ONE-YEAR PREMIUM RISK AND EMERGENCE PATTERN OF ULTIMATE LOSS BASED ON CONDITIONAL DISTRIBUTION

BY

ŁUKASZ DELONG AND MARCIN SZATKOWSKI

ABSTRACT

We study the relation between one-year premium risk and ultimate premium risk. In practice, the one-year risk is sometimes related to the ultimate risk by using a so-called emergence pattern formula which postulates a linear relation between both risks. We define the true emergence pattern of the ultimate loss for the one-year premium risk based on a conditional distribution of the ultimate loss derived from a multivariate distribution of the claims development process. We investigate three models commonly used in claims reserving and prove that the true emergence pattern formulas are different from the linear emergence pattern formula used in practice. We show that the one-year risk, when measured by VaR, can be under and overestimated if the linear emergence pattern formula is applied. We present two modifications of the linear emergence pattern formula. These modifications allow us to go beyond the claims development models investigated in the first part and work with an arbitrary distribution of the ultimate loss.

KEYWORDS

One-year risk, ultimate risk, premium risk, emergence pattern, Solvency II, claims development process.

1. INTRODUCTION

Insurance companies are exposed to premium and reserve risk and they quantify the so-called one-year and ultimate risk. By one-year premium risk, we understand the risk that the premiums earned in a given year are not sufficient to cover the payments made in the year and the reserve set at the end of the year for the losses incurred in this accident year (the one-year loss). By ultimate premium risk we understand the risk that the premiums earned in a given year are...
not sufficient to cover all the payments made in infinite time horizon for the losses incurred in this accident year (the ultimate loss). One-year and ultimate reserve risks are defined analogously, but the reserve risk is related to the adequacy of the current volumes of the claims reserves (the reserve risk concerns payments for the losses incurred in the past accident years). One-year risk is investigated for Solvency II risk capital, whereas ultimate risk is investigated for pricing and business strategies. The natural question is, what is the relation between the one-year risk and the ultimate risk and how to model this relation? We try to answer this question in this paper. We focus on the premium risk, that is, on losses from a new, single accident year.

We work in a discrete time setting. Let \( X = (X_0, X_1, X_2, \ldots, X_n) \) denote a stochastic process on a probability space \((\Omega, \mathcal{F}, P)\) with a filtration \( \mathcal{F} = (\mathcal{F}_k)_{k=0,1,\ldots,n} \). We set \( X_0 = 0 \). We assume that \( \mathcal{F} \) is the natural filtration generated by \( X \). Consequently, the process \( X \) is \( \mathcal{F} \)-adapted. We use \( X_0 \) to introduce an explicit notation for the information at time 0, which we later need to correctly define the true emergence pattern as a conditional distribution. The random variables \((X_1, X_2, \ldots, X_n)\) describe the claims development process for an accident year: \( X_1 \) denotes the claims paid in the first development year and \( X_n \) denotes the aggregate value of all claims paid. The ultimate premium risk is modelled with \( X_n \). The one-year premium risk is modelled with \( BE_1 = \mathbb{E}[X_n | \mathcal{F}_1] \).

The expected value \( BE_1 = \mathbb{E}[X_n | \mathcal{F}_1] = \mathbb{E}[X_n | X_1] \) is called the best estimate of the ultimate loss after the first year. For simplicity, we neglect discounting in the definition of the best estimate of the ultimate loss. We can see that the relation between the one-year premium risk and the ultimate premium risk depends on the relation between \( X_1 \) and \( X_n \). The relation between \( X_1 \) and \( X_n \) is usually specified in a claims reserving model which specifies the claims development process \((X_1, \ldots, X_n)\). In claims reserving models, see for example, Radtke et al. (2016), Hertig (1985) and Wüthrich and Merz (2008), we choose the distribution \( X_1 \) and the conditional distributions \((X_{i+1} | X_i, \ldots, X_1)_{i=1}^{n-1}\). We can model the one-year risk and the ultimate risk \((BE_1, X_n)\) by generating them in a forward way: we generate the first year loss \( X_1 \), we calculate the best estimate \( \mathbb{E}[X_n | X_1] \), using the development factors \((X_{i+1} | X_i, \ldots, X_1)_{i=1}^{n-1}\) we generate the losses in the next development years and we get the ultimate loss \( X_n \). Such a forward simulation scheme is investigated, for example, in Wüthrich and Merz (2015) where the relation between the one-year reserve risk and the ultimate reserve risk is discussed in detail. The new problem which we study in this paper is how to model the one-year premium risk and the ultimate premium risk \((BE_1, X_n)\) by generating them in a backward way starting with the ultimate loss \( X_n \) for the new accident year under consideration.

This paper is motivated by a real problem faced by insurance companies in Solvency II regime. From business point of view, the unconditional distribution of \( X_n \) is well understood by decision makers, it is used in all planning...
reports, pricing of products and long-term risk analysis. At the same time, the premium risk module in Solvency II focuses on the risk in the first year for the premiums earned, hence the distribution of $X_1$ must be used. Due to the business reasons mentioned above, insurance companies are less willing to use the distribution of $X_1$. They would usually prefer to perform a backward simulation of $(X_1, X_n)$, that is, simulate the ultimate loss from a well-understood unconditional distribution of $X_n$ and allocate the simulated ultimate loss $X_n$ to the first development year. This allocation should be done based on the conditional distribution of $X_1|X_n = x$, or based on the unconditional distribution of $BE_1|X_n = x$ since $BE_1$ describes the one-year premium risk in Solvency II.

A practical approach to a backward simulation of the cumulative payments $(X_1, \ldots, X_n)$ starting from the ultimate loss $X_n$ is discussed in England et al. (2012) and Bird and Cairns (2011), who introduce concepts of an emergence pattern and emergence factors in reserve risk. These concepts can also be applied to premium risk. The authors suggest that the ultimate loss $X_n$ can be mapped into the best estimate of the ultimate loss $BE_1$ by using a linear function. To our knowledge, this allocation rule is used in practice by many insurance companies to model one-year premium and reserve risks. We postulate a new approach to define the emergence pattern of the ultimate loss. Since we investigate the one-year premium risk, we only have to define the emergence pattern of $X_n$ in the first development year, which is different to the reserve risk where we have to define the emergence patterns of $X_n$ in all development years except the first one. In this paper, the emergence pattern of the ultimate loss for the one-year premium risk is based on the conditional distribution of the best estimate of the ultimate loss given the ultimate loss $BE_1|((F_0, X_n)$ where the conditional distribution is derived from the multivariate distribution of the claims development process $(X_0, X_1, \ldots, X_n)$ and the definition of the best estimate of the ultimate loss after the first year $BE_1$. The definition $BE_1|((F_0, X_n)$ emphasises that we should not view the emergence pattern of the ultimate loss as $BE_1|F_n$. In the sequel, we use $BE_1|X_n$ as a simplified notation for $BE_1|((F_0, X_n)$. The conditional distribution of $BE_1|X_n$ is interpreted as the true emergence pattern of the ultimate loss $X_n$. From the conditional distribution of $BE_1|X_n$, we next derive the unconditional distribution of the best estimate of the ultimate loss $BE_1$ used for quantifying the true (unconditional) one-year premium risk.

We would like to remark that similar ideas of using conditional distributions given the ultimate loss can be found in Chapter 4.2 in Wüthrich and Merz (2008) where the authors postulate a claims reserving model with a latent variable determined by the ultimate loss, and in Wüthrich and Merz (2010) where ultimate incurred losses are used as prior for payments in a claims reserving problem. The backward schemes relating the intermediate cumulative loss to the ultimate loss are stated as the assumptions of the models, and they are not derived from forward claims development dynamics as in this paper. In both papers, the goal is to derive the a posteriori distribution of the ultimate loss in
a Bayesian framework, and to find the prediction of the ultimate loss and the
mean square error of prediction.

We consider three common claims development models used for claims
reserving (Gaussian Incremental Loss Ratio (ILR), Hertig’s Lognormal and
Over-Dispersed Poisson models) and we establish the conditional distributions
of $BE_1|X_n$ and the unconditional distributions of $BE_1$ in these models. We
prove that the true emergence patterns in these three models are different from
the emergence pattern postulated by England et al. (2012) and Bird and Cairns
(2011). We identify that the true one-year premium risk, when measured with
Value-at-Risk, can be underestimated if the emergence pattern formula from England et al. (2012) and Bird and Cairns (2011) is applied, and the true one-year premium risk, when measured with Value-at-Risk, can be higher than the ultimate premium risk at some low confidence levels. Each of
the three investigated claims development models assumes a particular dis-
tribution of the ultimate loss and a particular claims development process
(respectively: normal, lognormal and Poisson). In order to investigate possible
relations between the one-year premium risk and the ultimate premium risk
in models with various distributions of the ultimate loss and various claims
development processes, we keep the conditional distributions $BE_1|X_n$ from the
well-known claims development models but use an arbitrary distribution $X_n$. The key conclusions from this part of the paper are that the distribution of the
one-year premium risk predicted by the emergence pattern formula by England et al. (2012) and Bird and Cairns (2011) can have incorrect tail behaviour and
the true one-year premium risk, when measured with Value-at-Risk, can be
higher than the ultimate premium risk at all high confidence levels. The last
result contradicts the common belief among actuaries that the one-year risk is
always lower than the ultimate risk if we go far in the right tail, see for example,
Lloyd’s (2014) or AISAM-ACME (2007). Yet, if standard deviation is used to
measure the risk, then the true emergence pattern formulas from this paper
give the same one-year premium risk as the emergence pattern formula from
England et al. (2012) and Bird and Cairns (2011), and the one-year premium
risk is always lower than the ultimate premium risk in our models.

We consider light-tailed, subexponential and Pareto-type distributions. Unless otherwise stated, we assume that the distribution $F$ of the ultimate loss
$X_n$ is absolutely continuous, supported on $(0, \infty)$ with infinite right-end point
and has finite second moment. By a light-tailed loss distribution we mean a
distribution $F$ such that
\[\lim_{x \to \infty} e^{tx}(1 - F(x)) = 0, \quad \text{for some } t > 0. \tag{1.1}\]

By a subexponential loss distribution we mean a distribution $F$ such that for
independent copies of $X \sim F$ we have
\[Pr(X_1 + \cdots + X_n > x) \sim nPr(X > x) \quad x \to \infty. \tag{1.2}\]
It is known that \( \lim_{x \to \infty} e^{tx}(1 - F(x)) = \infty \) for any \( t > 0 \). By a Pareto-type distribution with a tail index \( \theta \) we mean a distribution \( F \) such that

\[
1 - F(x) = L(x)x^{-\theta}, \quad x > 0,
\]

where \( L \) denotes a slowly varying function, that is, a function which satisfies \( \lim_{x \to \infty} \frac{L(tx)}{L(x)} = 1 \) for any \( t > 0 \). Since subexponential distributions include Pareto-type distributions, we focus on subexponential distributions with all moments finite.

We always assume that the risk is measured with Value-at-Risk, unless otherwise stated. We quantify the (unconditional) ultimate premium risk and the (unconditional) one-year premium risk with \( \text{VaR}_\gamma[X_n] \) and \( \text{VaR}_\gamma[\text{BE}_1 - \mathbb{E}[\text{BE}_1]] \) at some confidence level \( \gamma \). In reserve risk models, the so-called one-year claims development results \( \text{CDRs} \) are investigated, see Wüthrich and Merz (2015). The counterpart of the one-year \( \text{CDR} \), which we want to use for the one-year premium risk, is the technical result for the new accident year, defined as the difference between the premiums earned in the year and the best estimate of the ultimate loss at the end of the year for the losses incurred in this accident year. Since the premiums include an expected profit margin, we replace the premiums with the expected loss \( \mathbb{E}[X_n] \) in our definition of the premium risk. From insurance point of view, we only consider confidence levels \( \gamma > 0.5 \) such that \( \text{VaR}_\gamma[X_n] > \mathbb{E}[X_n] \) (but we may have \( \text{VaR}_\gamma[\text{BE}_1] < \mathbb{E}[\text{BE}_1] \)).

