Asymptotic Analysis of the Loss Given Default in the Presence of Multivariate Regular Variation

Qihe Tang† and Zhongyi Yuan‡∗
†Department of Statistics and Actuarial Science, University of Iowa
241 Schaeffer Hall, Iowa City, IA 52242, USA
E-mail: qihe-tang@uiowa.edu
‡Department of Risk Management, Pennsylvania State University
362 Business Building, University Park, PA 16802, USA
E-mail: ZhongyiYuan@psu.edu

September 5, 2013

Abstract

Consider a portfolio of n obligors subject to possible default. We propose a new structural model for the loss given default, which takes into account the severity of default. Then we study the tail behavior of the loss given default under the assumption that the losses of the n obligors jointly follow a multivariate regular variation structure. This structure provides an ideal framework for modeling both heavy tails and asymptotic dependence. Multivariate models involving Archimedean copulas and mixtures are revisited. As applications, we derive asymptotic estimates for the value at risk and conditional tail expectation of the loss given default and compare them with the traditional empirical estimates.

Keywords: asymptotic analysis; conditional tail expectation; extreme risks; large joint movements; loss given default; multivariate regular variation; value at risk

2010 MSC: Primary 91G40; Secondary 62P05; 60G70; 62H20

1 Introduction

1.1 Loss given default

Consider an obligor whose net loss over a fixed holding period, denoted by a real-valued random variable \( X \), drives its credit rating changes and default. Traditionally, the loss given default (LGD) due to the obligor is described as

\[
ed^\delta 1_{(X>a)};
\]

\( \delta \geq 0 \) is a constant, \( a \) is the threshold, and \( 1_{(X>a)} \) is the indicator function of the event \( X > a \).

*Corresponding author: Zhongyi Yuan; Tel.: 1-814-865-6211.
where \( e \) denotes the total exposure at default, \( \delta \in [0, 1] \) denotes the LGD as a percentage of the total exposure, \( a \) denotes the default threshold with default occurring when the loss \( X \) is greater than \( a \), and \( 1_E \) is the indicator function of an event \( E \). Such a threshold model descends from Merton’s (1974) firm-value model and has been widely used in practice. The exposure at default \( e \), the LGD rate \( \delta \), and the default probability \( \Pr (X > a) \) are key parameters to be estimated for measuring credit risk under certain internal ratings-based frameworks. Related discussions can be found in Schuermann (2004) and McNeil et al. (2005), among many others.

An obvious drawback of this traditional model is that, as in many other credit risk models, the LGD rate \( \delta \) is assumed to be deterministic, which ignores the fact that the LGD fluctuates with economic cycles. To fix such a drawback, Gupton et al. (1997, Chapter 7) and Gupton and Stein (2002) proposed to use a beta distribution to model the recovery rate \( 1 - \delta \). Schuermann (2004) suggested that the distribution of the recovery rate contain two modes. Recently, Farinelli and Shkolnikov (2012) gave a short review of this issue and they proposed two structural models for the LGD to be stochastic and to account for correlations between default events and associated losses. See also Acharya et al. (2003) and Chava et al. (2011), among others, for related discussions on the variability of recovery rates.

However, the strong tie of the LGD rate \( \delta \) to the severity of default has not been sufficiently emphasized. The severity of default of the obligor is described as

\[
S = \frac{X}{a} - 1,
\]

with \( S \leq 0 \) meaning no default, \( S > 0 \) meaning default, and a larger value of \( S \) meaning a more severe default. To take into account the severity of default, it becomes very natural to introduce a nondecreasing function \( G \) on \((-\infty, \infty)\), called loss settlement function, with \( G(s) = 0 \) for \( s \leq 0 \) and \( G(\infty) = 1 \), to define the LGD as a percentage of the total exposure. The limit \( G(\infty) = 1 \) corresponds to a catastrophic situation where no value is left to recover. As a result, the LGD due to the obligor becomes

\[
eG \left( \frac{X}{a} - 1 \right). \tag{1.2}
\]

In this model, the loss variable \( X \) drives the default and the loss settlement function \( G \) determines the pattern of settling the LGD. Clearly, the traditional model of LGD in (1.1) corresponds to a limit case, as \( t \to \infty \), of (1.2) with the loss settlement function being \( G(s) = \delta 1_{(0<s\leq t)} + 1_{(s>t)} \). To be consistent with the tradition of this study, the LGD in (1.2) can be rewritten as \( eG \left( \frac{X}{a} - 1 \right) 1_{(X/a - 1 > 0)} \), where the indicator function serves as the default indicator. However, this further decomposition is not necessary since we have assumed that \( G(s) = 0 \) for \( s \leq 0 \).

The idea of introducing such a loss settlement function can be found scattered in the literature of credit risk. For example, Pykhtin (2003) and Tasche (2004) implicitly employed
the idea in their discussions and even suggested an exponential loss settlement function. More realistically, even with the loss variable $X$ given, the LGD should still be stochastic due to its dependence also on the external economy environment. This can be achieved by, for example, introducing another source of randomness to represent the external economy environment and letting it be another argument of the loss settlement function $G$. The reader is referred to van Damme (2011) for some related discussions. We shall not pursue such an extension in this work, but we would like to point out that the methodologies developed here will largely work for the extension.

Based on this idea, we introduce below a static structural model for the LGD. Consider a portfolio of $n$ obligors such as loans, corporate bonds, and other instruments subject to possible default. The net loss of obligor $i$ over a fixed holding period is denoted by a real-valued random variable $X_i$. Obligor $i$ defaults once the loss $X_i$ is greater than some default threshold. For each $i = 1, \ldots, n$, the same as above, introduce a loss settlement function $G_i$ on $(−\infty, \infty)$ that is nondecreasing with $G_i(s) = 0$ for $s \leq 0$ and $G_i(\infty) = 1$. Then we model the LGD of the portfolio as

$$L = \sum_{i=1}^{n} e_i G_i \left( \frac{X_i}{a_i} - 1 \right), \quad (1.3)$$

where $e_1, \ldots, e_n$ are $n$ positive constants denoting scaled exposures, so that

$$\sum_{i=1}^{n} e_i = 1, \quad (1.4)$$

and $a_1, \ldots, a_n$ are another $n$ positive constants denoting default thresholds, specified according to exogenously given default probabilities.

### 1.2 Modeling loss variables

We are interested in the heavy-tailed case. In particular, we shall assume that the loss variables $X_1, \ldots, X_n$ follow Pareto-type distributions. It is generally acknowledged that loss variables should be mutually dependent because the obligors are exposed to common or similar macroeconomic factors. The intricate dependence structure is a main cause for credit contagion. In this subsection, we focus on the issue of modeling dependence among the loss variables.

An overwhelming approach of modeling the dependence is to use Gaussian copulas, which underlie the well-known CreditMetrics and Moody’s KMV models. However, recent empirical studies often reveal large joint movements among loss variables, resulting in a much more dangerous situation than described by Gaussian copulas. In this regard, we refer the reader to a colloquial-style article by Salmon (2012). Non-Gaussian copulas such as $t$ copulas have lately been introduced to credit risk modeling by, e.g. Frey et al. (2001), Frey and McNeil (2003), and Schloegl and O’Kane (2005).
Another prevailing practice in credit risk modeling is to introduce common systematic risk factors for all obligors and idiosyncratic risk factors for individual obligors, and then to model individual losses as functions of these risk factors. Formally, assume the stochastic structure

$$X_i = h_i(\eta_0, \eta_i), \quad i = 1, \ldots, n,$$

(1.5)

where each $h_i$ is a multivariate componentwise monotone function and $\eta_0, \eta_1, \ldots, \eta_n$ are independent random vectors. The vector $\eta_0$ consists of the systematic risk factors while the vectors $\eta_1, \ldots, \eta_n$ consist of the idiosyncratic risk factors not explained by the systematic risk factors. Such a model can easily capture large joint movements among the loss variables $X_1, \ldots, X_n$. A lot of popular models such as CreditMetrics, Moody’s KMV, $t$ copulas, and normal mean-variance mixtures can be retrieved from (1.5).

In this paper, we employ a different approach. We assume that the loss variables $X_1, \ldots, X_n$ jointly follow a multivariate regular variation (MRV) structure; see Section 3 for its definition. As demonstrated by Embrechts et al. (2009), the MRV structure yields an ideal modeling environment for quantitative risk management. In particular, it provides a unified framework for modeling both heavy tails of marginal distributions and asymptotic dependence at varying degree among loss variables. This framework has been extensively applied to insurance, finance, and risk management; in this regard, recent works include Mikosch (2003), Böcker and Klüppelberg (2010), Joe and Li (2011), Asimit et al. (2011), and Hua (2012). As remarked by McNeil et al. (2005, Section 8.1.2), the lack of public information and data is a substantial obstacle to the use of statistical methods in credit risk. From this point of view, the MRV framework has an advantage of not requiring a complete joint distribution as it targets the tail area only.

