Kuhn:

Tails of Credit Default Portfolios

Sonderforschungsbereich 386, Paper 410 (2005)

Tails of Credit Default Portfolios

Gabriel Kuhn*

December 21, 2004

Abstract

We derive analytic expressions for the tail behavior of credit losses in a large homogeneous credit default portfolio. Our model is an extended CreditMetrics model; i.e. it is a one-factor model with a multiplicative shock-variable. We show that the first order tail behavior is robust with respect to this shock-variable. In a simulation study we compare different models for the latent variables. We fix default probability and correlation of the latent variables and the first order tail behavior of the limiting credit losses in all models and observe a completely different tail behavior leading to very different VaR estimates. For three portfolios of different credit quality we suggest a pragmatic model selection procedure and compare the fit with that of the β-model.

J.E.L. Subject Classifications:	G11, G21, G39, C19.	
AMS 2000 Subject Classifications:	primary:	91B28, 60F05
	secondary:	91B70, 62E20, 60B10

Keywords: β-model, credit default portfolio, extreme value theory, heavy-tailed risk factor, latent variable model, multivariate t-distribution, one factor model, regular variation, tail behavior of portfolio credit loss, Value at Risk (VaR).

*Center for Mathematical Sciences, Munich University of Technology, D-85747 Garching, Germany, email: gabriel@ma.tum.de, http://www.ma.tum.de/stat/

1 Introduction

We consider a homogeneous portfolio $L^{(m)} = \frac{1}{m}\sum_{j=1}^{m} L_j$ of m bonds $L_j \in \{0,1\}$, where $L_j = 1$ indicates the default of the credit of company j. Each bond is characterized by the vector (S_j, s), where S_j is a latent variable, e.g. the equity value of company j. The number s denotes the default threshold in the sense that the bond of company j defaults, if $S_j < s$.

The credit loss of the portfolio is expressed as the fraction of defaulted bonds and the portfolio is homogeneous in the sense that all bonds have the same characteristics; i.e. the vector $(S_1, S_2, \dots, S_m)$ follows a factor model

$$S_j := W s^*(X, Y_j), \tag{1.1}$$

where $W > 0$, $X \in \mathbb{R}$ and $(Y_j)_{j\in\mathbb{N}}$ is an iid sequence of real random variables. The Y_j are interpreted as a company-specific risk factors, X is a common risk factor (which can be extended to a vector of common factors) and W is a global risk factor and allows for a tuning of the model.

A well-known example for $s^*(\cdot,\cdot)$ is the CreditMetrics model as described in Gupton, Finger and Bhatia (1997). We consider an *extended CreditMetrics model* given by

$$S_j = W(aX + bY_j), \quad a, b > 0 \text{ and } W > 0,\ X, Y_j \in \mathbb{R} \text{ random}. \tag{1.2}$$

The *CreditMetrics model* corresponds to $W = 1$, $X, Y_j \overset{iid}{\sim} \mathcal{N}(0,1)$ and $a = \sqrt{\rho}$, $b = \sqrt{1-\rho}$ for some $\rho \in (0,1)$, modelling the correlation between S_i and S_j for $i \neq j$. One popular extension of this model takes $W = \sqrt{\nu/\chi^2_\nu}$, which yields for $(S_1, \dots, S_m)$ a multivariate t_ν distribution, called the *multivariate t-model*.

A treatment of different credit portfolio models with a finite number of loans can be found in Frey and McNeil (2001, 2002, 2003) and in Frey, McNeil and Nyfeler (2001).

For the limiting portfolio $L := \lim_{m\to\infty} L^{(m)}$ it can be shown (see Theorem 2.3) that L is a random variable and the limit is in the almost sure sense. For model (1.2) with $W \equiv 1$ Lucas, Klaassen, Spreij and Straetmans (2003) show under weak regularity conditions that the tail behavior of L is *Weibull*-like, i.e. $P(L > q) = (1-q)^\alpha \mathcal{L}(1/(1-q))$, $q \in (0,1)$, for some $\alpha > 0$ and a *slowly varying* function $\mathcal{L}$ (see Definition 2.7 for the term Weibull-like and Definition 2.6 for the concepts of regular and slow variation).

For a random variable $W > 0$ the result remains true with the same α but a different slowly varying function $\mathcal{L}$ appears. We indicate the influence of W in Section 3 by simulation, showing that it has an important influence on the right-tail behavior of L. In Section 4 we fit four (extended) CreditMetrics models to three portfolios of different credit quality. We also investigate the fit of a simple β-model. This model, however, proves as being too simplistic in most real world credit portfolios. The extended CreditMetrics model proves to be superior provided the shock-variable W is chosen correctly.

All proofs are gathered in the Appendix.

2 Results

First, we give some notations used throughout the paper.

Notation 2.1 *(i) Random variables are always denoted by capital letters.*

(ii) $F_\cdot$ denotes the distribution function of the random variable '$\cdot$' and $f_\cdot$ denotes its density, e.g. F_X and f_X are the distribution function and density of X, respectively. Further, let $\overline{F}_\cdot := 1 - F_\cdot$ denote the tail-distribution of '$\cdot$'.

(iii) Let $h = h(x_1, x_2)$ be a function of two variables. Then $D_2 h := \partial h / \partial x_2$.

(iv) $\mathbf{1}_{\boldsymbol{A}}$ denotes the indicator function of the set $\boldsymbol{A}$.

(v) We write $a(x) \sim b(x)$ as $x \to x_0$, if $\lim_{x \to x_0} a(x)/b(x) = 1$.

(vi) We write $a(x) = o(1)$ as $x \to \infty$, if $\lim_{x \to \infty} a(x) = 0$. □

We shall investigate the tail-distribution of the limiting portfolio credit loss as defined in the following definition in combination with Theorem 2.3

Definition 2.2 *Let $L_j := \mathbf{1}_{\{S_j < s\}} = \mathbf{1}_{\{W s^*(X, Y_j) < s\}}$ denote the default indicator of the bond of company j and define the* portfolio credit loss *by*

$$L^{(m)} := \frac{1}{m} \sum_{j=1}^{m} L_j = \frac{1}{m} \sum_{j=1}^{m} \mathbf{1}_{\{W s^*(X, Y_j) < s\}}.$$

The (almost sure) limit of $L^{(m)}$ as $m \to \infty$ is called limiting portfolio credit loss *and denoted by L.* □

Theorem 2.3 *Consider the setting of Definition 2.2. Then*

$$\begin{aligned} \lim_{m \to \infty} L^{(m)} &= \lim_{m \to \infty} \frac{1}{m} \sum_{j=1}^{m} L_j \\ &\overset{\text{a.s.}}{=} \mathbb{E}(L_1 | W, X) = P(S_1 < s | W, X) =: L \overset{\text{d}}{=} F_Y \left(y^*(s/W, X) \right). \end{aligned}$$

□

Considering the variance of L, we observe the following lemma.

Lemma 2.4 *(i) Choose the setting of Definition 2.2 with* $p_{\text{loss}} := P(L_j = 1)$, *then* $0 \le \sqrt{\mathrm{Var}(L)} \le \sqrt{p_{\text{loss}}(1 - p_{\text{loss}})}$. *The upper bound is obtained for* $L_i \overset{\text{a.s.}}{=} L_j$ $\forall i, j$ *and the* $\mathrm{Var}(L) = 0$ *is obtained for* L_i *independent of* L_j $\forall i \ne j$.

(ii) In the extended CreditMetrics model (1.2) *the upper bound is obtained for* $a = 1$ *and* $b = 0$, *and the lower bound is obtained for* $a = 0$, $b = 1$ *and* $W \equiv \text{const}$.

□

Next, we introduce our key assumptions on the factor model (1.1) and the risk factors.

Assumption 2.5 *(i)* $0 < W \sim F_W$, $X \sim F_X$, $(Y_j)_{j \in \mathbb{N}}$ *are iid with* $Y_1 \sim F_Y$ *and all random variables are independent.*

(ii) Denote by $\mathcal{S}$, $\mathcal{W}$, $\mathcal{X}$ *and* $\mathcal{Y}$ *the supports of* S_j, W, X *and* Y_j, *respectively, and let* $\mathcal{W} \subseteq (0, \infty)$, $\inf \mathcal{X} = -\infty$ *and* $\sup \mathcal{Y} = +\infty$. *We further assume that* F_X *and* F_Y *have densities* f_X *and* f_Y, *respectively, and that* f_X *is monotone on some interval* $(-\infty, z_X)$ *and* f_Y *is monotone on some interval* (z_Y, ∞).

(iii) The factor model $s^*(x, y)$ *is strictly increasing, differentiable in both components and the inverse functions exist on its support; i.e. for all* $s \in \mathcal{S}$, $w \in \mathcal{W}$, *and* $x \in \mathcal{X}$ *there exists an inverse function* $y^*(s/w, x) \in \mathcal{Y}$ *and for all* $s \in \mathcal{S}$, $w \in \mathcal{W}$, *and* $y \in \mathcal{Y}$ *there exists an inverse function* $x^*(s/w, y) \in \mathcal{X}$, *so that*

$$s = ws^*(x^*(s/w, y), y) = ws^*(x, y^*(s/w, x)).$$

(iv) We assume $\lim_{y \to \infty} F_X\left(x^*(0, y)\right) / \overline{F}_Y(y) < \infty$.

(v) The default threshold s *is negative.* □

Assumption 2.5 is nothing but Assumption 1 and the comment before Assumption 2A of Lucas et al. (2003), amended by some further regularities.

Assumption 2.5(iii) says that we only consider factor models, where, given three components of (S_j, W, X, Y_j), the fourth is uniquely determined.

Assumption 2.5(iv) is needed since we extend the standard latent variable model $s^*(X, Y_j)$ by the multiplicative factor W. Note that the default probability $P(L_j = 1)$ is in general small and therefore, if $E(S_j) = 0$, we always have $s < 0$, hence Assumption 2.5(v) is not restrictive.

Assumptions 2.5 hold for a large number of factor models. For instance, they are satisfied by the CreditMetrics model as well as for the multivariate t-model. In the following we focus on the extended CreditMetrics model (1.2) and turn our attention to the right tail behavior of the limiting portfolio credit losses L. From the right tail behavior we can deduce the *riskyness* of the portfolio.

Before we specify the different types of distributions of X and Y_j further, we introduce the concept of *regular variation*.

Definition 2.6 *(i) A positive, Lebesgue measurable function r is called* regularly varying at infinity with index $\alpha \in \mathbb{R}$ *and we write $r \in \mathcal{R}_\alpha$, if*

$$r(tx)/r(x) \overset{x\to\infty}{\longrightarrow} t^\alpha,\ t > 0.$$

If $\mathcal{L} \in \mathcal{R}_0$, then $\mathcal{L}$ is called slowly varying at infinity *and we write $\mathcal{L} \in \mathcal{R}_0$.*

(ii) $r \in \mathcal{R}_\alpha$ if and only if $r(x) = x^\alpha \mathcal{L}(x)$ for $\mathcal{L} \in \mathcal{R}_0$.

(iii) If $X \sim F$ with $\overline{F} \in \mathcal{R}_{-\alpha}$ for some $\alpha \geq 0$ holds, then the random variable X is called regularly varying at infinity with index $-\alpha$ *and we write $X \in \mathcal{R}_{-\alpha}$.* □

For more details on the concept of regular variation we refer to Bingham, Goldie and Teugels (1987).

If we want to determine large losses of the limiting portfolio L we are interested in its right tail behavior near 1 and we use extreme value theory as the natural tool to describe this tail.

Definition 2.7 *We say that the random variable X or the distribution function F of X belongs to the* maximum domain of attraction *of the* Weibull *distribution*

$$\Psi_\kappa(x) = \exp\left(-\left(\max\{-x, 0\}\right)^\kappa\right), \quad \kappa > 0,$$

if for the iid sequence $X_1, X_2, \ldots \overset{iid}{\sim} F$ there exist norming constants $c_n > 0$, $d_n \in \mathbb{R}$ such that (as $n \to \infty$)

$$\left(\max\{X_1, \ldots, X_n\} - d_n\right)/c_n \overset{\mathrm{d}}{\to} \Psi_\kappa.$$

We write $X \in DA(\Psi_\kappa)$ or $F \in DA(\Psi_\kappa)$, and it can be shown that in this case F has a finite right endpoint $x_F := \sup\{x \in \mathbb{R} : F(x) < 1\} < \infty$. It also can be shown, that $F \in DA(\Psi_\kappa)$ if and only if $\overline{F}(x) = (x_F - x)^\kappa \cdot \mathcal{L}(1/(x_F - x))$ with $\mathcal{L} \in \mathcal{R}_0$, $x_F < \infty$ and $\kappa > 0$. □

For more details on extreme value theory we refer to Embrechts, Klüppelberg and Mikosch (1997) or to Resnick (1987).

The following two assumptions classify the different regimes of tail behavior of the risk factors $X \sim F_X$ and $Y_j \sim F_Y$. The first regime assumes polynomially decreasing tails of the risk factors.

Assumption 2.8 *(i)* $F_X(-\cdot) \in \mathcal{R}_{-\mu_X}$, $\mu_X > 0$, *i.e.* $F_X(-x) = x^{-\mu_X}\mathcal{L}_X(x)$, $x > 0$, *and* $\mathcal{L}_X \in \mathcal{R}_0$.

(ii) $Y_j \sim F_Y$ *and* $\overline{F}_Y \in \mathcal{R}_{-\nu_Y}$, $\nu_Y > 0$, *i.e.* $\overline{F}_Y(y) = y^{-\nu_Y}\mathcal{L}_Y(y)$, $y > 0$, *and* $\mathcal{L}_Y \in \mathcal{R}_0$.

(iii) Let $s < 0$ *be the threshold from Definition 2.2 and consider the function* $x^*(\cdot,\cdot)$ *defined in Assumption 2.5(iii). Define*

$$\zeta(w,y) \;:=\; \frac{yD_2x^*(s/w,y)}{x^*(s/w,y)}.$$

Assume $\lim_{y\to\infty}\zeta(w,y) = \zeta \in (0,\infty)$ *for any* $w \in (0,\infty)$ *pointwise. Further, assume that there exists an integrable (w.r.t.* F_W*) function* u *such that* $\zeta(w,y) \le u(w)$ *for all* $w \in (0,\infty)$, *for all* $y \in (y_0,\infty)$ *and some* y_0. □

The second regime assumes exponentially decreasing tails of the risk factors.