In Section 2, we describe the emergence pattern formula from England et al. (2012) and Bird and Cairns (2011). In Section 3, we investigate the true emergence patterns and the true one-year risks in Gaussian ILR, Hertig’s Lognormal and Over-Dispersed Poisson models. In Section 4, we study the one-year premium risk if we use an arbitrary distribution of the ultimate loss and allocate this loss to the first year by using the conditional distributions from Section 3. All proofs are presented in the last section.

2. The Linear Emergence Pattern

England et al. (2012) and Bird and Cairns (2011) introduced the concept of an emergence pattern of the ultimate loss for one-year reserve risk. They postulate a linear relation between the best estimate of the ultimate loss and the ultimate loss. If their approach is applied to one-year premium risk, then we have the relation:

\[
\text{BE}_{1\text{p}} = \alpha X_n + (1 - \alpha)\mathbb{E}[X_n], \tag{2.1}
\]

where \( \alpha \) is called an emergence factor, and \( \alpha \in (0, 1) \). It should be noted that \( \text{BE}_{1\text{p}} \) is a \( \sigma(X_n) \)-measurable random variable, in contrast to \( \text{BE}_1 = \mathbb{E}[X_n|X_1] \).
which is a $\sigma(X_1)$-measurable random variable. England et al. (2012) and Bird and Cairns (2011) propose to use

$$\alpha = \frac{SD[BE_1]}{SD[X_n]} = \frac{SD[BE_1 - \mathbb{E}[BE_1]]}{SD[X_n - \mathbb{E}[X_n]]}. \quad (2.2)$$

The emergence factor $\alpha$ in (2.2) measures the relation between the one-year premium risk and the ultimate premium risk for the new accident year under consideration, where the risk is measured with standard deviation.

We would like to remark that in this paper we don’t consider estimation error and we assume that all parameters of the claims development process are given. Consequently, the emergence factor $\alpha$ can be directly derived from the multivariate distribution of $(X_1, \ldots, X_n)$ assumed for the accident year. In practice, assuming a Markovian claims development process or a claims development process with independent increments, we would have to estimate the distribution of the loss in the first year $X_1$ and the distributions of the development factors $(X_{i+1} | X_i)_{i=1}^{n-1}$, or $(X_{i+1} - X_i)_{i=1}^{n-1}$, in consecutive years and, next, compare the deviations of $BE_1$ and $X_n$ calculated/simulated in the forward way. The idea of estimating $\alpha$ for the one-year premium risk is the same as for the one-year reserve risk, see England et al. (2012) and Bird and Cairns (2011). The distributions of the development factors $(X_{i+1} | X_i)_{i=1}^{n-1}$, or $(X_{i+1} - X_i)_{i=1}^{n-1}$, are estimated from the historical losses in a run-off triangle in a claims reserving model (e.g. Chain-Ladder, ILR and additive model). However, in order to derive (2.2) for the one-year premium risk, we also have to establish the unconditional distribution of $X_1$, which is usually related to the exposure (premium) in the accident year, see ILR and additive models in Chapter 8.3 in Wüthrich and Merz (2008). The parameters of the claims development process estimated from historical data in a claims reserving model can be used in the new accident year. In particular, $X_1$ can be simulated using the planned volume of the premiums to be earned in the new accident year, coming from the financial plan of the company. Due to the assumption that we know all parameters of the claims development process, the one-year premium risk is, and can be, investigated independently of the one-year reserve risk in the sense that the development factors observed in the new calendar year in the loss run-off triangle don’t impact the best estimate of the ultimate loss for the losses incurred in the new accident year, that is, we don’t have to consider the so-called re-reserving/bootstraping techniques for $\mathbb{E}[X_n | X_1]$. We are aware of the simplifications; however, the key goal of this paper is to analyse probabilistic properties of the one-year premium risk and the ultimate premium risk implied by various claims development processes.

We have a simple algorithm (2.1) for the simulation of $BE_1$ starting from $X_n$. Moreover, the parametrisation of the formula (2.1) is very appealing. We point out that we can use any distribution of the ultimate loss $X_n$ parameterised in the premium risk model, and we can switch from the ultimate premium risk to the one-year premium risk using a single scaling factor $\alpha$ parameterised in
the claims reserving model. The distribution of $X_n$ used in (2.1) does not have to coincide with the distribution of the ultimate loss from the claims reserving model where the parameter $\alpha$ is calibrated, see England et al. (2012) and Bird and Cairns (2011).

Let us investigate theoretical foundations of the emergence pattern formula (2.1). Since we want to model a pair of dependent random variables $(BE_1, X_n)$, the simplest approach would be to use a linear factor model for $(BE_1, X_n)$. It is natural to assume that $\mathbb{E}[BE_1] = \mathbb{E}[X_n]$ to guarantee that the estimation of the ultimate loss is unbiased. We get the relation:

$$BE_1 = \eta X_n + (1 - \eta)\mathbb{E}[X_n] + \xi, \quad \eta = \rho(BE_1, X_n) \frac{SD(BE_1)}{SD(X_n)}, \quad (2.3)$$

where $\mathbb{E}[\xi] = 0$, $Var[\xi] = (1 - \rho^2(BE_1, X_n))Var[B_E_1]$ and $\xi$ is independent of $\mathcal{F}_n$.

Let us recall that $BE_1 = \mathbb{E}[X_n | X_1]$ is a $\sigma(X_1)$-measurable random variable, the emergence pattern formula $BE_{ep} = \alpha X_n + (1 - \alpha)\mathbb{E}[X_n]$ is a $\sigma(X_n)$-measurable random variable, and we note that $BE_1$ is a $\sigma(X_n) \vee \sigma(\xi)$-measurable random variable, so we have to enlarge the filtration $\mathcal{F}$ to define (2.3) and we set $\tilde{\mathcal{F}}_k = \mathcal{F}_k \vee \sigma(\xi)$, $k = 0, 1, \ldots, n$. We can see that the linear factor model (2.3) leads to the emergence pattern (2.1) with $\alpha$ from (2.2) only if $\rho(BE_1, X_n) = 1$. We can immediately derive the following conclusions.

**Corollary 2.1.** 1. The emergence pattern formula (2.1) is built on the assumption that the best estimate of the ultimate loss $BE_1$ is perfectly linearly correlated with the ultimate loss $X_n$ and both are $\sigma(X_1)$-measurable random variables.

2. The emergence pattern formula (2.1) cannot describe the true emergence pattern of the ultimate loss, in the sense of the conditional distribution of $BE_1 | X_n = x$, and the true joint distribution of the one-year risk and the ultimate risk $(BE_1, X_n)$ in non-trivial models of claims development, including Gaussian ILR, Hertig’s Lognormal and Over-Dispersed Poisson models.

The emergence pattern formula (2.1) implies that the conditional distribution $BE_{ep} | X_n = x$ is degenerate. However, the true conditional distribution $BE_1 | X_n = x$ and the true conditional emergence pattern formula for the ultimate loss cannot be degenerated in a non-trivial claims development model, since we expect that there are many scenarios with many different possible values of $BE_1$ which can lead to the same value of $X_n$. Hence, there is a clear need for an improvement in the emergence pattern formula.

In many applications, including Solvency II calculation of the capital requirement, we only need an unconditional sample of the best estimate of the ultimate loss $BE_1$, instead of a joint sample $(BE_1, X_n)$ or a conditional sample $BE_1 | X_n = x$. Let us use the formula (2.1) to generate the unconditional sample of $BE_1$. We are interested in properties of the (unconditional) one-year risk versus the ultimate risk.
Theorem 2.1. Let us consider the emergence pattern formula (2.1) with an emergence factor $\alpha \in (0, 1)$.

1. $E[BE_1^{ep}] = E[X_n]$ and $Var[BE_1^{ep}] = \alpha^2 Var[X_n] < Var[X_n]$.

2. If $X_n$ has a light-tailed distribution (subexponential with all moments finite), then $BE_1^{ep}$ has a light-tailed distribution (subexponential with all moments finite).

3. If $X_n$ has a Pareto-type distribution with tail index $\theta$, then $BE_1^{ep}$ has a Pareto-type distribution with tail index $\theta$.

4. $VaR_\gamma[BE_1^{ep} - E[BE_1^{ep}]] = \alpha VaR_\gamma[X_n - E[X_n]] < VaR_\gamma[X_n - E[X_n]]$.

If we use (2.1), then the unconditional expected value of the best estimate of the ultimate loss is equal to the unconditional expected value of the ultimate loss. This property is desired since otherwise the estimation of the ultimate loss would be biased. If the emergence factor $\alpha$ is set in accordance with condition (2.2), then $Var[BE_1^{ep}] = Var[BE_1]$, which is again a desired property. Next, the one-year premium risk is a fraction of order $\alpha$ of the ultimate premium risk, if the risk is measured with standard deviation and Value-at-Risk. This property is less clear but it is in accordance with a common belief among actuaries that the one-year risk is lower than the ultimate risk, see for example, Lloyd’s (2014) or AISAM-ACME (2007). Finally, the distributions of the one-year premium risk and the ultimate premium risk have the same tail behaviour. Theorem 2.1 shows that the formula (2.1) has some desirable properties if applied to simulate the unconditional samples of the best estimates of the ultimate loss $BE_1$ from the ultimate losses $X_n$.

3. THE TRUE EMERGENCE PATTERNS IN CLAIMS DEVELOPMENT MODELS

In the next subsections, we derive the conditional distributions of $BE_1|X_n$ and the unconditional distributions of $BE_1$ in three claims development models used for claims reserving.

3.1. Gaussian ILR model

We consider ILR model with Gaussian incremental losses, see for example, Radtke et al. (2016) and Wüthrich and Merz (2008). We investigate cumulative payments given by

$$X_j = \sum_{i=1}^{j} \epsilon_i,$$

where $\epsilon_i \sim N(E\mu_i; E\sigma_i^2)$ for $i \in \{1, \ldots, n\}$,

(3.1)

and $(\epsilon_i)_{i=1}^n$ is a sequence of independent random variables. By $E$ we denote the exposure in the accident year, and $\epsilon_i$ represents the incremental loss in
development year $i$. We assume that $\mu_i \in \mathbb{R}$, $\sigma_i > 0$, $E > 0$. We denote

$$\mu = \sum_{i=1}^{n} \mu_i, \quad \sigma^2 = \sum_{i=1}^{n} \sigma_i^2.$$  

The best estimate of the ultimate loss is determined by

$$BE_1 = E[X_n | X_1] = X_1 + E(\mu - \mu_1). \quad (3.2)$$

The next theorem describes the distributions in which we are interested. The distributions have special parameterisations which mimic the emergence pattern formula (2.1).

**Theorem 3.1.** Let us consider the model (3.1) of claims development and the multivariate distribution of the claims development process $(X_1, \ldots, X_n)$. Let

$$\mu_{X_n} = E[X_n], \quad \sigma_{X_n}^2 = Var[X_n], \quad \alpha = \frac{SD[BE_1]}{SD[X_n]}.$$  

We have the following loss distributions:

$$X_n \sim N(\mu_{X_n}, \sigma_{X_n}^2), \quad (3.3)$$

$$BE_1 | X_n = x \sim N(\alpha^2 x + (1 - \alpha^2)\mu_{X_n}; \alpha^2(1 - \alpha^2)\sigma_{X_n}^2), \quad (3.4)$$

$$BE_1 \sim N(\mu_{X_n}; \alpha^2\sigma_{X_n}^2), \quad (3.5)$$

with the parameter $\alpha$ which satisfies $\alpha \in (0, 1)$.

We remark that $\mu_{X_n}$ coincides with $\mu$, but we prefer to have an explicit parameter for $E[X_n]$.

