

Sharp Probability Tail Estimates for Portfolio Credit Risk

Jeffrey F. Collamore ^{1,*}, Hasitha de Silva ² and Anand N. Vidyashankar ³

¹ Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark

² Department of Mathematics, George Mason University, 4000 University Drive, Fairfax, VA 22030, USA

³ Department of Statistics, George Mason University, 4000 University Drive, Fairfax, VA 22030, USA

* Correspondence: collamore@math.ku.dk

Abstract: Portfolio credit risk is often concerned with the tail distribution of the total loss, defined to be the sum of default losses incurred from a collection of individual loans made out to the obligors. The default for an individual loan occurs when the assets of a company (or individual) fall below a certain threshold. These assets are typically modeled according to a factor model, thereby introducing a strong dependence both among the individual loans, and potentially also among the multivariate vector of common factors. In this paper, we derive sharp tail asymptotics under two regimes: (i) a *large loss regime*, where the total number of defaults increases asymptotically to infinity; and (ii) a *small default regime*, where the loss threshold for an individual loan is allowed to tend asymptotically to negative infinity. Extending beyond the well-studied Gaussian distributional assumptions, we establish that—while the thresholds in the large loss regime are characterized by idiosyncratic factors specific to the individual loans—the rate of decay is governed by the common factors. Conversely, in the small default regime, we establish that the tail of the loss distribution is governed by systemic factors. We also discuss estimates for Value-at-Risk, and observe that our results may be extended to cases where the number of factors diverges to infinity.

Keywords: large deviations; tail approximations; Value-at-Risk; Expected Shortfall; model uncertainty; multiple types; risk management

MSC: 60F10; 91-10; 91G40; 91G70



Citation: Collamore, Jeffrey F., Hasitha de Silva, and Anand N. Vidyashankar. 2022. Sharp Probability Tail Estimates for Portfolio Credit Risk. *Risks* 10: 239. <https://doi.org/10.3390/risks10120239>

Academic Editor: Anita Behme

Received: 27 October 2022

Accepted: 1 December 2022

Published: 14 December 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Banks, insurance companies, and other financial institutions frequently maintain and manage large credit portfolios. The main risk associated with such portfolios is that debtors may default on their obligations. Portfolio credit risk is concerned with large but rare loss events induced by defaults. Factor models are commonly used to represent the asset returns of a company, and defaults are said to occur when these assets fall below a certain threshold. From a statistical perspective, correlation between defaults of distinct companies is introduced by allowing the assets to share systemic factors, also referred to as latent factors or random effects. Specifically, if Y_i denotes the value of the i -th firm (as determined by its assets), then this dependence can be modeled, in the simplest case, by setting

$$Y_i = aZ + b\epsilon_i, \quad i = 1, \dots, n, \quad (1)$$

where $a^2 + b^2 = 1$, $\{\epsilon_i\}$ are independent and identically distributed (i.i.d.) random variables and independent of Z . A useful feature of (1) is that, conditioned on the latent variable Z , the assets $\{Y_i\}$ are independent. When Z and $\{\epsilon_i\}$ are Normally distributed, (1) reduces to the widely used single-factor Gaussian model, referred to as the Gaussian copula model

in the literature. More generally, assuming that there are k , possibly dependent, factors influencing the asset variable Y_i , an additive factor model takes the form

$$Y_i = \sum_{j=1}^k a_j Z_j + b_i \epsilon_i, \quad i = 1, \dots, n. \tag{2}$$

The i -th firm is then said to default if $Y_i < d$, where d represents the threshold; in other words, setting $X_i = \mathbf{1}_{\{Y_i < d\}}$, and letting $L_n := \sum_{i=1}^n U_i X_i$, where U_i is the loss incurred when the i -th loan defaults, then L_n denotes the *total loss* due to defaults, and our primary focus in this paper is to provide sharp tail approximations of $\mathbb{P}(L_n > nx_n)$ under different asymptotic regimes, where either $x_n \uparrow 1$, or d is replaced by $d_n \downarrow -\infty$ as $n \rightarrow \infty$.

To motivate (2) from a financial perspective, it is helpful to briefly review the classical framework originally introduced by Merton (1974). To this end, beginning with a single loan to a firm with assets $V(t)$ at time t , it is natural to model $V(t)$ according to the stochastic differential equation

$$\frac{dV(t)}{V(t)} = mdt + \sigma dW(t), \tag{3}$$

for some mean drift m and volatility σ , where $\{W(t)\}$ is a standard Brownian motion. Now, if the firm is to repay a loan of size K at the future time T , then default occurs if $V(T) < K$. Upon integrating and applying Ito’s formula, we see that, viewed from the current time t , this default will happen if

$$Y := \frac{W(T) - W(t)}{\sigma\sqrt{T-t}} < \frac{\log K - \log V(t) - (\sigma/2 - m)(T-t)}{\sigma\sqrt{T-t}} := d, \tag{4}$$

where the threshold d denotes the “distance to default”. Now consider a portfolio of loans to a collection of firms. Let the values of the n firms be given by V_1, \dots, V_n ; then the model in (3) can be extended to

$$\frac{dV_i(t)}{V_i(t)} = m_i dt + \sum_{j=1}^l \sigma_{ij} dW_j(t), \quad i = 1, \dots, n, \tag{5}$$

where $(W_1(t), \dots, W_l(t))$ is l -dimensional Brownian motion, and μ_i and σ_{ij} are positive constants for all i, j . Using the reasoning leading to (4), it follows that the default occurs when $V_i(T) < K_i$; and this is equivalent to

$$Y_i < \frac{\log K_i - \log V_i(t) + (\sigma_i^2/2 - m_i)(T-t)}{\sigma_i\sqrt{T-t}} := d_i, \tag{6}$$

which, after conducting simple algebra, yields that

$$Y_i := \sum_{j=1}^k \frac{\sigma_{ij}}{\sigma_i} \left(\frac{W_j(T) - W_j(t)}{\sqrt{T-t}} \right) < d_i \tag{7}$$

for $\sigma_i^2 = \sum_{j=1}^k \sigma_{ij}^2$. Thus, choosing one component of the Brownian motion to be firm-specific for each loan, and the remaining factors to be common among all loans (so that $l = k + n$), we are led to the factor model

$$Y_i = \sum_{j=1}^k a_{ij} Z_j + b_i \epsilon_i, \quad i = 1, \dots, n, \tag{8}$$

where a_{ij} and b_i are positive constants for all i, j , and Z_1, \dots, Z_k and $\{\epsilon_i\}$ are i.i.d. standard Normal random variables, which is (2), but with (a_{ij}, b_i, d_i) in place of (a_j, b, d) and

an independent Normal assumption on the factors. The variables Z_1, \dots, Z_k represent the *common factors*, while the random variables $\{\epsilon_i\}$ represent the *idiosyncratic factors* (or individual factors).

The Gaussian assumption has often been criticized and other distributional assumptions have been suggested. For instance, Schönbucher (2000), Gordy (2003), Hull and White (2004), and Burtschell et al. (2009) propose extensions of factor models with non-normal distributions, including Archimedean and t-copula models. The reference Bush et al. (2011) analyzes a dynamic extension of Vasicek’s homogeneous single factor model in which the systematic risk factors follow a Brownian motion. Some large deviation estimates for sums of random variables have been provided in Maier and Wüthrich (2009), where the dependencies are modeled according to a copula (rather than through a threshold factor model).

The aim of this work is to develop *sharp* asymptotics for the tails of the total loss distribution. Such tail asymptotics were first introduced under independent Gaussian assumptions in Glasserman et al. (2007). However, in contrast, we will adopt a general framework where both the common and idiosyncratic factors are allowed to assume very general distributions (and not necessarily the same distributions). Our work seems to provide the first theoretical results that are nonlogarithmic and do not invoke convenient tail assumptions (as are specified in the Normal or regularly varying distributions, and their multivariate extensions via the Gaussian and t-copulas).

To describe our results, let $\{Y_n\}$ be given as in (2), and, as before, let

$$L_n = \sum_{i=1}^n U_i X_i, \quad \text{for } X_i = \mathbf{1}_{\{Y_i < d\}},$$

where $d \in \mathbb{R}$ and $\{U_i\}$ is an i.i.d. sequence, independent of $\{X_i\}$. As the random variable U_i represents the loss incurred when the i^{th} loan defaults, we have that $U_i = l_i V_i$, where l_i is the size of the i^{th} loan and $V_i \in [0, 1]$ represents the recovery rate. Assume that the sequence $\{U_i\}$ is i.i.d. and independent of $\{X_i\}$, and assume without loss of generality that $\mathbf{E}[U_i] = 1$. Furthermore, set $\mathcal{Z} = -\sum_{j=1}^k a_j Z_j$. Then conditional on \mathcal{Z} , the central tendency of $\{L_n/n\}$ is given by

$$\lim_{n \rightarrow \infty} \frac{1}{n} \mathbf{E}[L_n | \mathcal{Z}] = \mathbf{E}[U_i X_i | \mathcal{Z}] = p(\mathcal{Z}), \tag{9}$$

where

$$p(\mathcal{Z}) = \mathbf{P}(X_i = 1 | \mathcal{Z}) = \mathbf{P}(b\epsilon_i < d + \mathcal{Z} | \mathcal{Z}), \tag{10}$$

by (2). Then by the conditional Chebyshev inequality,

$$\mathbf{P}\left(\left|\frac{L_n}{n} - p(\mathcal{Z})\right| > \epsilon \mid \mathcal{Z}\right) \leq \frac{\text{Var}(L_n | \mathcal{Z})}{n^2 \epsilon^2} = \frac{p(\mathcal{Z})(1 - p(\mathcal{Z}))}{n \epsilon^2},$$

implying that $\{L_n/n\}$ converges in probability to $p(\mathcal{Z})$ conditioned on \mathcal{Z} . Hence,

$$\mathbf{P}\left(\frac{L_n}{n} > x\right) = \mathbf{E}\left[\mathbf{P}\left(\frac{L_n}{n} > x \mid \mathcal{Z}\right)\right] \rightarrow \mathbf{P}(p(\mathcal{Z}) > x) \quad \text{as } n \rightarrow \infty. \tag{11}$$

In particular, if $k = 1$ in (2) and $\mathcal{Z} \stackrel{d}{=} aZ \sim \text{Normal}(0, a^2)$, then it follows from (11) that

$$\mathbf{P}\left(\frac{L_n}{n} > x\right) = \mathbf{P}(aZ > p^{-1}(x)) = 1 - \Phi\left(\frac{b\Phi^{-1}(x) - d}{a}\right), \tag{12}$$

which is a formula originally established in Vasicek (1991).