Assumption 2.9 *(i)* $F_X(-x) = r_X(x)\exp\left(-\mu_X x^{\mu_2}(1+\varepsilon_X(x))\right)$, $x > 0$, *where* $\varepsilon_X(x) = o(1)$, $r_X \in \mathcal{R}_{\mu_1}$, $\mu_X, \mu_2 > 0$ *and* $\mu_1 \in \mathbb{R}$. *Further, let also the derivatives* ε_X' *and* r_X' *be ultimately monotone for* $x \to \infty$.

(ii) $\overline{F}_Y(y) = r_Y(y)\exp\left(-\nu_Y y^{\nu_2}(1+\varepsilon_Y(y))\right)$, $y > 0$, *where* $\varepsilon_Y(y) = o(1)$, $r_Y \in \mathcal{R}_{\nu_1}$, $\nu_Y, \nu_2 > 0$ *and* $\nu_1 \in \mathbb{R}$. *Further, let also the derivatives* ε_Y' *and* r_Y' *be ultimately monotone for* $y \to \infty$.

(iii) Let $s < 0$ *be the fixed threshold from Definition 2.2 and consider the function* $x^*(\cdot,\cdot)$ *defined in Assumption 2.5(iii). Define*

$$\zeta(w,y) \;:=\; \frac{\mu_2\left(-x^*(s/w,y)\right)^{\mu_2-1}\left(-D_2x^*(s/w,y)\right)}{\nu_2 y^{\nu_2-1}}$$

Assume $\lim_{y\to\infty}\zeta(w,y) = \zeta \in (0,\infty)$ *for any* $w \in (0,\infty)$ *pointwise. Further, assume that there exists an integrable (w.r.t.* F_W*) function* u *such that* $\zeta(w,y) \le u(w)$ *for all* $w \in (0,\infty)$, *for all* $y \in (y_0,\infty)$ *and some* y_0. □

Note that Assumptions 2.8 and 2.9 are slightly stronger than Assumptions 2A and 2B in Lucas et al. (2003), since we use the existence of a density of L in the proof of the following Theorem 2.10.

We now determine the right tail behavior of the limiting portfolio credit loss distribution.

Theorem 2.10 *Consider the setting of Assumption 2.5. If Assumptions 2.8 or 2.9 are satisfied, then* $L \in DA(\psi_\kappa)$ *with* $\kappa = \zeta\mu_X/\nu_Y > 0$, *i.e. there exists* $\mathcal{L} \in \mathcal{R}_0$ *such that*

$$P(L > q) = (1-q)^{\zeta\mu_X/\nu_Y}\mathcal{L}(1/(1-q)), \quad q \in (0,1). \tag{2.3}$$

□

For $W \equiv 1$ this result has been proved in Lucas et al. (2003), Theorems 2 and 3. Hence, our result shows that the tail of the portfolio loss is in first order robust with respect to a shock variable W. Consequently, any difference between $W \equiv 1$ and a random $W > 0$ can only be found in the second order tail expansion, the slowly varying function $\mathcal{L}(1/(1-q))$.

As an example, we derive an analytic expression of $\mathcal{L}(1/(1-q))$ in the extended CreditMetrics framework, both, in the setting of Assumptions 2.8 and 2.9:

Theorem 2.11 *Given the extended CreditMetrics model* (1.2) *with* $X \sim t_{\mu_X}$, $Y_j \sim t_{\nu_Y}$ *and* $W > 0$ *such that* $\mu_X \geq \nu_Y$. *Let Assumptions 2.5 hold. Then the distribution of* L *is of the form (2.3) with* $\kappa = \mu_X/\nu_Y$ *and* $\mathcal{L} \in \mathcal{R}_0$ *satisfies for* $q \to 1$ *the relation*

$$\mathcal{L}\left(\frac{1}{1-q}\right) \sim C_{\mu_X}\mu_X^{(\mu_X-1)/2}\int_0^\infty\left(-\frac{s}{aw}(1-q)^{1/\nu_Y} + \frac{b}{a}C_{\nu_Y}^{1/\nu_Y}\nu^{(\nu_Y-1)/(2\nu_Y)}\right)^{-\mu_X} \mathrm{d}F_W(w).$$

□

Theorem 2.12 *Given the extended CreditMetrics model* (1.2) *with* $X, Y_j \sim \mathcal{N}(0,1)$ *and* $W > 0$ *such that* $E(1/W) < \infty$ *and* $b \geq a$. *Let Assumptions 2.5 hold. Then the distribution of* L *is of the form (2.3) with* $\kappa = b^2/a^2$ *and* $\mathcal{L} \in \mathcal{R}_0$ *satisfies for* $q \to 1$ *the relation*

$$\mathcal{L}\left(\frac{1}{1-q}\right) \sim \int_0^\infty \exp\left(-\frac{s^2}{2aw} + \frac{sb}{a^2w}\sqrt{-2\ln(1-q)}\right)\frac{(-2\ln(1-q))^{b^2/(2a^2)}}{\left|\frac{s}{aw} - \frac{b}{a}\sqrt{-2\ln(1-q)}\right|}\,\mathrm{d}F_W(w).$$

□

Remark 2.13 (i) In the setting of Theorem 2.12 we require $b \geq a > 0$. The natural choice in this model is $a = \sqrt{\rho}$ and $b = \sqrt{1-\rho}$ for $\rho \in (0,1)$ modelling the correlation between S_i and S_j for $i \neq j$. Then, $b \geq a$ is equivalent to $\rho \leq 1/2$ and this is always given in practice.

(ii) The first order tail behavior is a function of the correlation ρ only.

(iii) As can be seen in the proof, for the CreditMetrics model Assumptions 2.5(iv) and 2.8(iii) or 2.9(iii) are superfluous. However, in the extended model, one can easily construct examples, where these restrictions are essential. □

Setting $W \equiv 1$ in Theorem 2.12 we immediately obtain Theorem 6 of Lucas et al. (2003).

Corollary 2.14 (Lucas et al. (2003), Theorem 6) *For the CreditMetrics model with* $b \geq a$ *the tail-distribution of* L *is of the form (2.3) with* $\kappa = b^2/a^2$ *and* $\mathcal{L} \in \mathcal{R}_0$ *satisfies for* $q \to 1$ *the relation*

$$\mathcal{L}\left(\frac{1}{1-q}\right) \sim \frac{a}{b}\exp\left(-\frac{s^2}{2a} + \frac{sb}{a^2}\sqrt{-2\ln(1-q)}\right)(-2\ln(1-q))^{(b^2-a^2)/(2a^2)}.$$

□

3 A simulation study

We focus on the extended CreditMetrics model (1.2). Denote the default probability by $p_{\text{loss}} = P(S_j \le s)$ and we assume that $p_{\text{loss}} < 1/2$. We consider different distributions of W, X and Y_j and show their influence on the tail-distribution of the limiting portfolio credit loss L. We consider the following examples.

Model 3.1 *(1)* $W \equiv 1$ *and* $X, Y_j \overset{iid}{\sim} \mathcal{N}(0,1)$ *and* $b \ge a$.

(2) $W \overset{\mathrm{d}}{=} \sqrt{4/\chi_4^2}$ *and* $X, Y_j \overset{iid}{\sim} \mathcal{N}(0,1)$ *and* $b \ge a$.

(3) $W \equiv 1$ *and* $X \sim t_{\mu_X}$, $Y_j \sim t_{\nu_Y}$ *and* $\mu_X \ge \nu_Y > 2$.

(4) $W \overset{\mathrm{d}}{=} \sqrt{4/\chi_4^2}$ *and* $X \sim t_{\mu_X}$, $Y_j \sim t_{\nu_Y}$ *and* $\mu_X \ge \nu_Y > 2$. □

As shown in Theorems 2.11 and 2.12, all these models fall into the framework of our assumptions, i.e. for $q \in (0,1)$ there are functions $\mathcal{L}_1, \dots, \mathcal{L}_4 \in \mathcal{R}_0$ such that

$$\begin{aligned} P(L > q) &= (1-q)^{b^2/a^2} \mathcal{L}_{1,2}(1/(1-q)), \text{ in case of model 1 and 2,} \\ P(L > q) &= (1-q)^{\mu_X/\nu_Y} \mathcal{L}_{3,4}(1/(1-q)), \text{ in case of model 3 and 4.} \end{aligned}$$

As indicated in Remark 2.13 the restriction $b \ge a$ for model 1 and 2 is quite natural corresponding to $\rho < 1/2$; see Table 1 for some scenarios. The restriction $\mu_X \ge \nu_Y$ for model 3 and 4 can be seen in the same spirit as we choose $\mu_X/\nu_Y = b^2/a^2 \ge 1$. The bound $\nu_Y, \mu_X > 2$ is needed to ensure finite variance of S_j.

To make the four models comparable, we fix the following parameters

- the default probability $p_{\text{loss}} := P(S_j \le s)$,
- the correlation-structure $\rho := \text{Corr}(S_i, S_j)$ $\forall i \ne j$ and
- the first order tail behavior $\kappa = \zeta \mu_X/\nu_Y$ of the limiting portfolio credit loss L, given by Theorem 2.10.

For all models we have $\text{Corr}(S_i, S_j) = a^2 EX^2/(a^2 EX^2 + b^2 EY_i^2)$ $\forall i \ne j$. Let $a = \sqrt{\rho}$, $b = \sqrt{1-\rho}$ in models 1, 2 and $a = \sqrt{\rho(\mu_X - 2)/\mu_X}$, $b = \sqrt{(1-\rho)(\nu_Y - 2)/\nu_Y}$, $\mu_X, \nu_Y > 2$ in models 3, 4. Then we have always the same correlation $\rho \in (0,1)$ in all models.

By Theorem 2.12 we have $\kappa = b^2/a^2 = (1-\rho)/\rho$ in model 1 and 2 as the parameter of the first order tail behavior. In model 3 and 4 we get $\kappa = \mu_X/\nu_Y$ (by Theorem 2.11), therefore we choose $\mu_X = 2/\rho$ and $\nu_Y = 2/(1-\rho)$ and this leads to $a = b = \sqrt{\rho(1-\rho)}$. Hence we have the same κ in all models.

The threshold s is the p_{loss}-quantile of S_j. Since $S_j \sim \mathcal{N}(0,1)$ in model 1 and $S_j \sim t_4$ in model 2, we can read off this quantile from standard tables. In model 3 and 4, we choose s as the empirical p_{loss}-quantile of S_j. The simulation run length is 10^7, which should suffice to obtain a reliable estimate.

In choosing the specific default probabilities and correlations we follow Frey, McNeil and Nyfeler (2001), i.e. we consider three rating groups of decreasing credit quality, which we label A, B and C; see Table 1. This leads to the (rounded) parameters given in Table 2.

group	A	B	C
p_{loss}	0.01%	0.50%	7.50%
ρ	2.58%	3.80%	9.21%

Table 1: Values for default probability and correlation of the three credit quality groups.

As stated in Theorem 2.3 we have $L \stackrel{\text{d}}{=} F_Y(s/(bW) - Xa/b)$ and we simulate L by this distributional equality, see Figures 1 to 3 corresponding to the three groups. Each of Figures 1 to 3 shows four graphs, each with four curves, corresponding to the different models (1)-(4) with parameters as given in Tables 1 and 2.

The upper left graph corresponds to the tail-distribution $\overline{L}(q)$ of the limiting portfolio, where the arguments q are chosen such that $0 \le \overline{L}(q) \le 0.1$ for all four models; the lower left graph is similar but zoomed in, i.e. q is such that $0 \le \overline{L}(q) \le 0.01$. The right graphs show the quantile functions or the *Value-at-Risk* $L^{\leftarrow}(p) = \text{VaR}_p$ of the portfolios with $0.9 \le p \le 1$ and $0.99 \le p \le 1$, respectively.

In Table 3 the VaR_p of all models in the three groups for p running through the different values $95\%, 99\%, 99.5\%, 99.9\%, 99.95\%$.

We observe in all groups that model 2 leads to a portfolio with larger quantiles than model 1 and, similarly, model 4 gives larger quantiles than model 3; this is obviously due to W. Although three parameters are the same in all models, we observe a completely different behavior of the four models in their right tails. As can be seen in Table 3, the 99.95%-quantile of model 2 in group A is 90 times larger than in case of model 1 and even 440 times larger than in case of model 3. In group B we observe in model 2 an up to 25 times larger 99.95%-quantile than in model 3 and in group C the riskyness of the models turn where model 4 shows up to 50% larger quantiles than model 2.

To quantify the different portfolio behavior further we also estimate empirically the standard deviation of L, see Table 4. The (rounded) 95% confidence intervals are, as usual, based on the asymptotic χ^2_{n-1} distribution of the empirical variance. We

Model 1: $X, Y_i \overset{iid}{\sim} \mathcal{N}(0,1)$ and $W \equiv 1$:							
		a	b	s	κ	μ_X	ν_Y
group	A	.161	.987	-3.73	37.8	.500	.500
	B	.195	.981	-2.58	25.3	.500	.500
	C	.303	.953	-1.44	9.86	.500	.500

Model 2: $X, Y_i \overset{iid}{\sim} \mathcal{N}(0,1)$ and $W \overset{\mathrm{d}}{=} \sqrt{4/\chi_4^2}$:							
		a	b	s	κ	μ_X	ν_Y
group	A	.161	.987	-13.0	37.8	.500	.500
	B	.195	.981	-4.60	25.3	.500	.500
	C	.303	.953	-1.78	9.86	.500	.500

Model 3: $X \sim t_{\mu_X}$, $Y_i \sim t_{\nu_Y}$ and $W \equiv 1$:							
		a	b	s	κ	μ_X	ν_Y
group	A	.159	.159	-10.1	37.8	77.5	2.05
	B	.191	.191	-1.81	25.3	52.6	2.08
	C	.289	.289	-.782	9.86	21.7	2.20

Model 4: $X \sim t_{\mu_X}$, $Y_i \sim t_{\nu_Y}$ and $W \overset{\mathrm{d}}{=} \sqrt{\chi_1^2}$:							
		a	b	s	κ	μ_X	ν_Y
group	A	.159	.159	-10.2	37.8	77.5	2.05
	B	.191	.191	-1.88	25.3	52.6	2.08
	C	.289	.289	-.679	9.86	21.7	2.20

Table 2: Parameter setting of the four models in the three groups (given by Table 1).

observe that model 2 has a larger empirical deviation than model 1 and, similarly, model 4 shows larger deviation than model 3. As in case of the VaR the differences of the standard deviations are not negligible: in group A model 2 shows 850 times more deviation than model 3; see Table 4. From Lemma 2.4 we get an upper bound for the standard deviation and observe in all our models a quite small standard deviation compared to the upper bound, see also Table 4. The meaning of the last line in Table 4 will be explained in the following section.