### 1.3 Our goals

Our main focus is on the behavior of $L$ approaching its upper endpoint $1$. This upper endpoint should be understood correctly. As shown by (1.4), it represents the sum of scaled exposures, which may practically correspond to a huge amount of potential losses. Thus, we must be very cautious with the behavior of $L$ approaching its catastrophic edge $1$. For this purpose, let us define

$$K = \frac{1}{1 - L},$$

(1.6)

which is a continuous and strictly increasing function of $L$ and, hence, can be thought of as a proxy for $L$. Actually,

$$\Pr \left( L > 1 - \frac{1}{x} \right) = \Pr ( K > x ), \quad x > 0,$$

which converts and magnifies the behavior of $L$ approaching $1$ to that of $K$ exploding to $\infty$. In view of the fact that the latter is more visible and easier to depict than the former, we
shall instead study the tail behavior of $K$ at $\infty$.

For the loss vector $X = (X_1, \ldots, X_n)$ following a general MRV framework, we derive an exact asymptotic formula for the tail probability of $K$, which unifies the cases with and without large joint movements among the loss variables. For the special cases that $X$ possesses an Archimedean copula and that $X$ follows a mixture structure, we improve the asymptotic formula to be more explicit. As applications, we establish asymptotic estimates for the value at risk (VaR) and conditional tail expectation (CTE) of $K$ (hence, for those of $L$) for a confidence level $q$ close to 1. In view of the fact that the traditional empirical estimates for the VaR and CTE need a large sample size when the confidence level $q$ is close to 1, our asymptotic estimates have apparent advantages in this situation.

In the study of credit risk, extreme risks used to be significantly underestimated because they are far out of the ordinary range of observations and the tail dependence is usually not correctly modeled. While extreme events occur with a relatively small probability, they usually lead to catastrophic economic and social consequences. The recent global recession triggered by the collapse of the sub-prime mortgage market in the United States in 2008 can be deemed as such a consequence of underestimating extreme risks. The reader is referred to Embrechts et al. (1998), Taleb (2007), and Salmon (2012) for similar remarks.

Fortunately, the study of extreme risks has recently attracted increasing attention in insurance, finance, and, in particular, risk management. Embrechts et al. (1999) presented a short review of extreme value theory as an important risk management tool for insurance and finance. Mashal et al. (2003) analyzed the impact of extreme events on the fair values and risk measures of popular multi-name credit derivatives. Based on a special case of model (1.5), Bassamboo et al. (2008) studied the asymptotic behavior of portfolio credit risk and analyzed implications of extreme dependence among obligors. Chavez-Demoulin and Embrechts (2011) gave a wealth of motivating discussions on the relevance of extreme value theory in the study of credit risk.

The rest of this paper consists of seven sections. After listing notational conventions in Section 2 and introducing the concept of MRV in Section 3, we present our main result for losses following a general MRV structure in Section 4 and examine two special cases involving Archimedean copulas and mixtures in Section 5. Then we propose applications of the study to estimating the VaR and CTE of the LGD in Section 6 and conduct numerical studies in Section 7. Finally, Section 8 contains concluding remarks.

## 2 Notational Conventions

For $x, y \in (-\infty, \infty)$, we write $x \vee y = \max\{x, y\}$, $x \wedge y = \min\{x, y\}$, and $x^+ = x \vee 0$. For $x, y \in (-\infty, \infty)^n$, we write $x^+ = (x_1^+, \ldots, x_n^+)$ and understand relations such as $x > y$, $x \leq y$, and $x \pm y$ as componentwise. As usual, $1/\infty$ is understood as 0. For a real number
For two positive functions \( h_1 \) and \( h_2 \), write \( h_1 \lesssim h_2 \) or \( h_2 \gtrsim h_1 \) if \( \limsup h_1 / h_2 \leq 1 \), and write \( h_1 \sim h_2 \) if \( \lim h_1 / h_2 = 1 \).

For a nondecreasing function \( h \) on \((-\infty, \infty)\), define

\[
h^\leftarrow(y) = \inf \{ x \in (-\infty, \infty) : h(x) \geq y \} \quad \text{and} \quad h^\rightarrow(y) = \sup \{ x \in (-\infty, \infty) : h(x) \leq y \}
\]
as the càglàd and càdlàg inverses of \( h \), respectively, where we follow the usual conventions \( \inf \emptyset = \infty \) and \( \sup \emptyset = -\infty \). If \( h \) is continuous and strictly increasing, then both \( h^\leftarrow \) and \( h^\rightarrow \) coincide with the usual inverse \( h^{-1} \) of \( h \). Clearly, for every fixed \( y \in (-\infty, \infty) \), we have

\[
\{ x : x > h^\rightarrow(y) \} \subseteq \{ x : h(x) > y \} \subseteq \{ x : h(x) \geq y \} \subseteq \{ x : x \geq h^\leftarrow(y) \}.
\]

For limit relationships according to \( x \to \infty \) (mainly in Sections 3–5) or \( q \uparrow 1 \) (mainly in Section 6), we often omit the limit procedure.

## 3 Multivariate Regular Variation

### 3.1 Definition

First we recall the concept of regular variation. A positive measurable function \( h \) on \([0, \infty)\) is said to be regularly varying at \( x = 0+ \) or \( \infty \) with regularity index \( r \in (-\infty, \infty) \), written as \( h \in \text{RV}_r(x_0) \), if

\[
\lim_{x \to x_0} \frac{h(xy)}{h(x)} = y^r, \quad y > 0.
\]

We often suppress the argument \( x_0 \) in \( \text{RV}_r(x_0) \) if there is no confusion in the context. When \( r = 0 \), the function \( h \) is said to be slowly varying at \( x_0 \). The reader is referred to Bingham et al. (1987) or Resnick (1987) for a comprehensive study of regular variation. It is straightforward to extend the definition to \( r = \pm \infty \) to cover the rapid variation case. For example, we have \( h \in \text{RV}_{-\infty}(\infty) \) if

\[
\lim_{x \to \infty} \frac{h(xy)}{h(x)} = 0, \quad y > 1.
\]

Next let us review the concept of vague convergence. Consider an \( n \)-dimensional cone \([0, \infty)^n \setminus \{0\}\) equipped with a Borel sigma-field \( \mathcal{B} \). A measure on the cone is called Radon if its value is finite for every compact set in \( \mathcal{B} \). For a sequence of Radon measures \( \{\mu, \mu_k, k = 1, 2, \ldots\} \) on \([0, \infty)^n \setminus \{0\}\), we say that \( \mu_k \) vaguely converges to \( \mu \), written as \( \mu_k \xrightarrow{v} \mu \), if the relation

\[
\lim_{k \to \infty} \int_{[0,\infty]^n \setminus \{0\}} h(z) \mu_k(\mathrm{d}z) = \int_{[0,\infty]^n \setminus \{0\}} h(z) \mu(\mathrm{d}z)
\]
holds for every nonnegative continuous function $h$ with compact support. It is known that $\mu_k^v \to \mu$ on $[0, \infty]^n \setminus \{0\}$ if and only if the convergence
\[
\lim_{k \to \infty} \mu_k [0, t]^c = \mu [0, t]^c
\]
holds for every continuity point $t \in [0, \infty]^n \setminus \{0\}$ of the limit $\mu [0, t]^c$. See Lemma 6.1 and Section 3.3.5 of Resnick (2007) for this assertion and related discussions.

Now we are ready to introduce the concept of MRV. A random vector $X = (X_1, \ldots, X_n)$ taking values in $[0, \infty]^n \setminus \{0\}$ is said to follow a distribution with a multivariate regularly varying tail if there exist a positive normalizing function $b(x) \uparrow \infty$ and a Radon measure $\nu$ on $[0, \infty]^n \setminus \{0\}$, which is not identically 0, such that
\[
x \Pr \left( \frac{X}{b(x)} \in \cdot \right) \nu \to \nu(\cdot) \quad \text{on } [0, \infty]^n \setminus \{0\}.
\] (3.1)
The normalizing function $b$ is not unique, but different choices may result in limit measures that differ by a constant factor only. See Section 5.4.2 of Resnick (1987) or Section 6.1.4 of Resnick (2007) for more discussions on this concept.