Formula (3.4) establishes the true emergence pattern of the ultimate loss $X_n$ in our claims development model. Moreover, it describes a backward simulation scheme for $BE_1$, starting from $X_n$, where we can switch (in accordance with the underlying probabilistic model of claims development) from the ultimate premium risk to the one-year premium risk. As in (2.1), we only use the distribution of the ultimate loss $X_n$ from the ultimate premium risk and a single emergence factor $\alpha$ which summarises the information from the reserve risk module about the claims development process required for the one-year premium risk. We follow the main idea behind the emergence pattern formula (2.1). However, we are able to improve the emergence pattern formula from England et al. (2012) and Bird and Cairns (2011), so that it yields the correct conditional distribution of $BE_1 | X_n = x$ and, consequently, the correct unconditional distribution of $BE_1$ in our claims development model, given by (3.5).

If we assume, as in Theorem 3.1, that the ultimate loss $X_n \sim N(\mu_{X_n}, \sigma_{X_n}^2)$, then the emergence pattern formula (2.1) yields the following unconditional
distribution of the best estimate of the ultimate loss:

$$BE_{1}^{ep} \sim N \left( \mu_{Xn}, \alpha^{2} \sigma_{Xn}^{2} \right).$$

(3.6)

Interestingly, the unconditional distributions of the best estimate of the ultimate loss (3.5) and (3.6) are the same. Consequently, there exists a claims development model where the unconditional distribution of the true one-year premium risk can be characterised with the emergence pattern formula from England et al. (2012) and Bird and Cairns (2011). However, as we see in the next subsections, this is an exceptional case.

We conclude with properties of the unconditional distribution of $BE_{1}$, which are special cases of the properties for $BE_{1}^{ep}$ from Theorem 2.1.

**Theorem 3.2.** Let us consider the model and the assumptions from Theorem 3.1.

1. $E[BE_{1}] = E[X_{n}]$ and $Var[BE_{1}] = \alpha^{2} Var[X_{n}] < Var[X_{n}]$.
2. $VaR_{\gamma}[BE_{1} - E[BE_{1}]] = \alpha VaR_{\gamma}[X_{n} - E[X_{n}]] < VaR_{\gamma}[X_{n} - E[X_{n}]]$, where

$$VaR_{\gamma}[BE_{1} - E[BE_{1}]] = \alpha \Phi^{-1}(\gamma) SD[X_{n}].$$

**3.2. Hertig’s Lognormal model**

We consider a multiplicative loss model where the development factors are modelled with lognormal distributions, see for example, Hertig (1985) or Wüthrich and Merz (2008). We investigate cumulative payments given by

$$X_{1} = \epsilon_{1}, \quad X_{i} = X_{i-1} \cdot \epsilon_{i}, \quad \text{where} \quad \epsilon_{i} \sim LogN(m_{i}, s_{i}^{2}) \quad \text{for} \quad i \in \{1, \ldots, n\},$$

(3.7)

and $(\epsilon_{i})_{i=1}^{n}$ is a sequence of independent random variables. We assume that $m_{i} \in \mathbb{R}, s_{i} > 0$. We denote

$$m = \sum_{i=1}^{n} m_{i}, \quad s^{2} = \sum_{i=1}^{n} s_{i}^{2}.$$ 

The best estimate of the ultimate loss is determined by

$$BE_{1} = E[X_{n} | X_{1}] = X_{1} e^{m - m_{1} + \frac{1}{2}(s^{2} - s_{1}^{2})}.$$ 

(3.8)

Using the same reasoning as in the previous section, we can derive the true emergence pattern for the ultimate loss which comes from the claims development process (3.7).

**Theorem 3.3.** Let us consider the model (3.7) of claims development and the multivariate distribution of the claims development process $(X_{1}, \ldots, X_{n})$. Let

$$\mu_{X_{n}} = E[X_{n}], \quad \psi_{X_{n}} = \frac{SD[X_{n}]}{E[X_{n}]}, \quad \alpha = \frac{SD[BE_{1}]}{SD[X_{n}]}.$$
We have the following loss distributions:

\[ X_n \sim \text{LogN}\left(\tilde{m}; \tilde{s}^2\right), \quad (3.9) \]

\[ BE_1|X_n = x \sim \text{LogN}\left(\tilde{\alpha}^2 \log(x) + \left(1 - \tilde{\alpha}^2\right)\left(\tilde{m} + \frac{\tilde{s}^2}{2}\right); \tilde{\alpha}^2(1 - \tilde{\alpha}^2)\tilde{s}^2\right), \quad (3.10) \]

\[ BE_1 \sim \text{LogN}\left(\tilde{m} + \left(1 - \tilde{\alpha}^2\right)\frac{\tilde{s}^2}{2}; \tilde{\alpha}^2\tilde{s}^2\right), \quad (3.11) \]

where

\[ \tilde{m} = \log(\mu_{X_n}) - \frac{1}{2} \log(1 + \psi_x^2), \quad \tilde{s}^2 = \log(1 + \psi_x^2), \]

\[ \tilde{\alpha}^2 = \frac{\log(1 + \alpha^2 \psi_x^2)}{\log(1 + \psi_x^2)}, \]

with the parameter \( \alpha \) which satisfies \( \alpha \in (0, 1) \).

**Theorem 3.4.** Let us consider the model and the assumptions from Theorem 3.3.

1. \( \mathbb{E}[BE_1] = \mathbb{E}[X_n] \) and \( \text{Var}[BE_1] = \alpha^2 \text{Var}[X_n] < \text{Var}[X_n] \).
2. \( \text{VaR}_\gamma[BX_1 - \mathbb{E}[BE_1]] < \text{VaR}_\gamma[X_n - \mathbb{E}[X_n]] \) for \( \gamma > \gamma^* \), \( \text{VaR}_\gamma[BX_1 - \mathbb{E}[BE_1]] > \text{VaR}_\gamma[X_n - \mathbb{E}[X_n]] \) for \( \gamma < \gamma^* \), where

\[
\text{VaR}_\gamma[BX_1] = \mu_{X_n}(1 + \alpha^2 \psi_x^2)^{-1/2} e^{\Phi^{-1}(\gamma)} \sqrt{\log(1 + \psi_x^2)} \Phi^{-1}(\gamma),
\]

\[ \gamma^* = \Phi\left(\frac{1}{2} \sqrt{\log(1 + \alpha^2 \psi_x)} + \frac{1}{2} \sqrt{\log(1 + \psi_x^2)}\right), \]

3. We have the limit

\[
\lim_{\gamma \to 1} \frac{\text{VaR}_\gamma[BX_1 - \mathbb{E}[BE_1]]}{\text{VaR}_\gamma[X_n - \mathbb{E}[X_n]]} = 0.
\]

If we make the assumptions from Theorem 3.3 and apply the emergence pattern formula (2.1) to the claims development process (3.7), then we can deduce the distribution:

\[ BE_{1\text{ep}} \sim \alpha \cdot \text{LogN}(\tilde{m}, \tilde{s}^2) + (1 - \alpha)\mu_{X_n}. \quad (3.12) \]

Let us compare Theorems 3.4 with 2.1 and the emergence pattern formula (3.10) with (2.1). Recalling the discussion from the previous sections, it is clear that the conditional distributions of the best estimate of the ultimate loss given the ultimate loss (3.10) and (2.1) are different. We can now observe that the unconditional distributions of the best estimate of the ultimate loss (3.11) and
(3.12) are also different if derived from the emergence pattern formula and the conditional distribution approach.

We now focus on the premium risk in the claims development model (3.7). We can notice that the supports of $BE_1$ and $BE_1^{\psi}$ generated with (3.11) and (3.12) are different. From point 2 from Theorem 3.4 we deduce that the true one-year risk can be higher than the ultimate risk if the confidence level $\gamma$ is too low. We point out that a confidence level $\gamma$ at which we may observe that $VaR_{\gamma}[BE_1] > VaR_{\gamma}[X_n]$ is above the lowest confidence level we allow for, that is, is above $\tilde{\gamma} = \Phi\left(\frac{1}{\sqrt{2}} \log(1 + \psi_{X_n})\right)$ at which $VaR_{\tilde{\gamma}}[X_n] = E[X_n]$. For a skewed distribution of $X_n$ with high $\psi_{X_n}$, a low confidence level $\gamma$ at which we may observe that $VaR_{\gamma}[BE_1] > VaR_{\gamma}[X_n]$ can be very high, if $\alpha$ is also high. However, at all sufficiently high confidence levels the true one-year risk is lower than the ultimate risk. This property stands in contrast with the property of the emergence pattern formula (2.1) where the one-year risk is lower than the ultimate risk. Let us recall that

$$E[X_n] = \frac{\alpha}{\psi_{X_n}} \left[ \frac{1}{\alpha - \psi_{X_n}} - 1 \right]$$

and overestimates the true one-year risk at low confidence levels $\alpha$ and $\psi_{X_n}$. Hence, if $VaR_{\gamma}[BE_1] - E[BE_1] > VaR_{\gamma}[X_n] - E[X_n]$ for any $\psi_{X_n}$ and $\gamma$. Hence, if $VaR_{\gamma}[BE_1] - E[BE_1] \gamma > VaR_{\gamma}[X_n] - E[X_n] > \alpha$, then the emergence pattern formula (2.1) underestimates the true one-year risk, and if $VaR_{\gamma}[BE_1] - E[BE_1] \gamma < \alpha$, then the emergence pattern formula (2.1) overestimates the true one-year risk. The results are presented in Figure 1. They confirm that the emergence pattern formula (2.1) can under and overestimate the true one-year premium risk in the claims development model (3.7). For example, for $\psi_{X_n} = 1$ and $\alpha = 50\%$, the emergence pattern formula (2.1) overestimates the true one-year risk at $\gamma = 99.5\%$ by about 25\%. For $\psi_{X_n} = 3$ and $\alpha = 50\%$, the emergence pattern formula (2.1) underestimates the true one-year risk at $\gamma = 99.5\%$ by about 9\%. We can see that the discrepancies between the linear emergence pattern formula and the true one-year risk can be much larger in both directions for different parameters. In practice, we would rarely expect $\psi_{X_n} > 3$. We also observe that the true one-year premium risk can be higher than the ultimate premium risk at low confidence levels $\gamma$. Even though point 3 from Theorem 3.4 shows that $VaR_{\gamma}[BE_1] - E[BE_1] = o\left(VaR_{\gamma}[X_n] - E[X_n]\right)$ for $\gamma \to 1$, we end up with
FIGURE 1: The ratios $\frac{\text{VaR}_\gamma[BE_1 - \mathbb{E}[BE_1]]}{\text{VaR}_\gamma[X_n - \mathbb{E}[X_n]]}$ in the Hertig’s Lognormal model.

$\text{VaR}_\gamma[BE_1 - \mathbb{E}[BE_1]] \approx \alpha^2 \text{VaR}_\gamma[X_n - \mathbb{E}[X_n]]$ for $\gamma$ in the range 99.99–99.999% in our experiments.