As observed in Glasserman et al. (2007), this last equation shows that a meaningful asymptotic result can only be obtained by studying the problem in a limiting sense, e.g.,

by letting $x_n \uparrow 1$, or by replacing the threshold d with d_n and letting $d_n \downarrow -\infty$. Notice that as $x \uparrow 1$, the right-hand side of (12) converges to zero, and using the formula for the tails of a Normal distribution (as given, e.g., in Chow and Teicher (1997)), this term decays asymptotically as

$$\frac{a}{b\Phi^{-1}(x)} \exp\left(-\left(\frac{b}{a}\Phi^{-1}(x)\right)^2\right). \tag{13}$$

In particular, taking $x \equiv x_n = \Phi(s\sqrt{\log n})$ and letting $n \rightarrow \infty$, this last expression is asymptotic to

$$\frac{a}{b} \frac{1}{\log n} \exp\left(-s^2\left(\frac{b}{a}\right)^2 \log n\right); \tag{14}$$

and hence

$$\lim_{n \rightarrow \infty} \frac{1}{\log n} \log\left(1 - \Phi\left(\frac{b\Phi^{-1}(x_n) - d}{a}\right)\right) = -\frac{s^2 b^2}{a^2}, \tag{15}$$

suggesting that $\mathbf{P}(L_n > nx_n)$ will exhibit a similar limit behavior to that of $\mathbf{P}(p(Z) > x_n)$; namely,

$$\lim_{n \rightarrow \infty} \frac{1}{\log n} \log \mathbf{P}(L_n > nx_n) = -\frac{s^2 b^2}{a^2}.$$

Asymptotics of this type were established on a logarithmic scale in Glasserman et al. (2007) for Gaussian k -factor models with finitely many types.

Letting $\{x_n\}$ be a sequence converging to one and satisfying certain regularity conditions, our first main result examines the decay of $\mathbf{P}(L_n > nx_n)$ as $n \rightarrow \infty$ for the factor model in (2). Setting $\mathcal{Z} = -\sum_{j=1}^k a_j Z_k$ and letting $F_{\mathcal{Z}}$ denote its distribution function, then we establish that

$$\mathbf{P}\left(\frac{L_n}{n} > x_n\right) \sim \bar{F}_{\mathcal{Z}}(p^{-1}(x_n)) \quad \text{as } n \rightarrow \infty, \tag{16}$$

where $p(\cdot)$ is given as in (10) and is thus determined by the distribution of the idiosyncratic factors $\{\epsilon_i\}$. On the other hand, the tail decay is determined by $\bar{F}_{\mathcal{Z}} := 1 - F_{\mathcal{Z}}$, i.e., by the common factors, thus exhibiting an interesting interplay between the roles of the common factors and the individual factors associated with the given loans. [Here and in the following, $f(x_n) \sim g(x_n)$ as $n \rightarrow \infty$ means that $\lim_{n \rightarrow \infty} (f(x_n)/g(x_n)) = 1$.] We establish this result under very general assumptions on the common factors Z_1, \dots, Z_k and the idiosyncratic sequence $\{\epsilon_i\}$, where, in principle, we do not even assume that any of these random variables have a common distribution. However, as in Glasserman et al. (2007) and the heuristic estimate (15), the rate at which $\{x_n\}$ tends to one must satisfy certain weak constraints (requiring that $\{x_n\}$ does not grow “too quickly”), where these constraints are dependent both on the distributions of Z_1, \dots, Z_k and on $\{\epsilon_i\}$. Specifically, we describe a range of possible values for $\{x_n\}$ (in contrast to a single sequence, as in Glasserman et al. (2007)), and also provide a uniform estimate for (16) within this range. We emphasize here that our result is the first to study sharp asymptotics for a general class of distributions and, in particular, in the context of the Gaussian model. In the process, we also introduce a new approach based on a simple conditioning argument combined with Hoeffding’s concentration inequality. We also relate our estimate to Value-at-Risk estimation, and describe an extension to multiple types, where the default level d and parameters $\mathbf{a} = (a_1, \dots, a_k)$ and b are allowed to belong to different classes, and hence are allowed to vary amongst the different loans.

The asymptotic regime described in (16) can be viewed as the *large loss regime*, where defaults occur simultaneously because, as $x_n \uparrow 1$, it becomes increasingly likely that the given loans default, and default occurs under the conditional law of large numbers

whenever \mathcal{Z} exceeds the threshold $p^{-1}(x_n)$. In Glasserman et al. (2007), a *small default regime*, is also considered, where the default threshold d is replaced with a sequence $\{d_n\}$, where $d_n \downarrow -\infty$, and default occurs when $X_n < d_n$, where X_n is again given as in (2). Note that as $n \rightarrow \infty$, we have that $\mathbf{P}(X_n < d_n) \rightarrow 0$, so as n increases, the quality of the credit increases. Thus, this regime considers high-quality credits, with small default probabilities, and thus the event $\{L_n > nx\}$ will decay to zero, even when $x \in (0, 1)$ is fixed. For the small-default regime, we once again adopt a rather general framework, where the common factors and idiosyncratic factors may assume general distributions, as in (16). Under some natural conditions on the decay of $d_n \downarrow -\infty$, we show that if the distribution of $\{\epsilon_i\}$ is symmetric, then

$$\mathbf{P}\left(\frac{L_n}{n} > x\right) \sim \bar{F}_{\mathcal{Z}}(|b|F_{\epsilon}^{-1}(x) - d_n) \quad \text{as } n \rightarrow \infty, \tag{17}$$

where F_{ϵ} denotes the distribution function of ϵ_i . [A corresponding result also holds under general assumptions on $\{\epsilon_i\}$.] Since x is fixed, we note that the quantity on the right-hand side is determined by the rate of decay of $\bar{F}_{\mathcal{Z}}(C - d_n)$ for some constant C , i.e., the rate of decay behaves roughly like $\bar{F}_{\mathcal{Z}}(|d_n|)$; and similarly, the tail behavior of \mathcal{Z} will also determine how large the sequence $\{d_n\}$ may be chosen. Thus, we see that the decay rate in (17) and the choice of $\{d_n\}$ are both essentially determined by \mathcal{Z} and hence the common factors, while the idiosyncratic factors plays no role in determining the rate of decay in this estimate.

We conclude by observing that estimates such as (16) and (17) also provide some insight into the role of *dependence* amongst the common factors Z_1, \dots, Z_k , which, in the existing literature, are generally assumed to be independent. From a practical perspective, these factors will generally be dependent, with their dependence described through the distribution of the random variable \mathcal{Z} . For example, in the Normal case, this dependence is characterized through a covariance matrix, while for general elliptical distributions, such as the t -distribution, this dependence is characterized through the corresponding dispersion matrix. As an example, we calculate the rate function in the Normal case, and illustrate how this dependence influences the rate of decay in our estimates.

The rest of the paper is organized as follows. Section 2 is concerned with the main results in the large loss and small default regimes, while Section 3 describes extensions to multiple types, and to a nonstandard formulation of the problem, where the number of factors is allowed to tend to infinity as the number of loans tends to infinity.

2. Sharp Tail Asymptotics

Given an arbitrary random variable R , denote the distribution function of R by F_R , the corresponding density by f_R , and its mean value by μ_R (assuming that these exist). Furthermore, let $\bar{F}_R(z) := 1 - F_R(z)$ denote the tail probability, $\forall z \in \mathbb{R}$. Finally let $\lambda_R(z) := f_R(z)/\bar{F}_R(z)$ denote the hazard function, and let $\Lambda_R(z) = \int_0^z \lambda_R(x)dx$ denote the cumulative hazard function. It is well known that $\mathbf{P}(R > z) = e^{-\Lambda_R(z)}$.

Throughout this section, assume that

$$L_n = \sum_{i=1}^n U_i X_i,$$

where $X_i = \mathbf{1}_{\{Y_i < d\}}$ denotes the default of the i th lender, and assume that X_i and U_i are mutually independent for each i and that the sequence $\{U_i\}$ is i.i.d. Without loss of generality, we assume that $\mathbf{E}[U_i] = 1$, and since U_i denotes the loss of an individual loan, we further assume that $U_i \leq l$, where $l < \infty$ denotes the loan size. As in the previous section, we model $\{Y_i\}$ according to the factor model (2), where it is also assumed that $(a_1^2 + \dots + a_k^2) + b^2 = 1$.

Throughout this article, we will always assume that the random variables Z_1, \dots, Z_k and $\epsilon_1, \epsilon_2, \dots$ have a density with unbounded support, and that F_{ϵ} is strictly increasing and

continuous, implying, in particular, that F_ϵ has a proper inverse. Furthermore, in (2), we assume that $\{\epsilon_i\}$ is independent of $\{Z_1, \dots, Z_k\}$.

2.1. Tail Approximation for the Large Loss Regime

For any given $x \in \mathbb{R}$, set $z(x) = p^{-1}(x)$, where we recall that $p(\cdot)$ denotes the conditional default probability of an obligor. In other words, $z(x)$ solves the equation $x = p(z(x))$, describing the threshold value where default becomes “likely” for $\mathcal{Z} \geq z$. In particular, the i^{th} obligor defaults if $Y_i = \sum_{i=1}^d a_i Z_i + b\epsilon_i < d$, or equivalently if $\text{sgn}(b)\epsilon_i < (d + \mathcal{Z})/|b|$, where $\text{sgn}(b) = b/|b|$ and

$$\mathcal{Z} := - \sum_{i=1}^k a_i Z_i.$$

Letting $\tilde{\epsilon} = \text{sgn}(b)\epsilon$, this leads to the equation

$$x = p(z(x)) = F_{\tilde{\epsilon}}\left(\frac{d + z(x)}{|b|}\right), \quad \text{or} \quad p^{-1}(x) \equiv z(x) = |b|F_{\tilde{\epsilon}}^{-1}(x) - d. \tag{18}$$

Of course, if F_ϵ has a symmetric distribution—as is often the case—then $F_{\tilde{\epsilon}}$ may be replaced with F_ϵ in the previous equation, and this comment holds for all of the results in the article.

