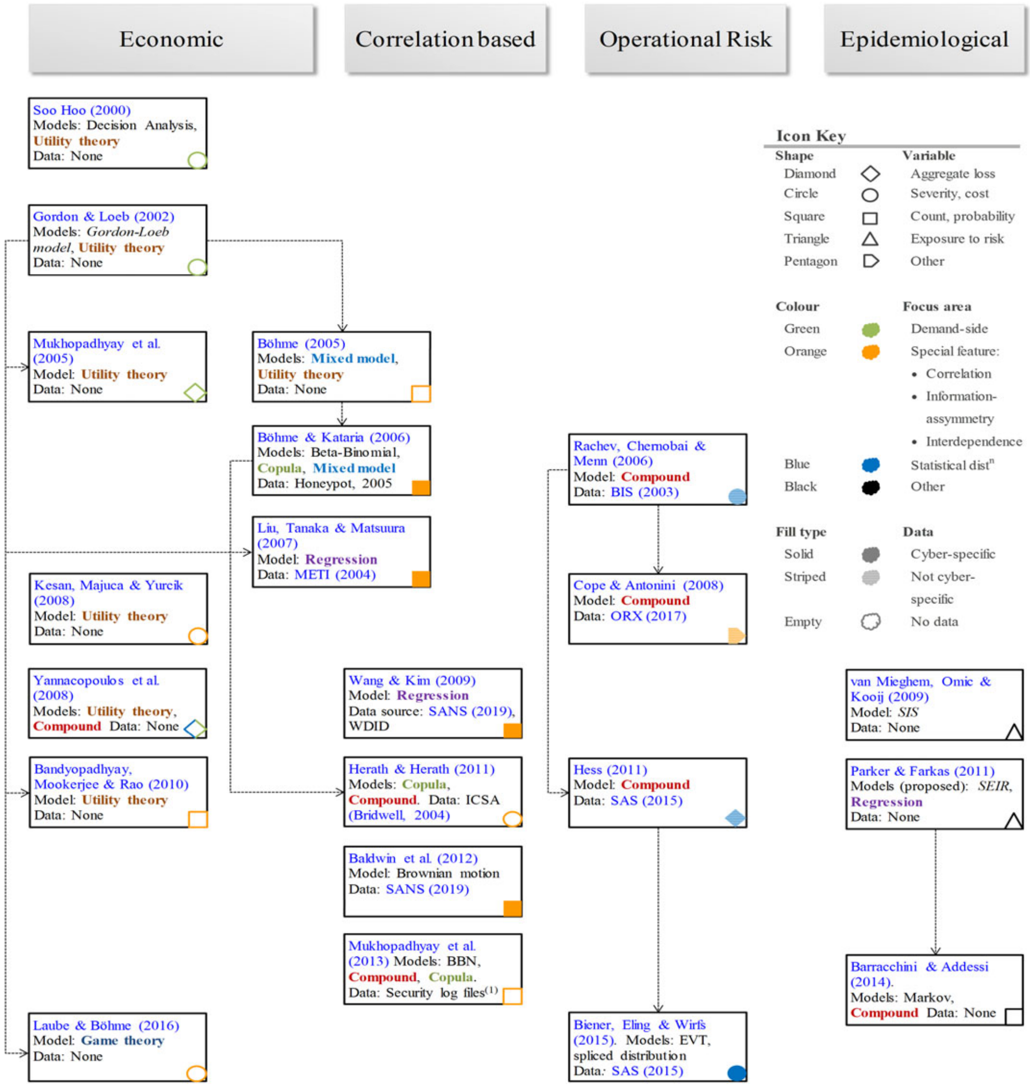


Figure A.2. Overview of cyber-risk models. Text colour: common model types. Abbreviations: Bank for International Settlements [BIS] (2013); Honeypot – Pouget et al. (2005); ICSA: International Computer Security Association – Bridwell (2004); Ministry of Economy Trade Industry [METI] (2004); Operational Riskdata eXchange Association [ORX] (2017); SysAdmin, Audit, Admin and Security [SANS] (2019); World Development Indicators Database (WDID): World Bank (2019). SEIR: Susceptible-Exposed-Infected-Recovered, SIS: Susceptible-Infected-Susceptible. Note (1): undisclosed source.