3.3. Over-Dispersed Poisson model

Finally, we consider ILR model with Over-Dispersed Poisson incremental losses, see for example, Wüthrich and Merz (2008) or Taylor and McGuire (2016). We investigate cumulative payments given by

$$X_j = \sum_{i=1}^{j} \epsilon_i, \quad \text{where} \quad \epsilon_i \sim ODP(\mu \omega_i, \psi) \quad \text{for} \quad i \in \{1, \ldots, n\}, \quad (3.13)$$

and $(\epsilon_i)_{i=1}^{n}$ is a sequence of independent random variables. We assume that $\mu > 0$, $\omega_i > 0$, $\psi > 0$. The parameter $\mu$ is interpreted as the expected ultimate loss, and $\omega_i$ is interpreted as the proportion of $\mu$ which is expected to be paid in development year $i$. Hence, we assume that $\sum_{i=1}^{n} \omega_i = 1$. We assume that the dispersion $\psi$ is independent of the development year so that we can derive analytical results. The assumption that $\epsilon \sim \text{ODPoiss}(\mu, \psi)$ means that $\epsilon/\psi \sim \text{Poiss}(\mu/\psi)$. The best estimate of the ultimate loss is determined by

$$BE_1 = \mathbb{E}[X_n | X_1] = X_1 + \mu(1 - \omega_1). \quad (3.14)$$

Theorem 3.5. Let us consider the model (3.13) of claims development and the multivariate distribution of the claims development process $(X_1, \ldots, X_n)$. Let

$$\mu_{X_n} = \mathbb{E}[X_n], \quad \psi_{X_n} = \frac{\text{Var}[X_n]}{\mathbb{E}[X_n]}, \quad \alpha = \frac{\text{SD}[BE_1]}{\text{SD}[X_n]}.$$
We have the following loss distributions:

\[ X_n \sim \psi_{X_n} \cdot \text{Poiss}(\mu_{X_n}/\psi_{X_n}) , \quad (3.15) \]

\[ BE_1|X_n = x \sim \psi_{X_n} \cdot \text{Bin}\left(\frac{x}{\psi_{X_n}}; \alpha^2\right) + (1 - \alpha^2)\mu_{X_n} , \quad (3.16) \]

\[ BE_1 \sim \psi_{X_n} \cdot \text{Poiss}(\alpha^2\mu_{X_n}/\psi_{X_n}) + (1 - \alpha^2)\mu_{X_n} , \quad (3.17) \]

with the parameter \( \alpha \) which satisfies \( \alpha \in (0, 1) \).

**Theorem 3.6.** Let us consider the model and the assumptions from Theorem 3.5.

1. \( \mathbb{E}[BE_1] = \mathbb{E}[X_n] \) and \( \text{Var}[BE_1] = \alpha^2\text{Var}[X_n] < \text{Var}[X_n] \).
2. \( \text{VaR}_\gamma[BE_1 - \mathbb{E}[BE_1]] < \text{VaR}_\gamma[X_n - \mathbb{E}[X_n]] \) for \( \gamma > \gamma^* \) where \( \gamma^* < 1 \) is the last point where the distribution functions of \( BE_1 \) and \( X_n \) intersect, or \( \text{VaR}_\gamma[BE_1 - \mathbb{E}[BE_1]] < \text{VaR}_\gamma[X_n - \mathbb{E}[X_n]] \) for all \( \gamma \) if the distribution functions of \( BE_1 \) and \( X_n \) don’t intersect. The relation between \( \text{VaR}_\gamma[BE_1 - \mathbb{E}[BE_1]] \) and \( \text{VaR}_\gamma[X_n - \mathbb{E}[X_n]] \) can change for \( \gamma < \gamma^* \) since the distribution functions of \( BE_1 \) and \( X_n \) can intersect more than once (if they intersect).
3. We have the limit

\[
\lim_{\gamma \to 1} \frac{\text{VaR}_\gamma[BE_1 - \mathbb{E}[BE_1]]}{\text{VaR}_\gamma[X_n - \mathbb{E}[X_n]]} = 1.
\]

If we make the assumptions from Theorem 3.5 and we apply the emergence pattern formula (2.1) to the claims development process (3.13), then we can derive the distribution:

\[ BE_{1ep} \sim \alpha\psi_{X_n} \cdot \text{Poiss}(\mu_{X_n}/\psi_{X_n}) + (1 - \alpha)\mu_{X_n} . \quad (3.18) \]

By point 2 from Theorem 3.6, there exists a confidence level \( \gamma^* < 1 \) such that the true one-year risk is lower than the ultimate risk at all sufficiently high confidence levels \( \gamma > \gamma^* \). At low confidence levels \( \gamma < \gamma^* \), the true one-year risk can be lower or higher than the ultimate risk. This property makes the interpretation of the one-year premium risk in relation to the ultimate premium risk difficult in the claims development model (3.13) since the relation between the true one-year premium risk and the ultimate premium risk can change multiple times as we change the confidence level. From point 3 from Theorem 3.6 we can deduce that the ratio between the Value-at-Risk measures for the one-year risk \( BE_1 \) and the ultimate risk \( X_n \) at a sufficiently high confidence level must be higher than \( \alpha \). Consequently, the emergence pattern formula (2.1) underestimates the true one-year premium risk at all high confidence levels. The emergence pattern formula (2.1) can also underestimate the true one-year premium risk at some low confidence levels, for example, when the true one-year risk is above the ultimate risk.
Example 3.2. We choose $\psi_{X_n} = 1$ and we consider Poisson distributions of $X_n$ with different $\mu_{X_n}$. The Over-Dispersed Poisson model approaches the Gaussian ILR model for $\mu_{X_n} \to \infty$, hence we are interested in low $\mu_{X_n}$. Our key results are presented in Figure 2. We remark that some ratios are negative since $VaR_\gamma[BE_1] < \mathbb{E}[BE_1]$ (but $VaR_\gamma[X_n] > \mathbb{E}[X_n]$ by our assumption). We can confirm that the emergence pattern formula (2.1) can under and overestimate the true one-year premium risk in the claims development model (3.13). For example, for $\mu_{X_n} = 1.5$ and $\alpha = 85\%$, the emergence pattern formula (2.1) underestimates the true one-year risk at $\gamma = 99.5\%$ by 24\% and the true one-year risk is above the ultimate risk by 12\%. For $\mu_{X_n} = 25$ and $\alpha = 15\%$, the emergence pattern formula (2.1) underestimates the true one-year risk at $\gamma = 99.5\%$ by 14\%. For $\mu_{X_n} = 1.5$ and $\alpha = 85\%$, the true one-year risk is higher than the ultimate risk at $\gamma = 75\%$ and 99.5\%, but the true one-year risk is lower than the ultimate risk for $\gamma = 90\%$ and 99.99\% – in this example the distribution functions of $BE_1$ and $X_n$ intersect nine times in the range from the median up to the quantile of order $\gamma^* = 99.555\%$. Moreover, the emergence pattern formula (2.1) overestimates the true one-year risk at $\gamma = 90\%$ by 39\%. Finally, even though point 3 from Theorem 3.6 shows that $VaR_\gamma[BE_1 - \mathbb{E}[BE_1]] \sim VaR_\gamma[X_n - \mathbb{E}[X_n]]$ for $\gamma \to 1$, the convergence for $\gamma$ in the range 99.99–99.999\% is slow in our experiments, yet it can be observed.

4. ONE-YEAR PREMIUM RISK WITH AN ARBITRARY ULTIMATE LOSS DISTRIBUTION

The models which we investigated in the previous section assume a particular form of the claims development process and lead to a particular distribution of
the ultimate loss. What is important for practical applications in Solvency II in premium risk module is the possibility of using any distribution of the ultimate loss. Yet, it may be difficult to specify a priori the joint multivariate distribution for cumulative payments \((X_1, \ldots, X_n)\) which lead to a pre-specified distribution of the ultimate loss \(X_n\). What we suggest is to use an arbitrary unconditional distribution of the ultimate loss \(X_n\) and a conditional distribution of \(BE_1|X_n\), extracted from the well-known claims development model, as a mechanism for allocating the ultimate loss \(X_n\) to the best estimate \(BE_1\). We focus on the conditional distributions \(BE_1|X_n\) from the Gaussian ILR model and the Hertig’s Lognormal model.

We use the following construction. Let us consider a vector of dependent random variables \((BE_1, X_n) \sim f\) with a joint density function of \(f\). The joint density \(f\) implies the marginal and conditional densities:

\[
X_n \sim f_{X_n}, \quad X_n|BE_1 = x \sim f_{X_n|BE_1=x}, \quad BE_1|X_n = x \sim f_{BE_1|X_n=x}.
\]

We define a new joint density of \((BE_1, X_n) \sim f_{\text{new}}\) by choosing \(BE_1|X_n=x \sim f_{BE_1|X_n=x}\) and \(X_n \sim f_{X_n}^{\text{new}}\) where \(f_{X_n}^{\text{new}}\) is a density absolutely continuous with respect to \(f_{X_n}\). We have the new marginal and conditional densities:

\[
f_{BE_1}^{\text{new}}(z) = \int f_{BE_1|X_n=x}(z) f_{X_n}^{\text{new}}(x) dx, \quad (4.1)
\]

\[
f_{X_n|BE_1=x}^{\text{new}}(z) = C \cdot f_{X_n|BE_1=x}(z) f_{X_n}^{\text{new}}(z), \quad (4.2)
\]

where \(C\) denotes a normalising constant which depends on \(x\) but is independent of \(z\).

We can construct new models of claims development by using the distribution of the ultimate loss \(X_n\) and the emergence pattern of \(X_n\) characterised with the conditional distribution of \(BE_1|X_n\). We recall that the emergence factor \(\alpha\) needed for \(BE_1|X_n\) is discussed in Section 2. Our approach has two advantages. Firstly, we have a flexible and interpretable probabilistic model where we can freely choose the distribution of the ultimate loss and switch from the ultimate premium risk to the one-year premium risk (the backward simulation scheme). This is a desired feature in Solvency II modelling. Secondly, we can investigate properties of the one-year premium risk versus the ultimate premium risk in various claims development models, beyond the models we know from the claims reserving literature. This is the topic of the next two subsections.

4.1. The conditional distribution from Gaussian ILR model

From (3.4), we can deduce that the unconditional distribution of \(BE_1\) is a mixture of normal distributions. We introduce the representation:

\[
BE_1^{\text{normal, ep}} = \alpha^2 X_n + (1 - \alpha^2) \mu_{X_n} + \sqrt{\alpha^2(1 - \alpha^2)} \sigma_{X_n} \xi , \quad (4.3)
\]
where $\xi \sim N(0, 1)$, and $\xi$ is independent of $\mathcal{F}_n$. The formula (4.3) is our new emergence pattern formula which allocates the ultimate loss with an arbitrary distribution to the best estimate of the ultimate loss. We call (4.3) the additive normal emergence pattern formula. As discussed when establishing the formula (2.3), the emergence pattern formula (4.3) is a $\sigma(X_n) \vee \sigma(\xi)$-measurable random variable. In the sequel, $BE_{1\text{normal, ep}}$ is denoted by $BE_1$, since we only study the unconditional distribution of $BE_{1\text{normal, ep}}$. The additive normal emergence pattern can be seen as a version of the classical emergence pattern formula (2.1) where we simply add a Gaussian noise in order to have a non-degenerate distribution of $BE_1|X_n = x$. If we choose a normal distribution for $X_n$ in (4.3), then we are in the framework of the Gaussian ILR model and the results from Section 3.1 apply.

**Theorem 4.1.** Let us assume that the ultimate loss is modelled with $X_n \sim F_{X_n}$. Let $\mu_{X_n} = \mathbb{E}[X_n]$, $\sigma^2_{X_n} = \text{Var}[X_n]$ and choose an emergence factor $\alpha \in (0, 1)$. The emergence pattern of the ultimate loss $X_n$ is described with the formula (4.3). We have the following properties:

1. $\mathbb{E}[BE_1] = \mathbb{E}[X_n]$ and $\text{Var}[BE_1] = \alpha^2 \text{Var}[X_n] < \text{Var}[X_n]$.
2. If $X_n$ has a light-tailed distribution (subexponential with all moments finite), then $BE_1$ has a light-tailed distribution (subexponential with all moments finite). Moreover, if $X_n$ has a light-tailed distribution such that $\lim_{x \to \infty} (1 - F_{X_n}(x))e^{v^2 x^2} = 0$ for all $v > 0$, then $\text{VaR}_\gamma[BE_1 - \mathbb{E}[BE_1]] > \text{VaR}_\gamma[X_n - \mathbb{E}[X_n]]$ for $\gamma > \gamma^*$, with some $\gamma^* < 1$.
3. If $X_n$ has a Pareto-type distribution with tail index $\theta$, then $BE_1$ has a Pareto-type distribution with tail index $\theta$, and we have the limit
   \[
   \lim_{\gamma \to 1} \frac{\text{VaR}_\gamma[BE_1 - \mathbb{E}[BE_1]]}{\text{VaR}_\gamma[X_n - \mathbb{E}[X_n]]} = \alpha^2.
   \]

From point 2 we deduce a new and interesting property that the true one-year premium risk can be higher than the ultimate premium risk at all high confidence levels. Points 2 and 3 indicate that the distributions of the true one-year premium risk and the ultimate premium risk have the same tail behaviour. This property holds in Gaussian ILR, Hertig’s Lognormal and Over-Dispersed Poisson models where the distribution of $X_n$ is uniquely specified by the underlying claims development model, as well as if we model the one-year risk with the emergence pattern formula (2.1). Finally, point 3 gives a new possible value of $\lim_{\gamma \to 1} \frac{\text{VaR}_\gamma[BE_1 - \mathbb{E}[BE_1]]}{\text{VaR}_\gamma[X_n - \mathbb{E}[X_n]]}$.