As stated in the introduction, it follows by Vacicek’s law of large numbers that, conditioned on \mathcal{Z} ,

$$\frac{L_n}{n} \rightarrow p(\mathcal{Z}) := F_{\tilde{\epsilon}}\left(\frac{d + \mathcal{Z}}{|b|}\right) \geq x \quad \text{when} \quad \mathcal{Z} \geq z(x), \tag{19}$$

where we have used the monotonicity of $F_\epsilon(\cdot)$. Thus,

$$\mathbf{P}\left(\frac{L_n}{n} > x \mid \mathcal{Z} = z\right) \rightarrow 0 \quad \text{for} \quad z < z(x),$$

while

$$\mathbf{P}\left(\frac{L_n}{n} > x \mid \mathcal{Z} = z\right) \rightarrow 1 \quad \text{for} \quad z > z(x).$$

Thus, letting $x_n \uparrow 1$, we expect by a simple conditioning argument that

$$\mathbf{P}\left(\frac{L_n}{n} > x_n\right) = \int_{\mathbb{R}} \mathbf{P}\left(\frac{L_n}{n} > x_n \mid \mathcal{Z} = z\right) dF_{\mathcal{Z}}(z) \sim \int_{z(x_n)}^{\infty} dF_{\mathcal{Z}}(z), \tag{20}$$

as $n \rightarrow \infty$, provided that

$$\int_{-\infty}^{z(x_n)} \mathbf{P}\left(\frac{L_n}{n} > x_n \mid \mathcal{Z} = z\right) dF_{\mathcal{Z}}(z) \sim o(\bar{F}_{\mathcal{Z}}(z(x_n))) \quad \text{as} \quad n \rightarrow \infty. \tag{21}$$

This leads to our first main result.

Theorem 1. *Let $\{x_n\}$ be chosen such that $x_n \uparrow 1$ as $n \rightarrow \infty$, and let $1 > y_n \geq x_n$ be chosen such that the cumulative hazard of \mathcal{Z} satisfies*

$$\frac{\Lambda_{\mathcal{Z}}(p^{-1}(x_n))}{n(1 - y_n)^2} = o(1) \quad \text{as} \quad n \rightarrow \infty. \tag{22}$$

Furthermore, assume that there exists a finite constant M such that $\{(\lambda_{\mathcal{Z}}(v)/\lambda_{\epsilon^*}(v)) : v \geq M\}$ is bounded from above for large M , where $\epsilon_i^* = b\epsilon_i - d$; and the density $f_{\mathcal{Z}}(z)$ is nonincreasing for $z \geq M$. Then

$$\lim_{n \rightarrow \infty} \sup_{x_n \leq x \leq y_n} \left| \frac{\mathbf{P}(L_n > nx)}{\bar{F}_{\mathcal{Z}}(p^{-1}(x))} - 1 \right| = 0. \tag{23}$$

The conditions of the theorem are widely satisfied, as we now illustrate through a few of examples. As a first example, suppose that the common factors Z_1, \dots, Z_k and idiosyncratic factors $\epsilon_1, \epsilon_2, \dots$ have the standard Normal distribution. If $\{Z_1, \dots, Z_k\}$ are independent (which is not required for our general result), then $\mathcal{Z} := -(a_1 Z_1 + \dots + a_k Z_k) \sim \text{Normal}(0, \|a\|^2)$, where $\|a\|^2 = a_1^2 + \dots + a_k^2$, and by (18), $z(x) = b\Phi^{-1}(x) - d$. Returning to the main result in Glasserman et al. (2007), now suppose that $x_n = y_n = \Phi(s\sqrt{\log n})$ for $s \in (0, 1)$, where Φ is the distribution function of a standard Normal distribution. Then $p^{-1}(x_n) \equiv z(x_n) = bs\sqrt{\log n} - d$. Since $\Phi(y) \sim (Ce^{-y^2/2})/y$ as $y \rightarrow \infty$ for some constant C , it follows by a direct calculation that

$$n(1 - x_n)^2 \sim \frac{C^2 n}{s^2 \log n} e^{-(s\sqrt{\log n})^2} \sim \frac{C^2}{s^2 \log n} n^{1-s^2}, \quad n \rightarrow \infty, \tag{24}$$

while

$$\Lambda_{\mathcal{Z}}(p^{-1}(x_n)) = -\log \bar{F}_{\mathcal{Z}}(p^{-1}(x_n)) \sim \frac{b^2 s^2}{2\|a\|^2} \log n, \quad n \rightarrow \infty, \tag{25}$$

using that $\mathcal{Z} \sim \text{Normal}(0, \|a\|^2)$. Thus, (22) holds, and hence it follows from (23) that

$$\mathbf{P}\left(\frac{L_n}{n} > x_n\right) \sim \bar{F}_{\mathcal{Z}}(p^{-1}(x_n)) \sim \frac{C}{(|b|s\sqrt{\log n} - d)/\|a\|} e^{-\frac{(|b|s\sqrt{\log n} - d)^2/2\|a\|^2}{2\|a\|^2}} \quad \text{as } n \rightarrow \infty,$$

using the definition of $p^{-1}(x_n)$ in the last step, which is a sharper version of Glasserman et al. (2007)'s main result specialized to the single-type case.

Alternatively, if $(Z_1, \dots, Z_k)' \sim \text{Normal}(0, \Sigma)$, where Σ is a positive definite matrix, then $\mathcal{Z} := -(a_1 Z_1 + \dots + a_k Z_k) \sim \text{Normal}(0, a'\Sigma a)$, where $a' = (a_1, \dots, a_k)$. Arguing as before, it can be seen that

$$\begin{aligned} \mathbf{P}\left(\frac{L_n}{n} > x_n\right) &\sim \bar{F}_{\mathcal{Z}}(p^{-1}(x_n)) \\ &\sim \left(\frac{C}{(|b|s\sqrt{\log n} - d)/a'\Sigma a}\right) e^{-\frac{(|b|s\sqrt{\log n} - d)^2/2a'\Sigma a}{2a'\Sigma a}} \quad \text{as } n \rightarrow \infty. \end{aligned}$$

Finally, taking the logarithm and scaling by $\log n$, we obtain the weaker result

$$\lim_{n \rightarrow \infty} \frac{1}{\log n} \log \mathbf{P}\left(\frac{L_n}{n} > x_n\right) = -\frac{s^2 b^2}{a'\Sigma a}, \tag{26}$$

which explicitly shows the effect of dependence (through the matrix Σ) on the risk estimate. Once again, note that by taking Σ to be the $k \times k$ identity matrix, we recover the result in Glasserman et al. (2007) for the single-type case.

As a second example, suppose that the common factors are mutually independent and have a symmetric two-sided subexponential distribution. Then $Z_i^+ := Z_i \vee 0$ has a subexponential distribution $\forall i$, and consequently by a slight modification of Albrecher and Asmussen (2010), Proposition IX.1.9, we have that

$$\mathbf{P}(\mathcal{Z} > u) \sim \sum_{j=1}^k \mathbf{P}\left(Z_j^+ > \frac{u}{|a_j|}\right) \quad \text{as } u \rightarrow \infty.$$

In particular, if $\{Z_i^+\}$ has a regularly varying distribution with parameter α (and the distribution is equally divided along the positive and negative axes), then

$$\mathbf{P}(\mathcal{Z} > u) \sim \frac{1}{2}(|a_1|^\alpha + \dots + |a_k|^\alpha)L(u)u^{-\alpha} \quad \text{as } u \rightarrow \infty,$$

for a slowly varying function $L(\cdot)$, where $L(x) = \beta^\alpha$ in the Pareto(α, β) case.

Now if the common factors $\{Z_i\}$ have a Pareto(α, β) distribution, and if the idiosyncratic factors $\{\epsilon_i\}$ have the standard Gaussian distribution, then the ratio $\lambda_{\mathcal{Z}}(v)/\lambda_{\epsilon^*}(v)$ tends to zero as $v \rightarrow \infty$, as can either be verified by direct calculation or by observing the Pareto(α, β) has heavier tails than the Gaussian distribution. Turning to the condition (22), suppose that as in the previous example, $x_n = y_n$ has been chosen such that $x_n = \Phi(s\sqrt{\log n})$ for $s \in (0, 1)$. Then, just as before, we have that (24) holds, while

$$\begin{aligned} \Lambda_{\mathcal{Z}}(p^{-1}(x_n)) &= -\log \bar{F}_{\mathcal{Z}}(p^{-1}(x_n)) \sim \log(\beta + p^{-1}(x_n))^\alpha \\ &= \alpha \log(\beta + bF_\epsilon^{-1}(\Phi(s\sqrt{\log n})) - d) \sim \alpha \log(bs\sqrt{\log n}), \quad n \rightarrow \infty. \end{aligned}$$

Thus, (22) is easily satisfied (and a much wider choice of $\{x_n\}, \{y_n\}$ is also possible). Finally, by (23), we conclude that

$$\mathbf{P}\left(\frac{L_n}{n} > x_n\right) \sim \bar{F}_{\mathcal{Z}}(p^{-1}(x_n)) \sim \frac{\beta^\alpha}{(\beta + bs\sqrt{\log n} - d)^\alpha} \quad \text{as } n \rightarrow \infty, \tag{27}$$

where we have used that $F_\epsilon^{-1}(x_n) = \Phi^{-1}(\Phi(s\sqrt{\log n})) = s\sqrt{\log n}$.

Additional examples involving Gamma distributions and stretched exponential distributions for the common factors can be derived along the lines of the above examples; see de Silva (2016) in the one-factor case. In the multi-factor case, where the factors are light-tailed, it can often be challenging to identify the precise distribution of the sum \mathcal{Z} . However, if Z_1, \dots, Z_k belong to a class which is closed under convolution, then this distribution can be identified as the same type of distribution as the individual factors Z_i . This is, for example, the case with the Normal distribution or the Gamma distribution.

Next we turn to the proof of Theorem 1. First, we begin by establishing a lemma needed in this proof.