4 Cutting Gordon's knot

Recall that for all models of section 3 the parameters where chosen such that default probability, correlation and first order tail behavior are the same for all models in

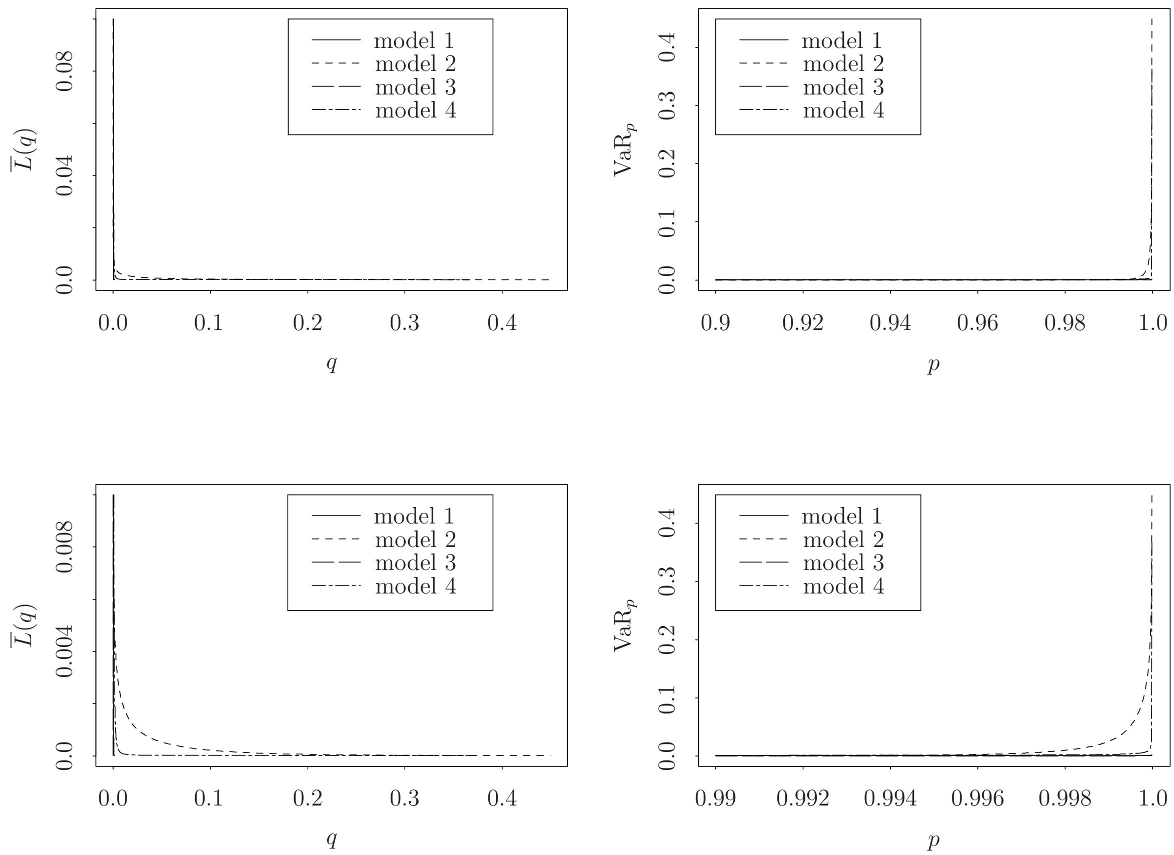

Figure 1: Tail-distribution and Value-at-Risk of the four models with group A-parameter setting.

each group A-C. Nonetheless, we observe completely different upper tails for the different models. This indicates that a naive quantile estimator based on extreme value theory may be grossly misleading. Such a method would concentrate on the parameter κ in Theorem 2.10 and replace the slowly varying function $\mathcal{L}$ by a constant, see Chapter 6 of Embrechts et al. (1997) for details. However, as can be seen in Theorem 2.12, $\mathcal{L}$ is far away from being constant and has a strong influence near the right endpoint $q = 1$.

To overcome the problem, which model to choose, we suggest in the following a pragmatic approach, which originates in the β-model. The β model is a simple model often used in practice, where the parameters are estimated by matching the first two moments; see e.g. Bluhm et al. (2003), p.39. The $\beta(c, d)$-distribution has density

$$f_{\beta(c,d)}(q) = \frac{\Gamma(c+d)}{\Gamma(c)\Gamma(d)} q^{c-1}(1-q)^{d-1}, \quad 0 < q < 1, \quad c, d > 0.$$

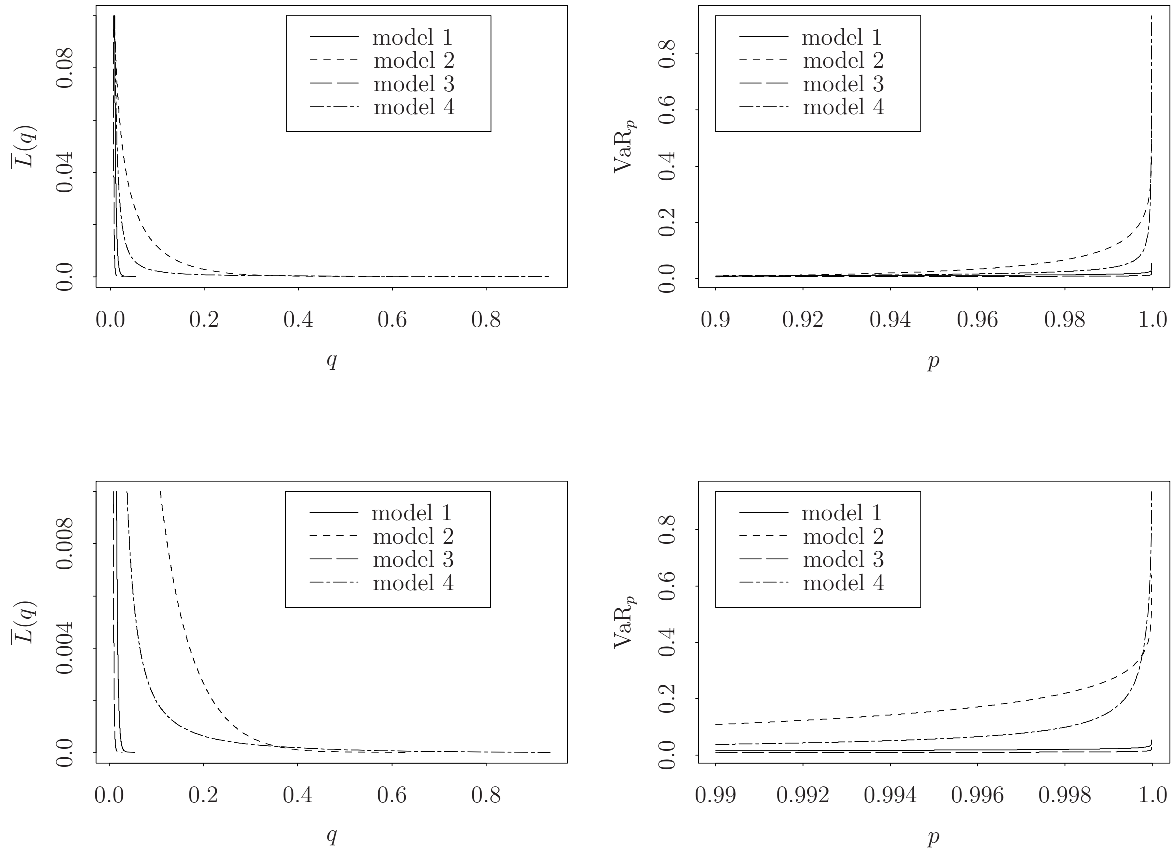

Figure 2: Tail-distribution and Value-at-Risk of the four models with group B-parameter setting.

From Example 3.3.17 of Embrechts et al. (1997) we know that the $\beta(c,d)$-distribution satisfies the weak requirement of being in $DA(\Psi_d)$. As our main focus is on VaR-estimation, we fit besides the location parameter the first order tail behavior. Since $\kappa = (1-\rho)/\rho$ and $E\beta(c,d) = c/(c+d)$ we obtain

$$c = \frac{1-\rho}{\rho}\frac{p_{\text{loss}}}{1-p_{\text{loss}}} \quad \text{and} \quad d = \frac{1-\rho}{\rho}. \tag{4.4}$$

This means we match the default probability p_{loss} and the correlation ρ.

We observe that VaR estimated from the β-model compared to our models 1-4 is slightly more moderate but roughly of the same order as for model 2 in all groups; see Table 3.

The question arises, if there is any further advantage of the latent variable models 1-4 in comparison to the simple and easy to fit β-model for VaR estimation, which after all, has the correct first order tail behaviour. One drawback of the β-model is that it has no economic interpretation in the credit risk context. From a statistical

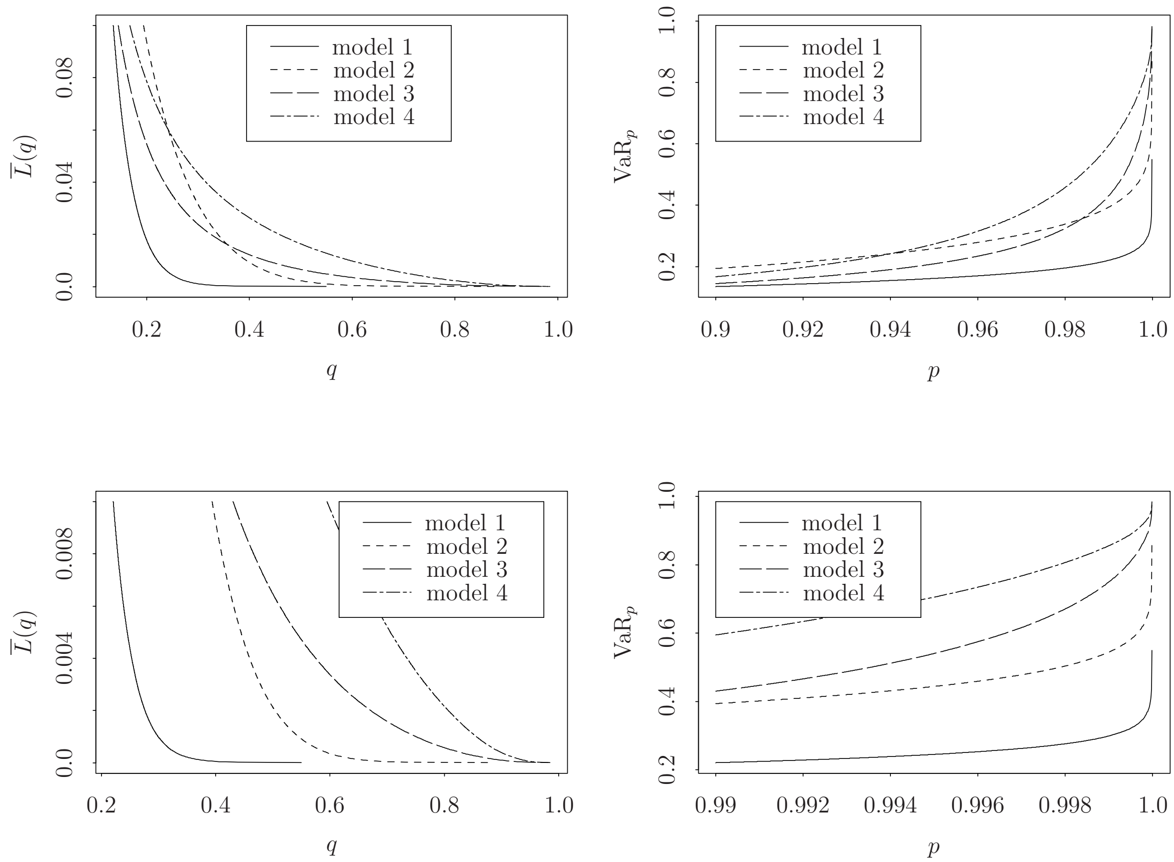

Figure 3: Tail-distribution and Value-at-Risk of the four models with group C-parameter setting.

point of view, models 2 and 4 constitute a much richer class of models in the sense that more parameters can be specified.

One parameter, which we have not considered up to now is the standard deviations (see Table 4) and here we can observe substantial differences between the models. As the first order tail behavior is determined by ρ solely, it is independent of W. As W acts as a random standard deviation of the factor models, it is natural to match the empirical standard deviation by choosing a proper W. In our simulations we observe for models 2 and 4 that the standard deviation of L_ν is decreasing in ν_W. For the normal factor model 2 we can proof this by asymptotic expansion. Consequently, we can estimate ν_W by matching the standard deviation.