### 3.2 Some implications

Relation (3.1) implies that the limit measure $\nu$ is homogeneous; that is, there exists some index $0 \leq \alpha < \infty$ such that
\[\nu(tB) = t^{-\alpha} \nu(B) \quad \text{for all } B \in \mathcal{B};\]
see Page 178 of Resnick (2007). Hence, we write $X \in \text{MRV}_{-\alpha}(\infty)$ or simply $X \in \text{MRV}_{-\alpha}$. The dimension $n$ is suppressed in this notation as it will always be clear from the context. Some useful implications of the homogeneity property of $\nu$ are listed below:

(a) $\nu [0, t]^c$ is continuous in $t$ for every $t > 0$;
(b) $\nu (t, \infty] > 0$ holds for some $t > 0$ if and only if it holds for every $t > 0$;
(c) $b(\cdot) \in \text{RV}_{1/\alpha}$.

Note that the MRV structure is usually defined for $0 < \alpha < \infty$ only, but it can be easily generalized for $\alpha = 0$. When $\alpha = 0$, the homogeneity property mentioned above implies that the limit measure $\nu$ does not assign any mass to the local area $[0, \infty)^n \setminus \{0\}$. Therefore, both $\nu [0, t]^c$ and $\nu (t, \infty]$ are constant in $t \in [0, \infty)^n \setminus \{0\}$. This case can be better understood under the polar coordinate transformation. Let $\mathcal{S}_{n-1}^+ = \{x \in [0, \infty]^n : |x| = 1\}$ be the unit sphere in the first quadrant, where $|\cdot|$ is the usual Euclidean norm. The polar coordinate transformation is a map from $[0, \infty)^n \setminus \{0\}$ to $(0, \infty] \times \mathcal{S}_{n-1}^+$ defined to be
\[
x \mapsto (R, \Theta) = \left( |x|, \frac{x}{|x|} \right).
\]
When \( \alpha = 0 \), the limit measure \( \nu \) is concentrated on the ultimate sphere of \( [0, \infty]^n \setminus \{0\} \), an area described by \( R = \infty \) and \( \Theta \in S_{n-1}^+ \) in the polar coordinate system.

Throughout the paper, denote by \( 1_i \) the vector with the \( i \)th element being 1 and the other elements being 0, \( i = 1, \ldots, n \). We assume that the random vector \( X^+ \) follows MRV\(-\alpha\) with index \( 0 \leq \alpha < \infty \) and limit measure \( \nu \) satisfying \( c_i = \nu(1_i, \infty) > 0 \) for \( i = 1, \ldots, n \). The assumption on \( \nu \) means that the marginal tails of \( X \) are proportionally equivalent and, hence, all the components of \( X \) are comparable in the right tail. Actually, by Lemma 4.2 below one easily sees that (3.1) gives \( \Pr(X_i > x) \sim c_i/b^- (x) \). Hence, introducing a distribution function \( F \) with \( F(x) = 1 - F(x) \sim (b^- (x))^{-1} \), we have

\[
\lim_{x \to \infty} \frac{\Pr(X_i > x)}{F(x)} = c_i, \quad i = 1, \ldots, n. \tag{3.2}
\]

Then we can rewrite (3.1) as

\[
\frac{1}{F(x)} \Pr\left( \frac{X}{x} \in \cdot \right) \Rightarrow \nu(\cdot) \quad \text{on } [0, \infty]^n \setminus \{0\}. \tag{3.3}
\]

Hereafter, we shall follow the MRV structure defined by (3.3) for some distribution function \( F \) and understand the limit measure \( \nu \) and the constants \( c_i = \nu(1_i, \infty) > 0, \ i = 1, \ldots, n \), accordingly.

The limit measure \( \nu \) carries all asymptotic dependence information of \( X \) in the upper-right tail, which is crucially important for credit risk modeling. Under (3.3), if \( \nu(1, \infty) > 0 \) then

\[
\lim_{x \to \infty} \frac{1}{F(x)} \Pr\left( \bigcap_{i=1}^{n} (X_i > x) \right) = \nu(1, \infty) > 0,
\]

which means that \( X^+ \) exhibits large joint movements. For this case, it is easy to check that these components are pairwise asymptotically dependent. In case the limit measure \( \nu \) is concentrated on a straight line \( \{t > 0 : l_1 t_1 = \cdots = l_n t_n \} \) for some \( l > 0 \), the components of \( X^+ \) are so-called asymptotically fully dependent. On the contrary, if \( \nu(1, \infty) = 0 \), then \( X^+ \) does not exhibit large joint movements. Nevertheless, in this case, some, but not all, components of \( X^+ \) may still exhibit large joint movements since \( \nu \) can assign positive masses to the corresponding planes. In other words, \( \nu(1, \infty) = 0 \) does not necessarily imply that the components of \( X^+ \) are pairwise asymptotically independent. See Section 6.5.1 of Resnick (2007) for related discussions. Only the case with \( \nu(1, \infty) > 0 \) is of interest to us because it is relevant for modeling credit contagion.

4 Main Result

Throughout the paper, we make the following Assumption 4.1 on the loss variables (or verify it for special cases) and Assumption 4.2 on the loss settlement functions:
Assumption 4.1 The random vector $X^+$ follows $\text{MRV}_-\alpha$ as defined by (3.3) with index $0 \leq \alpha < \infty$ and limit measure $\nu$ satisfying

$$c_i = \nu (1_i, \infty) > 0 \quad \text{for each } i = 1, \ldots, n.$$ 

Assumption 4.2 There is some distribution function $G$ concentrated on $[0, \infty)$ with $G \in \text{RV}_{-\beta}$, $0 < \beta \leq \infty$, such that

$$G_i(x) \sim d_i G(x) \quad \text{for some } d_i > 0 \text{ for each } i = 1, \ldots, n.$$ 

Assumption 4.1 implies that the relations in (3.2) hold, while Assumption 4.2 means that the loss settlement functions associated with the $n$ obligors are comparable in reflecting large losses. Note that the exponential loss settlement function $G$ suggested by Pykhtin (2003) and Tasche (2004) satisfies $G \in \text{RV}_{-\infty}$.

Recall the LGD of form (1.3) and the quantity $K$ defined by relation (1.6). Here comes our main result:

Theorem 4.1 Under Assumptions 4.1 and 4.2, it holds that

$$\lim_{x \to \infty} \frac{\Pr (K > x)}{F(G^{-1}(1 - 1/x))} = \tilde{\nu}(A), \quad (4.1)$$

where $A = \{y \in [0, \infty]^n : \sum_{i=1}^n e_i/y_i < 1\}$ and $\tilde{\nu}$ is a Radon measure on $[0, \infty]^n \setminus \{0\}$ defined by

$$\tilde{\nu} [0, t]^e = \nu \left[0, \left(a_1 (d_1 t_1)^{1/\beta}, \ldots, a_n (d_n t_n)^{1/\beta}\right)^e\right], \quad t > 0. \quad (4.2)$$

The proof of Theorem 4.1 is postponed to the end of this section after some lemmas are prepared.

Note that the denominator on the left-hand side of (4.1) as a composite function of $x$ belongs to $\text{RV}_{-\alpha/\beta}(\infty)$. Following the proof of the theorem, one sees that $\tilde{\nu}$ also satisfies

$$\tilde{\nu} (t, \infty) = \nu \left(\left(a_1 (d_1 t_1)^{1/\beta}, \ldots, a_n (d_n t_n)^{1/\beta}\right), \infty\right], \quad t > 0. \quad (4.3)$$

When $\alpha/\beta = 0$ (that is, $\alpha = 0$ or $\beta = \infty$), the limit measure $\tilde{\nu}$ is concentrated on the ultimate sphere $\{x \in [0, \infty]^n : |x| = \infty\}$ only, and the limit on the right-hand side of (4.1) reduces to

$$\tilde{\nu} (A) = \tilde{\nu} (e, \infty) = \nu \left(\left(a_1 (d_1 e_1)^{1/\beta}, \ldots, a_n (d_n e_n)^{1/\beta}\right), \infty\right] = \nu (a, \infty).$$

Relation (4.1) gives an exact asymptotic formula for $\Pr (K > x)$ only if $\tilde{\nu} (A) > 0$. Thus, it becomes necessary to find conditions under which $\tilde{\nu} (A) > 0$ holds. For this purpose, we establish the following simple lemma:
Lemma 4.1 In Theorem 4.1, \( \tilde{\nu}(A) > 0 \) if and only if \( \nu(1, \infty] > 0 \).

Proof. By relation (4.3), it is clear that \( \tilde{\nu}(t, \infty] > 0 \) holds for every \( t > 0 \) if and only if \( \nu(t, \infty] > 0 \) holds for every \( t > 0 \). Notice that \( (1, \infty] \subset A \subset (e, \infty] \).