Lemma 1. *Assume that there exists a finite constant M such that $\{(\lambda_{\mathcal{Z}}(v)/\lambda_{\epsilon^*}(v)) : vs. \geq M\}$ is bounded from above for large M , where $\epsilon_i^* = b\epsilon_i - d$; and the density $f_{\mathcal{Z}}(z)$ is nonincreasing for $z \geq M$. Then for any sequence $\{\delta_n\}$ such that $\delta_n = o(1 - y_n)$ as $n \rightarrow \infty$, and for any $x \in [x_n, y_n]$ (where $\{x_n, y_n\}$ is given as in Theorem 1), we have that*

$$\frac{\bar{F}_{\mathcal{Z}}(p^{-1}(x + \delta_n))}{\bar{F}_{\mathcal{Z}}(p^{-1}(x - \delta_n))} \geq 1 - \gamma_n \quad \text{for all } n \geq \text{some } N_0 \tag{28}$$

for some sequence $\{\gamma_n\} \downarrow 0$ as $n \rightarrow \infty$.

Proof of Lemma 1. Since $f_{\mathcal{Z}}(z)$ is assumed to be nonincreasing for large z ,

$$\bar{F}_{\mathcal{Z}}(p^{-1}(x - \delta_n)) - \bar{F}_{\mathcal{Z}}(p^{-1}(x + \delta_n)) \leq 2\delta_n f_{\mathcal{Z}}(p^{-1}(x - \delta_n))(p^{-1})'(x - \delta_n)$$

for sufficiently large n and $x \in [x_n, y_n] \uparrow 1$ (and hence $p^{-1}(x) \uparrow \infty$). Thus, to establish (28), we need to show that

$$\delta_n \frac{f_{\mathcal{Z}}(p^{-1}(x - \delta_n))(p^{-1})'(x - \delta_n)}{\bar{F}_{\mathcal{Z}}(p^{-1}(x - \delta_n))} \leq \frac{1}{2}\gamma_n, \quad n \geq N_0. \tag{29}$$

To verify (29), note that p is the distribution function of $\epsilon_i^* = b\epsilon_i - d$, and hence

$$\lambda_{\epsilon^*}(p^{-1}(v)) = \frac{f_{\epsilon^*}(p^{-1}(v))}{1 - F_{\epsilon^*}(p^{-1}(v))} = \frac{p'(p^{-1}(v))}{1 - v}, \quad \forall v,$$

and differentiating the equation $p(p^{-1}(v)) = v$ then yields

$$(p^{-1})'(v) = \frac{1}{p'(p^{-1}(v))} = \frac{1}{\lambda_{\epsilon^*}(p^{-1}(v))} \frac{1}{1-v}, \quad \forall v. \tag{30}$$

Substituting this last expression into the left-hand side of (29), we now see that (29) holds provided that

$$\left(\frac{\delta_n}{1-x+\delta_n}\right) \frac{\lambda_{\mathcal{Z}}(p^{-1}(x-\delta_n))}{\lambda_{\epsilon^*}(p^{-1}(x-\delta_n))} \leq \frac{1}{2}\gamma_n, \quad n \geq N_0, \tag{31}$$

Now $\lambda_{\mathcal{Z}}(z)/\lambda_{\epsilon^*}(z)$ is assumed to be bounded for sufficiently large z . Finally, choose $\delta_n = o(1-y_n)$ as $n \rightarrow \infty$. Then in the denominator on the right-hand side of (31), we have that $(1-x+\delta_n) \geq (1-y_n+\delta_n) \sim 1-y_n$ as $n \rightarrow \infty$, and the lemma follows. \square

Proof of Theorem 1. For any $v \in \mathbb{R}$, set $z(v) = p^{-1}(v)$. Throughout the proof, let $\mathbf{E}_{z(v)}[\cdot]$, $\text{Var}_{z(v)}(\cdot)$ denote conditional expectation and conditional variance given $\mathcal{Z} = z(v)$.

The idea of the proof is to study $\mathbf{P}(L_n > nx_n)$ over the regions

$$\{\mathcal{Z} \geq p^{-1}(x+\delta_n)\}, \quad \{\mathcal{Z} \leq p^{-1}(x-\delta_n)\}, \quad \text{and} \quad \{p^{-1}(x-\delta_n) < \mathcal{Z} < p^{-1}(x+\delta_n)\},$$

respectively and show that $\mathbf{P}(L_n > nx; \mathcal{Z} \geq p^{-1}(x+\delta_n)) \rightarrow 1$, while the probabilities in the other regions converge to zero.

Let $\{\delta_n\}$ be a positive-valued sequence, where $\delta_n = o(1-y_n)$. For any $v \in \mathbb{R}$, $p(z(v)) := p \circ p^{-1}(v) = v$, and thus $\mathbf{E}_{z(v)}[U_i X_i] = \mathbf{E}[U_i]p(z(v)) = v$ (using the independence of $\{U_i\}$ and $\{X_i\}$ and the fact that $\{X_i\}$ has a Bernoulli distribution). Hence by Chebyshev’s inequality,

$$\mathbf{P}\left(\left|\frac{L_n}{n} - v\right| > \delta_n \mid \mathcal{Z} = z(v)\right) \leq \frac{1}{n\delta_n^2} \text{Var}_{z(v)}(U_i X_i) \leq \frac{C}{n\delta_n^2}, \tag{32}$$

where $\text{Var}_{z(v)}(U_i X_i) \leq \mathbf{E}[U_i^2] := C < \infty$. Next, given x , set $v_n = x + \delta_n$. Then setting v to be equal to v_n in the previous equation yields

$$\mathbf{P}\left(\frac{L_n}{n} > x \mid \mathcal{Z} = z(v_n)\right) = \mathbf{P}\left(\frac{L_n}{n} - v_n > -\delta_n \mid \mathcal{Z} = z(v_n)\right) \geq 1 - \frac{C}{n\delta_n^2}. \tag{33}$$

First observe that $\mathbf{P}(L_n(z) > nx \mid \mathcal{Z} = z)$ is monotonically increasing in z since, if $z_1 < z_2$, then $X_i(z_1) = \mathbf{1}_{\{be_i < d+z_1\}} = 1$ implies $X_i(z_2) = 1$. Thus, $L_n(z_1) \leq L_n(z_2)$. Consequently,

$$\begin{aligned} \mathbf{P}\left(\frac{L_n}{n} > x; \mathcal{Z} \geq z(v_n)\right) &= \int_{z(v_n)}^{\infty} \mathbf{P}\left(\frac{L_n}{n} > x \mid \mathcal{Z} = z\right) dF_{\mathcal{Z}}(z) \\ &\geq \bar{F}_{\mathcal{Z}}(z(v_n)) \mathbf{P}\left(\frac{L_n}{n} > x \mid \mathcal{Z} = z(v_n)\right) \\ &\geq \bar{F}_{\mathcal{Z}}(z(v_n)) \left(1 - \frac{C}{n\delta_n^2}\right). \end{aligned} \tag{34}$$

Recalling that $p^{-1}(v_n) = z(v_n)$, this leads to the estimate

$$1 \geq \frac{\mathbf{P}(L_n > nx; \mathcal{Z} \geq z(v_n))}{\bar{F}_{\mathcal{Z}}(p^{-1}(v_n))} \geq \left(1 - \frac{C}{n\delta_n^2}\right), \tag{35}$$

uniformly in $x \in [x_n, 1 - \delta_n]$ (where the left inequality follows after noticing that the middle term is a conditional probability). Then by Lemma 1,

$$\frac{\mathbf{P}(L_n > nx; \mathcal{Z} \geq z(v_n))}{\bar{F}_{\mathcal{Z}}(p^{-1}(x))} \rightarrow 1 \quad \text{as} \quad n \rightarrow \infty, \tag{36}$$

uniformly for $x \in [x_n, y_n]$.

Next, for any $w \in \mathbb{R}$, again observe that $\mathbf{E}_{z(w)}[U_i X_i] = w$, and consequently by Hoeffding’s inequality

$$\mathbf{P}\left(\frac{L_n}{n} - w > \delta_n \mid \mathcal{Z} = z(w)\right) \leq e^{-2n\delta_n^2/l^2}, \tag{37}$$

where l is the maximum value of the bounded random variable U . Given x , now set $w_n = x - \delta_n$. Since $\mathbf{P}(L_n(z) > nx \mid \mathcal{Z} = z)$ is monotonically increasing in z ,

$$\begin{aligned} \mathbf{P}\left(\frac{L_n}{n} > x; \mathcal{Z} \leq z(w_n)\right) &= \int_{-\infty}^{z(w_n)} \mathbf{P}\left(\frac{L_n}{n} > x \mid \mathcal{Z} = z\right) dF_{\mathcal{Z}}(z) \\ &\leq \mathbf{P}\left(\frac{L_n}{n} > x \mid \mathcal{Z} = z(w_n)\right) \leq e^{-2n\delta_n^2/l^2}. \end{aligned} \tag{38}$$

Then

$$\frac{\mathbf{P}(L_n > nx; \mathcal{Z} \leq z(w_n))}{\bar{F}(p^{-1}(x))} \leq \exp\left\{-\frac{2n\delta_n^2}{l^2} + \Lambda_{\mathcal{Z}}(p^{-1}(x))\right\}, \tag{39}$$

where we have used the representation $\bar{F}_{\mathcal{Z}}(v) = e^{-\Lambda_{\mathcal{Z}}(v)}$ for any v . Then the right-hand side of (39) tends to zero provided that

$$\frac{\Lambda_{\mathcal{Z}}(p^{-1}(x))}{n\delta_n^2} \leq \frac{\Lambda_{\mathcal{Z}}(p^{-1}(x_n))}{n\delta_n^2} = o(1) \quad \text{as } n \rightarrow \infty, \tag{40}$$

which holds under the assumption (22) and the choice of $\delta_n = o(1 - y_n)$.