Theorem 4.1 *In the setting of model 2, let* $W = W_\nu \stackrel{\mathrm{d}}{=} \sqrt{\nu_W/\chi^2_{\nu_W}}$ *and denote* $L_j = L_{j,\nu}$ *and* $L = L_\nu$. *Then, the standard deviation of* L_ν *is decreasing in* ν *for sufficiently small default probability* $p_{\text{loss}} = P(L_{j,\nu} = 1) = P(S_j < s)$. $\square$

We conclude this section by a comparison of the extended CreditMetrics models

VaR$_p$ for group A, $p_{\text{loss}} = 0.0001$ and $\rho = 0.0258$.						
	p	95%	99%	99.5%	99.9%	99.95%
model	1	$2.18 \cdot 10^{-4}$	$3.29 \cdot 10^{-4}$	$3.82 \cdot 10^{-4}$	$5.14 \cdot 10^{-4}$	$5.76 \cdot 10^{-4}$
	2	$1.54 \cdot 10^{-8}$	$1.76 \cdot 10^{-4}$	$1.46 \cdot 10^{-3}$	$2.48 \cdot 10^{-2}$	$5.12 \cdot 10^{-2}$
	3	$1.09 \cdot 10^{-4}$	$1.12 \cdot 10^{-4}$	$1.13 \cdot 10^{-4}$	$1.15 \cdot 10^{-4}$	$1.16 \cdot 10^{-4}$
	4	$2.98 \cdot 10^{-4}$	$7.30 \cdot 10^{-4}$	$1.06 \cdot 10^{-3}$	$2.46 \cdot 10^{-3}$	$3.52 \cdot 10^{-3}$
	β	$1.91 \cdot 10^{-8}$	$1.10 \cdot 10^{-3}$	$4.73 \cdot 10^{-3}$	$2.37 \cdot 10^{-2}$	$3.47 \cdot 10^{-2}$

VaR$_p$ for group B, $p_{\text{loss}} = 0.005$ and $\rho = 0.038$.						
	p	95%	99%	99.5%	99.9%	99.95%
model	1	0.0107	0.0152	0.0173	0.0221	0.0242
	2	0.0254	0.108	0.155	0.265	0.308
	3	0.00715	0.00871	0.00942	0.0113	0.0122
	4	0.0143	0.0376	0.0568	0.151	0.226
	β	0.0285	0.069	0.0886	0.135	0.155

VaR$_p$ for group C, $p_{\text{loss}} = 0.075$ and $\rho = 0.0921$.						
	p	95%	99%	99.5%	99.9%	99.95%
model	1	0.162	0.221	0.245	0.299	0.321
	2	0.259	0.394	0.444	0.544	0.581
	3	0.209	0.431	0.541	0.750	0.810
	4	0.274	0.595	0.706	0.856	0.889
	β	0.233	0.345	0.388	0.478	0.513

Table 3: VaR$_p$, $p = 95\%, 99\%, 99.5\%, 99.9\%, 99.95\%$, for the four models and the fitted β-distribution in the three groups.

2 and 4 with the β-model estimated from the parameters p_{loss} and ρ given in Table 1. The estimated parameters for group A-C are given in Table 2. We estimate σ from the parameters (4.4) of the β-model; the estimates are given in the last line in Table 4. We see that in case of group C the standard deviation of model 3 is already slightly larger than in the β-case, therefore we set $\nu_W = \infty$ (corresponding to $W \equiv 1$) for model 4 in group C. All results are summarized in Table 5.

In Figure 4 we plot the tail-distribution (left column) and the VaR$_p$ (right column), where the upper, middle and lower row correspond to group A, B and C, respectively. In Table 6 we also give the VaR$_p$ estimates of model 2, 4 and β-model in the three groups for certain values of p. We observe now that model 2 and the β-model are very similar in all groups, indicating that the β-model gives a reason-

	group		
	A	B	C
$\sigma_{\max}$	0.01	0.0705	0.263
	$\widehat{\sigma}_L$		
model 1	$6.54 \cdot 10^{-5} \pm 2 \cdot 10^{-6}$	$3.00 \cdot 10^{-3} \pm 9 \cdot 10^{-5}$	$4.50 \cdot 10^{-2} \pm 2 \cdot 10^{-3}$
2	$2.93 \cdot 10^{-3} \pm 9 \cdot 10^{-5}$	$2.16 \cdot 10^{-2} \pm 7 \cdot 10^{-4}$	$8.72 \cdot 10^{-2} \pm 3 \cdot 10^{-3}$
3	$3.38 \cdot 10^{-6} \pm 1 \cdot 10^{-7}$	$1.18 \cdot 10^{-3} \pm 4 \cdot 10^{-5}$	$7.84 \cdot 10^{-2} \pm 3 \cdot 10^{-3}$
4	$3.28 \cdot 10^{-4} \pm 1 \cdot 10^{-5}$	$1.24 \cdot 10^{-2} \pm 4 \cdot 10^{-4}$	$1.07 \cdot 10^{-1} \pm 4 \cdot 10^{-3}$
β	$1.61 \cdot 10^{-3}$	$1.37 \cdot 10^{-2}$	$7.71 \cdot 10^{-2}$

Table 4: Estimated standard deviations $\widehat{\sigma}_L$ with 95%CI of the estimator for the four models in the three groups. The last line shows the standard deviation of the fitted β-model.

	group		
$\widehat{\nu_W}$	A	B	C
model 2	7.77	8.22	5.74
4	2.65	3.72	0

Table 5: Estimated ν_W for model 2 and 4 with $W \sim \sqrt{\nu_W/\chi^2_{\nu_W}}$.

able approximation for model 2, provided the standard deviations of both models coincide. In other words, the similarity of model 2 and the β-model suggests model 2 as a substantial improvement of the β-model.

As to model 4, we see that in group A the quantiles of model 2 and β are roughly three times larger than in model 4. In group B, all three models are comparable and in group C model 4 behaves roughly 50% riskier than the other models. We shall further comment on model 4 in the next section.

5 A word of warning

In the heavy-tailed models 3 and 4 we restrict the parameters to $\mu_X \geq \nu_Y$, i.e. we consider only X being not heavier-tailed than Y_j. As can be seen in Table 2 we always have $\mu_X > \nu_Y$ with a rather large ratio $\mu_X/\nu_Y > 9.8$. We did this for good reasons. Because, if $\nu = \mu_X = \nu_Y$, then this models a very extreme economic situation, the more extreme, the smaller ν is. In this case X (and Y_j) have extremely heavy tails and, thus, have with very high probability extremely large realizations. Consequently, it can happen that a large negative observation of X dominates all Y_j such that almost the whole portfolio defaults. This would model an economy which

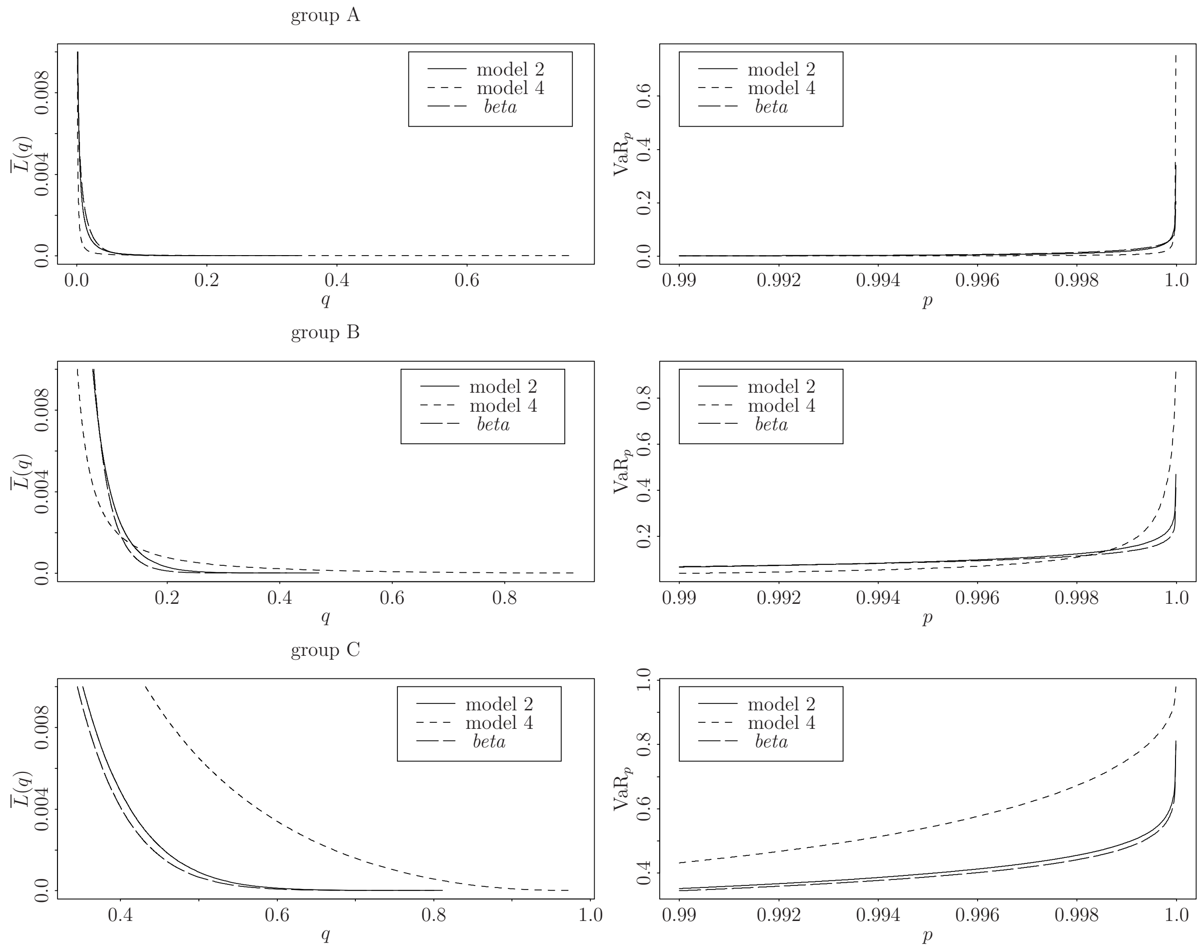

Figure 4: Tail-distribution and Value-at-Risk of model 2, 4 and β-model in the three groups.

fluctuates wildly. In that case the limiting portfolio credit loss behaves like the model built on $S_j = \min\{\sqrt{\rho}X, \sqrt{1-\rho}Y_j\}$.

Corollary 5.1 *Define* $L_{\wedge}^{(m)} := \frac{1}{m}\sum_{j=1}^{m} L_j^{\wedge}$ *with* $L_j^{\wedge} := \mathbf{1}_{\left\{\min\{\sqrt{\rho}X,\sqrt{1-\rho}Y_j\}\leq s\right\}}$. *Let* $X, Y_1, \ldots, Y_m \overset{iid}{\sim} t_\nu$. *Then*

$$L^{\wedge} := \lim_{m\to\infty} L_{\wedge}^{(m)} \overset{\text{a.s.}}{=} \begin{cases} 1, & \textit{with probability } F_{t_\nu}\left(s/\sqrt{\rho}\right), \\ F_{t_\nu}\left(s/\sqrt{1-\rho}\right), & \textit{with probability } \overline{F}_{t_\nu}(s/\sqrt{\rho}). \end{cases}$$

□

Theorem 5.2 *Choose the model* $L_{\wedge}^{(m)}$ *and* $L^{\wedge}$ *as in the setting above. Let* $L^{(m)}$ *and* L *correspond to model 3 with* $\mu_X = \nu_Y =: \nu$. *Further, choose the same default threshold* s *for both models.*

(i) Let m *be fixed. Then,* $\lim_{s\to-\infty} P\left(L^{(m)} = q \middle| L_{\wedge}^{(m)} = q\right) = 1$, *for any* $q \in \{0, 1/m, 2/m \ldots, 1\}$.

VaR_p for group A, $p_{\text{loss}} = 0.0001$, $\rho = 0.0258$ and $\sigma = 1.61 \cdot 10^{-3}$.						
	p	95%	99%	99.5%	99.9%	99.95%
model	2	$6.91 \cdot 10^{-5}$	$1.69 \cdot 10^{-3}$	$4.19 \cdot 10^{-3}$	$1.88 \cdot 10^{-2}$	$2.99 \cdot 10^{-2}$
	4	$2.66 \cdot 10^{-4}$	$9.62 \cdot 10^{-4}$	$1.65 \cdot 10^{-3}$	$5.81 \cdot 10^{-3}$	$1.00 \cdot 10^{-2}$
	beta	$1.91 \cdot 10^{-8}$	$1.10 \cdot 10^{-3}$	$4.73 \cdot 10^{-3}$	$2.37 \cdot 10^{-2}$	$3.47 \cdot 10^{-2}$

VaR_p for group B, $p_{\text{loss}} = 0.005$, $\rho = 0.038$ and $\sigma = 1.37 \cdot 10^{-2}$.						
	p	95%	99%	99.5%	99.9%	99.95%
model	2	0.0249	0.0673	0.0906	0.151	0.180
	4	0.0145	0.0400	0.0617	0.170	0.257
	beta	0.0285	0.0693	0.0886	0.135	0.155

VaR_p for group C, $p_{\text{loss}} = 0.075$, $\rho = 0.0921$ and $\sigma = 7.71 \cdot 10^{-2}$.						
	p	95%	99%	99.5%	99.9%	99.95%
model	2	0.234	0.352	0.398	0.495	0.531
	4	0.209	0.432	0.542	0.750	0.810
	beta	0.233	0.345	0.388	0.478	0.513

Table 6: VaR_p, $p = 95\%, 99\%, 99.5\%, 99.9\%, 99.95\%$ for model 2, 4 and β-model in the three groups.

(ii) Let $\varepsilon > 0$. Then, $\lim_{s \to -\infty} P\left(|L - L^\wedge| < \varepsilon \mid L^\wedge\right) \overset{\text{a.s.}}{=} 1.$

□

From Theorem 5.2(ii) conclude that in the setting of model 3, where s, ρ and $\mu_X = \nu_Y = \nu$ are small, the limiting portfolio credit loss L degenerates in the sense that most of the mass is near the point $F_{t_\nu}\left(s/\sqrt{1-\rho}\right)$ (when the Y_j's dominate the portfolio) and some very rare events can be observed close to 1 (when X dominates the portfolio). From Theorem 5.2(i) conclude that this behavior also can be observed for portfolios with a finite number of loans. Of course, model 4 has the same structure; the difference to model 3 being that large fluctuations are multiplied by a random W.

6 Conclusion

In this paper we derived the tail behavior of aggregate credit losses extending results of Lucas et al. (2003). We enriched the one factor latent variable model by a positive multiplicative shock variable W. In the models, where the latent variables follow a multivariate normal or t-distribution, we observed that first order tail behavior is a function of the correlation between the latent variables. In particular, W has no influence on the first order tail behavior of the limiting credit loss portfolio.

In a simulation study we observed an impact of the second order tail behavior on the quantiles by comparing four different models. We fitted the models by matching default probability, correlation between latent variables and first order tail behavior. In some credit scenario we observed quantiles that were up to 440 times larger than in another scenario.