Appendix B. Results

B.1. Identifying Severity CDFs (Algorithm 4.1)

Table B.1. Large-loss CDFs and scores. Final selections (percentiles: coloured font, A–E; CDFs: boxed) correspond to maximum overall scores (boxed). Weibull (shifted; asterisked: light-tailed), Burr (type III: Dagum), and Pearson: 3, 4, and 6 parameter CDFs respectively. Coloured bars: models – quantile divided by maximum (empirical severity); scores – relative magnitude. Criteria for ✓ (failing which, ✗): percentile deemed to be acceptable (in terms of ME plots); spliced CDF yields consistent ILFs over a given set of limits ($10k, $100m). Underlying costs: Ponemon Institute (2012a–2012i, 2013a–2013j, 2014a–2014k, 2015, inflated to 2016

Threshold percentile	Large-loss models (AIC^c selections; final selections are boxed)					Goodness-of-fit score (KS critical: 5%; minimum *KS-ratio by* class, divided by *KS-ratio*)					Overall scores and consistency check (are limit factors consistent over all limits considered?)				
	Class A	Class B	Class C	Class D	Class E	Class A	Class B	Class C	Class D	Class E	Class A	Class B	Class C	Class D	Class E
0.25%	Weibull	Weibull	Pearson	Weibull	Weibull*	0.63	0.36	0.32	0.71	0.39	✓ 0.31	✓ 0.19	✓ 0.16	✓ 0.35	✓ 0.20
65.5%	Weibull	Weibull	Weibull	Weibull	Weibull	0.73	0.62	0.41	0.79	0.56	✓ 0.71	✓ 0.63	✓ 0.47	✓ 0.76	✓ 0.58
66.5%	Weibull	Weibull	Weibull	Weibull	Weibull	0.73	0.48	0.56	0.75	0.62	✓ 0.72	✓ 0.53	✓ 0.59	✓ 0.73	✓ 0.63
67.5%	Weibull	Weibull	Weibull	Weibull	Weibull	0.80	0.51	0.49	0.76	0.53	✓ 0.77	✓ 0.55	✓ 0.53	✓ 0.74	✓ 0.55
68.5%	Weibull	Weibull	Weibull	Weibull	Weibull*	0.73	0.39	0.52	0.77	0.77	✓ 0.72	✓ 0.46	✓ 0.56	✓ 0.75	✓ 0.76
69.5%	Weibull	Weibull	Weibull	Weibull	Weibull*	0.60	0.49	0.50	0.79	0.73	✓ 0.62	✓ 0.54	✓ 0.54	✓ 0.77	✓ 0.72
70.5%	Weibull	Burr	Weibull	Weibull	Weibull*	0.47	0.25	0.54	0.67	0.74	✓ 0.53	✓ 0.36	✓ 0.58	✓ 0.68	✓ 0.73
71.5%	Weibull	Burr	Weibull	Weibull	Weibull	0.50	0.28	0.53	0.65	0.53	✓ 0.55	✓ 0.37	✓ 0.57	✓ 0.66	✓ 0.56
72.5%	Weibull	Burr	Weibull	Weibull	Weibull*	0.43	0.36	0.40	0.69	0.59	✓ 0.49	✓ 0.44	✓ 0.47	✓ 0.70	✓ 0.62
73.5%	Weibull	Burr	Weibull	Weibull	Weibull	0.43	0.40	0.37	0.59	0.50	✓ 0.49	✓ 0.47	✓ 0.45	✓ 0.62	✓ 0.55
74.5%	Weibull	Burr	Weibull	Weibull	Weibull	0.43	0.36	0.44	0.54	0.56	✓ 0.49	✓ 0.44	✓ 0.50	✓ 0.58	✓ 0.60
75.5%	Weibull	Burr	Weibull	Weibull	Weibull	0.47	0.52	0.38	0.49	0.57	✓ 0.53	✓ 0.57	✓ 0.45	✓ 0.55	✓ 0.61
