VaR and Expected Shortfall in Portfolios of Dependent Credit Risks: Conceptual and Practical Insights

Rüdiger Frey
Swiss Banking Institute
University of Zurich
Plattenstrasse 14
CH-8032 Zurich
Tel: +41 1 634 29 57
Fax: +41 1 634 49 03
freyr@isb.unizh.ch

Alexander J. McNeil
Department of Mathematics
Federal Institute of Technology
ETH Zentrum
CH-8092 Zurich
Tel: +41 1 632 61 62
Fax: +41 1 632 15 23
mcneil@math.ethz.ch

January 23, 2002

Abstract

In the first part of this paper we address the non-coherence of value-at-risk (VaR) as a risk measure in the context of portfolio credit risk, and highlight some problems which follow from this theoretical deficiency. In particular, a realistic demonstration of the non-subadditivity of VaR is given and the possibly nonsensical consequences of VaR-based portfolio optimisation are shown. The second part of the paper discusses VaR and expected shortfall estimation for large balanced credit portfolios. All standard industry models (Creditmetrics, KMV, CreditRisk$^+$) are presented as Bernoulli mixture models to facilitate their direct comparison. For homogeneuous groups it is shown that measures of tail risk for the loss distribution may be approximated in large portfolios by analysing the tail of the mixture distribution in the Bernoulli representation. An example is given showing that, for portfolios of lower quality, choice of model has some impact on measures of extreme risk.

J.E.L. Subject Classification: G31, G11, C15

Keywords: risk measures, value-at-risk, coherence, expected shortfall, portfolio credit risk models, Bernoulli mixture models

1 Introduction

This paper is concerned with risk measurement in large credit portfolios. The nature of the dependence between obligors in such portfolios critically determines the tail of the overall credit loss distribution and makes the practical estimation of measures of tail risk a challenging task. Moreover, due to the skewness of the loss distribution for a typical dependent portfolio, it is essential that any risk measure should have reasonable theoretical properties, particularly with regard to agregation. In this paper we take up these two distinct issues.

In the first part we revive the debate on the theoretical properties of risk measures, but we do this in the particular context of credit portfolios. Since the pathbreaking work of Artzner, Delbaen, Eber, and Heath (1999) it is now well known that value-at-risk (VaR) is not a *coherent* risk measure, since it lacks the property of subadditivity. Two possibly pernicious aspects of this theoretical deficiency are: a decentralised risk management system may fail because VaRs calculated for individual portfolios may not be summed to produce an upper bound for the VaR of the combined portfolios; a risk manager who optimises

his portfolio to minimise VaR may (intentionally or unintentionally) produce an allocation which is highly risky by any more rational analysis. It is often argued that the practical implications of these theoretical criticisms are likely to be limited and, if our portfolios consist of common financial instruments which are subject only to market risks, this may also be true. However, when we turn to portfolios which are subject to credit risk, and which therefore give rise to highly skewed loss distributions, realistic examples show that the dangers indicated above should not be dismissed out of hand. We take up this theme in Section 2 of the paper.

In the second part of the paper (Section 3) we leave the theoretical debate to one side and consider what drives VaR, as well as the coherent alternative risk measure known as expected shortfall, when one models portfolio credit risk using one of the standard industry solutions such as the model proposed by the KMV corporation (KMV-Corporation 1997), the model proposed by the RiskMetrics group (RiskMetrics-Group 1997), or CreditRisk$^+$, developed by Credit Suisse Financial Products (Credit-Suisse-Financial-Products 1997). To address this question we consider stylized versions of these industry models for large, homogeneous groups of dependent credit risks. The similarities between the industry models have been noted in a number of recent papers including Koyluoglu and Hickman (1998), Gordy (2000) and Crouhy, Galai, and Mark (2000). It has been observed that the mathematical structures of these models can be mapped into each other, and this theme is taken up in detail in Frey and McNeil (2001). In the present paper we summarise how all standard models may be recast as *Bernoulli mixture models* and in this way we obtain a common mathematical representation that greatly facilitates their comparison. We show that the tail of the portfolio loss distribution is driven essentially by the mixing distribution in the Bernoulli mixture representation, and that VaR and expected shortfall may be estimated in large portfolios by calculating quantiles and conditional tail expectations for this mixture distribution and scaling them appropriately. We provide some numerical examples of the use of this technique.