To offer some decision support to the risk manager on which model to choose, we compared the VaR estimated from the β-model with the VaR estimated from the four extended CreditMetrics models. Fixing default probability and first order tail behavior we observed a similar (slightly more moderate) performance of the β-model and the multivariate t-model. From Section 5 we learned to be aware of the influence of heavy-tailed latent variables as the limiting credit loss portfolio may degenerate. This suggests the β-model as a simple model based on the fit of two quantities of interest, either matching the first two moments, or, perhaps more advisable in the context of risk management and VaR estimation, loss probability and correlation.

The multivariate t-model offers an improved fit by the shock variable W. We showed that W can influence the standard deviation without having influence on the other parameters. As for small loss probabilities the standard deviation of the limiting credit loss distribution decreases in ν, we estimate ν by matching the standard deviation. Consequently, the multivariate t-model improves the fit of the β-model.

Appendix

Proof of Lemma 2.4: As in Definition 2.2 set $L_i := \mathbf{1}_{\{Ws^*(X,Y_i)\le s\}}$ with $EL_i = p_{\text{loss}}$, $L^{(m)} = \frac{1}{m}\sum_{i=1}^m L_j$. We observe

$$\operatorname{Var}\left(\sum_{i=1}^m L_j\right) = E\left(\left(\sum_{i=1}^m L_j\right)^2\right) - \left(E\left(\sum_{i=1}^m L_j\right)\right)^2 \tag{A.1}$$

$$= \sum_{i,j=1}^m E(L_iL_j) - m^2p_{\text{loss}}^2 \le m^2p_{\text{loss}}(1-p_{\text{loss}}), \tag{A.2}$$

since $E(L_iL_j) \le E(L_i) = p_{\text{loss}}$. Hence, $\operatorname{Var}L^{(m)} \le p_{\text{loss}}(1-p_{\text{loss}})$ for all m, and, obviously, $\operatorname{Var}L^{(m)} = p_{\text{loss}}(1-p_{\text{loss}})$ holds for $L_i \overset{\text{a.s.}}{=} L_j$.

By Theorem 2.3 (independent of Lemma 2.4), we have $\lim_{m\to\infty} L^{(m)} \overset{\text{a.s.}}{=} L$ and L has bounded support $(0,1)$, hence $\operatorname{Var}L = \operatorname{Var}\left(\lim_{m\to\infty} L^{(m)}\right) = \lim_{m\to\infty}\operatorname{Var}L^{(m)} \le p_{\text{loss}}(1-p_{\text{loss}})$.

In the extended CreditMetrics setting $S_j = W(aX+bY_j)$ obviously we have for $a=1$, $b=0$ that $L_i \overset{\text{a.s.}}{=} L_j$, for all i,j, hence $L^{(m)} \overset{\text{a.s.}}{=} L_1$ therefore $\operatorname{Var}L = \operatorname{Var}L_1 = p_{\text{loss}}(1-p_{\text{loss}})$ and for $W = \text{const} \in (0,\infty)$, $a=0$, $b=1$ we have $L_1, L_2, \ldots \overset{iid}{\sim} \operatorname{Ber}(p_{\text{loss}})$, therefore $\operatorname{Var}L = 0$. $\square$

Proof of Theorem 2.3: Given W and X, the indicator variables $L_j = \mathbf{1}_{\{S_j\le s\}}$ are iid, hence a conditional law of large numbers holds as $m\to\infty$ with

$$L^{(m)} = \frac{1}{m}\sum_{j=1}^m \mathbf{1}_{\{S_j\le s\}} = \frac{1}{m}\sum_{j=1}^m \mathbf{1}_{\{Ws^*(X,Y_j)\le s\}}$$

$$\xrightarrow{\text{a.s.}} \quad E\left(\mathbf{1}_{\{S_1\le s\}}\,|W,X\right) = P\left(S_1 \le s\,|W,X\right) =: L.$$

Furthermore, by Assumption 2.5(iv), $s^*(\cdot,\cdot)$ is increasing and invertible with respect to the second component, therefore

$$L = P(Ws^*(X,Y_1)\le s|W,X) = P(Y_1 \le y^*(s/W,X)|W,X) \overset{\text{d}}{=} F_Y\left(y^*(s/W,X)\right).$$

$\square$

For the proof of Theorem 2.10 we need the following Lemmas A.1, A.2 and A.3.

Lemma A.1 (Smith (1983), Theorem 10.3, Chapter 13) *Let μ be a finite measure on $\boldsymbol{A} \subset \mathbb{R}^m$, $\boldsymbol{B}$ an open interval in $\mathbb{R}$ and $h : \boldsymbol{A}\times\boldsymbol{B} \to \mathbb{R}^n$ defined by $(w,t) \mapsto h(w,t)$. Assume that the following holds.*

(i) For almost every $t \in \boldsymbol{B}$, the function $h(\cdot, t)$ is measurable on $\boldsymbol{A}$ and for some t it is integrable.

(ii) For almost every $w \in \boldsymbol{A}$, the function $h(w, \cdot)$ is $\mathcal{C}^1$ on $\boldsymbol{B}$.

(iii) There is an integrable $u : \boldsymbol{A} \to \mathbb{R}$ such that

$$|D_2 h(w,t)| \leq u(w) \quad \text{for all } t \in \boldsymbol{B} \text{ and almost all } w \in \boldsymbol{A}.$$

Then the function $\int h(w, \cdot)\, \mathrm{d}\mu(w)$ is $\mathcal{C}^1$ and satisfies

$$\frac{\partial}{\partial t} \int h(w,t)\, \mathrm{d}\mu(w) \;=\; \int D_2 h(w,t)\, \mathrm{d}\mu(w).$$

□

Lemma A.2 *(i) Choose the setting of Assumption 2.8, then*

$$f_X(-x) \sim \frac{\mu_X}{x} F_X(-x) \text{ as } x \to \infty \quad \text{and} \quad f_Y(y) \sim \frac{\nu_Y}{y} \overline{F}_Y(y) \text{ as } y \to \infty.$$

(ii) Choose the setting of Assumption 2.9, then

$$\begin{aligned} f_X(-x) &\sim \mu_X \mu_2 x^{\mu_2 - 1} F_X(-x), \quad \text{as } x \to \infty, \text{ and} \\ f_Y(y) &\sim \nu_Y \nu_2 y^{\nu_2 - 1} \overline{F}_Y(y), \quad \text{as } y \to \infty. \end{aligned}$$

Proof: In the setting of Assumption 2.8 just apply the *Monotone Density Theorem*, e.g. Theorem 1.7.2 in Bingham et al. (1987), since the densities f_X and f_Y are ultimately monotone. In the setting of Assumption 2.9 we have (the asymptotic behavior of f_Y is shown similarly)

$$F_X(-x) \;=\; r_X(x) \exp\left(-\mu_X x^{\mu_2}(1 + \varepsilon_X(x))\right).$$

We obtain

$$f_X(-x) = F_X(-x) \mu_X \mu_2 x^{\mu_2 - 1} \left(1 + \varepsilon_X(x) - \frac{x \varepsilon'_X(x)}{\mu_2} + \frac{x^{1-\mu_2}}{\mu_X \mu_2} \frac{r'_X(x)}{r_X(x)}\right).$$

As r'_X is ultimately monotone, the monotone density theorem yields $r'_X(x)/r_X(x) \sim c/x$ as $x \to \infty$. Since $\mu_2 > 0$, it follows that

$$\frac{x^{1-\mu_2}}{\mu_X \mu_2} \frac{r'_X(x)}{r_X(x)} \sim \frac{c}{\mu_X \mu_2} x^{-\mu_2} \overset{x \to \infty}{\longrightarrow} 0.$$

Considering $x \varepsilon'_X(x)$ choose x such that $\varepsilon'_X(\xi)$ is monotone for all $\xi \geq x$. Note that $\varepsilon_X(x) = o(1)$ and monotonicity of ε'_X implies $\varepsilon'_X(x) = o(1)$. Without loss of generality let $\varepsilon'_X > 0$ be decreasing. Hence there exists $\delta > 0$ such that

$$-\varepsilon_X(x) = \int_x^\infty \varepsilon'_X(\xi)\, \mathrm{d}\xi \geq \sum_{i=\lfloor x \rfloor + 1}^{\infty} \int_i^{i+1} \varepsilon'_X(\xi)\, \mathrm{d}\xi \geq \sum_{i=\lfloor x \rfloor + 2}^{\infty} \varepsilon'_X(i) \in [0, \delta). \quad \text{(A.3)}$$

Therefore, $i\varepsilon'_X(i) \overset{i\to\infty}{\longrightarrow} 0$, hence (by monotonicity of ε'_X), $x\varepsilon'_X(x) = o(1)$ holds. $\square$

Lemma A.3 *Consider the setting of Assumption 2.5 and let q_0 be close to 1. If Assumptions 2.8 or 2.9 are satisfied, then, for $q \in (q_0, 1)$, L has density*

$$f_L(q) \quad = \quad -\frac{1}{f_Y\left(F_Y^{\leftarrow}(q)\right)} \int_0^\infty f_X\left(x\left(s/w, F_Y^{\leftarrow}(q)\right)\right) D_2 x\left(s/w, F_Y^{\leftarrow}(q)\right) \mathrm{d}F_W(w)\,.$$

Proof: From Theorem 2.3 we have

$$L \overset{\mathrm{d}}{=} F_Y\left(y^*(s/W, X)\right).$$

Let q be close to 1; then $L(W, X)$ is larger than q, if $y^*(s/W, X)$ is close to ∞, since Y has support unbounded to the right. By Assumption 2.5(iii) F_Y is strictly increasing near its right endpoint, hence the (continuous) inverse $F_Y^{\leftarrow}$ exists. By independence of W and X we have

$$\begin{aligned} P(L > q) &= P\left(F_Y\left(y^*(s/W, X)\right) > q\right) \\ &= P\left(y^*(s/W, X) F_Y^{\leftarrow}(q)\right) = P\left(X < x\left(s/W, F_Y^{\leftarrow}(q)\right)\right) \\ &= \int_0^\infty F_X\left(x\left(s/w, F_Y^{\leftarrow}(q)\right)\right) \mathrm{d}F_W(w), \end{aligned} \tag{A.4}$$

where the inequality sign is reversed since $y^*(s/W, X)$ is decreasing in X. By Assumption 2.5(iii) X and Y_j have ultimately monotone densities f_X and f_Y, respectively.

To show existence and to derive an analytic representation of $f_L(q)$, we set

$$h(w, q) := F_X\left(x\left(s/w, F_Y^{\leftarrow}(q)\right)\right) \tag{A.5}$$

and show that h satisfies the conditions of Lemma A.1. Since $x^*(\cdot, \cdot)$ is continuous we have that $h(\cdot, t)$ is measurable on $(0, \infty)$ and, since $|h| \le 1$, it is also integrable with respect to F_W for some $q \in (0, 1)$. Therefore Lemma A.1(i) is applies. Next we have to show that $h(w, \cdot) : (q_0, 1) \to (0, 1)$ is $\mathcal{C}^1$ on $(q_0, 1)$ (as we consider $q \to 1$ we do not need continuity of h for all $q \in (0, 1)$). We choose q_0 large enough such that $F_Y^{\leftarrow}$ is $\mathcal{C}^1$ and denote $y := F_Y^{\leftarrow}(q)$. To show that $F_X(x^*(s/w, y))$ is $\mathcal{C}^1$ first note that $x^*(s/w, \cdot)$ is $\mathcal{C}^1$ and decreasing (by Assumption 2.5(iv)). Therefore, $\lim_{y\to\infty} x^*(s/w, y) = c \ge -\infty$ and, by Assumption 2.8(iii), $c < 0$. Assuming $c > -\infty$ implies $y^*(s/w, c-1) = \infty \notin \mathcal{Y}$. This contradicts Assumption 2.5(iii) as $c - 1 \in \mathcal{X}$. Therefore $\lim_{y\to\infty} x^*(s/w, y) = -\infty$, hence $h(w, \cdot)$ is $\mathcal{C}^1$ and Lemma A.1 applies.

To show that Lemma A.1(iii) holds observe that

$$D_2 h(w, q) = D_2 x^*(s/w, F_Y^{\leftarrow}(q)) f_X(x^*(s/w, F_Y^{\leftarrow}(q)))/f_Y(F_Y^{\leftarrow}(q)),$$

as $x^*(s/\cdot, y)$ is increasing. Define $y_0 := F_Y^{\leftarrow}(q_0)$, $y := F_Y^{\leftarrow}(q)$, and $x_{w,y} := x^*(s/w, y)$ and choose the setting of Assumption 2.8. Then

$$|D_2 h(w,q)| = \frac{f_X\left(x^*\left(s/w, F_Y^{\leftarrow}(q)\right)\right)}{f_Y\left(F_Y^{\leftarrow}(q)\right)} \left|D_2 x^*\left(s/w, F_Y^{\leftarrow}(q)\right)\right| = \frac{f_X\left(x_{w,y}\right)}{f_Y\left(y\right)} |D_2 x_{w,y}|$$

$$= \left|\frac{y D_2 x_{w,y}}{x_{w,y}}\right| \frac{|x_{w,y}| f_X(x_{w,y})}{F_X(x_{w,y})} \frac{\overline{F}_Y(y)}{y f_Y(y)} \frac{F_X(x_{w,y})}{\overline{F}_Y(y)}. \tag{A.6}$$

By Lemma A.2(i) we have $f_Y(y) \sim \nu_Y / y\overline{F}_Y(y)$ and therefore $\overline{F}_Y(y)/(y f_Y(y)) \leq 1/\nu_Y + \varepsilon_Y$ for all $y \geq y_0$ and an $\varepsilon_Y > 0$. Similarly, $|x_{w,y}| f_X(x_{w,y})/F_X(x_{w,y}) \to \mu_X$ as $y \to \infty$. As $x_{w,y}$ is increasing in w, $|x_{\infty,y}| f_X(x_{\infty,y})/F_X(x_{\infty,y}) \leq \mu_X + \varepsilon_X$ for all $y \geq y_0$ and an $\varepsilon_X > 0$ implies $|x_{w,y}| f_X(x_{w,y})/F_X(x_{w,y}) \leq \mu_X + \varepsilon_X$ for all $y \geq y_0$, an $\varepsilon_X > 0$ and for all $w \in (0, \infty)$. By Assumption 2.5(iv),

$$\lim_{y\to\infty} \frac{F_X(x_{w,y})}{\overline{F}_Y(y)} \leq \lim_{y\to\infty} \frac{F_X(x^*(0,y))}{\overline{F}_Y(y)} < \infty,$$

i.e. $\sup_{w\in(0,\infty),\, y\in(y_0,\infty)} F_X(x_{w,y})/\overline{F}_Y(y) = C < \infty$. Note that there exists a function $u(w)$, integrable with respect to F_W, such that $y D_2 x_{w,y}/x_{w,y} \leq u(w)$ for all y (Assumption 2.8(iii)). Hence $|D_2 h(w,q)|$ is dominated by an integrable function $u(w)$ and Lemma A.1(iii) is satisfied. Showing that there is an integrable upper bound $u(w)$ such that $|D_2 h(w,q)| \leq u(w)$ $\forall q \in (q_0, 1)$ in the setting of Assumption 2.9 is proved similarly using the asymptotic behavior of f_X and f_Y, given in Lemma A.2(ii).