76.5%	Weibull	Burr	Weibull	Weibull	Weibull	0.50	0.27	0.36	0.59	0.48	✓ 0.55	✓ 0.37	✓ 0.43	✓ 0.62	✓ 0.54
B: 77.50%	Weibull	Burr	Weibull	Weibull	Weibull	0.68	**1.00**	0.35	0.53	0.51	✓ 0.70	✓ **0.96**	✓ 0.43	✓ 0.58	✓ 0.57
78.5%	Weibull	Burr	Weibull	Weibull	Weibull*	0.64	0.80	0.29	0.55	0.65	✓ 0.66	✓ 0.80	✓ 0.37	✓ 0.59	✓ 0.68
79.5%	Weibull	Burr	Weibull	Weibull	Weibull*	0.63	0.31	0.29	0.49	0.75	✓ 0.66	✓ 0.39	✓ 0.37	✓ 0.54	✓ 0.76
C: 81.00%	Weibull	Burr	Burr	Weibull	Weibull*	0.56	0.34	**1.00**	0.62	0.73	✓ 0.60	✓ 0.42	✓ **0.97**	✓ 0.65	✓ 0.75
81.5%	Weibull	Burr	Burr	Weibull	Weibull*	0.53	0.51	0.28	0.59	0.85	✓ 0.58	✓ 0.56	✓ 0.36	✓ 0.63	✓ 0.84
82.5%	Weibull	Burr	Weibull	Weibull	Weibull*	0.52	0.51	0.33	0.54	0.84	✓ 0.56	✓ 0.56	✓ 0.41	✓ 0.58	✓ 0.83
E: 83.91%	Weibull	Weibull	Weibull	Weibull	Weibull*	0.49	0.28	0.42	0.53	**1.00**	✓ 0.54	✓ 0.36	✓ 0.48	✓ 0.57	✓ **0.98**
84.5%	Weibull	Burr	Weibull	Weibull	Weibull	0.57	0.36	0.41	0.54	0.66	✓ 0.61	✓ 0.42	✓ 0.47	✓ 0.58	✓ 0.71
85.5%	Weibull	Burr	Weibull	Weibull	Weibull	0.69	0.40	0.35	0.65	0.58	✓ 0.71	✓ 0.45	✓ 0.41	✓ 0.68	✗
86.5%	Weibull	Weibull	Weibull	Weibull	Weibull	0.82	0.48	0.41	0.68	0.58	✓ 0.82	✓ 0.52	✓ 0.47	✓ 0.70	✗
A: 87.25%	Weibull	Weibull	Weibull	Weibull	Weibull	**1.00**	0.46	0.39	0.75	0.66	✓ **0.99**	✓ 0.50	✓ 0.44	✓ 0.76	✗
88.5%	Weibull	Weibull	Weibull	Weibull	Weibull	0.80	0.48	0.39	0.63	0.56	✓ 0.81	✓ 0.52	✓ 0.44	✓ 0.66	✗
89.5%	Weibull	Weibull	Weibull	Weibull	Weibull	0.91	0.34	0.48	0.49	0.52	✓ 0.91	✓ 0.39	✓ 0.52	✓ 0.53	✗
90.5%	Weibull	Burr	Weibull	Weibull	Weibull	0.75	0.63	0.41	0.51	0.58	✓ 0.76	✓ 0.65	✓ 0.45	✓ 0.54	✗
91.5%	Weibull	Weibull	Weibull	Weibull*	Weibull	0.88	0.28	0.49	0.65	0.49	✓ 0.88	✓ 0.33	✓ 0.52	✓ 0.67	✗
D: 92.12%	Weibull	Burr	Weibull	Weibull*	Weibull	0.67	0.30	0.31	**1.00**	0.49	✓ 0.68	✓ 0.35	✓ 0.35	✓ **0.99**	✗
93.5%	Weibull	Burr	Burr	Weibull	Burr	0.52	0.26	0.25	0.67	0.28	✓ 0.55	✗	✗	✗	✗
94.5%	Weibull	Burr	Weibull	Weibull	Burr	0.44	0.30	0.48	0.77	0.26	✗	✗	✗	✗	✗
95.5%	Weibull	Burr	Burr	Weibull	Weibull	0.41	0.28	0.27	0.64	0.30	✗	✗	✗	✗	✗
96.5%	Fatigue	Burr	Burr	Fatigue	Burr	0.32	0.30	0.20	0.21	0.29	✗	✗	✗	✗	✗
97.5%	Fatigue	Fatigue	Fatigue	Fatigue	Weibull	0.24	0.19	0.26	0.30	0.61	✗	✗	✗	✗	✗
98.5%	Fatigue	Fatigue	Fatigue	Fatigue	Fatigue	0.39	0.20	0.25	0.67	0.15	✗	✗	✗	✗	✗
99.5%	Fatigue	Fatigue	Fatigue	Fatigue	Fatigue	0.47	0.31	0.35	0.71	0.34	✗	✗	✗	✗	✗