**Example 4.1.** First, we assume that $X_n \sim \text{Exp}(\lambda)$ with $1 - F_{X_n}(x) = e^{-\lambda x}$. In this case, the claims development process for the pair $(BE_1, X_n)$ with the ultimate loss
\(X_n \sim \text{Exp}(\lambda)\) and the emergence pattern of the ultimate loss given by the conditional distribution (3.4) can be fully characterised. From (4.1) to (4.2), using classical calculus, we can derive the distributions:

\[
Pr(BE_1 \leq x) = \Phi\left(\frac{\lambda x - (1 - \alpha^2)}{\sqrt{\alpha^2(1 - \alpha^2)}}\right) - \exp\left(\frac{3(1 - \alpha^2) - 2\lambda x}{2\alpha^2}\right) \times \Phi\left(\frac{\lambda x - 2(1 - \alpha^2)}{\sqrt{\alpha^2(1 - \alpha^2)}}\right),
\]

\[X_n | BE_1 = x \sim N_0\left(\frac{\lambda x - 2(1 - \alpha^2)}{\alpha^2 \lambda}, \frac{1 - \alpha^2}{\alpha^2 \lambda^2}\right),\]

where \(N_0\) denotes normal distribution left truncated at 0 and \(\Phi\) denotes the standard normal distribution function. To have more flexibility in our numerical study, we consider \(X_n \sim \text{Weibull}(\lambda, \tau)\) with \(1 - F_{X_n}(x) = e^{-\lambda x^\tau}\). We choose \(\lambda = 1\) since \(\lambda\) is scaling parameter for the distribution of \(X_n\). If \(\tau = 1\), then \(X_n \sim \text{Exp}(1)\). If \(\tau > 2\), then \(\lim_{x \to \infty} (1 - F_{X_n}(x)) e^{vx^2} = 0\) for any \(v > 0\). In order to calculate \(\text{VaR}_\gamma[BE_1]\), we use Monte Carlo simulations and we generate a sample of size \(10^6\). The results are presented in Figure 3. For \(\tau < 2\), the one-year risk is lower than the ultimate risk at sufficiently high confidence levels but the one-year risk can be higher than the ultimate risk at low confidence levels. For \(\tau > 2\), we observe the new property: the one-year risk can be higher than the ultimate risk at all high confidence levels. For example, for \(\alpha = 85\%\) and \(\tau = 20\), we can observe that the one-year risk is higher than the ultimate risk at \(\gamma = 99.5\%\) and 99.99%, and at
all higher confidence levels. The higher \( \alpha \) and \( \tau \), the lower the minimal confidence level \( \gamma^* \) from which the one-year risk dominates the ultimate risk. If \( \tau > 3.6 \), then the Weibull distribution is negatively skewed. Negatively skewed distributions can appear in situations where data are clustered near an upper limit, and the most well-known case where we use a negatively skewed Weibull distribution is modelling of the future lifetime in developed countries where people tend to survive to older ages.

4.2. The conditional distribution from Hertig’s Lognormal model

From (3.10), we can deduce that the unconditional distribution of \( BE_1 \) is a mixture of lognormal distributions. We introduce the representation:

\[
BE_1^{\text{lognormal}_ep} = (X_n)^\tilde{\alpha}e^{(1-\tilde{\alpha}^2)(\tilde{m}+\tilde{s}^2)} + \sqrt{\tilde{\alpha}^2(1-\tilde{\alpha}^2)\xi},
\]

where \( \xi \sim N(0, 1) \), and \( \xi \) is independent of \( \mathcal{F}_n \). The formula (4.4) is our new emergence pattern formula which allocates the ultimate loss with an arbitrary distribution to the best estimate of the ultimate loss. We call (4.4) the multiplicative lognormal emergence pattern formula. As in the previous section, the emergence pattern formula (4.4) is a \( \sigma(BE_1) \)-measurable random variable. In the sequel, \( BE_1^{\text{lognormal}_ep} \) is denoted by \( BE_1 \), since we only study the unconditional distribution of \( BE_1^{\text{lognormal}_ep} \). The multiplicative lognormal emergence pattern can be seen as a version of the classical emergence pattern formula (2.1) where we allocate \( X_n \) to \( BE_1 \) with a random scaling factor in order to have a non-degenerate distribution of \( BE_1|X_n=x \). If we choose a lognormal distribution for \( X_n \) in (4.4), then we are in the framework of the Hertig’s model and the results from Section 3.2 apply.

**Theorem 4.2.** Let us assume that the ultimate loss is modelled with \( X_n \sim F_{X_n} \). Let \( \mu_{X_n} = \mathbb{E}[X_n] \), \( \psi_{X_n} = \frac{SD[X_n]}{\mu_{X_n}} \) and choose an emergence factor \( \alpha \in (0, 1) \). We assume that there exists a unique solution \( \tilde{\alpha} \in (0, 1) \) to the equation

\[
\frac{\mathbb{E}[X_n^{2\tilde{\alpha}^2}]}{(\mathbb{E}[X_n^{\tilde{\alpha}^2}])^2} \left( 1 + \psi_{X_n}^2 \right)^\tilde{\alpha}^2(1-\tilde{\alpha}^2) = 1 + \alpha^2 \psi_{X_n}^2. \tag{4.5}
\]

We set

\[
\tilde{m} = \frac{\log(\mu_{X_n}) - \log(\mathbb{E}[X_n^{\tilde{\alpha}^2}]) - (1 - \tilde{\alpha}^4)^{\tilde{s}^2}}{1 - \tilde{\alpha}^2},
\]

\[
\tilde{s}^2 = \log(1 + \psi_{X_n}^2).
\]

The emergence pattern of the ultimate loss \( X_n \) is described with the formula (4.4). We have the following properties:
1. \( \mathbb{E}[BE_1] = \mathbb{E}[X_n] \) and \( \text{Var}(BE_1) = \alpha^2 \text{Var}[X_n] < \text{Var}[X_n] \).

2. If \( X_n \) has a light-tailed distribution, then \( BE_1 \) has a subexponential distribution with all moments finite. Moreover, \( \text{VaR}_\gamma[BE_1 - \mathbb{E}[BE_1]] > \text{VaR}_\gamma[X_n - \mathbb{E}[X_n]] \) for \( \gamma > \gamma^* \), with some \( \gamma^* < 1 \).

3. If \( X_n \) has a subexponential distribution with all moments finite such that \( \limsup_{x \to \infty} 1 - \frac{1 - F_{X_n}(vx)}{1 - F_{X_n}(x)} < 1 \) for some \( v > 1 \) and \( X_n^{\alpha^2} \) is light-tailed or subexponential, then \( BE_1 \) has a subexponential distribution with all moments finite.

4. If \( X_n \) has a Pareto-type distribution with tail index \( \theta \) and regularly varying function \( L \) such that \( L(x^{\frac{\alpha^2}{\alpha^2}}) \sim C \cdot L(x) \) as \( x \to \infty \), then \( BE_1 \) has a Pareto-type distribution with tail index \( \frac{\theta}{\alpha^2} > \theta \), and

\[
\lim_{\gamma \to 1} \frac{\text{VaR}_\gamma[BE_1 - \mathbb{E}[BE_1]]}{\text{VaR}_\gamma[X_n - \mathbb{E}[X_n]]} = 0.
\]

**Remark 4.1.** Point 3 is satisfied for example for lognormal, Weibull, Benktander type I and type II distributions, see for example, the remark before Theorem 2.1 in Tang (2006). Point 4 is satisfied for example if \( L(x) \sim a \log(x)^b \), \( x \to \infty \), which is the case for Pareto, Burr and log-gamma distributions.

**Lemma 4.1.** Let us consider the equation (4.5) with \( \alpha \in (0, 1) \).

1. There exists at least one solution \( \tilde{\alpha} \in (0, 1) \) to (4.5).

2. There exists a unique solution \( \tilde{\alpha} \in (0, 1) \) to (4.5) for Gamma, Weibull, Pareto (for shape parameter greater than 2.15) and log-gamma (for shape parameter greater than 2.15) distributions of \( X_n \).

**Remark 4.2.** For Pareto and log-gamma distributions with shape parameters close to 2, the equation (4.5) may have several solutions. Low values of shape parameter are unlikely to be used in practice in our model since the emergence factor \( \alpha \) is based on variance and variance of Pareto or log-gamma explodes when the shape parameter approaches 2.

We have found another interesting property of the one-year premium risk versus the ultimate premium risk: the distribution of the true one-year premium risk can have a different tail behaviour than the distribution of the ultimate premium risk. If the tail of the one-year risk is heavier than the tail of the ultimate risk, as in point 2 from Theorem 4.2, then we expect that the one-year risk is higher than the ultimate risk at all high confidence levels, and we have already observed such a property in the previous section. If the tail of the one-year risk is lighter than the tail of the ultimate risk, as in point 4 from Theorem 4.2, then we expect that the one-year risk is lower than the ultimate risk at all high confidence levels.

**Example 4.2.** We consider \( X_n \sim \text{Weibull}(\tau) \). We want to illustrate point 2 from Theorem 4.2, hence we are interested in \( \tau \geq 1 \). In order to calculate \( \text{VaR}_\gamma[BE_1] \),

\[
\text{VaR}_\gamma[BE_1] = \mathbb{E}[BE_1] - \lambda \text{Var}(BE_1) \leq \mathbb{E}[X_n] - \lambda \text{Var}(X_n),
\]
we use Monte Carlo simulations and we generate a sample of size $10^6$. The results are presented in Figure 4. For $\tau = 10$ and $\alpha = 85\%$, the one-year risk is higher than the ultimate risk at $\gamma = 99.5\%$ by about 0.5\%, and at $\gamma = 99.99\%$ by about 10\%. If the emergence pattern formula (2.1) is applied in this claims development model to measure the one-year premium risk, then the true one-year risk is underestimated, respectively, by about 15\% and 23\% at $\gamma = 99.5\%$ and 99.99\%. For $\alpha = 85\%$, the one-year risk is higher than the ultimate risk at the confidence level $\gamma = 99.5\%$, and also at all confidence levels $\gamma > 99.5\%$, if we consider Weibull distributions with $\tau > 7.5$. As discussed in Example 4.2, negatively skewed Weibull distributions are not common in applications but they should not be excluded a priori from considerations.

We illustrate consequences of point 4 from Theorem 4.2. We consider $X_n \sim \text{Pareto}(1, \theta)$ with $1 - F_{X_n}(x) = x^{-\theta}$, $x > 1$. Then, log ($X_n$) has exponential distribution with parameter $\theta$ and we can use the results from Example 4.1 to calculate $\text{VaR}_\gamma[BE_1]$. The results are presented in Figure 5. For $\theta = 5$ and $\alpha = 50\%$, the one-year risk is lower than the ultimate risk at $\gamma = 99.5\%$ by about 68\% and at $\gamma = 99.99\%$ by about 81\%. If the emergence pattern formula (2.1) is applied in this claims development model to measure the one-year premium risk, then the true one-year risk is overestimated, respectively, by about 58\% and 165\% at $\gamma = 99.5\%$ and 99.99\%. We can also observe that the one-year risk can be higher than the ultimate risk at low confidence levels. Finally, even though point 4 from Theorem 4.2 shows that $\text{VaR}_\gamma[BE_1 - E[BE_1]] = o(\text{VaR}_\gamma[BE_1 - E[BE_1]])$ for $\gamma \to 1$, we end up with $\text{VaR}_\gamma[BE_1 - E[BE_1]] \approx \alpha^3 \text{VaR}_\gamma[BE_1 - E[BE_1]]$ for $\gamma$ in the range 99.99–99.999\% in our experiments.