Therefore, by Lemma A.1, we can interchange integration and differentiation and get the result. □

Proof of Theorem 2.10: By Lemma A.3, L has density f_L. If we observe $\lim_{q\to1}(1-q) f_L(q)/\overline{F}_L(q) = \kappa$ then we know from Corollary 3.3.13 in Embrechts et al. (1997) that $L \in DA(\Psi_\kappa)$. Substituting $y = F_Y^{\leftarrow}(q)$ (hence $1 - q = \overline{F}_Y(y)$), we obtain

$$\begin{aligned}
&\lim_{q\to1} \frac{(1-q) f_L(q)}{\overline{F}_L(q)} \\
&= \lim_{q\to1} \frac{(1-q)\int_0^\infty -f_X\left(x\left(s/w, F_Y^{\leftarrow}(q)\right)\right) D_2 x\left(s/w, F_Y^{\leftarrow}(q)\right) \mathrm{d}F_W(w)}{f_Y\left(F_Y^{\leftarrow}(q)\right)\int_0^\infty F_X\left(x\left(s/w, y\right)\right) \mathrm{d}F_W(w)} \\
&= \lim_{y\to\infty} \frac{\int_0^\infty -\overline{F}_Y(y)/f_Y(y) f_X\left(x\left(s/w, y\right)\right) D_2 x\left(s/w, y\right) \mathrm{d}F_W(w)}{\int_0^\infty F_X\left(x\left(s/w, y\right)\right) \mathrm{d}F_W(w)}.
\end{aligned} \tag{A.7}$$

Now we consider the setting of Assumption 2.8 and denote the integrands of (A.7) by

$$\begin{aligned}
h^*(w,y) &:= -\overline{F}_Y(y)/f_Y(y) f_X\left(x\left(s/w, y\right)\right) D_2 x\left(s/w, y\right) \quad \text{and} \\
h(w,y) &:= F_X\left(x\left(s/w, y\right)\right).
\end{aligned}$$

Choose $w \in (0,\infty)$ fixed, then Lemma A.2(i) yields

$$\lim_{y\to\infty} \frac{h^*(w,y)}{h(w,y)} = \frac{\mu_X}{\nu_Y} \lim_{y\to\infty} \frac{F_X\left(x^*(s/w,y)\right) y D_2 x^*(s/w,y)/x^*(s/w,y)}{F_X\left(x^*(s/w,y)\right)}.$$

By Assumption 2.8(iii) we have $\lim_{y\to\infty} y D_2 x^*(s/w,y)/x^*(s/w,y) = \zeta$, hence

$$\lim_{y\to\infty} \frac{h^*(w,y)}{h(w,y)} = \zeta \frac{\mu_X}{\nu_Y},$$

for almost every $w \in (0,\infty)$. Similarly to (A.6), $h^*(w,y)$ is dominated by an integrable function $u(w)$ for all $q \in (q_0, 1)$. Therefore we can apply the *Dominated Convergence Theorem* and with (A.7) we get

$$\lim_{q\to 1} \frac{(1-q) f_L(q)}{\overline{F}_L(q)} = \lim_{y\to\infty} \frac{\int_0^\infty h^*(w,y)\,\mathrm{d}F_W(w)}{\int_0^\infty h(w,y)\,\mathrm{d}F_W(w)} = \zeta \frac{\mu_X}{\nu_Y}.$$

Hence, $L \in DA(\Psi_{\zeta\mu_X/\nu_Y})$.

In the setting of Assumption 2.9 the same result is obtained similarly using the asymptotic behavior of f_X and f_Y given in Lemma A.2(ii). □

Proof of Theorem 2.11: We have $X \sim t_{\mu_X}$ and $Y_j \sim t_{\nu_Y}$. The t_ν density f_ν is given by

$$f_\nu(x) = C_\nu \left(1 + \frac{x^2}{\nu}\right)^{-(\nu+1)/2} = C_\nu |x|^{-\nu-1} \left(\frac{1}{\nu} + \frac{1}{x^2}\right)^{-(\nu+1)/2}, \quad C_\nu = \frac{\Gamma((\nu+1)/2)}{\sqrt{\nu\pi}\,\Gamma(\nu/2)}.$$

We immediately obtain that for $x > 0$ the t_ν distribution function F_ν is bounded by

$$\begin{aligned} \underline{f}_\nu(x) &:= \frac{C_\nu}{\nu} x^{-\nu} \left(\frac{1}{\nu} + \frac{1}{x^2}\right)^{-(\nu+1)/2} = \frac{C_\nu}{\nu} \left(x^{-2/(\nu+1)} + \frac{1}{\nu} x^{2-2/(\nu+1)}\right)^{-(\nu+1)/2} \\ &\leq \; F_\nu(-x) = \overline{F}_\nu(x) \quad \leq \quad \frac{C_\nu}{\nu} x^{-\nu} \left(\frac{1}{\nu}\right)^{-(\nu+1)/2} =: \overline{f}_\nu(x). \end{aligned} \tag{A.8}$$

Note that $\underline{f}_\nu(x) \sim \overline{f}_\nu(x)$ as $x \to \infty$. To get the asymptotic behavior of $\overline{F}_\nu$, we show that $\underline{f}_\nu^{\leftarrow}(q) \sim \overline{f}_\nu^{\leftarrow}(q)$ as $q \to 0$. First, we obtain

$$f_{\nu,U}^{\leftarrow}(q) = C_\nu^{1/\nu} \nu^{(\nu-1)/(2\nu)} q^{-1/\nu}. \tag{A.9}$$

Straightforward calculation yields

$$\underline{f}_\nu\left(\overline{f}_\nu^{\leftarrow}(q)\right) \;=\; q\left(1 + \left(\frac{\nu}{C_\nu}\right)^{2/\nu} q^{2/\nu}\right) \sim q, \quad \text{as } q \to 0,$$

hence

$$\underline{f}_\nu^{\leftarrow}(q) \sim \overline{f}_\nu^{\leftarrow}(q) \sim \overline{F}_\nu^{\leftarrow}(q) \quad \text{as } q \to 0. \tag{A.10}$$

Note that $x^*(s/w,y) = s/(aw) - yb/a$ and $F_Y^{\leftarrow}(q) = \overline{F}_Y^{\leftarrow}(1-q)$, therefore

$$F_X\left(x^*\left(\frac{s}{w}, F_Y^{\leftarrow}(q)\right)\right) \sim C_{\mu_X}\mu_X^{(\mu_X-1)/2}\left(-\frac{s}{aw} + \frac{b}{a}C_{\nu_Y}^{1/\nu_Y}\nu^{(\nu_Y-1)/(2\nu_Y)}(1-q)^{-1/\nu_Y}\right)^{-\mu_X}$$
$$= (1-q)^{\mu_X/\nu_Y}C_{\mu_X}\mu_X^{(\mu_X-1)/2}\left(-\frac{s}{aw}(1-q)^{1/\nu_Y} + \frac{b}{a}C_{\nu_Y}^{1/\nu_Y}\nu^{(\nu_Y-1)/(2\nu_Y)}\right)^{-\mu_X}.$$

Note that this asymptotic behavior holds uniformly for all $w \in (0,\infty)$, since $x^*(s/w,y)$ is decreasing in w and the asymptotic behavior also holds for $w=\infty$. Applying (A.4) yields

$$P(L>q) = \int_0^\infty F_X\left(x^*\left(\frac{s}{w}, F_Y^{\leftarrow}(q)\right)\right)\mathrm{d}F_W(w) \tag{A.11}$$
$$\sim (1-q)^{\mu_X/\nu_Y}C_{\mu_X}\mu_X^{(\mu_X-1)/2}\int_0^\infty\left(-\frac{s}{aw}(1-q)^{1/\nu_Y} + \frac{b}{a}C_{\nu_Y}^{1/\nu_Y}\nu^{(\nu_Y-1)/(2\nu_Y)}\right)^{-\mu_X}\mathrm{d}F_W(w).$$

We observe that

$$\lim_{y\to\infty}\frac{F_X(x^*(0,y))}{\overline{F}_Y(y)} = C\lim_{y\to\infty}(y)^{\nu_Y-\mu_X} = \begin{cases}\infty, & \mu_X < \nu_Y,\\ C, & \mu_X = \nu_Y,\\ 0, & \mu_X > \nu_Y,\end{cases}$$

for some constant $C < \infty$, i.e. the limit is finite, if X is not heavier-tailed than Y_j. Since $\mu_X \geq \nu_Y$, the upper limit is finite, hence Assumption 2.5(iv) is satisfied. We observe that the t-distribution falls into the setting of Assumption 2.8(i) and (ii). Considering Assumption 2.8(iii) we obtain $yD_2x^*(s/w,y)/x^*(s/w,y) = 1/(1-s/(bwy)) \nearrow 1$ for all w pointwise and 1 is an integrable upper bound. Hence we can apply Theorem 2.10 and obtain $\kappa = \mu_X/\nu_Y$. Comparing this to (A.11) gives the desired result. □

For the proof of Theorem 2.12 we need the following lemma.

Lemma A.4 *Let Φ and ϕ denote the distribution function and density of the standard normal distribution, respectively. Taking $0<C<1$, then for $x \geq \sqrt{1/(1-C)}$,*

$$C\phi(x)/x \leq \Phi(-x) \leq \phi(x)/x \quad \textit{and} \quad C\phi(x)/x \leq \overline{\Phi}(x) \leq \phi(x)/x.$$

Moreover, Mill's Ratio holds: $\Phi(-x) = \overline{\Phi}(x) \sim \phi(x)/x$ as $x\to\infty$.

Proof: Lemma 1.19.2 in Gännsler and Stute (1977) shows $(1/x - 1/x^3)\phi(x) \leq \Phi(-x) \leq \phi(x)/x$ for all $x>0$. Then, $1/x - 1/x^3 \geq C/x$ holds for $x>0$ if and only if $x \geq \sqrt{1/(1-C)}$. The limit relation follows from this as well. □

Proof of Theorem 2.12: We have $X, Y_j \overset{iid}{\sim} \mathcal{N}(0,1)$, $S_j = W(aX+bY_j)$. Let Φ and ϕ be the distribution function and density of the standard normal distribution, respectively. Applying (A.4) yields

$$P(L > q) = \int_0^\infty \Phi\left(s/(aw) - \Phi^{-1}(q)b/a\right) \,\mathrm{d}F_W(w). \tag{A.12}$$

Let $\varepsilon \in (0,1)$ and $x \le -\sqrt{1/\varepsilon}$, then, by Lemma A.4,

$$(1-\varepsilon)\phi(x)/|x| \le \Phi(x) \le \phi(x)/|x|. \tag{A.13}$$

The integrand of (A.12) increases in w, hence, for $\Phi^{-1}(q)b/a \ge \sqrt{1/\varepsilon}$,

$$\frac{\int_0^\infty \Phi\left(s/(aw) - \Phi^{-1}(q)b/a\right) \mathrm{d}F_W(w)}{\int_0^\infty \phi\left(s/(aw) - \Phi^{-1}(q)b/a\right) / \left|s/(aw) - \Phi^{-1}(q)b/a\right| \,\mathrm{d}F_W(w)} \in [1-\varepsilon, 1].$$

Therefore, as $q \to 1$,

$$\begin{aligned} P(L > q) &= \int_0^\infty \Phi\left(s/(aw) - \Phi^{-1}(q)b/a\right) \mathrm{d}F_W(w) \\ &\sim \int_0^\infty \phi\left(s/(aw) - \Phi^{-1}(q)b/a\right) / \left|s/(aw) - \Phi^{-1}(q)b/a\right| \,\mathrm{d}F_W(w) \\ &= \left(\frac{\phi\left(\Phi^{-1}(q)\right)}{\Phi^{-1}(q)}\right)^{b^2/a^2} \int_0^\infty \exp\left(-\frac{s^2}{2aw} + \frac{sb\Phi^{-1}(q)}{a^2 w}\right) \frac{\left(\Phi^{-1}(q)\right)^{b^2/a^2}}{\left|\frac{s}{aw} - \frac{b}{a}\Phi^{-1}(q)\right|} \,\mathrm{d}F_W(w). \end{aligned}$$

Note that $\Phi^{-1}(q) = \overline{\Phi}^{-1}(1-q)$. Then, again by Lemma A.4, $\phi\left(\Phi^{-1}(q)\right)/\Phi^{-1}(q) \sim 1-q$, as $q \to 1$.