B.2 Mean Excess Plots (A-D)

Markers in Figure B.1 indicate the apparent onset of volatility, or other such irregularity due to having too few data points. These correspond to maximum permissible thresholds for use in Algorithm 4.1 (Appendix B.1).

As can be seen, MEs for classes B and C initially decrease before assuming upward concavity (possibly indicating a Burr type CDF), and ultimately, continue to increase beyond the indicated percentiles (i.e. 94%, 93% respectively). This could also be indicative of a heavy-tailed Weibull, possibly a Pareto. In contrast, MEs for class A and D reduce after the threshold of 93% (sharply so, in class D), which undermines a CDF such as the Pareto, and may even imply a short-tailed CDF for D.

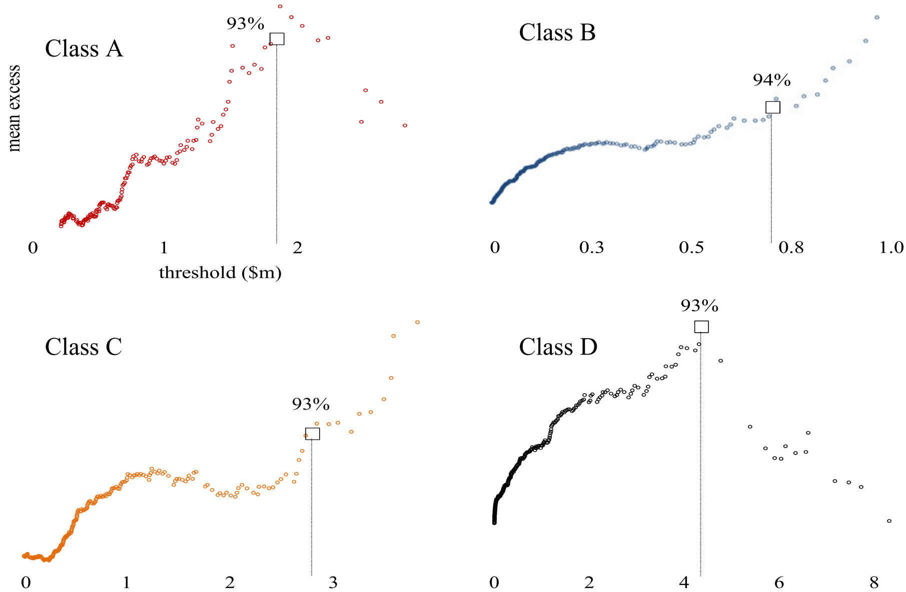


Figure B.1. Empirical ME plots. Axes: x (threshold, $m), y (mean excess, values omitted as they are unnecessary for this exercise). Data: costs sourced from Ponemon Institute (2012a–2012i, 2013a–2013j, 2014a–2014k, 2015, inflated to 2016. Square markers (i.e. 94th, 96th, 93rd, and 92nd percentiles: A–D respectively) indicate the onset volatile or irregular trends (used as maximum percentiles for).