□
Theorems 4.1 and 4.2 show that the common belief among actuaries that the one-year risk is always lower than the ultimate risk does not always hold to be true. We point out that the relation between the one-year risk and the ultimate risk also depends on the risk measure, not only on the claims development process. If standard deviation is used to measure the risk, then point 1 in Theorems 3.2, 3.4, 3.6, 4.1 and 4.2 shows that the true one-year premium risk is always lower than the ultimate premium risk in all our claims development models (in Theorems 3.2, 3.4 and 3.6, we prove that $\alpha \in (0, 1)$ and, consequently, we assume that $\alpha \in (0, 1)$ in Theorems 4.1 and 4.2). This conclusion agrees with the results from Wüthrich and Merz (2015) where the authors use mean square error of prediction to compare the true one-year reserve risk with the ultimate reserve risk. However, in this paper, we focus on Value-at-Risk. If Value-at-Risk is used to measure the risk, then the relation between the one-year risk and the ultimate risk depends on the confidence level and the claims development process. In many claims development models, including the models presented in Section 3 which are commonly used in claims reserving, if a sufficiently high confidence level for Value-at-Risk is chosen, then the true one-year premium risk is lower than the ultimate premium risk, which agrees with the common belief. However, there are claims development models, and examples of such model are discussed in Sections 4.1 and 4.2, where the true one-year premium risk can be higher than the ultimate premium risk at all high confidence levels. Actuaries should be aware of this property in order not to misunderstand and underestimate the one-year risk. On the other side, the true one-year premium risk can be much less dangerous than it is believed by users of the linear emergence pattern formula (2.1). Point 4 from Theorem 4.2 proves that, in a claims development model with the emergence pattern of the
ultimate loss given by (4.4), the tail index of the Pareto-type distribution of the true one-year risk is higher than the tail index of the Pareto-type distribution of the ultimate risk, whereas from Theorem 2.1 we know that the tail index of the Pareto-type distribution of the one-year risk predicted with (2.1) remains equal to the tail index of the distribution of the ultimate loss.

5. Proofs

By $f(x) \sim Cg(x)$ we mean the limit $\lim_{\lambda \to \infty} \frac{f(x)}{g(x)} = C$, by $f(x) = o(g(x))$ we mean the limit $\lim_{\lambda \to \infty} \frac{f(x)}{g(x)} = 0$, unless different limits are specified.

First, we collect some results which we use in the proofs below.

**Lemma 5.1.** We consider two independent r.v. $Z_1$ and $Z_2$. Let $\mathcal{S}$ denote the class of subexponential distributions, and $\mathcal{P}(\theta)$ denotes the class of Pareto-type distributions with tail index $\theta$.

1. If $Z_1$ and $Z_2$ are light-tailed, then $aZ_1 + bZ_2$ is light-tailed, for any $a, b \in \mathbb{R}$.
2. If $Z_1 \in \mathcal{S}$ and $1 - F_{Z_1}(x) = o(1 - F_{Z_1}(x))$, then $Z_1 + Z_2 \in \mathcal{S}$ and $1 - F_{Z_1 + Z_2}(x) \sim 1 - F_{Z_1}(x)$ [Lemma 9 in Geluk and De Vries (2006)].
3. If $Z_1 \in \mathcal{S}$ and $Z_2$ is a positive, bounded r.v., then $Z_1 \cdot Z_2 \in \mathcal{S}$ [Corollary 2.5 in Cline and Samorodnitsky (1994)]. If $Z_1 \in \mathcal{S}$ such that $\limsup_{\lambda \to \infty} \frac{1 - F_{Z_1}(\lambda v)}{1 - F_{Z_1}(\lambda)} < 1$ for some $v < 1$ and $Z_2$ is a positive r.v. such that $1 - F_{Z_2}(\lambda v) = o(1 - F_{Z_2}(\lambda))$ for some $v > 0$, or $1 - F_{Z_2}(\lambda v) = o(1 - F_{Z_2}(\lambda))$ for some $v > 1$, then $Z_1 \cdot Z_2 \in \mathcal{S}$ [Theorem 2.1 and Corollary 2.1.2 in Tang (2006)].
4. If $Z_1 \in \mathcal{P}(\theta)$ is a positive r.v. and $Z_2$ is a positive r.v. with a light-tailed distribution, then $Z_1 \cdot Z_2 \in \mathcal{P}(\theta)$ and $1 - F_{Z_1 \cdot Z_2}(x) \sim \mathbb{E}[Z_2^\theta](1 - F_{Z_1}(x))$ [Proposition 1.3.9.b in Mikosch (1999)].
5. If $Z_1, Z_2$ have all moments finite, then $Z_1 + Z_2$ and $Z_1 \cdot Z_2$ have all moments finite.

The proof of Theorem 2.1. Points 1 and 4 are obvious. Points 2 and 3 follow from Lemma 5.1 if we choose a degenerate distribution for $Z_2$.

The proofs of Theorems 3.1, 3.3 and 3.5. We use the probabilistic assumptions behind the claims development processes, derive the distributions and use appropriate parameterisations.

**Gaussian ILR.** By Result 4.6 in Johnson and Wichern (2007), if $(Z_1, Z_2)$ has a bivariate normal distribution with $Z_1 \sim \mathcal{N}(\mu_{Z_1}, \sigma_{Z_1}^2)$, $Z_2 \sim \mathcal{N}(\mu_{Z_2}, \sigma_{Z_2}^2)$ and zero correlation, then

$$Z_1 | Z_1 + Z_2 = z \sim \mathcal{N}\left(\frac{\sigma_{Z_2}^2}{\sigma_{Z_1}^2 + \sigma_{Z_2}^2}(z - \mu_{Z_2}); \frac{\sigma_{Z_2}^2}{\sigma_{Z_1}^2 + \sigma_{Z_2}^2}\right).$$ (5.1)
We set $Z_1 = X_1 = \epsilon_1 \sim N(\mu_1, \sigma^2_1)$, $Z_2 = X_n - X_1 = \sum_{i=2}^n \epsilon_i \sim N(\mu - \mu_1, E(\sigma^2 - \sigma^2_1))$, $Z = Z_1 + Z_2 = X_n \sim N(E\mu, E\sigma^2)$. The pair $(Z_1, Z_2)$ has a bivariate normal distribution with zero correlation. We apply (5.1) and get the conditional distribution:

$$X_1|X_n = x \sim N \left( E\mu_1 + \frac{\sigma^2_1}{\sigma^2}(x - E\mu); E\frac{\sigma^2_1(\sigma^2 - \sigma^2_1)}{\sigma^2} \right).$$

(5.2)

Since $BE_1 = X_1 + E(\mu - \mu_1)$, we derive our key conditional distribution:

$$BE_1|X_n = x \sim N \left( \frac{\sigma^2_1}{\sigma^2}x + \left(1 - \frac{\sigma^2_1}{\sigma^2}\right)E\mu; E\frac{\sigma^2_1(\sigma^2 - \sigma^2_1)}{\sigma^2} \right).$$

(5.3)

We note that $X_n \sim N(E\mu, E\sigma^2)$ and $BE_1 \sim N(E\mu, E\sigma^2_1)$. We set

$$\mu_{x_n} = E[X_n] = E\mu, \quad \sigma^2_{x_n} = Var[X_n] = E\sigma^2, \quad \alpha^2 = \frac{Var[BE_1]}{Var[X_n]} = \frac{\sigma^2_1}{\sigma^2},$$

(5.4)

and we can now prove the formulas for the distributions of $X_n$, $BE_1$ and $BE_1|X_n$ by using the new parameters from (5.4). Since $0 < \sigma^2_1 < \sigma^2$, we conclude that $\alpha \in (0, 1)$.

**Hertig’s model.** The development process (3.7) satisfies the following dynamics:

$$\log (X_i) = \log (X_{i-1}) + \log (\epsilon_i) = \log (X_{i-1}) + \eta_i,$$

where $\eta_i \sim N(m_i, s_i^2), \quad i = 2, \ldots, n,$

which fits the Gaussian ILR model. By (5.2), we state the conditional distribution:

$$\log (X_1)| \log (X_n) = x \sim N \left( m_1 + \frac{s_i^2}{s^2}(x - m); \frac{s_i^2(s^2 - s_i^2)}{s^2} \right),$$

and, using the definition of $BE_1$, see (3.8), we can derive our key conditional distribution:

$$BE_1|X_n = x \sim LogN \left( \frac{s_i^2}{s^2} \log (x) + \left(1 - \frac{s_i^2}{s^2}\right)(m + \frac{s^2}{2}); \frac{s_i^2(s^2 - s_i^2)}{s^2} \right).$$

(5.5)

We deduce that $X_n \sim LogN(m, s^2)$ and $BE_1 \sim LogN(m + \frac{1}{2}(s^2 - s_i^2), s_i^2)$. We can set

$$\mu_{x_n} = E[X_n] = e^{\mu + \frac{1}{2}s^2}, \quad \psi_{x_n} = \frac{Var[X_n]}{\left(E[X_n]\right)^2} = e^{\psi^2} - 1, \quad \alpha^2 = \frac{Var[BE_1]}{Var[X_n]} = \frac{e^{\psi^2} - 1}{e^{\alpha^2} - 1},$$

(5.6)
and we solve (5.6) to represent the parameters:

\[ s^2 = \log (1 + \psi_{X_n}^2), \quad s_1^2 = \log (1 + \alpha^2\psi_{X_n}^2), \quad m = \log (\mu_{X_n}) - \frac{1}{2} \log (1 + \psi_{X_n}^2). \]

(5.7)

Finally, we introduce three new parameters: \( \tilde{s} = s, \tilde{m} = m \) and \( \tilde{\alpha}^2 = \frac{s_1^2}{s^2} \). We can now deduce the formulas for the distributions of \( X_n, BE_1 \) and \( BE_1|X_n \). Since \( 0 < s_1^2 < s^2 \), we conclude that \( \alpha \in (0, 1) \).

**Over-Dispersed Poisson:** It is a known result that if \( Z_1 \sim \text{Poiss}(\mu_{Z_1}) \), \( Z_2 \sim \text{Poiss}(\mu_{Z_2}) \) and \( Z_1 \) is independent of \( Z_2 \), then

\[ Z_1|Z_1 + Z_2 = z \sim \text{Bin} \left( z, \frac{\mu_{Z_1}}{\mu_{Z_1} + \mu_{Z_2}} \right). \]  

(5.8)

Using the definitions of the Over-Dispersed Poisson distribution and the best estimate of the ultimate loss in this model, we can deduce the conditional distributions:

\[ X_1|X_n = x \sim \psi \cdot \text{Bin} \left( \frac{x}{\psi}; \omega_1 \right), \quad BE_1|X_n = x \sim \psi \cdot \text{Bin} \left( \frac{x}{\psi}; \omega_1 \right) + \mu(1 - \omega_1). \]  

(5.9)

Moreover, \( X_n \sim \psi \text{Poiss}(\mu/\psi) \) and \( BE_1 \sim \psi \text{Poiss}(\mu\omega_1/\psi) + \mu(1 - \omega_1) \). We set

\[ \mu_{X_n} = \mathbb{E}[X_n] = \mu, \quad \psi_{X_n} = \frac{\text{Var}[X_n]}{\mathbb{E}[X_n]} = \psi, \quad \alpha^2 = \frac{\text{Var}[BE_1]}{\text{Var}[X_n]} = \omega_1. \]

(5.10)

The formulas for the distributions \( X_n, BE_1 \) and \( BE_1|X_n \) can now be established by using the new parameters from (5.10). Since \( 0 < \omega_1 < 1, \alpha \in (0, 1) \).  