By Mill's Ratio, Lemma A.2(ii) holds, hence $X, Y_j \stackrel{iid}{\sim} \mathcal{N}(0,1)$ fall into the setting of Assumption 2.9(i) and (ii) with $\mu_X = \nu_Y = 1/2$. We have $\zeta(w,y) = (b/a)(s/(aw) - yb/a)/y$ and $\lim_{y\to\infty} \zeta(w,y) = b^2/a^2$. Note that $|\zeta(w,y)| = c_1/(wy) + c_2 \le c_1/(wy_0) + c_2$ for some $c_1, c_2 < \infty$ and all $y \ge y_0$. As $E(1/W) < \infty$, $(c_1/(wy_0) + c_2)$ is an integrable upper bound, hence Assumption 2.9(iii) is satisfied. From Lemma A.4 we obtain

$$\begin{aligned} \lim_{y\to\infty} \frac{F_X(x^*(0,y))}{\overline{F}_Y(y)} &= \lim_{y\to\infty} \frac{\overline{\Phi}(yb/a)}{\overline{\Phi}(y)} \\ &= \frac{a}{b} \lim_{y\to\infty} \exp\left(\frac{y^2}{2}\left(1 - \frac{b^2}{a^2}\right)\right) = \begin{cases} \infty, & b < a, \\ a/b, & b = a, \\ 0, & b > a. \end{cases} \end{aligned}$$

Hence, Assumption 2.5(iv) is satisfied and, obviously, also Assumptions 2.5(i)-(iii) and (v) are. Therefore, by Theorem 2.3, $P(L > q) = (1-q)^{b^2/a^2} \mathcal{L}(1/(1-q))$, where $\mathcal{L}$ satisfies for $q \to 1$ the relation

$$\mathcal{L}\left(\frac{1}{1-q}\right) \sim \int_0^\infty \exp\left(-\frac{s^2}{2aw} + \frac{sb\Phi^{-1}(q)}{a^2 w}\right) \frac{\left(\Phi^{-1}(q)\right)^{b^2/a^2}}{\left|\frac{s}{aw} - \frac{b}{a}\Phi^{-1}(q)\right|} \,\mathrm{d}F_W(w). \tag{A.14}$$

Choose $c > 1$ fixed and $x \geq (\sqrt{c}/(\sqrt{c}-1))^{1/2}$, then

$$f_c(x) := c^{-1/2}\phi(x)/x \leq \overline{\Phi}(x) \leq \phi(x)/x =: f_1(x).$$

Therefore, since $\overline{\Phi}(x)$ is decreasing, $f_c^{-1}(q) \leq \overline{\Phi}^{-1}(q) \leq f_1^{-1}(q)$, $q_0 \leq q < 1$, and some q_0. Taking logarithm, we want a solution of $\ln f_c(x) = \ln(1-q)$, i.e.

$$\frac{1}{2}x^2 + \ln x + \frac{1}{2}\ln(2\pi c) = -\ln(1-q).$$

By asymptotic expansion, similarly to Example 2, Section 1.5 in Resnick (1987) we obtain

$$f_c^{-1}(1-q) = \sqrt{-2\ln(1-q)} - \frac{\ln(-\ln(1-q)) + \ln(4\pi c)}{2\sqrt{-2\ln(1-q)}} + o\left(1/\sqrt{-\ln(1-q)}\right). \tag{A.15}$$

Since $f_c^{-1}(1-q) \sim f_1^{-1}(1-q)$, as $q \to 1$, $f_1^{-1}(1-q) \sim \overline{\Phi}^{-1}(1-q)$ holds. Note that $f_1^{-1}(1-q) - f_c^{-1}(1-q) = O\left(1/\sqrt{-\ln(1-q)}\right) \overset{q\to 1}{\longrightarrow} 0$, hence also $\exp\left(-f_c^{-1}(1-q)\right) \sim \exp\left(-\overline{\Phi}^{-1}(1-q)\right)$ holds.

Substituting $\Phi^{-1}(q)$ in (A.14) by (A.15), we obtain the desired result. $\square$

Proof of Theorem 4.1: From (A.1) we have

$$\mathrm{Var}\sum_{j=1}^{m} L_j = 2\sum_{i\neq j} E(L_i L_j) + m p_{\mathrm{loss}}(1 - m p_{\mathrm{loss}}), \tag{A.16}$$

where, since $F_\nu(s) := P(W_\nu s^*(X, Y_i) \leq s) = p_{\mathrm{loss}}$ and $s < 0$,

$$E(L_i L_j) = P\left(W_\nu s^*(X, Y_i) \leq F_\nu^{\leftarrow}(p_{\mathrm{loss}}), W_\nu s^*(X, Y_j) \leq F_\nu^{\leftarrow}(p_{\mathrm{loss}})\right).$$

As mentioned in the introduction, $(S_i^\nu, S_j^\nu) := (W_\nu s^*(X, Y_i), W_\nu s^*(X, Y_j))$ has a bivariate t_ν-distribution with correlation ρ. We apply now a dependence measure, called *lower tail-dependence coefficient*, defined by

$$\lambda := \lim_{p\to 0} P\left(S_j^\nu \leq F_\nu^{\leftarrow}(p) \,|\, S_i^\nu \leq F_\nu^{\leftarrow}(p)\right).$$

Hult and Lindskog (2001) observed, that in case of a multivariate t_ν-distribution and with $\rho := \mathrm{Corr}(S_i^\nu, S_j^\nu)$

$$\lambda = \lambda(\nu) = \left(\int_{\arccos((1+\rho)/2)}^{\pi/2} \cos^\nu(v)\,\mathrm{d}v\right) \Big/ \left(\int_0^{\pi/2} \cos^\nu(v)\,\mathrm{d}v\right).$$

Lemma 2.2 in Kostadinov (2004) shows that $\lambda(\nu)$ is strictly decreasing in ν. Let $\nu_1 < \nu_2$, hence $\lambda(\nu_1) - \lambda(\nu_2) =: \varepsilon > 0$. Since $P\left(S_j^{\nu_i} \le F_{\nu_i}^{\leftarrow}(p) \,|\, S_i^{\nu_i} \le F_{\nu_i}^{\leftarrow}(p)\right) \to \lambda(\nu_i)$, $i = 1, 2$, there exists $p_\varepsilon > 0$ such that for all $p \le p_\varepsilon$

$$\left|\lambda(\nu_i) - P\left(S_j^{\nu_i} \le F_{\nu_i}^{\leftarrow}(p) \,|\, S_i^{\nu_i} \le F_{\nu_i}^{\leftarrow}(p)\right)\right| < \frac{\varepsilon}{2}, \quad i = 1, 2.$$

Hence, for all $p \le p_\varepsilon$,

$$P\left(S_j^{\nu_1} \le F_{\nu_1}^{\leftarrow}(p) \,|\, S_i^{\nu_1} \le F_{\nu_1}^{\leftarrow}(p)\right) \;>\; P\left(S_j^{\nu_2} \le F_{\nu_2}^{\leftarrow}(p) \,|\, S_i^{\nu_2} \le F_{\nu_2}^{\leftarrow}(p)\right).$$

Since

$$\begin{aligned} & P\left(S_j^{\nu} \le F_{\nu}^{\leftarrow}(p_{\text{loss}}) \,|\, S_i^{\nu} \le F_{\nu}^{\leftarrow}(p_{\text{loss}})\right) \\ = \; & \frac{1}{p_{\text{loss}}} P\left(S_j^{\nu} \le F_{\nu}^{\leftarrow}(p_{\text{loss}}),\, S_i^{\nu} \le F_{\nu}^{\leftarrow}(p_{\text{loss}})\right) = \frac{1}{p_{\text{loss}}} E(L_i L_j), \end{aligned}$$

also $E(L_i L_j)$ is decreasing in ν, if p_{loss} is sufficiently small. Applying (A.16), Var $\left(L^{(m)}\right)$ is decreasing in ν, hence Var$L$ is. □

Proof of Corollary 5.1: We can rewrite

$$\begin{aligned} L_j^{\wedge} &= \mathbf{1}_{\min\{\sqrt{\rho}X, \sqrt{1-\rho}Y_j\} \le s} = \mathbf{1}_{\{\sqrt{\rho}X \le s\} \cup \{\sqrt{1-\rho}Y_j \le s\}} \\ &= \mathbf{1}_{\{\sqrt{1-\rho}Y_j \le s\}} + \left(1 - \mathbf{1}_{\{\sqrt{1-\rho}Y_j \le s\}}\right) \mathbf{1}_{\{\sqrt{\rho}X \le s\}}. \end{aligned}$$

Hence

$$\begin{aligned} L_{\wedge}^{(m)} &= \frac{1}{m}\sum_{j=1}^{m} L_j^{\wedge} = \frac{1}{m}\sum_{j=1}^{m} \mathbf{1}_{\{\sqrt{1-\rho}Y_j \le s\}} + \mathbf{1}_{\{\sqrt{\rho}X \le s\}} \frac{1}{m}\sum_{j=1}^{m} \left(1 - \mathbf{1}_{\{\sqrt{1-\rho}Y_j \le s\}}\right) \\ &=: \frac{1}{m}\sum_{j=1}^{m} B_j + \mathbf{1}_{\{\sqrt{\rho}X \le s\}} \frac{1}{m}\sum_{j=1}^{m} (1 - B_j), \end{aligned}$$

where

$$\begin{aligned} B_1, B_2, \ldots &\overset{iid}{\sim} \text{Ber}\left(F_X\left(s/\sqrt{1-\rho}\right)\right) \text{ and} \\ (1 - B_1), (1 - B_2), \ldots &\overset{iid}{\sim} \text{Ber}\left(1 - F_X\left(s/\sqrt{1-\rho}\right)\right) \end{aligned}$$

are iid Bernoulli sequences. Therefore, for $m \to \infty$, $L^{(m)}$ converges almost surely to $F_X\left(s/\sqrt{1-\rho}\right) + \left(1 - F_X\left(s/\sqrt{1-\rho}\right)\right) \mathbf{1}_{\{\sqrt{\rho}X \le s\}}$. □

Proof of Theorem 5.2:

<u>*(i):*</u> $X, Y_1 \overset{iid}{\sim} t_\nu$ are regularly varying on $\mathbb{R}^-$, i.e. $F_{t_\nu}(-\cdot) \in \mathcal{R}_{-\nu}$. Hence

$$P\left(\sqrt{\rho}X + \sqrt{1-\rho}Y_1 \le s\right) \sim P\left(\min\left\{\sqrt{\rho}X, \sqrt{1-\rho}Y_1\right\} \le s\right), \quad s \to -\infty,$$

see for instance Example 3.2 in combination with Definition 1.1 in Goldie & Klüppelberg (1998). For convenience we define $A := \sqrt{\rho}X$ and $B := \sqrt{1-\rho}Y_1$. Then, A and B are independent in $\mathcal{R}_{-\nu}$ satisfying $P(A > s)/P(B > s) \stackrel{s\to\infty}{\longrightarrow} c \in (0,\infty)$ and $P(A \leq s)/P(B \leq s) \stackrel{s\to-\infty}{\longrightarrow} c \in (0,\infty)$. Let F_A and F_B denote the dfs of A and B, respectively. Writing $x \wedge y := \max\{x, y\}$ and $x \vee y := \max\{x, y\}$ we have

$$
\begin{aligned}
&P\left(L_1^{\wedge} = 1 \,|L_1 = 1\right) = P\left(A \wedge B < -s \,|A + B < -s\right) \\
&= \quad P\left(A \vee B > s \,|A + B > s\right) = \frac{P\left(A \vee B > s,\ A + B > s\right)}{P\left(A + B > s\right)} \\
&= \quad 1 - \frac{P\left(A \vee B \leq s,\ A + B > s\right)}{P\left(A + B > s\right)}. \qquad \text{(A.17)}
\end{aligned}
$$

For illustration purposes see Figure 5. The set $\{(a, b) \in \mathbb{R}^2 :\ a \vee b > s,\ a + b > s\}$ is the hatched area in Figure 5 above the lines $\{a + b = s\}$ and $\{a \vee b = s\}$; the set $\triangle := \{(a, b) :\ a \vee b \leq s,\ a + b > s\}$ is the triangle with edges $(s, 0)$, (s, s) and $(0, s)$. Let $\|(a, b)\|_1 := |a| + |b|$ denote the 1-norm, then for any $\varepsilon > 0$ it holds that

$$
\begin{aligned}
\triangle &= \{(a, b) :\ a \vee b \leq s,\ a + b > s\} \\
&\subset \quad \left\{(a, b) :\ \varepsilon < \frac{b}{a} < \frac{1}{\varepsilon},\ \|(a, b)\|_1 > s\right\} \\
&\quad\ \cup \{(a, b) :\ (1 - \varepsilon)s < a < s,\ 0 < b < \varepsilon s\} \\
&\quad\ \cup \{(a, b) :\ 0 < a < \varepsilon s,\ (1 - \varepsilon)s < b < s\} \\
&=: \quad \boldsymbol{S}_1 \cup \boldsymbol{S}_2 \cup \boldsymbol{S}_3, \qquad \text{(A.18)}
\end{aligned}
$$

where $\boldsymbol{S}_1$ can be identified in Figure 5 as the set between the two lines through the points $(0, 0)$, $(s, \varepsilon s)$ and $(0, 0)$, $(\varepsilon s, s)$ and above the line $\{a + b = s\}$; the sets $\boldsymbol{S}_2$ and $\boldsymbol{S}_3$ represent in the figure the two small rectangles.