Finally observe that

$$\mathbf{P}\left(\frac{L_n}{n} > x; z(w_n) < \mathcal{Z} < z(v_n)\right) \leq \bar{F}_{\mathcal{Z}}(p^{-1}(v_n)) - \bar{F}_{\mathcal{Z}}(p^{-1}(w_n));$$

hence by Lemma 1,

$$\frac{\mathbf{P}(L_n > nx; z(w_n) < \mathcal{Z} < z(v_n))}{\bar{F}_{\mathcal{Z}}(p^{-1}(x))} = \frac{\bar{F}_{\mathcal{Z}}(p^{-1}(x + \delta_n))}{\bar{F}_{\mathcal{Z}}(p^{-1}(x))} - \frac{\bar{F}_{\mathcal{Z}}(p^{-1}(x - \delta_n))}{\bar{F}_{\mathcal{Z}}(p^{-1}(x))} = 0, \tag{41}$$

uniformly for $x \in [x_n, y_n]$. By combining (36), (39), (40), and (41), we obtain the statement of the theorem. \square

2.2. Some Consequences of Theorem 1

A natural application of Theorem 1 is to risk management, where one evaluates Value-at-Risk and Expected Shortfall for a credit risk portfolio. Recall that for a random variable R , the Value-at-Risk at level $\alpha \in (0, 1)$ is defined by

$$\text{VaR}_{\alpha}(R) = \inf\{x : \mathbf{P}(R > x) \leq 1 - \alpha\} = q_{\alpha}(F_R),$$

where q_{α} denotes the α -level quantile, while Expected Shortfall is defined by

$$\text{ES}_{\alpha}(R) = \frac{1}{1 - \alpha} \int_{\alpha}^1 \text{VaR}_v(R) dv. \tag{42}$$

When the distribution function F_R has a nonzero density function (as we assume here), so F_R is strictly increasing and continuous, these definitions simplify to the following:

$$\mathbf{P}(R > \text{VaR}_{\alpha}(R)) = 1 - \alpha; \quad \text{ES}_{\alpha}(R) = \frac{1}{1 - \alpha} \int_{\text{VaR}_{\alpha}(R)}^{\infty} x dF_R(x).$$

We first focus our discussion on Value-at-Risk estimates. Now by Theorem 1, it follows that for $x \in [x_n, y_n]$,

$$\left| \mathbf{P}\left(\frac{L_n}{n} > x\right) - \bar{F}_{\mathcal{Z}}\left(p^{-1}(x)\right) \right| \leq \gamma_n \bar{F}_{\mathcal{Z}}\left(p^{-1}(x)\right), \tag{43}$$

where γ_n is a sequence which tends to zero as $n \rightarrow \infty$ (whose magnitude can be inferred from the proof of Theorem 1). Note that if this estimate were *exact* (i.e., the right-hand side of (43) were zero, then it would suffice to compute the Value-at-Risk of \mathcal{Z} , where

$$1 - \alpha = \bar{F}_{\mathcal{Z}}(\text{VaR}_{\alpha}(\mathcal{Z})) = \bar{F}_{\mathcal{Z}}\left(p^{-1} \circ p(\text{VaR}_{\alpha}(\mathcal{Z}))\right),$$

suggesting that $\text{VaR}_{\alpha}(L_n/n) \approx p(\text{VaR}_{\alpha}(\mathcal{Z}))$. For a precise upper bound, let n be fixed and choose β such that $(1 - \beta)(1 - \gamma_n) \leq 1 - \alpha$. Now choose $x = p(\text{VaR}_{\beta}(\mathcal{Z}))$. Then

$$\mathbf{P}\left(\frac{L_n}{n} > x\right) \leq (1 + \gamma_n)\bar{F}_{\mathcal{Z}}\left(p^{-1}(x)\right) = \bar{F}_{\mathcal{Z}}\left(p^{-1} \circ p(\text{VaR}_{\beta}(\mathcal{Z}))\right) = (1 + \gamma_n)(1 - \beta) \leq 1 - \alpha,$$

implying that $\text{VaR}_{\alpha}(L_n/n) \leq p(\text{VaR}_{\beta}(\mathcal{Z}))$.

Turning to the expected shortfall, notice from (42) that one can obtain tail behavior from the tails of $F_R(\cdot)$. Using the tail estimates from previous subsection, one can obtain estimates for the expected shortfall.

Next we turn to the problem of characterizing the path which leads to the rare event $\{L_n > nx_n\}$, where $x_n \uparrow 1$ as $n \rightarrow \infty$. In Dembo et al. (2004), it is suggested that, if we fix $x_n = x \in (\mu, 1)$ for $\mu > \mathbf{E}[U_1 X_1]$ and condition on $\mathcal{Z} = z$, then $\{L_n/n\}$ attains a high level $x > \mu$ by exponentially shifting the distribution $U_i X_i$, as mandated by the Gibbs conditioning principle. However, Theorem 1 suggests that the event $\{L_n/n > x_n\}$ (or $\{L_n/n > x\}$) is actually triggered by an unusually large value for \mathcal{Z} , and that under this choice of \mathcal{Z} , the event $\{L_n/n > x_n\}$ occurs by following the expected mean trajectory of the increments $\{U_i X_i\}$ similar to the branching process case as discussed in Ney and Vidtyashankar (2008). Thus, it is more natural to focus on the random variable \mathcal{Z} , and to characterize which value of \mathcal{Z} is most likely to occur when the event $\{L_n > nx_n\}$ is observed.

For this purpose, let $\{\Delta_n\}$ be any nondecreasing sequence, and consider $\mathbf{P}(|\mathcal{Z} - z(x_n)| > \Delta_n | L_n > nx_n)$, where $z(x_n) = p^{-1}(x_n)$ describes the threshold, where the conditional event $\{L_n > nx_n | \mathcal{Z} = z\}$ transitions from being a rare event (for $z < z(x_n)$) to a likely event (for $z > z(x_n)$). Then

$$\begin{aligned} & \mathbf{P}\left(\mathcal{Z} > z(x_n) + \Delta_n \mid \frac{L_n}{n} > x_n\right) \\ &= \frac{\mathbf{P}(L_n > nx_n | \mathcal{Z} > z(x_n) + \Delta_n) \mathbf{P}(\mathcal{Z} > z(x_n) + \Delta_n)}{\mathbf{P}(\mathcal{Z} > z(x_n))} \\ &\sim \frac{\mathbf{P}(\mathcal{Z} > z(x_n) + \Delta_n)}{\mathbf{P}(\mathcal{Z} > z(x_n))} \quad \text{as } n \rightarrow \infty, \end{aligned} \tag{44}$$

where the last step follows from Theorem 1. Conversely, applying Hoeffding’s inequality as in the proof of Theorem 1, we also have that

$$\begin{aligned} & \mathbf{P}\left(\mathcal{Z} < z(x_n) - \Delta_n \mid \frac{L_n}{n} > x_n\right) \\ &= \frac{\mathbf{P}(L_n > nx_n | \mathcal{Z} < z(x_n) - \Delta_n) \mathbf{P}(\mathcal{Z} < z(x_n) - \Delta_n)}{\mathbf{P}(\mathcal{Z} > z(x_n))} = o(1) \end{aligned} \tag{45}$$

as $n \rightarrow \infty$, since by reasoning as in (39) (and the discussion following this equation), we obtain under the assumption (22) that

$$\mathbf{P}\left(\frac{L_n}{n} > x_n \mid \mathcal{Z} < z(x_n) - \Delta_n\right) = o(\mathbf{P}(\mathcal{Z} > z(x_n))) \text{ as } n \rightarrow \infty,$$

for any nondecreasing sequence $\{\Delta_n\}$ (and, in fact, the previous equation also holds for a sequence $\{\Delta_n\}$ which decreases to zero). Combining (44) and (45), we arrive at the following.

Proposition 1. *Let $\{\Delta_n\}$ be a nondecreasing sequence, and let $z(x_n) = p^{-1}(x_n)$. Then under the conditions of Theorem 1,*

$$\mathbf{P}\left(\left|\mathcal{Z} - z(x_n)\right| > \Delta_n \mid \frac{L_n}{n} > x_n\right) \sim \frac{\mathbf{P}(\mathcal{Z} > z(x_n) + \Delta_n)}{\mathbf{P}(\mathcal{Z} > z(x_n))} \text{ as } n \rightarrow \infty. \tag{46}$$

The main goal is to show that the right-hand side of (46) converges to zero, and thus to identify $z(x_n)$ as the most likely value of \mathcal{Z} under the rare event $\{L_n > nx_n\}$. However, to show that the right-hand side converges to zero, we need to make a correct choice of Δ_n , which, in turn, will depend on the distribution of \mathcal{Z} . For example, if \mathcal{Z} has a Normal distribution and $\Delta_n = \Delta \in (0, \infty)$ for each n , then one easily calculates that

$$\lim_{n \rightarrow \infty} \frac{\mathbf{P}(\mathcal{Z} > z(x_n) + \Delta)}{\mathbf{P}(\mathcal{Z} > z(x_n))} = 0.$$

However, for distributions with heavier tails, a larger sequence $\{\Delta_n\}$ will be needed to obtain a similar result. For subexponential distributions, it was suggested in Goldie and Resnick (1988) and Asmussen and Collamore (1999) that one normalize according to the so-called auxiliary function. In particular, if \mathcal{Z} is subexponential with auxiliary function $h(\cdot)$, then

$$\mathbf{P}\left(\frac{\mathcal{Z} - y}{h(y)} > \mid \mathcal{Z} > y\right) = \mathbf{P}(W > y), \text{ for any } y,$$

and for some random variable W , where

$$h(y) = \mathbf{E}[\mathcal{Z} - y \mid \mathcal{Z} > y] = \frac{1}{\bar{F}_{\mathcal{Z}}(y)} \int_y^\infty \bar{F}_{\mathcal{Z}}(w) dw.$$

For example, if \mathcal{Z} is regularly varying with parameter α , then $h(x) = x/\alpha$. Then choosing $\Delta_n = \gamma_n h(z(x_n))$ for an arbitrary sequence $\gamma_n \rightarrow \infty$, we obtain that

$$\lim_{n \rightarrow \infty} \frac{\mathbf{P}(\mathcal{Z} - z(x_n) > \gamma_n h(z(x_n)))}{\mathbf{P}(\mathcal{Z} > z(x_n))} \geq \lim_{k \rightarrow \infty} \mathbf{P}(W > \gamma_k) = 0,$$

as desired.