**The proofs of Theorems 3.2, 3.4 and 3.6.**

**Point 1.** The results can be deduced from the distributions derived in Theorems 3.1, 3.3 and 3.5 and the property that \( \alpha \in (0, 1) \).

**Point 2 from Theorem 3.2.** We use the well-known formula for the quantile of normal distribution together with \( \alpha \in (0, 1) \) and \( \text{VaR}_Y[X_n] > \mathbb{E}[X_n] \).

**Points 2 and 3 from Theorem 3.4.** We use the distribution (3.11) and the formulas

\[ \text{VaR}_Y[X_n] = e^{\bar{\gamma} + \frac{\alpha^2}{2} \Phi^{-1}(\gamma)}, \quad \text{VaR}_Y[BE_1] = e^{\bar{\gamma} + (1 - \tilde{\alpha}^2)\frac{\alpha^2}{2} \Phi^{-1}(\gamma)}, \]

(5.11)

which describe the quantiles of the lognormal distributions of \( BE_1 \) and \( X_n \). We can now solve the inequality \( \text{VaR}_Y[BE_1] < \text{VaR}_Y[X_n] \), which is equivalent to \( \text{VaR}_Y[BE_1 - \mathbb{E}[BE_1]] < \text{VaR}_Y[X_n - \mathbb{E}[X_n]] \), and find \( \gamma^* \). The limit \( \lim_{\gamma \to 1} \frac{\text{VaR}_Y[BE_1 - \mathbb{E}[BE_1]]}{\text{VaR}_Y[X_n - \mathbb{E}[X_n]]} \) can be directly calculated using the formula (5.11).

**Point 2 from Theorem 3.6.** Let \( Y \sim \text{Poiss}(\alpha^2\mu) \) and \( X \sim \text{Poiss}(\mu) \). It is known that \( F_Y(x) \leq F_Y(x) \) for all \( x \), see for example, the remark at the end of
Section 1.2 in Klenke and Mattner (2010). Let us choose arbitrary $k > 0$. Our first goal is to show that $F_X(x + k) < F_Y(x)$ for all sufficiently high $x$. We only consider $x > \mu$. By Theorem 2 from Short (2013), we have the bounds:

$$\Phi\left(\sqrt{2H(\mu, x)}\right) < F_X(x) < \Phi\left(\sqrt{2H(\mu, x + 1)}\right),$$

where $H(\mu, x) = \mu - x + x \log \left(\frac{x}{\mu}\right)$. Consequently, our assertion holds if $H(\alpha^2 \mu, x) > H(\mu, x + k + 1)$ for all sufficiently high $x$. Using the definition of $H$, we have to investigate the inequality:

$$x \log \left(\frac{x}{x + k + 1}\right) + k + 1 - (k + 1) \log (x + k + 1) - x \log (\alpha^2) > \mu(1 - \alpha^2) - (k + 1) \log \mu. \quad (5.12)$$

We can prove that $\lim_{x \to +\infty} \left( x \log \left(\frac{x}{x + k + 1}\right) + k + 1 \right) = 0$. Moreover, $\lim_{x \to +\infty} \left( -(k + 1) \log (x + k + 1) - x \log (\alpha^2) \right) = +\infty$, since $\alpha \in (0, 1)$. Hence, the inequality (5.12) is satisfied for all sufficiently high $x$.

Let us consider our claims development model. For simplicity we assume that $\psi_{X_n} = 1$. Let $Y = BE_1 - (1 - \alpha^2)\mu$. By the arguments from above, we know that $F_Y(x) > F_{X_n}(x + (1 - \alpha^2)\mu)$ for all sufficiently high $x$. Since the distribution $F_{BE_1}$ is the distribution $F_Y$ shifted to the right by $(1 - \alpha^2)\mu$, we deduce that $F_{BE_1}(x) > F_{X_n}(x)$ for all sufficiently high $x$. The result is proved and can be extended to arbitrary $\psi_{X_n}$.

**Point 3 from Theorem 3.6.** By Lemma 9.1 from Steutel and Van Harn (2004), if $X \sim \text{Poiss}(\mu)$, then $-\log (1 - F_X(x)) \sim x \log (x)$. Consequently, we can derive the limit:

$$\lim_{t \to \infty} \frac{\log (1 - F_{BE_1}(tx))}{\log (1 - F_{X_n}(t))} = x, \quad x > 0. \quad (5.13)$$

We use the arguments from Propositions 2.2 and 2.6.vi from Resnick (2007). Since the limit (5.13) holds for a sequence of functions, the sequence of inverse functions also converges. We can establish the limit:

$$\lim_{t \to \infty} \frac{F_{BE_1}^{-1}\left(1 - (1 - F_{X_n}(t))^\nu\right)}{t} = y, \quad y > 0. \quad (5.14)$$

By the property that $F_{X_n}(F_{X_n}^{-1}(t)) \sim t, t \to 1$, we end up with

$$\lim_{t \to \infty} \frac{F_{BE_1}^{-1}(1 - 1/t^\nu)}{F_{X_n}^{-1}(1 - 1/t)} = y, \quad y > 0. \quad (5.15)$$

Finally, we choose $\nu = 1$ in (5.15). We conclude that $VaR_{\gamma}[BE_1 - \mathbb{E}[BE_1]] \sim VaR_{\gamma}[X_n - \mathbb{E}[X_n]]$ for $\gamma \to 1$. □
The proof Theorem 4.1. Point 1 is obvious.

Point 2. The first statement follows from Lemma 5.1. The result for light-tailed distributions is obvious. We consider a subexponential distribution with all moments finite. We know that \(Z_1 = \alpha^2 X_n + (1 - \alpha^2)\mu X_n\) is subexponential with all moments finite. The random variable \(Z_2 = \sqrt{\alpha^2(1 - \alpha^2)}\sigma X_n \xi\) has normal distribution and is light-tailed. We observe that

\[
\lim_{x \to \infty} \frac{1 - F_{Z_2}(x)}{1 - F_{Z_1}(x)} = \lim_{x \to \infty} e^{tx}(1 - F_{Z_2}(x)) = 0, \quad \text{for some } t > 0, \quad (5.16)
\]

since the numerator converges to zero \((Z_2 \text{ is light-tailed})\) and the denominator converges to infinity \((Z_1 \text{ is subexponential})\). We deduce that \(BE_1 = Z_1 + Z_2\) has a subexponential distribution with all moments finite.

We prove the second statement. We can derive the lower bound for the convolution:

\[
1 - F_{BE_1}(x) = \Pr(\alpha^2 X_n + (1 - \alpha^2)\mu X_n + Z_2 > x) \\
\geq \Pr((1 - \alpha^2)\mu X_n + Z_2 > x, \alpha^2 X_n \leq x) \\
= \Pr((1 - \alpha^2)\mu X_n + Z_2 > x) F_{X_n}(\frac{x}{\alpha^2}), \quad x > 0.
\]

Using a bound for the tail of normal distribution, see Gordon (1941), the assumption that \(\lim_{x \to \infty} (1 - F_{X_n}(x))e^{w^2} = 0\) and \(\lim_{x \to \infty} e^{wx} = +\infty\) for all \(w > 0, v > 0\), we can show that

\[
\lim_{x \to \infty} \frac{1 - F_{BE_1}(x)}{1 - F_{X_n}(x)} \geq \lim_{x \to \infty} \frac{\Pr((1 - \alpha^2)\mu X_n + Z_2 > x) F_{X_n}(\frac{x}{\alpha^2})}{1 - F_{X_n}(x)} \\
\geq \frac{1}{\sqrt{2\pi}} \lim_{x \to \infty} \frac{e^{-\frac{(ux)^2}{2}} F_{X_n}(\frac{x}{\alpha^2})}{(1 - F_{X_n}(x))} \\
\geq K \lim_{x \to \infty} \frac{e^{wx}(u(x) + \frac{1}{u(x))})(1 - F_{X_n}(x))}{e^{wx}(u(x) + \frac{1}{u(x))})(1 - F_{X_n}(x))} = +\infty,
\]

where \(u(x) = \frac{x - (1 - \alpha^2)\mu X_n}{\sqrt{\alpha^2(1 - \alpha^2)e^{x^2}}}\) and \(K, v, w\) denote some positive constants. We can deduce that \(1 - F_{BE_1}(x) > 1 - F_{X_n}(x)\) for all sufficiently high \(x\), and \(VaR_p[B\xi] > VaR_p[X_n]\) for all sufficiently high \(\gamma\).

Point 3. \(Z_1\) and \(Z_2\) are still defined as in point 1. By Lemma 5.1 and similar arguments as in (5.16), \(1 - F_{BE_1}(x) \sim 1 - F_{Z_1}(x) \sim \alpha^{2\theta}(1 - F_{X_n}(x))\), and \(1 - F_{Z_2}(x) = o(1 - F_{Z_1}(x))\). We can deduce that \(BE_1\) inherits the tail from \(X_n\). Using Lemma 5.1 again, we prove that \(1 - F_{BE_1 - \mu}(x) \sim \alpha^{2\theta}(1 - F_{X_n - \mu}(x))\). The result for VaR follows from the arguments used in Proposition 2.6.vi from Resnick (2007).
The proof Theorem 4.2. Points 1. Using independence between $X_n$ and $\xi \sim N(0, 1)$, as well as moments for log-normal distribution, we can calculate the moments:

$$\mathbb{E}[BE^k] = \mathbb{E}[(X_n)^{\delta^2}]e^{k(1-\delta^2)(\bar{m} + \frac{\xi^2}{2}) + \frac{k^2}{2}(1-\delta^2)m^2}, \quad k = 1, 2. \quad (5.17)$$

We require that $\mathbb{E}[BE_1] = \mathbb{E}[X_n]$ and $Var[BE_1] = \alpha^2 Var[X_n]$. If we use (5.17), then we derive the equation:

$$\mathbb{E}[(X_n)^{\delta^2}]_\delta e^{\delta^2(1-\delta^2)m^2} = 1 + \alpha^2 \psi_{X_n}^2. \quad (5.18)$$

We can see that we can choose any $\delta^2 > 0$ and look for $\bar{a}$ which satisfies (5.18). The most obvious choice is $\delta^2$ from Theorem 3.3 so that our new emergence pattern formula coincides with the true emergence pattern formula in the Hertig’s model if $X_n$ has lognormal distribution. The parameter $\bar{m}$ is next derived from $\mathbb{E}[BE_1] = \mathbb{E}[X_n]$.

Point 2. Since $X_n$ has a moment generating function and $\bar{a} \in (0, 1)$, then $Z_1 = X_n^{\delta^2}$ has a moment generating function and is light-tailed. The lognormal distribution $Z_2 = e^{(1-\delta^2)^2(\bar{m} + \frac{\xi^2}{2}) + \sqrt{\delta^2(1-\delta^2)m^2}}$ is subexponential and satisfies \(\lim_{x \to \infty} \frac{1-F_{Z_2}(vx)}{1-F_{Z_2}(x)} < 1\) for some $v > 1$, see the remark before Theorem 2.1 in Tang (2006). Moreover, we observe that

$$\lim_{x \to \infty} \frac{1-F_{Z_2}(vx)}{1-F_{Z_2}(x)} = \lim_{x \to \infty} \frac{e^{tv}(1-F_{Z_2}(vx))}{e^{t(x-1)}e^{vx}(1-F_{Z_2}(x))} = 0,$$

for some $t > 0$, for all $v \geq 1$, \((5.19)\)

since the numerator converges to zero and the denominator converges to infinity. By Lemma 5.1, $BE_1 = Z_1Z_2$ is subexponential with all moments finite. Finally, we prove that

$$\lim_{x \to \infty} \frac{1-F_{X_n}(x)}{1-F_{BE_1}(x)} = \lim_{x \to \infty} \frac{e^{tx}(1-F_{X_n}(x))}{e^{tx}(1-F_{BE_1}(x))} = 0, \quad \text{for some} \ t > 0,$$

since the numerator converges to zero and the denominator converges to infinity. We can deduce that $1 - F_{BE_1}(x) > 1 - F_{X_n}(x)$ for all sufficiently high $x$, and $VaR_x[BE_1] > VaR_x[X_n]$ for all sufficiently high $y$.