2 Measures of Risk for Credit Portfolios

In this section we present the essential ideas of Artzner, Delbaen, Eber, and Heath (1999) in a slightly different notation tailored to an application to portfolio credit losses.

2.1 Measures of Risk

Fix a probability space $(\Omega, \mathcal{F}, P)$ and denote by $L^0(\Omega, \mathcal{F}, P)$ the set of almost surely finite random variables on that space. Financial risks are represented by a *convex cone* $\mathcal{M} \subseteq L^0(\Omega, \mathcal{F}, P)$ of random variables. Any random variable L in this set will be interpreted as a possible loss of some credit portfolio over a given time horizon. Recall that $\mathcal{M}$ is a convex cone if $L_1 \in \mathcal{M}$ and $L_2 \in \mathcal{M}$ implies that $L_1 + L_2 \in \mathcal{M}$ and $\lambda L_1 \in \mathcal{M}$ for every $\lambda > 0$.

Definition 2.1. Given some convex cone $\mathcal{M}$ of random variables, a measure of risk with domain $\mathcal{M}$ is a mapping $\rho : \mathcal{M} \to \mathbb{R}$.

In economic terms we interprete $\rho(L)$ as the amount of capital that should be added as buffer to a portfolio with loss given by L, so that the portfolio becomes acceptable to an external or internal risk controller. Our presentation differs here slightly from Artzner, Delbaen, Eber, and Heath (1999) who interpret a rv $L \in \mathcal{M}$ as the future value (instead of loss) of a currently held portfolio.

Denote the loss distribution of loss L by $F_L(l) = P(L \leq l)$. In this paper we are concerned solely with two risk measures which are based on the loss distribution F_L, namely VaR and expected shortfall. We recall the definition of these risk measures and, in the following subsection, the definition of a coherent risk measure.

Definition 2.2 (Value-at-Risk). Given some confidence level $\alpha \in (0,1)$, the value-at-risk (VaR) of our portfolio at the confidence level α is given by the smallest number l such that the probability that the loss L exceeds l is no larger than $(1 - \alpha)$. Formally,

$$\text{VaR}_\alpha = \inf\{l \in \mathbb{R}, \ P(L > l) \leq 1 - \alpha\}. \tag{1}$$

This definition of VaR coincides with the definition of an α-quantile of the distribution of L in terms of a generalised inverse of the distribution function F_L. We observe this by noting

$$\text{VaR}_\alpha = \inf\{l \in \mathbb{R}, \ 1 - F_L(l) \leq 1 - \alpha\} = \inf\{l \in \mathbb{R}, \ F_L(l) \geq \alpha\}.$$

For a random variable X with df F_X we will denote the α-quantile of the distribution by $q_\alpha(F_X)$, or sometimes $q_\alpha(X)$, and write $\text{VaR}_\alpha(X)$ when we wish to stress that the quantile should be interpreted as a VaR number.

A simple definition of expected shortfall which suffices for continuous loss distributions is as follows.

Definition 2.3 (Expected shortfall, continuous loss distribution). Consider a loss L with continuous df F_L satisfying $\int_\mathbb{R} |l| dF_L(l) < \infty$. Then the expected shortfall at confidence level $\alpha \in (0,1)$, is defined to be

$$\text{ES}_\alpha = E(L \mid L \geq \text{VaR}_\alpha) = \frac{E(L; L \geq \text{VaR}_\alpha(L))}{P(L \geq \text{VaR}_\alpha(L))}. \tag{2}$$

A more general definition has been proposed by Acerbi and Tasche (2001) as follows.

Definition 2.4 (Generalized expected shortfall). Given an integrable rv L and $\alpha \in (0,1)$. The generalized expected shortfall at confidence level α is given by

$$\text{GES}_\alpha = \frac{1}{1 - \alpha}\Big(E(L; L \geq \text{VaR}_\alpha(L)) + q_\alpha(1 - \alpha - P(L \geq \text{VaR}_\alpha(L)))\Big). \tag{3}$$

For a rv with continuous distribution the second term disappears and (3) reduces to (2).

2.2 Axioms of Coherence

We enumerate the four axioms of coherence in the form that we require them. We differ from Artzner, Delbaen, Eber, and Heath (1999) in that we interpret a random variable L as a loss rather than a future value; this leads to different signs in certain axioms. Moreover, in order to simplify the presentation we assume that the risk capital $\rho(L)$ earns no interest. For further discussion of the axioms for coherent risk measures see Fritelli and Gianin (2002) in the same volume.