Therefore, if \( \tilde{\nu}(A) > 0 \), then \( \tilde{\nu}(e, \infty] > 0 \), which implies, in turn, \( \tilde{\nu}(t, \infty] > 0 \) and \( \nu(t, \infty] > 0 \) for every \( t > 0 \). Hence, \( \nu(1, \infty] > 0 \). The other implication follows similarly. ■

Theorem 4.1 and Lemma 4.1 together explain why some models based on Gaussian copulas may lead to a significant underestimate of extreme risks. Formally, if the random vector \( X^+ \) exhibits large joint movements, namely, \( \nu(1, \infty] > 0 \), then \( \tilde{\nu}(A) > 0 \) and relation (4.1) gives

\[
\Pr(K > x) \sim \tilde{\nu}(A) F(G \leftarrow (1 - 1/x)). \tag{4.4}
\]

On the contrary, if \( X^+ \) does not exhibit large joint movements, namely, \( \nu(1, \infty] = 0 \), as in the Gaussian copula case, then \( \tilde{\nu}(A) = 0 \) and relation (4.1) gives

\[
\Pr(K > x) = o(1) F(G \leftarrow (1 - 1/x)). \tag{4.5}
\]

The first case with large joint movements is crucial for credit risk modeling. The asymptotic formula (4.4) gives the exact rate at which the tail probability of \( K \) decays to 0. However, in order to apply this formula, the computation of \( \tilde{\nu}(A) \) will be a key issue, which we shall address in Section 6. For the second case in which \( X^+ \) does not exhibit large joint movements, relation (4.5) does not capture the decaying rate of the tail probability of \( K \). For this case, further conditions on \( X^+ \), such as hidden regular variation, have to be imposed. For details of hidden regular variation, the reader is referred to Section 9.4 of Resnick (2007). We shall not pursue such an extension in this work.

We prepare two lemmas that will be used to prove Theorem 4.1.

Lemma 4.2 For a nondecreasing function \( h \in RV_r(\infty) \), \( 0 < r \leq \infty \), its two inverse functions satisfy \( h^{-}(x) \sim h^{\rightarrow}(x) \) and both belong to \( RV_{1/r}(\infty) \).

Proof. By Proposition 0.8(v) of Resnick (1987), for every \( 0 < r \leq \infty \), we have \( h \in RV_r(\infty) \) if and only if \( h^{-} \in RV_{1/r}(\infty) \). Therefore, it suffices to verify that \( h^{-}(x) \sim h^{\rightarrow}(x) \). Trivially, \( h^{-}(x) \leq h^{\rightarrow}(x) \). On the other hand, it holds for every fixed \( \varepsilon > 0 \) and \( x > 0 \) that

\[
h^{\rightarrow}(x) = \sup \{ y : h(y) \leq x \} \\
\leq \inf \{ y : h(y) \geq (1 + \varepsilon)x \} \\
= h^{-}((1 + \varepsilon)x) \\
\sim (1 + \varepsilon)^{1/r} h^{-}(x),
\]

where the last step is due to the fact that \( h^{-} \in RV_{1/r}(\infty) \). By the arbitrariness of \( \varepsilon \) we obtain \( h^{\rightarrow}(x) \lesssim h^{-}(x) \) and this concludes the proof. ■
Lemma 4.3 Define \( Y_i = 1/G_i(X_i/a_i - 1) \) for \( i = 1, \ldots, n \). Under Assumptions 4.1 and 4.2, the random vector \( Y \) follows MRV\(_{-\alpha/\beta} \) with

\[
\Pr \left( \frac{Y}{x} \in \cdot \right) \xrightarrow{F(G^+ (1 - 1/x))} \tilde{\nu} \left( \cdot \right) \quad \text{on} \ [0, \infty)^n \setminus \{0\},
\]

(4.6)

where \( \tilde{\nu} \) is defined by relation (4.2).

**Proof.** For each \( i = 1, \ldots, n \), write \( h_i(x) = 1/G_i(x/a_i - 1) \) and \( h_0(x) = 1/G(x) \). Then \( Y = h(X) = (h_1(X_1), \ldots, h_n(X_n)) \), each \( h_i \) is a nondecreasing function belonging to \( RV_{\beta} \), and \( h_i(x) \sim h_0(x/a_i - 1) / d_i \). Therefore, applying Lemma 4.2 we have

\[
h_i^-(x) \sim h_i^+(x) \sim a_i d_i^{1/\beta} h_i^+(x) \sim a_i d_i^{1/\beta} h_0^+(x).
\]

(4.7)

Let \( 0 < \varepsilon < 1 \) and \( \mathbf{t} > 0 \) be fixed. By relations (2.1), (4.7), and (3.3), and the continuity property of \( \nu \), it holds that

\[
\frac{\Pr (h(X) \in x[0, \mathbf{t}]^c)}{F(h_0^+(x))} \leq \frac{1}{F(h_0^+(x))} \Pr \left( \bigcup_{i=1}^n \left( X_i \geq h_i^+(xt_i) \right) \right)
\]

\[
\lesssim \frac{1}{F(h_0^+(x))} \Pr \left( \bigcup_{i=1}^n \left( X_i^+ > (1 - \varepsilon) a_i (d_i t_i)^{1/\beta} h_0^+(x) \right) \right)
\]

\[
\rightarrow (1 - \varepsilon)^{-\alpha} \nu \left[ 0, (a_1 (d_1 t_1)^{1/\beta}, \ldots, a_n (d_n t_n)^{1/\beta}) \right]^c.
\]

Symmetrically,

\[
\frac{\Pr (h(X) \in x[0, \mathbf{t}]^c)}{F(h_0^+(x))} \gtrsim \frac{1}{F(h_0^+(x))} \Pr \left( \bigcup_{i=1}^n \left( X_i > h_i^+(xt_i) \right) \right)
\]

\[
\gtrsim (1 + \varepsilon)^{-\alpha} \nu \left[ 0, (a_1 (d_1 t_1)^{1/\beta}, \ldots, a_n (d_n t_n)^{1/\beta}) \right]^c.
\]

By the arbitrariness of \( \varepsilon \), the identity \( h_0^+(x) = G^+ (1 - 1/x) \), and relation (4.2), we have, for \( \mathbf{t} > 0 \),

\[
\lim_{x \to \infty} \frac{\Pr (h(X) \in x[0, \mathbf{t}]^c)}{F(G^+ (1 - 1/x))} = \nu \left[ 0, (a_1 (d_1 t_1)^{1/\beta}, \ldots, a_n (d_n t_n)^{1/\beta}) \right]^c = \tilde{\nu} [0, \mathbf{t}]^c.
\]

(4.8)

For the ordinary case with \( 0 < \alpha/\beta < \infty \), the vague convergence in (4.6) follows from (4.8) since the limit \( \tilde{\nu} [0, \mathbf{t}]^c \) is discontinuous at every point \( \mathbf{t} \) with one or more components equal to 0. For the other case with \( \alpha/\beta = 0 \) (that is, \( \alpha = 0 \) or \( \beta = \infty \)), the limit measure \( \tilde{\nu} \) is concentrated on the ultimate sphere \( \{ \mathbf{x} \in [0, \infty]^n : |\mathbf{x}| = \infty \} \) only. Then by definition, the vague convergence in (4.6) still holds.

In case \( \alpha/\beta = 0 \), relation (4.8) cannot be used to evaluate \( \tilde{\nu} \) for a subset of the ultimate sphere \( \{ \mathbf{x} \in [0, \infty]^n : |\mathbf{x}| = \infty \} \). For this question, one may appeal to the polar coordinate transformation.
**Proof of Theorem 4.1.** Recall relations (1.6) and (1.3) and the definition of $Y$ in Lemma 4.3. We have, for every $x > 0$,

$$
\Pr(K > x) = \Pr \left( L > 1 - \frac{1}{x} \right)
= \Pr \left( \sum_{i=1}^{n} e_i \left( 1 - \frac{1}{Y_i} \right) > 1 - \frac{1}{x} \right)
= \Pr \left( \sum_{i=1}^{n} e_i \frac{1}{Y_i} < \frac{1}{x} \right)
= \Pr \left( \frac{Y}{x} \in A \right).
$$

It is easy to verify that $\tilde{\nu}(\partial A) = 0$. Therefore, an application of Lemma 4.3 concludes the proof.

By Assumption 4.1 on the MRV structure of the random vector $X^+$, we have assumed tail equivalence for the loss variables $X_1, \ldots, X_n$, namely, relation (3.2) holds with some $c_i > 0$ for each $i = 1, \ldots, n$. In case these losses are not tail equivalent, say, some (but not all) constants $c_i$ are zero, it is easy to see that the proof of Theorem 4.1 is still valid. However, we have $\nu(1, \infty] = 0$ by (3.3) and, hence, $\tilde{\nu}(A) = 0$ by Lemma 4.1. Thus, for this case, Theorem 4.1 fails to obtain an exact asymptotic formula for $\Pr(K > x)$.