By Resnick (2004), section 4.1 and 4.3, the vector (A, B) is bivariate regularly varying. More precisely, let $\|\cdot\|$ be any norm on $\mathbb{R}^2$, then

$$
\frac{P\left(\|(A, B)\| \geq x,\ (A, B)/\|(A, B)\| \in \cdot\right)}{P\left(\|(A, B)\| \geq x\right)} \stackrel{x\to\infty}{\longrightarrow} \Theta(\cdot).
$$

Θ is a measure on the unit simplex $\mathfrak{S} = \{\boldsymbol{s} \in \mathbb{R}^2 :\ \|\boldsymbol{s}\| = 1\}$ called *spectral measure*. Since A and B are independent, Θ is concentrated on $(1, 0)$, $(-1, 0)$, $(0, 1)$ and $(0, -1)$. Note that, using the 1-norm, symmetry of A and B yields

$$
\frac{P\left(\|(A, B)\|_1 > s\right)}{P\left(A + B > s\right)} < \frac{P\left(\|(A, B)\|_1 > s\right)}{P\left(A + B > s,\ A, B > 0\right)} = 4.
$$

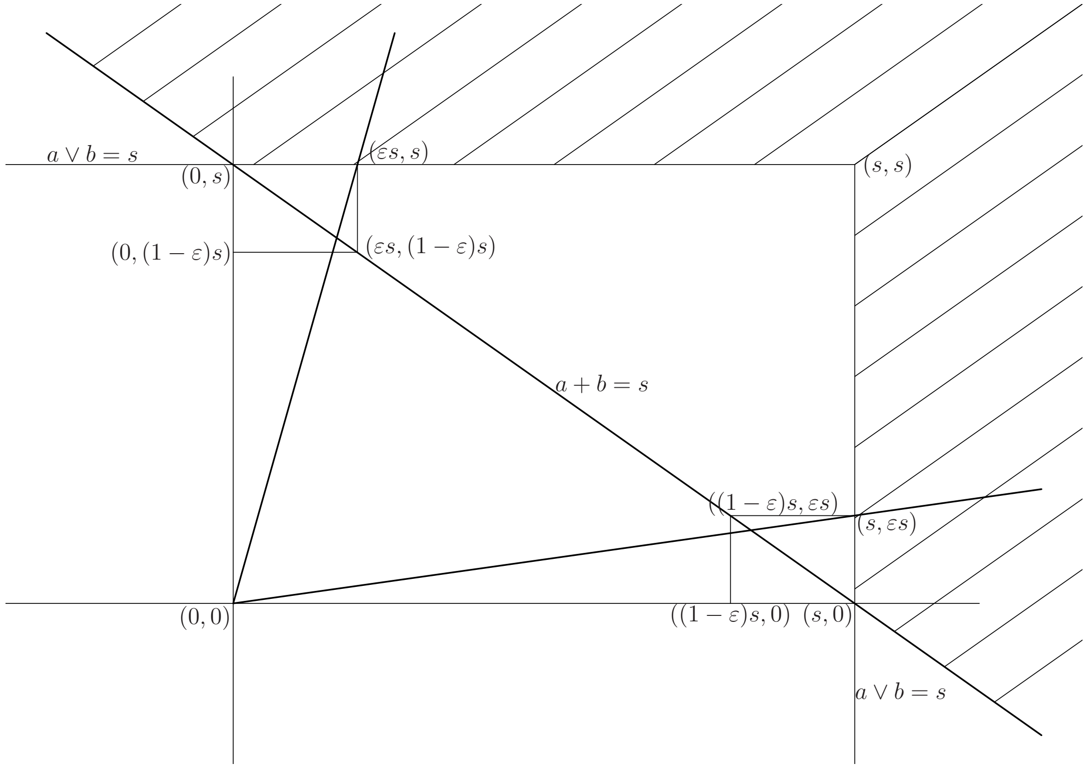

Figure 5: Illustration of (A.17) and (A.18).

Hence, for any $\varepsilon > 0$, we have

$$
\begin{aligned}
P\left((A, B) \in \boldsymbol{S}_1 \,|A + B > s\right) &= \frac{P\left(A + B > s,\ \varepsilon < B/A < 1/\varepsilon\right)}{P\left(A + B > s\right)} \qquad \text{(A.19)} \\
&= \frac{P\left(A, B > 0,\ \varepsilon < B/A < 1/\varepsilon,\ \|(A, B)\|_1 > s\right)}{P\left(\|(A, B)\|_1 > s\right)} \frac{P\left(\|(A, B)\|_1 > s\right)}{P\left(A + B > s\right)} \\
&< 4 \frac{P\left(A, B > 0,\ \varepsilon < B/A < 1/\varepsilon,\ \|(A, B)\|_1 > s\right)}{P\left(\|(A, B)\|_1 > s\right)} \overset{s \to \infty}{\longrightarrow} 4\Theta\left(\mathcal{S}_1^n\right) = 0,
\end{aligned}
$$

since $\mathcal{S}_1^n := \{\boldsymbol{s}/\|\boldsymbol{s}\|_1 : \boldsymbol{s} \in \mathcal{S}_1\}$ has no points on the axes. Considering rectangle $\boldsymbol{S}_2$

we obtain as $s \to \infty$ using again $P(A+B>s) \sim P(A \vee B > s)$

$$
\begin{aligned}
P\left((A,B) \in \boldsymbol{S}_2 \,|A+B>s\right) &\leq \frac{P\left((1-\varepsilon)s < A < s\right) P(0<B<\varepsilon s)}{P\left(A+B>s\right)} \\
\sim \quad & \frac{P\left((1-\varepsilon)s < A < s\right) P(0<B<\varepsilon s)}{P\left(A \vee B > s\right)} \\
= \quad & \frac{\left(F_A(s) - F_A((1-\varepsilon)s)\right)\left(F_B(\varepsilon s) - F_B(0)\right)}{\overline{F}_A(s) + \overline{F}_B(s) - \overline{F}_A(s)\overline{F}_B(s)} \\
= \quad & \frac{s^{-\nu}\left((1-\varepsilon)^{-\nu}\mathcal{L}_A((1-\varepsilon)s) - \mathcal{L}_A(s)\right)\left(\frac{1}{2} - s^{-\nu}\varepsilon^{-\nu}\mathcal{L}_B(\varepsilon s)\right)}{s^{-\nu}\left(\mathcal{L}_A(s) + \mathcal{L}_B(s)\right) - s^{-2\nu}\mathcal{L}_A(s)\mathcal{L}_B(s)} \\
\longrightarrow \quad & \frac{1}{2(1+c)}\left((1-\varepsilon)^{-\nu} - 1\right) < \frac{\nu\varepsilon}{(1+c)}\,. \qquad \text{(A.20)}
\end{aligned}
$$

The last convergence holds since $F_A(s) = 1-(s)^{-\nu}\mathcal{L}_A(s)$, $F_B(s) = 1-(s)^{-\nu}\mathcal{L}_B(s)$, $\mathcal{L}_B(s)/\mathcal{L}_A(s) \stackrel{s\to\infty}{\longrightarrow} c$ and $\mathcal{L}_A, \mathcal{L}_B \in \mathcal{R}_0$; $(1-\varepsilon)^{-\nu} - 1 < 2\nu\varepsilon$ holds for ε small enough since $(1-\varepsilon)^{-\nu} - 1 \sim \nu\varepsilon$ as $\varepsilon \to 0$. Combining (A.19) and (A.20), (A.18) yields

$$
\lim_{s\to\infty} \frac{P\left(A \vee B \leq s,\ A+B>s\right)}{P\left(A+B>s\right)} < K\varepsilon\,, \quad \forall \varepsilon > 0\,,
$$

for some constant $0 < K < \infty$. Hence the latter limit equals 0 and applying this to (A.17) yields

$$
P\left(L_1 = 1\,|L_1^{\wedge} = 1\right) \stackrel{s\to-\infty}{\longrightarrow} 1.
$$

Therefore, for all m and any $q \in \{0, 1/m, 2/m, \ldots, 1\}$, we conclude

$$
P\left(L^{(m)} = q\,|L_{\wedge}^{(m)} = q\right) \stackrel{s\to-\infty}{\longrightarrow} 1.
$$

<u>*(ii):*</u> Now we consider the limiting case $m \to \infty$. Recall that $L \stackrel{\mathrm{d}}{=} F_{t_\nu}(y^*(s,X)) = F_{t_\nu}\left((s-\sqrt{\rho}X)/\sqrt{1-\rho}\right)$ and from Corollary 5.1 we know $L^{\wedge} \in \left\{F_{t_\nu}\left(s/\sqrt{1-\rho}\right), 1\right\}$. Further, $L^{\wedge} = F_{t_\nu}\left(s/\sqrt{1-\rho}\right)$ if $s-\sqrt{\rho}X < 0$. Define $\mathcal{B}_\varepsilon(x) := (x-\varepsilon, x+\varepsilon)$, then

$$
\begin{aligned}
& P\left(L \in \mathcal{B}_\varepsilon\left(F_{t_\nu}\left(s/\sqrt{1-\rho}\right)\right) \,\middle|\, L^{\wedge} = F_{t_\nu}\left(s/\sqrt{1-\rho}\right)\right) \qquad \text{(A.21)} \\
\geq \quad & P\left(s-\sqrt{\rho}X \leq \sqrt{1-\rho}F_{t_\nu}^{\leftarrow}\left(F_{t_\nu}\left(s/\sqrt{1-\rho}\right)(1+\varepsilon)\right) \,\middle|\, s-\sqrt{\rho}X < 0\right) \\
& -P\left(s-\sqrt{\rho}X < \sqrt{1-\rho}F_{t_\nu}^{\leftarrow}\left(F_{t_\nu}\left(s/\sqrt{1-\rho}\right)(1-\varepsilon)\right) \,\middle|\, s-\sqrt{\rho}X < 0\right).
\end{aligned}
$$

By (A.9) and (A.10) we have $F_{t_\nu}^{\leftarrow}(qc) \sim F_{t_\nu}^{\leftarrow}(q)c^{-1/\nu}$ as $q \to 0$, hence

$$
\sqrt{1-\rho}F_{t_\nu}^{\leftarrow}\left(F_{t_\nu}\left(s/\sqrt{1-\rho}\right)(1\pm\varepsilon)\right) \sim (1\pm\varepsilon)^{-1/\nu}s \quad \text{as } s \to -\infty.
$$

From $s - \sqrt{\rho}X \leq (1 \pm \varepsilon)^{-1/\nu} s$ follows $\sqrt{\rho}X \geq \left(1 - (1 \pm \varepsilon)^{-1/\nu}\right) s$. Since $1 - (1 - \varepsilon)^{-1/\nu} < 0$, we conclude (as $s \to -\infty$)

$$\begin{aligned} & P\left(s - \sqrt{\rho}X < \sqrt{1-\rho} F_{t_\nu}^{\leftarrow}\left(F_{t_\nu}\left(s/\sqrt{1-\rho}\right)(1-\varepsilon)\right) \Big| \, s - \sqrt{\rho}X < 0\right) \\ \sim \quad & P\left(\sqrt{\rho}X > \left(1 - (1-\varepsilon)^{-1/\nu}\right) s \big| \sqrt{\rho}X > s\right) \to 0. \end{aligned}$$

Since $1 - (1 + \varepsilon)^{-1/\nu} > 0$, we conclude (as $s \to -\infty$)

$$\begin{aligned} & P\left(s - \sqrt{\rho}X < \sqrt{1-\rho} F_{t_\nu}^{\leftarrow}\left(F_{t_\nu}\left(s/\sqrt{1-\rho}\right)(1+\varepsilon)\right) \Big| \, s - \sqrt{\rho}X < 0\right) \\ \sim \quad & P\left(\sqrt{\rho}X \geq \left(1 - (1+\varepsilon)^{-1/\nu}\right) s \big| \sqrt{\rho}X > s\right) \to 1. \end{aligned}$$

Therefore, (A.21) converges to 1 as $s \to -\infty$ for all $\varepsilon > 0$.

We have $L^\wedge = 1$ if $\sqrt{\rho}X \leq s$, hence, similarly to (A.21),

$$\begin{aligned} P(L > 1 - \varepsilon | \, L^\wedge = 1) &= P\left(\sqrt{\rho}X \leq s - \sqrt{1-\rho} F_{t_\nu}^{\leftarrow}(1-\varepsilon) \Big| \sqrt{\rho}X \leq s\right) \\ &= \frac{F_{t_\nu}\left(\left(s - \sqrt{1-\rho} F_{t_\nu}^{\leftarrow}(1-\varepsilon)\right)/\sqrt{\rho}\right)}{F_{t_\nu}\left(s/\sqrt{\rho}\right)}. \end{aligned}$$

From (A.8) we know $F_{t_\nu}(-x) \sim Cx^{-\nu}$ as $x \to \infty$ for some constant C. Hence, $P(L > 1 - \varepsilon | \, L^\wedge = 1) \to 1$ as $s \to -\infty$ for all $\varepsilon > 0$. $\square$

References

[1] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1987) *Regular Variation.* Cambridge University Press, Cambridge.

[2] Bluhm, C., Overbeck, L. and Wagner, C. (2003) *An Introduction to Credit Risk Modeling.* Chapman&Hall, London.

[3] Embrechts, P., Klüppelberg, C. und Mikosch, T. (1997) *Modelling Extremal Events for Insurance and Finance.* Springer, Berlin.

[4] Frey, R. and McNeil, A. (2001) Modelling dependent defaults. Preprint, ETHZ. Available at `http://www.defaultrisk.com/pp_corr_17.htm`

[5] Frey, R. and McNeil, A. (2002) VaR and expected shortfall in portfolios of dependent credit risks: conceptual and practical insights. *Journal of Banking and Finance* **26**, 1317-1334.

[6] Frey, R. and McNeil, A. (2003) Dependent defaults in models of portfolio credit risk. *Journal of Risk* (6)1, 59-92.

[7] Frey, R., McNeil, A. and Nyfeler, M. (2001) Copulas and credit models. *RISK*, October, pp. 111-113.

[8] Gänssler, P. and Stute, W. (1977) *Wahrscheinlichkeitstheorie.* Springer, Berlin

[9] Goldie, C.M. and Klüppelberg, C. (1998). Subexponential distributions. In: R. Adler, R. Feldmann, M.S. Taqqu (Eds.) *A Practical Guide to Heavy Tails: Statistical Techniques for Analysing Heavy Tailed Distributions, pp. 435-459.* Birkhäuser, Boston.

[10] Gupton, G., Finger, C. and Bhatia, M. (1997) CreditMetrics – technical document. Morgan Guaranty Trust Company. Available at `http://www.defaultrisk.com/pp_model_20.htm`

[11] Hult, H. and Lindskog, F. (2001) Multivariate extremes, aggregation and dependence in elliptical distributions. Research paper, RiskLab. Available at `www.risklab.ch/ftp/papers/ExtremesEllipticalDistr.pdf`

[12] Lucas, A., Klaasen, P., Spreij, P. and Straetmans, S. (2003). Tail behaviour of credit loss distribution for general latent factor models. *Applied Mathematical Finance 10(4), 337-357.*

[13] Kostadinov, K. (2004) Non-parametric estimation of elliptical copula with application to credit risk. In preparation.

[14] Resnick, S.I. (1987) *Extreme Values, Regular Variations and Point Processes.* Springer, New York.

[15] Resnick, S.I. (2004) On the foundations of multivariate heavy tail analysis. *J. Appl. Probab.* 41A, 191-212.

[16] Smith, K. (1983) *Primer of Modern Analysis.* Springer, New-York.