2.3. Tail Approximation in the Small-Default Regime

An alternative formulation introduced in Glasserman et al. (2007) is to create a rare event by letting the level of default (or “distance to default”) tend to infinity along the negative axis, so that default occurs when $Y_i = \sum_{j=1}^k a_j Z_i + b \varepsilon_i$ falls below a threshold d_n (rather than d); that is, we study $X_i^{(n)} = \mathbf{1}_{\{Y_i < d_n\}}$. Since $d_n \downarrow -\infty$, it follows that $\mathbf{P}(Y_i < d_n) \rightarrow 0$ as $n \rightarrow \infty$. As noted previously, this formulation corresponds to the setting of “high quality” credits. It is then natural to adapt the standard large deviation framework and consider

$$\mathbf{P}\left(\frac{L_n}{n} > x\right) \text{ as } n \rightarrow \infty, \text{ for any } x \in (0, 1).$$

As in the previous section, we condition on $\mathcal{Z} = -\sum_{j=1}^d a_j Z_j$, and consider two regions, namely the regions $(-\infty, z_n(x))$ and $(z_n(x), \infty)$, where $z_n(x)$ is again the threshold where, conditional on $\mathcal{Z} = z_n(x)$, the conditional probability that L_n exceeds nx transitions from an “unlikely” to a “likely” event. To make this idea precise, set

$$p_n(z) = \mathbf{P}(Y_i < d_n | \mathcal{Z} = z), \tag{47}$$

where we observe (as in the argument leading to (18)) that with $\tilde{\epsilon} = \text{sgn}(b)\epsilon$,

$$\mathbf{P}(Y_i < d_n | \mathcal{Z} = z) = F_{\tilde{\epsilon}}\left(\frac{d_n + z}{|b|}\right), \tag{48}$$

where $F_{\tilde{\epsilon}}$ may again be replaced by F_{ϵ} if we further assume that F_{ϵ} has a symmetric distribution. Finally, let

$$z_n(x) = p_n^{-1}(x) = |b|F_{\tilde{\epsilon}}^{-1}(x) - d_n. \tag{49}$$

Then

$$x = p_n(z_n(x)) = \mathbf{P}\left(X_i^{(n)} = 1 | \mathcal{Z} = z_n(x)\right), \tag{50}$$

and thus by the law of large numbers for triangular arrays, as $n \rightarrow \infty$,

$$\mathbf{P}\left(\frac{L_n}{n} > x \mid \mathcal{Z} = z\right) \rightarrow 0 \quad \text{for } z < z_n(x),$$

while

$$\mathbf{P}\left(\frac{L_n}{n} > x \mid \mathcal{Z} = z\right) \rightarrow 1 \quad \text{for } z > z_n(x).$$

Consequently, as in (20), we expect that

$$\mathbf{P}\left(\frac{L_n}{n} > x\right) = \int_{\mathbb{R}} \mathbf{P}\left(\frac{L_n}{n} > x_n \mid \mathcal{Z} = z\right) dF_{\mathcal{Z}}(z) \sim \int_{z_n(x)}^{\infty} dF_{\mathcal{Z}}(z), \tag{51}$$

as $n \rightarrow \infty$, provided that

$$\int_{-\infty}^{z_n(x)} \mathbf{P}\left(\frac{L_n}{n} > x \mid \mathcal{Z} = z\right) dF_{\mathcal{Z}}(z) \sim o(\bar{F}_{\mathcal{Z}}(z_n(x))) \quad \text{as } n \rightarrow \infty. \tag{52}$$

Examining the proof of Lemma 1, in particular, we see that the entire argument can be repeated, but with $p_n(x)$ in place of $p(x)$, and with $\epsilon_{i,n}^* := b\epsilon_i - d_n$ in place of $\epsilon_i^* = b\epsilon_i - d$. Repeating the same argument that led to (31), but now with $(p_n(x), \epsilon_{i,n}^*)$ in place of $(p(x), \epsilon_i^*)$, we obtain an analog of (31), namely the requirement that δ_n must be chosen such that

$$\left(\frac{\delta_n}{1 - x + \delta_n}\right) \frac{\lambda_{\mathcal{Z}}(p_n^{-1}(x - \delta_n))}{\lambda_{\epsilon_{i,n}^*}(p_n^{-1}(x - \delta_n))} \leq \frac{1}{2}\gamma_n, \quad n \geq N_0, \tag{53}$$

for some sequence $\gamma_n \downarrow 0$ as $n \rightarrow \infty$. However, at this stage, the situation is a little different than in the previous discussion, since in the denominator of (53), $\epsilon_{i,n}^* = b\epsilon_i - d_n \implies \lambda_{\epsilon_{i,n}^*}(y) = \lambda_{\epsilon}((y + d_n)/b)$, which by (49) implies that

$$\lambda_{\epsilon_{i,n}^*}(p_n^{-1}(x - \delta_n)) = \frac{1}{|b|} \lambda_{\tilde{\epsilon}}(F_{\tilde{\epsilon}}^{-1}(x - \delta_n)),$$

where the main observation is that the term in parentheses on the right-hand side does not tend to infinity as $n \rightarrow \infty$, in contrast to the numerator in (53), where, in particular, we do

have that $p_n^{-1}(x) = |b|F_{\bar{\epsilon}}^{-1}(x) - d_n \uparrow \infty$. Thus, in contrast to Theorem 1, we now require that $\{\delta_n\}$ be chosen such that

$$\delta_n = o\left(\frac{1}{\lambda_{\mathcal{Z}}(|b|F_{\bar{\epsilon}}^{-1}(x) - d_n)} \wedge 1\right) \text{ as } n \rightarrow \infty. \tag{54}$$

Next, turning to the proof of Theorem 1, we see that, once again, the argument can be repeated provided that $n\delta_n^2 \rightarrow \infty$ (in the application of Chebyshev’s inequality), and provided that $\Lambda_{\mathcal{Z}}(p_n^{-1}(x)) = o(n\delta_n^2)$ (in the application of Hoeffding’s inequality). Since the second subsumes the first, we concentrate on the application of Hoeffding’s inequality. To this end, let δ_n be chosen so that (54) holds. Then we need

$$n\delta_n^2 \left(1 - \frac{\Lambda_{\mathcal{Z}}(|b|F_{\bar{\epsilon}}^{-1}(x) - d_n)}{n\delta_n^2}\right) \rightarrow \infty \text{ as } n \rightarrow \infty,$$

which reduces to the requirement that

$$\frac{\Lambda_{\mathcal{Z}}(|b|F_{\bar{\epsilon}}^{-1}(x) - d_n)}{n\delta_n^2} = o(1) \text{ as } n \rightarrow \infty. \tag{55}$$

Now if the minimum in (54) is attained at $\lambda_{\mathcal{Z}}(\cdot)$, then δ_n is chosen such that

$$\delta_n \left(\lambda_{\mathcal{Z}}(|b|F_{\bar{\epsilon}}^{-1}(x) - d_n)\right) = \gamma_n$$

for some sequence $\gamma_n \downarrow 0$. Then

$$\frac{\Lambda_{\mathcal{Z}}(|b|F_{\bar{\epsilon}}^{-1}(x) - d_n)}{n\delta_n^2} = \frac{1}{n\gamma_n^2} \left(\lambda_{\mathcal{Z}}(|b|F_{\bar{\epsilon}}^{-1}(x) - d_n)\right)^2 \Lambda_{\mathcal{Z}}(|b|F_{\bar{\epsilon}}^{-1}(x) - d_n). \tag{56}$$

Since we are free to choose $\{\delta_n\}$, hence $\{\gamma_n\}$ as any sequence which decays to zero, the right-hand side of (56) will be $o(1)$ provided that

$$\frac{1}{n\gamma_n^2} \left(\lambda_{\mathcal{Z}}(|b|F_{\bar{\epsilon}}^{-1}(x) - d_n)\right)^2 \Lambda_{\mathcal{Z}}(|b|F_{\bar{\epsilon}}^{-1}(x) - d_n) = o(1) \text{ as } n \rightarrow \infty. \tag{57}$$

Conversely, if the minimum in (54) is attained at one, then $\delta_n = \gamma_n$ for some $\gamma_n \downarrow 0$, and then

$$\frac{\Lambda_{\mathcal{Z}}(|b|F_{\bar{\epsilon}}^{-1}(x) - d_n)}{n\delta_n^2} = \frac{1}{n\gamma_n^2} \Lambda_{\mathcal{Z}}(|b|F_{\bar{\epsilon}}^{-1}(x) - d_n), \tag{58}$$

so we then need

$$\frac{1}{n} \Lambda_{\mathcal{Z}}(|b|F_{\bar{\epsilon}}^{-1}(x) - d_n) = o(1) \text{ as } n \rightarrow \infty. \tag{59}$$

This leads to the requirement that

$$\frac{1}{n} \left\{ \left(\lambda_{\mathcal{Z}}(|b|F_{\bar{\epsilon}}^{-1}(x) - d_n)\right)^2 \vee 1 \right\} \Lambda_{\mathcal{Z}}(|b|F_{\bar{\epsilon}}^{-1}(x) - d_n) = o(1) \text{ as } n \rightarrow \infty, \tag{60}$$

which, we observe, depends only on \mathcal{Z} and not on ϵ . This leads to the following theorem.

Theorem 2. Let $x \in (0, 1)$. Assume that the density $f_{\mathcal{Z}}(z)$ is nonincreasing for $z \geq M$, for some $M < \infty$, and assume that λ_{ϵ} is finite everywhere. Furthermore, let d_n be chosen such that (60) holds. Then

$$\lim_{n \rightarrow \infty} \frac{\mathbf{P}(L_n > nx)}{\bar{F}_{\mathcal{Z}}(|b|F_{\bar{\epsilon}}^{-1}(x) - d_n)} = 1. \tag{61}$$

As with the previous theorem, the main technical condition (given in (60)) is widely satisfied, but again needs to be verified on a case-by-case basis. For example, if the common factors are mutually independent with a symmetric two-sided Pareto(α, β) distribution, then $Z_j^+ = Z_j \vee 0$ is subexponential, and, as before, it follows from a slight modification of Albrecher and Asmussen (2010, Proposition IX.1.9), that

$$\mathbf{P}(\mathcal{Z} > u) \sim \sum_{j=1}^k \mathbf{P}\left(Z_j^+ > \frac{u}{|a_j|}\right) \sim \frac{\beta}{2}(|a_1|^\alpha + \dots + |a_k|^\alpha)u^{-\alpha} \quad \text{as } u \rightarrow \infty.$$

Since $\lambda_{\mathcal{Z}}(y) \sim B/y$ as $y \rightarrow \infty$ in the Pareto case, the maximum in (60) is attained at one for large n . Then (60) requires that $\log(|d_n|^\alpha)/n \rightarrow 0$ as $n \rightarrow \infty$, or $(\log |d_n|)/n \rightarrow 0$. Then Theorem 2 leads to the estimate

$$\mathbf{P}\left(\frac{L_n}{n} > x\right) \sim \frac{\beta}{2}(|a_1|^\alpha + \dots + |a_k|^\alpha)|d_n|^{-\alpha} \quad \text{as } n \rightarrow \infty, \tag{62}$$

which is asymptotically independent of $x \in (0, 1)$. It is worth noticing here that $d_n \downarrow -\infty$ at an exponential rate; this allows for a large range of values to be classified as “good credit” when the common factors are from a two-sided Pareto distribution.