## 5 Two Special Cases

### 5.1 The Archimedean copula case

In this subsection we propose to use an Archimedean copula to model the dependence among the loss variables. See Nelsen (2006) for a textbook treatment of copulas and see Frees and Valdez (1998) for discussions of copulas in the actuarial context.

An Archimedean copula has the form

$$
C(u_1, \ldots, u_n) = \varphi^{-1} \left( \varphi(u_1) + \cdots + \varphi(u_n) \right),
$$

(5.1)

where $\varphi : (0, 1) \to (0, \infty)$, called the generator, is a strictly decreasing and convex function with $\varphi(0+) = \infty$ and $\varphi(1-) = 0$, and the function $\varphi^{-1}$ is the usual inverse of $\varphi$. As summarized by Theorem 4.6.2 of Nelsen (2006), the expression in (5.1) defines a proper $n$-dimensional copula for all $n = 2, 3, \ldots$ if and only if $\varphi^{-1}$ is completely monotone; that is,

$$
(-1)^k \frac{d^k}{dt^k} \varphi^{-1}(t) \geq 0, \quad k = 0, 1, \ldots
$$
Assume that the random vector $X$ possesses an Archimedean copula of form (5.1). Moreover, following Charpentier and Segers (2009), we assume that the generator $\varphi$ satisfies

$$
\lim_{u \downarrow 0} \frac{\varphi(1-tu)}{\varphi(1-u)} = t^r, \quad t > 0,
$$

for some constant $r$. Note that the constant $r$, given its existence, should satisfy $1 \leq r \leq \infty$. Table 2 of Charpentier and Segers (2009) contains a list of bivariate Archimedean copulas satisfying condition (5.2).

In the rest of this subsection, for $t \in [0, \infty)^n$, the expression $(\sum_{i=1}^{n} t_i^r)^{1/r}$ with $r = \infty$ is interpreted as $\bigvee_{i=1}^{n} t_i$. We state the first corollary of Theorem 4.1 as follows:

**Corollary 5.1** Let $X$ possess an Archimedean copula of form (5.1) with generator $\varphi$ satisfying relation (5.2) for some $1 \leq r \leq \infty$, let the components of $X$ satisfy the tail equivalence condition (3.2) for some distribution function $F$ with $F \in \text{RV}_{-\alpha}$, $0 \leq \alpha < \infty$, and some positive constants $c_1, \ldots, c_n$, and let Assumption 4.2 be valid. Then relation (4.1) holds with $\tilde{\nu}$ defined by

$$
\tilde{\nu} [0, t]^c = \left( \sum_{i=1}^{n} \left( c_i t_i^{-\alpha} d_i^{-\alpha/\beta} t_i^{-\alpha/\beta} \right)^r \right)^{1/r}, \quad t > 0.
$$

Corollary 5.1 follows directly from the following two lemmas and Theorem 4.1. The first lemma below is excerpted from the proof of Theorem 4.1 of Charpentier and Segers (2009):

**Lemma 5.1** Suppose that a uniform random vector $U = (U_1, \ldots, U_n)$ is jointly distributed by an Archimedean copula of form (5.1) with generator $\varphi$ satisfying relation (5.2) for some $1 \leq r \leq \infty$. Then it holds that

$$
\lim_{x \to \infty} x \Pr \left( \bigcup_{i=1}^{n} \left( U_i > 1 - \frac{t_i}{x} \right) \right) = \left( \sum_{i=1}^{n} t_i^r \right)^{1/r}, \quad t > 0.
$$

We verify the MRV structure of the random vector $X^+$ in the next lemma. See Theorem 1.1 and Proposition 2.6 of Li and Wu (2013) for this result with $1 < r < \infty$; see also Lemma A.3 of Huang et al. (2013) for a similar result.

**Lemma 5.2** Let the random vector $X$ and the Archimedean copula $C$ be as specified in Corollary 5.1. Then $X^+$ follows MRV $-\alpha$ with

$$
\frac{1}{F(x)} \Pr \left( \frac{X^+}{x} \in \cdot \right) \overset{\nu}{\rightarrow} \nu(\cdot) \quad \text{on } [0, \infty)^n \setminus \{0\},
$$

where $\nu$ is defined by

$$
\nu [0, t]^c = \left( \sum_{i=1}^{n} \left( c_i t_i^{-\alpha} \right)^r \right)^{1/r}, \quad t > 0.
$$

(5.3)
Proof. Let $U = (U_1, \ldots, U_n)$ be a uniform random vector jointly distributed by the Archimedean copula $C$. It holds for every $t > 0$, $0 < \varepsilon < 1$, and $x > 0$ that

$$\frac{1}{F(x)} \Pr \left( \frac{X^+}{x} \in [0, t]^c \right) = \frac{1}{F(x)} \left( 1 - C \left( F_1(x), \ldots, F_n(x_n) \right) \right)$$

$$= \frac{1}{F(x)} \Pr \left( \bigcup_{i=1}^n \left( U_i > 1 - F_i(x_i) \right) \right)$$

$$\lesssim \frac{1}{F(x)} \Pr \left( \bigcup_{i=1}^n \left( U_i > 1 - (1 + \varepsilon) c_i t_i^{-\alpha} F_i(x) \right) \right)$$

$$\rightarrow (1 + \varepsilon) \left( \sum_{i=1}^n (c_i t_i^{-\alpha})^r \right)^{1/r},$$

where the last step is due to Lemma 5.1. By the arbitrariness of $\varepsilon$ we have

$$\limsup_{x \to \infty} \frac{1}{F(x)} \Pr \left( \frac{X^+}{x} \in [0, t]^c \right) \leq \left( \sum_{i=1}^n (c_i t_i^{-\alpha})^r \right)^{1/r}.$$}

Establishing the corresponding lower bound in a similar way concludes the proof.

The next lemma gives a sufficient and necessary condition for $X^+$ to exhibit large joint movements:

**Lemma 5.3** Under the assumptions of Lemma 5.2, the limit measure $\nu$ defined by (5.3) satisfies $\nu(1, \infty] > 0$ if and only if $1 < r \leq \infty$.

**Proof.** First we prove the necessity. Recall that $1 \leq r \leq \infty$. In case $r = 1$, relation (5.3) shows that $\nu$ is concentrated on the axes only and, thus, $\nu(1, \infty] = 0$.

Next we prove the sufficiency. For $r = \infty$, relation (5.3) becomes

$$\nu[0, t]^c = \bigvee_{i=1}^n c_i t_i^{-\alpha}, \quad t > 0,$$

implying that $\nu$ is concentrated on the straight line $\{ t > 0 : c_1^{-1/\alpha} t_1 = \cdots = c_n^{-1/\alpha} t_n \}$ and, therefore, $\nu(1, \infty] = \bigwedge_{i=1}^n c_i > 0$. For $1 < r < \infty$, applying the inclusion-exclusion formula to (5.3) leads to

$$\nu(1, \infty] = \sum_{\varnothing \neq I \subseteq \{1, \ldots, n\}} (-1)^{|I|-1} \left( \prod_{i \in I} c_i^r \right)^{1/r},$$

where $|I|$ denotes the cardinality of $I$. Denote the right-hand side above by $h(c_1, \ldots, c_n)$ and understand it as a function of $(c_1, \ldots, c_n) \in (0, \infty)^n$. Taking $\frac{\partial^n}{\partial c_1 \cdots \partial c_n}$ on $h(c_1, \ldots, c_n)$, it
is easy to see that only the term corresponding to \( I = \{1, \ldots, n\} \) remains. Thus,
\[
\frac{\partial^n}{\partial c_1 \cdots \partial c_n} h(c_1, \ldots, c_n) = \frac{\partial^n}{\partial c_1 \cdots \partial c_n} (-1)^{n-1} \left( \sum_{i=1}^{n} c_i^r \right)^{1/r}
= (-1)^{n-1} \prod_{i=0}^{n-1} \left( \frac{1}{r} - i \right) \cdot \left( \sum_{i=1}^{n} c_i^r \right)^{1/r-n} \prod_{i=1}^{n-1} r c_i^{r-1} > 0.
\]
This, together with the observation that \( h(0) = 0 \), verifies the positivity of \( h(c_1, \ldots, c_n) \) over \((0, \infty)^n\).