B.3. Densities, Limited Moments

For beta and gamma families (Table B.2: B.6–B.7) gamma (Γ) and beta (B) functions, and respective lower incomplete variations are defined as follows:

$$\begin{aligned} \Gamma(a) &= \int_0^\infty u^{a-1}\exp(-u)du, & \Gamma(a;b) &= \int_0^b u^{a-1}\exp(-u)du\,f \\ B(a,b) &= \int_0^1 u^{a-1}(1-u)^{b-1}du = \tfrac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}, & B(a,b;c) &= B(a,b)\int_0^c u^{a-1}(1-u)^{b-1}du \end{aligned} \tag{B.1}$$

where $a, b, c > 0; c < 1$ (Klugman et al., 2004: 102, 627–629), noting that in this case, the incomplete gamma, $\Gamma(a, b)$, is not "standardised" with divisor $\Gamma(a)$. In this table, limited moments for continuous distributions do not incorporate a shift (i.e. location parameter). For this, an adjustment can be applied as described in the following. Suppose random variable $Y = X + \phi$ has a shifted *CDF*, based on (non-negative) random variable X with location (i.e. "shift") parameter $\phi > 0$ (i.e. $Y \geq \phi$). Then limited moments for Y, when limit $l > \phi$ applies, can be determined analytically using $\mathrm{E}\big(Y^{(l)k}\big) = \mathrm{E}\big((X^{(l-\phi)} + \phi)^k\big)$, assuming respective limited moments for X exist. This follows from the fact that $\min(X + \phi, l) = \min(X, l - \phi) + \phi$. For $\phi > l \geq 0$, $\mathrm{E}Y^{(l)k} = l^k$ by definition.

https://doi.org/10.1017/S1357321724000205 Published online by Cambridge University Press

https://doi.org/10.1017/S1357321724000205 Published online by Cambridge University Press

Table B.2. Discrete and continuous distributions. Limit $l > 0$ applies to random variable X for limited moments B.5–B.7 (Klugman et al., 2004, sec. A.2.1.1, A3.1.1). *Dagum is represented as Burr(b,c,d) – (i.e. $a = 1$) throughout the present research to align with Vose (2019) parameterisation of Burr (ordinarily d = 1 for Burr). Location parameter, for a shifted CDF, is included after other applicable parameters a-d (limited moments, B.5–B.7, based on need to be adjusted accordingly)