Point 3. If the distribution of $X_n$ satisfies the assumption that $\lim_{x \to \infty} \frac{1-F_{X_n}(vx)}{1-F_{X_n}(x)} < 1$ for some $v > 1$, then the distribution of $X_n^{\delta^2}$, given by $F_{X_n^{\delta^2}}(x) = F_{X_n}(x^{\frac{1}{\delta^2}})$, also satisfies this assumption. If $X_n^{\delta^2}$ has a light-tailed distribution, then the result follows from point 2. Let $X_n^{\delta^2}$ be subexponential.
We recall that we put \( Z_2 = e^{(1-\tilde{\alpha}^2)(\tilde{\alpha}^2 + \sqrt{\tilde{\alpha}^2(1-\tilde{\alpha}^2)}E)\xi} \). We show that

\[
\lim_{x \to \infty} \frac{1 - F_{Z_2}(vx)}{1 - F_{Z_2}(x)} \leq \lim_{x \to \infty} \frac{(\log(x) - a)^2 + b^2}{(\log(vx) - a)(\log(x) - a)} \left( e^{\frac{-\log(vx)-a}{2\tilde{\alpha}^2}} + \sqrt{\tilde{\alpha}^2(1-\tilde{\alpha}^2)}} \right) = 0, \quad \text{for all } v > 1,
\]

where \( a, b \) are the parameters of \( Z_2 \), and we use bounds for the tail of normal distribution, see Gordon (1941). The statement now follows from Lemma 5.1.

**Point 4.** We can immediately prove the relation:

\[
1 - F_{X_n^{\tilde{\alpha}^2}}(x) = L(x^{\frac{1}{\tilde{\alpha}^2}})x^{-\frac{\theta}{\tilde{\alpha}^2}} \sim CL(x)x^{-\frac{\theta}{\tilde{\alpha}^2}}.
\]

By Lemma 5.1, we derive the asymptotic formula:

\[
1 - F_{BE_1}(x) \sim CE\left[Z_2^{\frac{\theta}{\tilde{\alpha}^2}}\right]L(x)x^{-\frac{\theta}{\tilde{\alpha}^2}}. \tag{5.20}
\]

Since \( X_n \) has a Pareto-type distribution \( 1 - F(x) = L(x)x^{-\theta} \), we have the asymptotic formula:

\[
\text{VaR}_\gamma[X_n] \sim (1 - \gamma)^{-1/\theta} \tilde{L}\left(\frac{1}{1-\gamma}\right), \quad \gamma \to 1, \tag{5.21}
\]

where \( \tilde{L} \) denotes the de Bruyn conjugate of \( L \), see (2.7) and Proposition 2.5 from Berlaint et al. (2004). Using (5.20)–(5.21) and the arguments from Proposition 2.6.vi from Resnick (2007), we can establish that

\[
\text{VaR}_\gamma[BE_1] \sim \left(CE\left[Z_1^{\frac{\theta}{\tilde{\alpha}^2}}\right]\right)^{\frac{\theta^2}{\tilde{\alpha}^2}} (1 - \gamma)^{-\frac{\theta^2}{\tilde{\alpha}^2}} \tilde{L}\left(\frac{1}{1-\gamma}\right)
\]

\[
\sim \left(CE\left[Z_1^{\frac{\theta}{\tilde{\alpha}^2}}\right]\right)^{\frac{\theta^2}{\tilde{\alpha}^2}} (1 - \gamma)^{\frac{\theta^2}{\tilde{\alpha}^2}} \text{VaR}_\gamma[X_n], \quad \gamma \to 1,
\]

and the result is proved. □

**The proof of Lemma 4.1.** Let \( k = \tilde{\alpha}^2 \). We simply denote \( X = X_n, \psi = \psi_{X_n} \). We choose \( \alpha \in (0, 1) \). We introduce the functions:

\[
f(k) = \mathbb{E}[X^k] = \mathbb{E}[e^{k\log(X)}], \quad h(k) = \log f(2k) - 2 \log f(k) + k(1 - k) \log (1 + \psi^2).
\]

It is easy to show that \( h(k) \) is continuous on \((0, 1)\). We look for \( k \in (0, 1) \) such that

\[
h(k) = \log (1 + \alpha^2 \psi^2).
\]
We have \( h(0) = 0 \) and \( h(1) = \log (1 + \psi^2) > \log (1 + a^2 \psi^2) \). It is known that \( k \mapsto f(k) \) is log-convex. Hence
\[
\frac{d}{dx} h(k) = 2(\frac{d}{dx} \log f(x)|_{x=2k} - \frac{d}{dx} \log f(x)|_{x=k}) + (1 - 2k) \log (1 + \psi^2) > 0, \quad \text{and} \quad k \mapsto h(k) \text{ is increasing on } (0, \frac{1}{2}). \]
We conclude that there exists at least one solution to (5.22).

\section*{Gamma distribution.}

Let \( X \sim \text{Gamma}(a, b) \). By straightforward calculations, we show that
\[
f(k) = \mathbb{E}[X^k] = b^k \frac{\Gamma(k + a)}{\Gamma(a)}, \quad \frac{d^2}{dk^2} \log f(k) = \frac{d^2}{dk^2} \log \Gamma(k + a),
\]
and
\[
\frac{d^2}{dk^2} h(k) = 4 \frac{d^2}{dx^2} \log \Gamma(x + a)|_{x=2k} - 2 \frac{d^2}{dx^2} \log \Gamma(x + a)|_{x=k} - 2 \log (1 + \psi^2)
\]
\[
= \frac{d^2}{dk^2} \log \left( \frac{\Gamma(2k + a)}{(\Gamma(k + a))^2} \right) - 2 \log (1 + \psi^2) = \frac{d^2}{dk^2} g(k) - 2 \log (1 + \psi^2),
\]
where we define
\[
g(k) = \log \left( \frac{\Gamma(2k + a)}{(\Gamma(k + a))^2} \right).
\]

Let \( a \geq \frac{1}{2} \). By Theorem 5 from Qi et al. (2008), the function \( g \) is 3-log-concave, which means that \( \frac{d^3}{dx^3} \log g(k) \leq 0 \). We can conclude that the function \( k \mapsto \frac{d}{dk} \log g(k) \) is decreasing. We have to consider three cases: (1) \( \frac{d}{dk} g(k) > 2 \log (1 + \psi^2) \) for all \( k \in (0, 1) \): the function \( k \mapsto h(k) \) is convex on \((0, 1)\); (2) \( \frac{d}{dk} g(k) < 2 \log (1 + \psi^2) \) for all \( k \in (0, 1) \): the function \( k \mapsto h(k) \) is concave on \((0, 1)\); (3) There exists \( k_0 \in (0, 1) \) such that \( \frac{d}{dk} g(k) > 2 \log (1 + \psi^2) \) for \( 0 < k < k_0 \) and \( \frac{d}{dk} g(k) < 2 \log (1 + \psi^2) \) for \( 1 > k > k_0 \): the function \( k \mapsto h(k) \) is convex on \((0, k_0)\) and concave on \((k_0, 1)\). In all three cases, the properties imply, together with the properties mentioned in the beginning, that there exists a unique solution to (5.22).

Let \( a < \frac{1}{2} \). From series representation of trigamma function, we deduce that the second derivative of log-gamma function is positive. By inequality (1.10) from Guo et al. (2015), which gives upper and lower bounds on the second derivative of log-gamma function, we get
\[
\frac{d^2}{dk^2} g(k) \leq \frac{2a + 3}{(2k + a + \frac{1}{2})(k + a + 1)} + \frac{2a^2 - 4k^2}{(2k + a)^2(k + a)^2}. \tag{5.23}
\]
Since \( k \mapsto h(k) \) is increasing on \((0, \frac{1}{2})\), we only have to consider \( k \in (\frac{1}{2}, 1) \). Using (5.23) and the assumption that \( k \in (\frac{1}{2}, 1) \), we derive the inequality
\[ \frac{d^2}{dk^2} h(k) = \frac{d^2}{dk^2} g(k) - 2 \log \left( 1 + \frac{1}{a} \right) \]

\[ \leq 2 \left( \frac{1}{a + \frac{1}{2}} - \log \left( 1 + \frac{1}{a} \right) + \frac{a^2 - 2k^2}{(2k + a)^2(k + a)^2} \right). \quad (5.24) \]

Let us introduce the function \( p(a) = \frac{1}{a + \frac{1}{2}} - \log \left( 1 + \frac{1}{a} \right) \). We have \( p(0^+) = -\infty \) and \( p\left(\frac{1}{2}\right) = \frac{1}{2} - \log 3 < 0 \). Moreover, by straightforward calculation, we can check that \( p'(a) > 0 \) for \( a \in (0, \frac{1}{2}) \). Consequently, the first two terms on the right-hand side of (5.24) are negative. It is clear that the last term in (5.24) is also negative. Hence, we have proved that the function \( k \mapsto h(k) \) is concave on \((\frac{1}{2}, 1)\). This property implies, together with the properties mentioned in the beginning, that there exists a unique solution to (5.22).

**Pareto distribution.** Let \( X \sim \text{Pareto}(x_0, \theta) \) with \( 1 - F(x) = \frac{x_0^\theta}{x - x_0^\theta} \) for \( x > x_0 \). We assume that \( \theta > 2 \) so that \( \text{Var}[X] < \infty \). By straightforward calculations, we can show that

\[ f(k) = \mathbb{E}[X^k] = \frac{\theta x_0^k}{\theta - k}, \quad \frac{d}{dk} \log f(k) = \log (x_0) + \frac{1}{\theta - k}. \]

We can also calculate that

\[ \frac{d}{dk} h(k) = \frac{2}{\theta - 2k} - \frac{2}{\theta - k} - (2k - 1) \log (1 + \psi^2). \]

Since \( \psi^2 = \frac{1}{\theta(\theta - 2)} \) and \( \log (1 + x) < x \), we can prove the inequality:

\[ \frac{d}{dk} h(k) > \frac{\theta^2 + \theta(6k^2 - 7k) + 2k^2 - 4k^3}{(\theta - 2)(\theta - 2k)(\theta - k)}. \quad (5.25) \]

The denominator of the fraction on the right-hand side of (5.25) is always positive and the numerator is a quadratic function in \( \theta \). The larger root of this quadratic function is given by \( \tilde{\theta}(k) = \frac{1}{2}k(\sqrt{36k^2 - 68k + 41} - 6k + 7) \). One can check that the maximal value of the function \( k \mapsto \tilde{\theta}(k) \) on \((0, 1)\) is equal to \( \theta^* \approx 2.15 \). Consequently, we can conclude that for \( \theta > \theta^* \) the first derivative (5.25) is positive for any \( k \in (0, 1) \), meaning that the function \( k \mapsto h(k) \) is increasing on \((0, 1)\). This property implies, together with the properties mentioned in the beginning, that there exists a unique solution to (5.22).

**Weibull distribution.** The proof is analogous as for Gamma distribution.

**Log-gamma distribution.** The proof is analogous as for Pareto distribution. \( \square \)

**Acknowledgments**

We would like thank the referees and the Editor-in-Chief for very valuable comments which we use to improve the paper. This research is financially supported with grant NCN 2018/31/B/HS4/02150.
REFERENCES

ŁUKASZ DELONG (Corresponding author)
SGH Warsaw School of Economics, Institute of Econometrics
Niepodległości 162, Warsaw 02-554, Poland
E-mail: lukasz.delong@sgh.waw.pl

MARcin SzATKOWSKi
SGH Warsaw School of Economics, Institute of Econometrics
Niepodległości 162, Warsaw 02-554, Poland and
Risk Department, STU ERGO Hestia SA
Hestii 1, Sopot 81-731, Poland
E-mail: marcin.szatkowski@ergohestia.pl