We close this subsection with a remark. Assume \( 0 < \alpha, \beta < \infty \) and \( r = \infty \). By Corollary 5.1, we have
\[
\tilde{\nu}[0, t]^c = \bigvee_{i=1}^{n} c_i a_i^{-\alpha} d_i^{-\alpha/\beta} t_i^{-\alpha/\beta}, \quad t > 0,
\]
(5.4)
implying that the limit measure \( \tilde{\nu} \) is concentrated on the straight line
\[
\left\{ t > 0 : c_1^{-\beta/\alpha} a_1^{\beta} d_1 t_1 = \cdots = c_n^{-\beta/\alpha} a_n^{\beta} d_n t_n \right\}.
\]
Therefore, the term \( \tilde{\nu}(A) \) admits the following significant simplification:
\[
\tilde{\nu}(A) = \left( \sum_{i=1}^{n} c_i a_i^{-\beta/\alpha} d_i \right)^{-\alpha/\beta}.
\]
(5.5)
One would expect that the limits of (5.4) and (5.5) be equal as \( \alpha \downarrow 0 \) or \( \beta \uparrow \infty \). However, as \( \alpha \downarrow 0 \), the limit of (5.4) is \( \bigvee_{i=1}^{n} c_i \) while the limit of (5.5) is \( \bigwedge_{i=1}^{n} c_i \), which are not equal unless \( c_1 = \cdots = c_n \). This indicates that only a mass of amount \( \bigwedge_{i=1}^{n} c_i \) is attracted to the ultimate sphere \( \{ x \in (0, \infty)^n : |x| = \infty \} \) while the remaining mass of amount \( \bigvee_{i=1}^{n} c_i - \bigwedge_{i=1}^{n} c_i \) is attracted to the ultimates of the axes or planes. The divergence of \( \tilde{\nu} \) as \( \beta \uparrow \infty \) is in a similar fashion.

### 5.2 Mixtures

Now assume that \( X = (X_1, \ldots, X_n) \) follows a mixture structure
\[
X = u + \sqrt{W} Z = u + \sqrt{W} \left( \sqrt{\rho \eta_0} 1 + \sqrt{1 - \rho \eta} \right),
\]
(5.6)
where \( u \in (-\infty, \infty)^n \) and \( 0 < \rho < 1 \) are nonrandom, \( W \) is a nonnegative random variable, independent of \( Z \), with \( \sqrt{W} \) distributed by \( F \) such that \( F \in \text{RV}_{-\alpha}, 0 \leq \alpha < \infty \), and \( \eta_0, \eta_1, \ldots, \eta_n \) are real-valued independent random variables with finite moments of order \( \gamma \) for some \( \gamma > \alpha \).
In model (5.6), the random variable \( W \) is usually interpreted as a common shock, the random variable \( \eta_0 \) as a systematic risk factor, the random variables \( \eta_1, \ldots, \eta_n \) as idiosyncratic risk factors, and the parameter \( \rho \) adjusts the impacts of both kinds of factors.

This model has been extensively applied to credit risk management; see Frey and McNeil (2003), Bassamboo et al. (2008), and Brereton et al. (2013), to name several. If \( W \) follows an inverse gamma distribution with both shape and scale parameters equal to \( \alpha/2 \), and \( Z \) follows a multivariate normal distribution, then \( X \) follows a multivariate \( t \) distribution; see, e.g. Section 3.2 of McNeil et al. (2005).

Trivially, the \( n \)-dimensional comonotonic vector \( \sqrt{W} \mathbf{1} \) follows \( \text{MRV}_{-\alpha} \) with

\[
\frac{1}{F(x)} \Pr \left( \frac{\sqrt{W} \mathbf{1}}{x} \in \cdot \right) \overset{\nu}{\to} \mu_W(\cdot) \quad \text{on } [0, \infty)^n \setminus \{0\},
\]

where \( \mu_W \) is defined by \( \mu_W[0,t]^c = \bigvee_{i=1}^n t^{-\alpha}_i, \ t > 0 \), implying that it is concentrated on the main diagonal of the first quadrant. Recall the multivariate version of Breiman’s theorem as established in Proposition A.1 of Basrak et al. (2002); see also Theorem 1 of Fougéres and Mercadier (2012). By this result, we know that \( \sqrt{W}Z^+ \) follows \( \text{MRV}_{-\alpha} \) with

\[
\frac{1}{F(x)} \Pr \left( \frac{\sqrt{W}Z^+}{x} \in \cdot \right) \overset{\nu(\cdot)}{\to} \mathbb{E} \left[ \mu_W \left( (Z^+)^{-1} \cdot \right) \right] \quad \text{on } [0, \infty)^n \setminus \{0\}, \tag{5.7}
\]

where \( (Z^+)^{-1} \cdot = \{ x \in [0, \infty)^n \setminus \{0\} : (Z^+_1 x_1, \ldots, Z^+_n x_n) \in \cdot \} \). By the continuity property of \( \nu \), it is easy to see that \( X^+ = \left( u + \sqrt{W}Z^+ \right)^+ \) follows the same \( \text{MRV}_{-\alpha} \) structure; that is,

\[
\frac{1}{F(x)} \Pr \left( \frac{X^+}{x} \in \cdot \right) \overset{\nu(\cdot)}{\to} \nu(\cdot) \quad \text{on } [0, \infty)^n \setminus \{0\}.
\]

This leads to the following corollary of Theorem 4.1:

**Corollary 5.2** Let \( X \) be defined by (5.6). Then under Assumption 4.2, relation (4.1) holds with \( \tilde{\nu} \) defined by

\[
\tilde{\nu}[0,t]^c = \nu \left[ 0, \left( a_1 (d_1 t_1)^{1/\beta}, \ldots, a_n (d_n t_n)^{1/\beta} \right) \right]^c, \quad t > 0,
\]

and \( \nu \) defined by (5.7).

We refer the reader to Li and Sun (2009), who verified the MRV structure of more general mixture models than (5.6).

Note that, due to the common shock \( W \), the random vector \( X^+ \) defined by (5.6) exhibits
large joint movements. Actually,

\[ \nu(1, \infty) = E \left[ \mu_W \left( (Z^+)^{-1} (1, \infty) \right) \right] \]

\[ = E \left[ \mu_W \left\{ x \in [0, \infty]^n \setminus \{0\} : \bigwedge_{i=1}^n Z_i^+ x_i > 1 \right\} \right] \]

\[ = E \left[ \left( \bigwedge_{i=1}^n Z_i^+ \right)^\alpha \right] > 0, \]

where in the third step we used the fact that \( \mu_W(t, \infty) = \bigwedge_{i=1}^n t_i^{-\alpha} \) for \( t > 0 \).

Note also that, due to the presence of the idiosyncratic risk factors \( \eta_1, \ldots, \eta_n \), the limit measure \( \nu \) is not concentrated on a straight line unless the adjusting parameter \( \rho \) reduces to 1. This means that the loss variables, though asymptotically dependent, are not asymptotically fully dependent.

6 Estimating the VaR and CTE

For a loss variable \( X \) distributed by \( F \) and a confidence level \( q \in (0, 1) \), the VaR and CTE of \( X \) at level \( q \) are defined to be

\[ \text{VaR}_q[X] = F_{c^{-}}(q) \quad \text{and} \quad \text{CTE}_q[X] = E \left[ X \mid X > \text{VaR}_q[X] \right], \]

respectively. The two risk measures are fundamentals in risk management. For their economic justifications, the reader is referred to Duffie and Pan (1997), Artzner et al. (1999), Hardy (2003), Hardy and Wirch (2004), McNeil et al. (2005), and Dhaene et al. (2006), among many others.

6.1 Asymptotic estimates for the VaR and CTE

Notice that the recent financial instability has caused regulators to become very prudent and require a confidence level \( q \) close to 1 in both VaR and CTE frameworks. In this subsection, we apply Theorem 4.1 to establish asymptotic estimates for the VaR and CTE of the quantity \( K \) defined by (1.6) as \( q \uparrow 1 \). Wüthrich (2003), Embrechts et al. (2009), and Zhu and Li (2012) carried out some asymptotic studies of the VaR and CTE in multivariate settings, which have a similar flavor to our work.