	Model or family	Notation, parameters	Density, distribution, support	Discrete: PGF, $P[t]$; mean, μ; variance, σ^2 Continuous: limited moments – $E(X^{(l)k})$; $l > 0$, $k \in \mathbf{Z}^+$	
Discrete	Binomial	$Bin(n,p)$ $n \in \mathbf{Z}^+, p \in (0,1)$	$f(x) = C_{(n,x)} p^x (1-p)^{n-x}$, $x = 0, 1, \ldots, n$	$\mathrm{P}[t] = (1 - p + pt)^n$ $\mu = np$; $\sigma^2 = np(1-p)$	B.2
	Poisson	$\mathrm{Pois}(\lambda)$ $\lambda > 0$	$f(x) = \lambda^x \exp(-\lambda)(x!)^{-1}$, $x = 0, 1, \ldots, n$	$\mathrm{P}[t] = \exp(\lambda(t-1))$; $\mu = \sigma^2 = \lambda$	B.3
	Negative binomial	$NB(a,b)$ $a, b > 0$	$f(x) = \frac{\Gamma(a+x) b^x}{\Gamma(a) x! (1+b)^{a+x}}, x = 0, 1, \ldots$	$P[t] = (1 - b(t-1))^{-a}$ $\mu = ab; \sigma^2 = ab(1+b)$	B.4
Continuous	Lognormal	$LN(\mu, \sigma)$, $\mu \in R, \sigma > 0$	$f(x) = \frac{\exp\left(-\frac{1}{2}s^2\right)}{x\sigma(2\pi)^{\frac{1}{2}}}; s = \sigma^{-1}(\ln(x) - \mu)$ $F(x) = 1 - S(x) = \Phi(s), x > 0$	$\mathrm{E}\left(X^{(l)k}\right) = \exp\left(k\mu + \frac{1}{2}k^2\sigma^2\right)\Phi(s - k\sigma) + l^k S(l)$	B.5
	Transformed beta (four parameter excluding shift)	$a, b, c, d > 0$ • Dagum: $a = 1$, Burr(b,c,d)* • GPD (a,b,d): $c = 1$ • Pareto (a,b): $c = d = 1$ • Log-logistic (b,c): $a = d = 1$	$f(x) = \frac{\Gamma(a+d) c x^{cd-1} b^{-cd}}{\Gamma(a)\Gamma(d)\left(1+\left(xb^{-1}\right)^c\right)^{a+d}}$ $F(x) = 1 - S(x) = \mathrm{B}(d, a; p(x))$, $p(x) = \left(1 + \left(x^{-1}b\right)^c\right)^{-1}; x > 0$	$\mathrm{E}\left(X^{(l)k}\right) = \frac{b^k \Gamma(m)\Gamma(q) B(m, q; p(l))}{\Gamma(d)} + l^k S(l)$, $k > -cd$; $m = d + kc^{-1}$; $q = a - kc^{-1}$	B.6
	Transformed gamma (three parameters, excluding shift)	$a, b, c > 0$ • Gamma: $c = 1$, G(a,b) • Weibull: $a = 1$, Weib(b,c) • Exponential: $a = c = 1$, Exp(b)	$f(x) = \frac{c x^{ac-1}}{b^{ac}\Gamma(a)} \exp(-x^c b^{-c})$ $F(x) = 1 - S(x) = \Gamma(a, x^c b^{-c}); x > 0$	$\mathrm{E}\left(X^{(l)k}\right) = \frac{b^k \Gamma\left(a + kc^{-1}; x^c b^{-c}\right)}{\Gamma(a)} + l^k S(l)$, $k > -ac$	B.7