The case of Gaussian factors is actually more intricate. Now suppose that the common factors Z_1, \dots, Z_k are independent and have the standard Gaussian distribution, so that $\mathcal{Z} = -(a_1Z_1 + \dots + a_kZ_k) \sim \text{Normal}(0, \|a\|^2)$, where $\|a\|^2 = a_1^2 + \dots + a_k^2$. Since the hazard function of a Normal random variable grows linearly in the asymptotic limit, (60) requires that $d_n^2 \bar{F}_{\mathcal{Z}}(\text{const.} + d_n) \sim d_n^2 \cdot \text{const.} \cdot (\log |d_n|)^{-1/2} d_n^2 = o(n)$ as $n \rightarrow \infty$, which is satisfied for any sequence $\{|d_n|\}$ which increases to infinity and satisfies $|d_n| \leq n^{1/4}$. For this choice of d_n , Theorem 2 yields that as $n \rightarrow \infty$,

$$\mathbf{P}\left(\frac{L_n}{n} > x\right) \sim \frac{C}{\|a\|y_n} e^{-y_n^2/2\|a\|^2}, \quad \text{where } y_n := bF_e^{-1}(x) + d_n, \tag{63}$$

and C is a constant.

We observe that, if one were to settle for logarithmic asymptotics rather than sharp asymptotics, then in the proof of Theorem 2 for the Gaussian case, one could take $|d_n| = s\sqrt{n}$ for $s \in (0, 1]$, and $\delta_n = 1$. Then the proof in Theorem 1 applies on the logarithmic scale, provided that a logarithmic analog of Lemma 1 holds, namely

$$\frac{1}{n} \log\left(\bar{F}_{\mathcal{Z}}(p_n^{-1}(x+1)) / \bar{F}_{\mathcal{Z}}(p_n^{-1}(x-1))\right) \rightarrow -\infty$$

as $n \rightarrow \infty$. Using that $p_n^{-1}(v) = bF_e^{-1}(v) - s\sqrt{n}$ for $|d_n| = s\sqrt{n}$, we then obtain by direct calculation that the left-hand side of the previous equation decays as

$$\frac{1}{n} \log\left(e^{-n(1-\frac{b}{s\sqrt{n}})} / e^{-n(1+\frac{b}{s\sqrt{n}})}\right) \rightarrow -\infty \quad \text{as } n \rightarrow \infty,$$

for some positive constant b . Consequently we obtain the weaker logarithmic analog of (63), namely

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbf{P}\left(\frac{L_n}{n} > x\right) = -\frac{s^2}{2\|a\|^2}, \tag{64}$$

which is Theorem 2.1 of Glasserman et al. (2007) in the single-type case. As discussed previously, when $\mathbf{Z} = (Z_1, \dots, Z_k)$ is a set of dependent common factors with covariance matrix Σ , one obtains

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbf{P}\left(\frac{L_n}{n} > x\right) = -\frac{s^2}{2\mathbf{a}'\Sigma\mathbf{a}}, \tag{65}$$

which again illustrates the role of the dependence in this asymptotic estimate. It is pertinent to notice here that there is a gap in the range of values of $\{d_n\}$ between the sharp asymptotics and logarithmic asymptotics. Thus, the logarithmic asymptotics cover a wider range of higher quality credits than what can be concluded from Theorem 2.

Referring back to the Pareto example, note that we could also introduce dependence in this model by considering an elliptical distribution for $\mathbf{Z} = (Z_1, \dots, Z_k)$. Then the dependence is characterized by the dispersion matrix, rather than the covariance matrix, and this will then play a role in the asymptotic decay, similar to what is seen in (65), but with a polynomial decay rather than exponential decay. Of course, further dependence structures amongst the common factors could also be introduced using copulas. For more details on elliptical distributions and copulas, see McNeil et al. (2015).

3. Some Extensions and Refinements

3.1. Multiple Types

In realistic problems, each loan will have its own default threshold d (the distance to default), and there may be differences in the constants $\mathbf{a} = (a_1, \dots, a_k)$ and b among the different loans. To address this problem, it is customary to divide the loans into different classes (or “types”), where the constants d , \mathbf{a} and b are fixed within a given class. More precisely, suppose that there are m types of loans, where the number of loans of each type are given by n_1, \dots, n_m , respectively. We assume that $\lim_{n \rightarrow \infty} (n_j/n) = \kappa_j \in (0, 1)$. Then, for all loans of type j ($j \in \{1, \dots, m\}$), set

$$Y_i^{(j)} = \langle \mathbf{a}_j, \mathbf{Z} \rangle + b_j \epsilon_i^{(j)}, \quad i = 1, \dots, n_j,$$

where $\mathbf{Z} = (Z_1, \dots, Z_k)$ is the vector of common factors described in the previous sections, $\mathbf{a}_j \in \mathbb{R}^k$ and $b_j \in \mathbb{R}$. Finally, assume that default occurs if $Y_i^{(j)} < d_j$, where $d_j \in \mathbb{R}$. Let $X_i^{(j)} = \mathbf{1}_{\{Y_i^{(j)} < d_j\}}$, and set

$$L_n^{(j)} = \sum_{i=1}^{n_j} U_i X_i^{(j)}, \quad i = 1, \dots, n_j,$$

where, for simplicity, we assume that $\{U_i\}$ is an i.i.d. sequence independent of j (i.e., the same among all types) and $\mathbf{E}[U_i] = 1$. Once again, our objective is to study $\mathbf{P}(L_n > nx_n)$ as $n \rightarrow \infty$, but where we now have $L_n = L_n^{(1)} + \dots + L_n^{(m)}$.

As in the previous section, we seek to identify those choices of \mathbf{Z} for which, conditional on \mathbf{Z} ,

$$\mathbf{P}\left(\frac{L_n}{n} > x_n \mid \mathbf{Z}\right) \rightarrow 1,$$

where we have kept the conditioning on all of the factors Z_1, \dots, Z_k (rather than on $\mathcal{Z} = -(a_1 Z_1 + \dots + a_k Z_k)$), since the values of \mathbf{Z} satisfying the previous equation will lie in different regions for different j , and the problem is genuinely multidimensional.

First notice that by the Vacicek’s law of large numbers applied to each type,

$$\frac{L_{n_j}(\mathbf{Z})}{n_j} \rightarrow p_j(\mathbf{Z}) \quad \text{in probability,}$$

where, for any $\mathbf{z} \in \mathbb{R}^d$,

$$p_j(\mathbf{z}) = \mathbf{E}\left[U_i X_i^{(j)}\right] = \mathbf{P}\left(X_i^{(j)} = 1 \mid \mathbf{Z} = \mathbf{z}\right) = F_{\epsilon_j}\left(\frac{d_j - \langle \mathbf{a}_j, \mathbf{z} \rangle}{b_j}\right).$$

Hence, as $n \rightarrow \infty$,

$$\frac{L_n(\mathbf{Z})}{n} \rightarrow p(\mathbf{Z}) := \sum_{j=1}^m \kappa_j p_j(\mathbf{Z}) \quad \text{in probability.} \tag{66}$$

Now, for a fixed x , define a region

$$G(x) = \left\{ \mathbf{z} \in \mathbb{R}^k : p(\mathbf{Z}) > x \right\}, \quad x \in [0, 1].$$

This set describes the values of \mathbf{z} where $(L_n(\mathbf{z})/n) \rightarrow y > x$. Then for any $x \in [0, 1]$, we have that $\mathbf{P}(L_n > nx | \mathbf{Z} = \mathbf{z}) \rightarrow 0$ for $\mathbf{z} \in (G(x))^c$, while $\mathbf{P}(L_n > nx | \mathbf{Z} = \mathbf{z}) \rightarrow 1$ for $\mathbf{z} \in G(x)$. Now let $\{x_n\}$ be an increasing sequence such that $x_n \uparrow 1$ as $n \rightarrow \infty$, and let ν denote the probability measure of \mathbf{Z} . Then arguing as in the beginning of Section 2.1, we have by a simple conditioning argument and an application of Chebyshev’s inequality (as in the proof of Theorem 1) that

$$\mathbf{P}\left(\frac{L_n}{n} > x_n\right) = \int_{\mathbb{R}^d} \mathbf{P}\left(\frac{L_n}{n} > x_n \mid \mathbf{Z} = \mathbf{z}\right) d\nu(\mathbf{z}) \sim \int_{G(x_n)} d\nu(\mathbf{z}), \tag{67}$$

as $n \rightarrow \infty$, provided that

$$\int_{(G(x_n))^c} \mathbf{P}\left(\frac{L_n}{n} > x_n \mid \mathbf{Z} = \mathbf{z}\right) d\nu(\mathbf{z}) = o(\nu(G(x_n))) \quad \text{as } n \rightarrow \infty. \tag{68}$$

To develop an analog of Lemma 1, observe that the change in the function $p(\mathbf{z}) := \sum_{j=1}^m \kappa_j p_j(\mathbf{z})$ is determined by its partial derivatives, namely

$$\frac{\partial}{\partial z_i} p(\mathbf{z}) = \sum_{j=1}^m \kappa_j f'_{\epsilon_j} \left(\frac{d_j - \langle \mathbf{a}_j, \mathbf{z} \rangle}{b_j} \right) \left(-\frac{a_{ij}}{b_j} \right) \quad \text{for } \mathbf{a}_j = (a_{1j}, \dots, a_{kj}). \tag{69}$$

Thus, a version of Lemma 1 can be established, but with the density function of a single variable replaced with the elements of (69) reflected along the different directional vectors, while the tail probabilities are replaced by the decay of $p(\mathbf{z})$ in the various directions. This leads to the estimate

$$\mathbf{P}\left(\frac{L_n}{n} > x_n\right) \sim \nu(G(x_n)) \quad \text{as } n \rightarrow \infty, \tag{70}$$

where ν is the probability measure of \mathbf{Z} , and the next step is to show that on the right-hand side of the above expression,

$$\nu(G(x_n)) \sim \nu\left(\bigcap_{j=1}^m G_j(x_n)\right) \quad \text{as } n \rightarrow \infty,$$

where $G_j(x) := \left\{ \mathbf{z} \in \mathbb{R}^k : p_j(\mathbf{Z}) > x \right\}$ (so the default occurs for every type). The details are addressed in a forthcoming work.