We only consider the scenario with large joint movements; namely, \( \nu(1, \infty) > 0 \). Therefore, \( \tilde{\nu}(A) > 0 \) by Lemma 4.1 and the exact asymptotic formula (4.4) holds. In the following result, the condition \( \alpha/\beta > 1 \) is imposed to ensure the finiteness of \( \text{CTE}_q[K] \):
Proposition 6.1 In addition to the conditions of Theorem 4.1, assume that $\alpha/\beta > 1$ and $\nu(1, \infty) > 0$. Then

$$
CTE_q[K] \sim \frac{\alpha}{\alpha - \beta} \text{VaR}_q[K] \sim \frac{\alpha}{\alpha - \beta} \frac{\hat{\nu}(A)^{\beta/\alpha}}{G(F^{-1}(q))}.
$$

(6.1)

Proof. Relation (4.4) implies that $K$ is regularly varying tailed with index $-\alpha/\beta$. Thus, the second relation in (6.1) follows immediately from relation (4.4) and Proposition 0.8(vi) of Resnick (1987). Moreover, note that

$$
CTE_q[K] = \text{VaR}_q[K] + \int_{\text{VaR}_q[K]}^{\infty} \frac{\Pr(K > y)}{\Pr(K > \text{VaR}_q[K])} dy.
$$

(6.2)

A simple application of Karamata’s theorem (see Theorem 1.5.11 of Bingham et al. 1987) to the integral above gives the first relation in (6.1).

The asymptotic estimates for the VaR and CTE of $L$ are readily known since, as shown in the following result, the VaR and CTE of $K$ are acting for those of $L$:

Lemma 6.1 Under the conditions of Proposition 6.1, it holds that

$$
\text{VaR}_q[K] = \frac{1}{1 - \text{VaR}_q[L]} \quad \text{and} \quad CTE_q[K] \sim \frac{\alpha^2}{\alpha^2 - \beta^2} \frac{1}{1 - CTE_q[L]}.
$$

(6.3)

Proof. The first relation in (6.3) holds trivially since $K$, via (1.6), is a continuous and strictly increasing function of $L$. Moreover, expressing $CTE_q[L]$ in terms of $\text{VaR}_q[L]$ as in (6.2), we have

$$
1 - CTE_q[L] = 1 - \text{VaR}_q[L] - \int_0^{1 - \text{VaR}_q[L]} \frac{\Pr(L > 1 - y)}{\Pr(L > 1 - (1 - \text{VaR}_q[L]))} dy.
$$

Note that $\Pr(L > 1 - \cdot)$ belongs to $\text{RV}_{\alpha/\beta}(0+)$. A similar argument to what we used on (6.2) yields

$$
1 - CTE_q[L] \sim \frac{\alpha}{\alpha + \beta} (1 - \text{VaR}_q[L])
$$

(6.4)

The second relation in (6.3) follows from a simple combination of the first relation in (6.1), the first relation in (6.3), and relation (6.4).

6.2 A general description

Let us address, in a general setting, the issue how large a sample needs to be for a statistical estimate to achieve a prescribed accuracy. A similar description was given by Tang and Yuan (2012). Suppose that we are to estimate a positive parameter $\theta$ by an estimator $\hat{\theta}_N$, which satisfies the central limit theorem

$$
\sqrt{N} \left( \hat{\theta}_N - \theta \right) \overset{d}{\to} \mathcal{N} \left( 0, \sigma^2_\theta \right), \quad N \to \infty,
$$

(6.5)
where \( \overset{d}{\rightarrow} \) denotes convergence in distribution and \( \mathcal{N}(0, \sigma_\theta^2) \) denotes the normal distribution with mean 0 and variance \( \sigma_\theta^2 \).

The accuracy of this estimation is usually characterized as follows. For some \( \varepsilon > 0 \) close to 0 and \( 0 < p < 1 \) close to 1, the estimator \( \hat{\theta}_N \) is within \( 100\varepsilon\% \) of the true value \( \theta \) with probability not smaller than \( p \); that is,
\[
\Pr \left( \left| \hat{\theta}_N - \theta \right| \leq \varepsilon \theta \right) \geq p.
\] (6.6)

By relation (6.5), approximately (6.6) means that
\[
\frac{\varepsilon \theta}{\sigma_\theta/\sqrt{N}} \geq z_p, \quad \text{or, equivalently,} \quad N \geq \left( \frac{z_p \sigma_\theta}{\varepsilon \theta} \right)^2,
\] (6.7)
where \( z_p = \Phi^{-1}((1 + p)/2) \) is the quantile of the standard normal distribution at \((1 + p)/2\).

Thus, the sample size \( N \) needed for the prescribed accuracy depends largely on \( \sigma_\theta \).

### 6.3 Empirical estimate for the VaR

Now we discuss the empirical estimate for the VaR of \( K \). Let \( K_1, \ldots, K_N \) be a random sample from \( K \), and denote by \( K_{(1)} \leq \cdots \leq K_{(N)} \) the order statistics. Traditionally, the VaR of \( K \) at level \( q \in (0, 1) \) is estimated as
\[
\hat{\text{VaR}}_{q,N}[K] = K_{\lceil Nq \rceil},
\] (6.8)
where \( \lceil x \rceil \) denotes the largest integer not exceeding \( x \in (-\infty, \infty) \).

Let \( q \in (0, 1) \) be fixed and assume that the distribution function of \( K \) has a positive density \( \kappa \) at \( \text{VaR}_q[K] \). Then by Corollary 21.5 of van der Vaart (1998), the empirical estimator \( \hat{\text{VaR}}_{q,N}[K] \) satisfies the central limit theorem
\[
\sqrt{N} \left( \hat{\text{VaR}}_{q,N}[K] - \text{VaR}_q[K] \right) \overset{d}{\rightarrow} \mathcal{N}(0, \sigma_1^2(q)), \quad N \to \infty,
\] (6.9)
with \( \sigma_1^2(q) = q(1 - q)/\kappa^2(\text{VaR}_q[K]) \).

Assume that the density function \( \kappa \) exists and is monotone ultimately. By the monotone density theorem (see Theorem 1.7.2 of Bingham et al. 1987),
\[
\kappa(x) \sim \frac{\alpha \Pr(K > x)}{\beta x}.
\]

Substituting \( x = \text{VaR}_q[K] \) to the above and then applying the second relation in (6.1), we have
\[
\kappa(\text{VaR}_q[K]) \sim \frac{\alpha(1 - q)}{\beta \text{VaR}_q[K]} \sim \frac{\alpha(1 - q)}{\beta} \frac{1}{\text{VaR}_q[K]} \sim \frac{\alpha(1 - q)}{\beta} \frac{1}{\tilde{\nu}(A) \beta/\alpha} \frac{\bar{G}(F \leftarrow (q))}{\nu(A)^{2/\alpha}}.
\]

Hence by (6.9), the asymptotic variance of \( \hat{\text{VaR}}_{q,N}[K] \) is
\[
\frac{\sigma_1^2(q)}{N} \sim \frac{\beta^2 \text{VaR}^2_q[K]}{\alpha^2 N(1 - q)} \sim \frac{\beta^2 \tilde{\nu}(A)^{2/\alpha}}{\alpha^2 N(1 - q) \left( \bar{G}(F \leftarrow (q)) \right)^2},
\]
Note that the asymptotic variance of $\widehat{\text{VaR}}_{q,N}[K]$ explodes as $q \uparrow 1$. This means that, for a confidence level $q$ close to 1 the sample size $N$ needs to be very large in order for $\widehat{\text{VaR}}_{q,N}[K]$ to achieve the prescribed accuracy.

### 6.4 Empirical estimate for the CTE

Now we extend the discussion to the empirical estimate for the CTE of $K$ at level $q \in (0, 1)$:

$$\widehat{\text{CTE}}_{q,N}[K] = \frac{1}{N(1-q)} \sum_{i=[Nq]+1}^{N} K_{(i)}. \quad (6.10)$$

Related discussions on the asymptotic variances of this estimate can be found in numerous references such as Manistre and Hancock (2005) and Brazauskas et al. (2008).

Assume that $\alpha/\beta > 2$ (so as to guarantee the finiteness of the variance of $K$) and that the distribution function of $K$ is continuous at $\text{VaR}_q[K]$. Then by Theorem 3.1 of Brazauskas et al. (2008), the empirical estimator $\widehat{\text{CTE}}_{q,N}[K]$ satisfies the central limit theorem

$$\sqrt{N} \left( \widehat{\text{CTE}}_{q,N}[K] - \text{CTE}_q[K] \right) \xrightarrow{d} \mathcal{N} \left( 0, \sigma_2^2(q) \right), \quad N \to \infty, \quad (6.11)$$

with

$$\sigma_2^2(q) = \frac{1}{(1-q)^2} \text{VAR} \left[ K I_{(K > \text{VaR}_q[K])} \right].$$

By the first relation in (6.1), we have

$$\mathbb{E} \left[ K I_{(K > \text{VaR}_q[K])} \right] = (1-q) \text{CTE}_q[K] \sim (1-q) \frac{\alpha}{\alpha - \beta} \text{VaR}_q[K].$$

Replacing $K$ with $K^2$ in the above and noticing that $\Pr(K^2 > x)$ is regularly varying with index $-a/(2\beta)$, we obtain

$$\mathbb{E} \left[ K^2 I_{(K > \text{VaR}_q[K])} \right] \sim (1-q) \frac{\alpha}{\alpha - 2\beta} \text{VaR}_q^2[K].$$

Hence by (6.11), the asymptotic variance of $\widehat{\text{CTE}}_{q,N}[K]$ is

$$\frac{\sigma_2^2(q)}{N} \sim \frac{1}{N(1-q)} \frac{\alpha}{\alpha - 2\beta} \text{VaR}_q^2[K] \sim \frac{1}{N(1-q)} \frac{\alpha}{\alpha - 2\beta} \left( \frac{\tilde{\nu}(A)^{\beta/\alpha}}{G(F^{-}(q))} \right)^2, \quad (6.12)$$

where the last step is due to the second relation in (6.1). Thus, similarly as in the VaR case, the asymptotic variance of $\widehat{\text{CTE}}_{q,N}[K]$ explodes as $q \uparrow 1$.