B.4. Risk-Adjusted ILFs

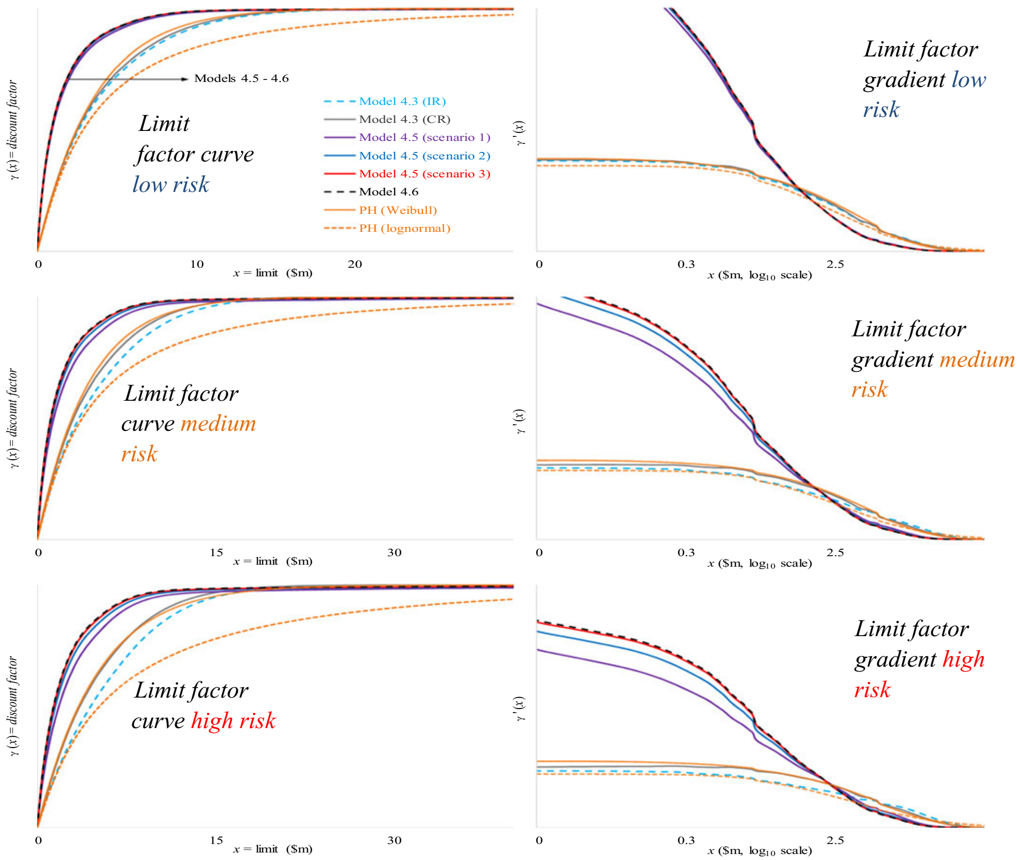


Figure B.2. Limit factor and gradient curves. Base limit: $100m. Risk margin (Model 4.3 (CR) in low (1–2), medium (3–4), and high environments achieve a risk margin of 5% at $10m, $100k, and $10k limits, respectively (based upon variance principle, which also applies to Models 4.5–4.6. PH transform applies to a compound Poisson-Weibull and lognormal CDF, fit to Ponemon Institute (2012a–2012i, 2013a–2013j, 2014a–2014k, 2015 costs, inflated to end of 2016). Loss count ~ Poisson(10) (all CR models), and 10 (deterministic for Model 4.3 IR) Model 4.3 (IR).

Key observations from the risk-adjusted ILFs and associated gradients in Figure B.2 include:

- PH (Weibull) limit factors are closely aligned to (variance-adjusted) Model 4.3 (CR), as is the case for Models 4.5 (scenario 3) and 4.6; PH (lognormal) and Model 4.3 ILFs crossover at a limit between the $15m-$20m (due to the underlying CDFs)
- Variance principle risk-adjusted limit factors, in this case, are generally consistent (i.e. positive and decreasing gradients, which is always the case for PH), although a subtle initial increase can be seen for Model 4.3 (i.e. closing the gap between CR and PH Weibull in medium–high risk, Figure B.2: 4, 6)
- Increasing the risk parameter leads to a greater risk adjustment at higher limits than lower limits for a given model (i.e. discount factor reduces, whilst ILFs increase at limits greater than $1m); a similar effect can be achieved through the correlation parameter in Models 4.5–4.6 (although this is partially offset by equalising risk margins at the $2.5m limit)

https://doi.org/10.1017/S1357321724000205 Published online by Cambridge University Press

Appendix C. Validation

Numerous checks have been performed by comparing *ALD*s and their moments against alternative derivations. One such example is depicted in Figure C.1 and Table C.1, which illustrates the close correspondence between Model 4.3 (CR) and an MC simulation based upon an algorithm developed by Homer and Rosengarten (2011).