Another issue which may be useful in applications is to incorporate migration probabilities in the decision making. Allowing the factor model to change between rating classes (and potentially allowing U_i to depend on j), where the index j above may be viewed as a rating class, one could obtain the tail estimates using

$$\sum_{j=1}^r \mathbf{P}_k \left(\frac{L_n^{(j)}}{n} > x_n; R = j \right) \mathbf{P}_k(R = r). \tag{71}$$

In the above expression, k is the initial class and j is the terminal class and $\mathbf{P}_k(R = r)$ represents the migration probabilities. Alternatively, one could study the tail probability associated with the *entire* portfolio under migration. Namely, allow the proportion of loans in class j to change from κ_j to κ'_j over a fixed time interval. Let $\mathbf{r} = (\kappa_1, \dots, \kappa_m)$ denote the proportion at the initial time, and $\mathbf{r}' = (\kappa'_1, \dots, \kappa'_m)$ the proportion at the end of the time interval, and let $\mathbf{P}(\mathbf{r}, d\mathbf{r}')$ denote the transition probability from state \mathbf{r} to \mathbf{r}' . Then, letting $L_n(\mathbf{r}')$ denote the losses under \mathbf{r}' , one could obtain tail estimate for the default probability after migration using

$$\int_{\mathbf{r}'} \mathbf{P}\left(\frac{L_n(\mathbf{r}')}{n} > x_n\right) \mathbf{P}(\mathbf{r}, d\mathbf{r}').$$

3.2. Tail Asymptotics for Divergent Number of Factors

In this section, we provide tail approximations in the large loss regime when the number of factors, k_n , diverges to infinity. Specifically, we consider the model

$$X_i^{(n)} = a_1^{(n)}Z_1^{(n)} + a_2^{(n)}Z_2^{(n)} + \dots + a_{k_n}^{(n)}Z_{k_n}^{(n)} + b^{(n)}\epsilon_i^{(n)}, \quad i = 1, 2, \dots, n, \tag{72}$$

where $\{a_j^{(n)} : j = 1, \dots, k_n\}$ and $b^{(n)}$ are constants satisfying

$$\sum_{j=1}^{k_n} (a_j^{(n)})^2 + (b^{(n)})^2 = 1, \text{ for all } n \geq 1. \tag{73}$$

Let $\mathcal{Z}_{k_n} = -\sum_{j=1}^{k_n} a_j^{(n)}Z_j^{(n)}$. In this context, the definition of the loss function L_n requires a slight modification. Following the discussions in the previous sections, the loss can be expressed as

$$L_n = \sum_{i=1}^n U_i^{(n)} X_i^{(n)}, \tag{74}$$

where $X_i^{(n)} = \mathbf{1}_{\{Y_i^{(n)} < d\}}$. Furthermore, let $\{U_i^{(n)} : 1 \leq i \leq n\}$ be i.i.d. with $\mathbf{E}[U_i^{(n)}] = 1$. We observe that conditional on \mathcal{Z}_{k_n} , the total loss is given by

$$L_n(\mathcal{Z}_{k_n}) = \sum_{i=1}^n U_i^{(n)} \mathbf{1}_{\{b^{(n)}\epsilon_i^{(n)} < d + \mathcal{Z}_{k_n}\}}. \tag{75}$$

Assuming that $\mathcal{Z}_\infty = \lim_{n \rightarrow \infty} \mathcal{Z}_{k_n}$ in probability, it follows by an application of the conditional Chebychev inequality that $\{L_n(\mathcal{Z}_{k_n})/n\}$ converges to $p(\mathcal{Z}_\infty)$ in probability. Now, the rate of the convergence in the law of large numbers will depend on the rate of increase of k_n relative to n . There are multiple ways in which this relative increase can occur. One approach involves investigating the behavior of $\mathbf{P}(\mathcal{Z}_{k_n} > p^{-1}(x_n))$, as before. Under mild conditions as in [Glasserman et al. \(2007\)](#), the above probability can be approximated, as in [Theorem 1](#), by $\mathbf{P}(\mathcal{Z}_\infty > p^{-1}(x_n))$. We state this formally as a theorem, whose proof can, in principle, be constructed along the lines of [Theorem 1](#). However, we state a logarithmic-level result, whose proof can be found in the thesis of [de Silva \(2016\)](#) (where an m -type extension is also given).

Theorem 3. Assume that $\{Z_j^{(n)} : j = 1, 2, \dots, k_n, n \geq 1\}$ and are i.i.d. standard Gaussian random variables and are mutually independent. Additionally, assume that

$$\liminf_{n \rightarrow \infty} \sum_{j=1}^{k_n} (a_j^{(n)})^2 > 0; \quad \limsup_{n \rightarrow \infty} \frac{1}{\log n} \sum_{j=1}^{k_n} (a_j^{(n)})^2 < 1;$$

$$\text{and} \quad \lim_{n \rightarrow \infty} \frac{b_n^2}{\sum_{j=1}^{k_n} (a_j^{(n)})^2} = \gamma^2, \quad 0 < \gamma < 1.$$

Then, for any $0 < s < 1$,

$$\lim_{n \rightarrow \infty} \frac{1}{\log n} \log \mathbb{P}(L_n > sn \log n) = -s^2 \gamma^2. \quad (76)$$

The conditions of the theorem essentially reduce the problem to a single-type factor model, and sharper asymptotics can be described. However, more interesting situations arise when these conditions are violated. For instance, if $p^{-1}(x_n) \sim k_n$ as $n \rightarrow \infty$, then under Gärtner–Ellis type conditions from large deviation theory, the rate function of the common factors and the tails of ϵ will play a role. These and other interesting extensions, such as when the number of types also diverges to infinity, are studied in a forthcoming work. Furthermore, another issue concerning model uncertainty is also addressed in a forthcoming work.

Author Contributions: Conceptualization, A.N.V.; Methodology, J.F.C., H.d.S., A.N.V.; Formal analysis, J.F.C., H.d.S., A.N.V.; Investigation, J.F.C., H.d.S., A.N.V.; Writing—original draft, J.F.C., A.N.V.; Writing—review and editing, J.F.C., A.N.V.; Supervision, A.N.V. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Asmussen, Søren, and Hansjörg Albrecher. 2010. *Ruin Probabilities*, 2nd ed. Singapore: World Scientific Publishing Co.
- Asmussen, Søren, and Jeffrey F. Collamore. 1999. Exact asymptotics for a large deviations problem for the GI/G/1 queue. *Markov Processes and Related Fields* 5: 451–76.
- Burtschell, Xavier, Jonathan Gregory, and Jean-Paul Laurent. 2009. A comparative analysis of CDO pricing models under the factor copula framework. *The Journal of Derivatives* 16: 9–37. [CrossRef]
- Bush, Nick, Ben M. Hambly, Helen Haworth, Lei Jin, and Christoph Reisinger. 2011. Stochastic evolution equations in portfolio credit modelling. *SIAM Journal on Financial Mathematics* 2: 627–64. [CrossRef]
- Chow, Yuan Shih, and Henry Teicher. 1997. *Probability Theory. Independence, Interchangeability, Martingales*, 3rd ed. New York: Springer.
- de Silva, Hasitha. 2016. Large Deviations and Rare Event Simulation for Portfolio Credit Risk. Ph.D. thesis, George Mason University, Fairfax, VA, USA.
- Dembo, Amir, Jean-Dominique Deuschel, and Darrell Duffie. 2004. Large portfolio losses. *Finance and Stochastics* 8: 3–16. [CrossRef]
- Glasserman, Paul, Wanmo Kang, and Perwez Shahabuddin. 2007. Large deviations in multifactor portfolio credit risk. *Mathematical Finance* 17: 345–79. [CrossRef]
- Goldie, Charles M., and Sidney Resnick. 1988. Distributions that are both subexponential and in the domain of attraction of an extreme-value distribution. *Advances in Applied Probability* 20: 706–18. [CrossRef]
- Gordy, Michael B. 2003. A risk-factor model foundation for ratings-based bank capital rules. *Journal of Financial Intermediation* 12: 199–232. [CrossRef]
- Hull, John C., and Alan D. White. 2004. Valuation of a CDO and an n-th to default CDS without Monte Carlo simulation. *The Journal of Derivatives* 12: 8–23. [CrossRef]
- Maier, Ramona, and Mario V. Wüthrich. 2009. Law of large numbers and large deviations for dependent risks. *Quantitative Finance* 9: 207–15. [CrossRef]
- McNeil, Alexander J., Rüdiger Frey, and Paul Embrechts. 2015. *Quantitative Risk Management: Concepts, Techniques and Tools*, Revised ed. Princeton: Princeton University Press.
- Merton, Robert C. 1974. On the pricing of corporate debt: The risk structure of interest rates. *The Journal of Finance* 29: 449–70.
- Ney, P. E., and Vidyashankar, A. N. 2008. Conditional limit laws and inference for generation sizes of branching processes. In *Markov Processes and Related Topics: A Festschrift for Thomas G. Kurtz*. Toronto: Institute of Mathematical Statistics, Volume 4, pp. 17–31.

Schönbucher, Philipp J. 2000. *Factor Models for Portfolio Credit Risk*. Bonn: University of Bonn.

Vasicek, Oldrich A. 1991. Limiting loan loss probability distribution. In *Finance, Economics and Mathematics*. Hoboken: Wiley, pp. 147–48.