As an example, according to relation (6.7) with $\varepsilon = 0.05$ and $p = 0.95$ and relation (6.12), for $\alpha = 2$, $\beta = 0.8$, and $q = 0.999$, at least $2,800,000$ samples are needed to obtain the prescribed accuracy. In particular, the situation becomes even much worse in case $\alpha/\beta < 2$ because the usual central limit theorem as described by (6.11) does not hold any more. In conclusion, the empirical estimator $\widehat{\text{CTE}}_{q,N}[K]$ becomes unreliable in various situations.
7 Numerical Studies

In this section we conduct numerical studies to examine the accuracy of the asymptotic relations. We consider the Archimedean copula case only. Recall that, with $r > 1$, Corollary 5.1 and Lemma 5.3 give that

$$\Pr(K > x) \sim \tilde{\nu}(A) F'(G^{-1}(1 - 1/x))$$

(7.1)

with $A = \{y \in [0, \infty]^n : \sum_{i=1}^n e_i/y_i < 1\}$.

We choose a Gumbel copula with generator $\varphi(u) = (-\ln u)^r$ for $1 < r < \infty$ and $0 < u < 1$. It is easy to verify that $\varphi$ satisfies relation (5.2) and $\varphi^{-1}$ is completely monotone. We remark that the Gumbel copula is particularly relevant in credit risk modeling because it reflects a situation that the strength of dependence increases gradually from the lower-left corner to the upper-right corner, as depicted by Figure 5.3 of McNeil et al. (2005).

Let the loss variables $X_1, \ldots, X_n$ be identically distributed with common Pareto distribution

$$F(x) = 1 - \left(\frac{\theta}{x + \theta}\right)^{\alpha}, \quad x, \alpha, \theta > 0,$$

and introduce a loss settlement function

$$G(s) = 1 - (s + 1)^{-\beta}, \quad s, \beta > 0.$$

The various parameters are set to be

$n = 3,$
$e = (e_1, e_2, e_3) = (0.2, 0.4, 0.4),$
$a_1 = a_2 = a_3 = a = 5,$
$c_1 = c_2 = c_3 = d_1 = d_2 = d_3 = 1,$
$\alpha = 2,$
$\theta = 2,$
$r = 5,$ and
$\beta = 0.8.$

The marginal default probability of each obligor is calculated to be 0.082. We compare the values of the tail probability, VaR, and CTE obtained by crude Monte Carlo (CMC) simulation with those obtained by the asymptotic formulas.

For the CMC estimates, we generate $N$ samples $X_1, \ldots, X_N$ in the R environment using the copula package; see Yan (2007) for discussions on this package. The tail probability $\Pr(K > x)$ is estimated by

$$\frac{1}{N} \sum_{i=1}^N 1(K_i > x),$$

where the random variables $K_i, i = 1, \ldots, N$, are independent copies of $K$, while the VaR and CTE are estimated by (6.8) and (6.10), respectively.
For the asymptotic estimates, how to compute $\tilde{\nu}(A)$ is a key issue. In our numerical studies, this value is computed by numerical integration; that is,

$$
\tilde{\nu}(A) = a^{-\alpha}(r - 1)(2r - 1)\left(\frac{\alpha}{\beta}\right)^3 \int_A \left(\sum_{i=1}^{3} t_i^{-\alpha r/\beta}\right)^{1/r-3} (t_1t_2t_3)^{-\alpha r/\beta - 1} dt_1dt_2dt_3.
$$

We use Mathematica 8 to compute it. However, we point out that this numerical integration becomes significantly more difficult for a higher dimension $n$.

For the high-dimensional case, one may think of the following alternative proposed by Resnick (2004). Recall the random vector $Y$ defined in Lemma 4.3, which possesses an MRV structure with limit measure $\tilde{\nu}$. We have the weak convergence

$$
\frac{1}{k} \sum_{i=1}^{N} \epsilon_{\frac{Y_i}{N/k}}(\cdot) \overset{w}{\rightarrow} \tilde{\nu}(\cdot), \quad N \rightarrow \infty,
$$

for every sequence $k = k(N) \rightarrow \infty$ such that $N/k \rightarrow \infty$, where $\epsilon$ is the Dirac measure and the random vectors $Y_1, \ldots, Y_N$ are independent copies of $Y$. This readily provides a consistent estimate for $\tilde{\nu}(A)$. However, this estimate is sensitive to the choice of $k$. An exploratory solution using the Stáricá plot works reasonably well for this issue, but it requires repeated trials; see Stáricá (1999) and Section 9.2.4 of Resnick (2007) for more details. We remark that, although this method still uses simulations to estimate $\tilde{\nu}(A)$, it works more efficiently than using rare event simulations to estimate the tail probability of $K$ directly.

In the next graph, we compare the CMC estimate for $\Pr(K > x)$ and the asymptotic estimate given by (7.1) on the left and show their ratio on the right. The CMC simulation is conducted with a sample of size $N = 10,000,000$.

Graph 7.1 is here.

From Graph 7.1 we see that, the ratio of the two estimates does not seem to converge to 1 as $q \uparrow 1$. However, this is an illusion due to the poor performance of the CMC estimate and there is nothing wrong with the asymptotic estimate. With a sample size increased to $N = 100,000,000$, we see a much improved convergence of the ratio in Graph 7.2.

Graph 7.2 is here.

Note that the ratio starts to fluctuate after staying around 1 for a while. Our explanation for this phenomenon is that the tail probability $\Pr(K > x)$ becomes too small as $q$ becomes close to 1, leading to a poor performance of simulation again even with a sample size as large as 100,000,000. However, the asymptotic estimate is always stable.

In Graphs 7.3 and 7.4, we compare the empirical estimates for the VaR and CTE of $K$ given by (6.8) and (6.10) with those asymptotic estimates given by (6.1). The empirical estimates are obtained with a sample of size $N = 10,000,000$. 

22
8 Concluding Remarks

The contribution of this paper is threefold. First, we proposed a new structural model for the LGD, which takes into account the severity of default, and we advocated to use the MRV structure to model loss variables because it can easily provide both heavy tails and asymptotic dependence of losses at varying degree. Second, under the MRV framework we derived an exact asymptotic formula for the tail probability of the LGD $L$ through its proxy $K$ defined by (1.6). Numerical studies show that this formula provides an accurate estimate. Third, as applications, we established asymptotic estimates for the VaR and CTE of $K$ (hence, for those of $L$) and compared them with the traditional empirical estimates.

We remark that the present study applies mainly to a portfolio with a fixed number of obligors. For this case, the MRV framework has proven to be useful in modeling credit risks because the microstructure of its limit measure can be easily adjusted to meet the need of modeling various asymptotic dependence structures among loss variables. However, the MRV framework becomes less relevant for a large portfolio. Then we may instead consider mixture models for loss variables and employ large deviation techniques. We shall extend our study in this direction in a future project.

Acknowledgments The authors would like to thank Johan Segers at Université catholique de Louvain for providing them with the proof of Lemma 5.3, thank Alexandru V. Asimit at City University London for his help on the numerical studies, and thank two anonymous referees for their useful comments. The research for this project was sponsored by The Actuarial Foundation (TAF) and the Casualty Actuarial Society (CAS) through the 2012 Individual Grants Competition.

References


[38] Pykhtin, M. Unexpected recovery risk. Risk 16 (2003), no. 8, 74–78.


Graph 7.1 Comparision of the estimates for the tail probability of $K$ ($N = 10,000,000$)
Graph 7.2  Comparision of the estimates for the tail probability of $K$ ($N = 100,000,000$)
Graph 7.3 Comparision of the estimates for the VaR of K

- Simulated
- Asymptotic
Graph 7.4 Comparision of the estimates for the CTE of K