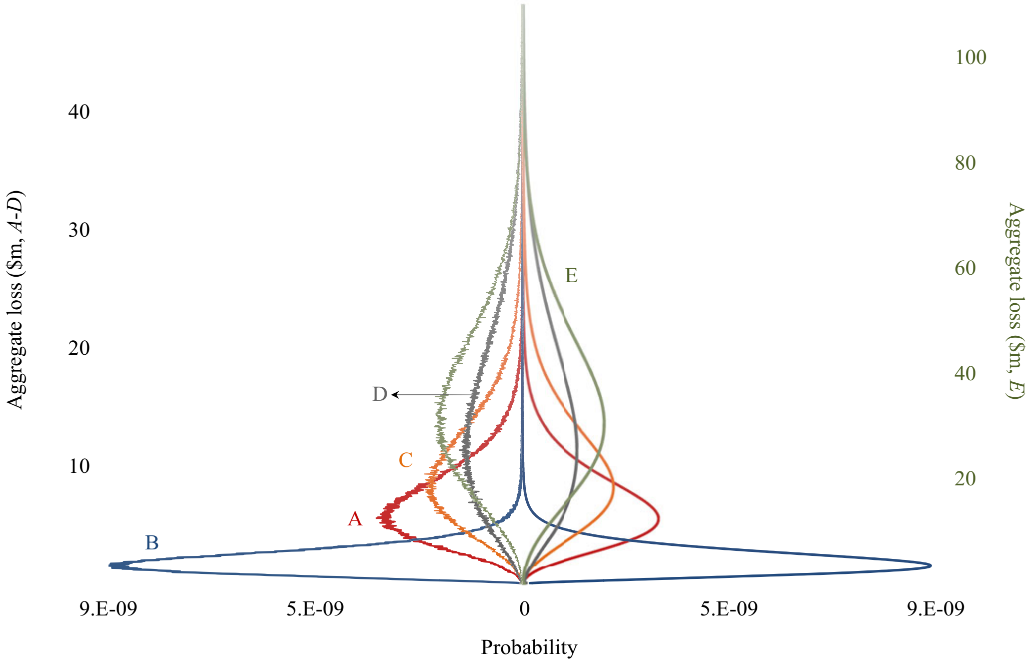


Figure C.1. ALDs: Monte Carlo versus FFT (Model 4.3, CR) – (*1*) *Left (of probability* $=0$*)*: *MC* simulation with 500k iterations; (*2*) *Right*: Model 4.3 (*CR*) with *FFT* (*truncation*, *span*) – *A–D*: ($96.2m, $23.5k), *E*: ($287.1, $70.1k). Limits: *A–D* ($20m), *E* ($80m); *Poisson* loss count with mean 10. Vertical axes – left (*A–D*); right (*E*). Underlying data: Ponemon Institute (2012a–2012i, 2013a–2013j, 2014a–2014k, 2015), costs inflated to year 2016.

Table C.1. Moments: Monte Carlo versus FFT. *MC* simulation with 500k iterations; Model 4.3 (*CR*) with *FFT* (*truncation*, *span*) – *A–D*: ($96.2*m*, $23.5k), *E*: ($287.1, $70.1*k*). Means: $m. Limits: *A–D* ($20*m*), *E* ($80*m*); *Poisson* loss count with mean 10. Underlying data based on Ponemon Institute (2012a–2012i, 2013a–2013j, 2014a–2014k, 2015, with costs inflated to end of 2016 year

Class	Method	Mean	Min	Std dev	Kurt	Skew
A	Monte Carlo	7.080	0	3.485	4.683	0.998
	Model 4.3	7.050	0	3.478	4.666	0.996
B	Monte Carlo	2.137	0	1.397	27.695	2.985
	Model 4.3	2.119	0	1.388	28.599	3.029
C	Monte Carlo	10.362	0	5.238	5.204	1.095
	Model 4.3	10.346	0	5.242	5.250	1.103
D	Monte Carlo	14.895	0	7.961	3.852	0.831
	Model 4.3	14.989	0	8.080	3.875	0.840
E	Monte Carlo	35.806	0	15.710	3.521	0.643
	Model 4.3	35.836	0	15.834	3.556	0.657

https://doi.org/10.1017/S1357321724000205 Published online by Cambridge University Press

Abbreviations

Akaike Information Criteria, AIC 14
Anderson Darling, AD 22
characteristic function, cf 7
Extreme Value Theory, EVT 3
Fast Fourier Transform, FFT 7
Increased Limit Factors, ILF 2
Internet Service Providers, ISP 2
Kullback-Leibler, KB 16
Limited Aggregate Severity, LAS 8
Maximum Likelihood Estimation, MLE 14
moment generating function, mgf 7
Multivariate Negative Binomial, MNB 12
Operational Risk, OR 3
probability generating function, pgf 7
Proportional-Hazard PH, 10

Cite this article: Bardopoulos J. (2025). Cyber-insurance pricing models. *British Actuarial Journal*. https://doi.org/10.1017/S1357321724000205