Axiom 1 (Translation invariance). For all $L \in \mathcal{M}$ and every $l \in \mathbb{R}$ a translation-invariant risk measure satisfies $\rho(L + l) = \rho(L) + l$.

Axiom 2 (Subadditivity). For all $L_1, L_2 \in \mathcal{M}$ a subadditive risk measure satisfies $\rho(L_1 + L_2) \leq \rho(L_1) + \rho(L_2)$.

Following Artzner et. al. the rationale behind Axiom 2 can be summarized by the statement "A merger does not create extra risk." Subadditivity reflects the idea that risk can be reduced by diversification, a time-honoured principle in finance and economics. We will see in Subsection 2.4, that the use of non-subadditive risk measures in a Markowitz-type portfolio-optimization problem may lead to optimal portfolios which are very concentrated and would be deemed risky by normal economic standards.

Axiom 3 (Positive homogeneity). For all $L \in \mathcal{M}$ and every $\lambda > 0$ a positive homogeneous risk measure satisfies $\rho(\lambda L) = \lambda \rho(L)$.

Note that subadditivity and positive homogeneity imply that the functional ρ is *convex* on $\mathcal{M}$.

Axiom 4 (Monotonicity). For L_1 and $L_2 \in \mathcal{M}$ such that $L_1 \leq L_2$ a.s. a monotonic risk measure satisfies $\rho(L_1) \leq \rho(L_2)$.

Definition 2.5 (Coherent risk measure). Given a risk measure ρ whose domain includes the convex cone $\mathcal{M}$. ρ is called coherent (on $\mathcal{M}$) if it satisfies Axiom 1, 2, 3 and 4.

It is immediately seen from the representation of VaR as quantile of the loss distribution obtained in Section 2.1 that VaR is translation invariant, positively homogeneous and monotonic. However, it is well-known that VaR is not a subadditive risk measure and therefore not coherent. In the following subsection we give a credit-related demonstration of this fact.

Generalised expected shortfall is a coherent risk measure; see Acerbi and Tasche (2001) for a proof. Expected shortfall is coherent only if we restrict ourselves to a convex cone of random variables with continuous dfs. However, if we allow distributions with atoms it is possible to construct examples that show that expected shortfall is also not always subadditive.

2.3 The Inconsistency of VaR in Credit Portfolio Management: An Example

Our example expands on an idea sketched out in Artzner, Delbaen, Eber, and Heath (1999). Consider a portfolio of $m = 50$ defaultable corporate bonds. We assume that defaults of different bonds are independent; the default probability is identical for all bonds and equal to 2%. The face value of the bonds is 100; this amount is paid back at $T = t + \Delta t$ if there is no default; otherwise there is no repayment whatsoever. The current (time t) price of the bond equals 95. The loss of bond i is hence given by the rv

$$L_i := -(100(1 - Y_i) - 95) = 100 Y_i - 5,$$

where the default indicator Y_i is equal to one if default occurs and equal to zero otherwise. $(L_i)_{1 \leq i \leq 50}$ form a sequence of iid rv's with $P(L_i = -5) = 0.98$ and $P(L_i = 95) = 0.02$.

We compare two portfolios, both with current value equal to 9500. Portfolio A is fully concentrated and consists of 100 units of bond 1. Portfolio B is completely diversified; it consists of two units of each of the bonds. Now let us consider VaR at a confidence level of 95% for both portfolios.

For portfolio A the portfolio loss is given by $L = 100 L_1$ and hence

$$\text{VaR}_{0.95}(L) = 100 \text{VaR}_{0.95}(L_1) = -500,$$

(as $P(L_1 \leq -5) = 0.98 > 0.95$). This means that even after a withdrawal of 500 the portfolio is still acceptable to a risk controller working with VaR at the 95% level.

For portfolio B we have

$$L = \sum_{i=1}^{50} 2 L_i = 200 \sum_{i=1}^{50} Y_i - 500,$$

and hence $\text{VaR}_\alpha(L) = 200 q_\alpha(\sum_{i=1}^{50} Y_i) - 500$. The sum $M := \sum_{i=1}^{50} Y_i$ has a binomial distribution with success probability $p = 0.02$. We get by inspection $q_{0.95}(M) = 3$, so that $\text{VaR}_{0.95}(L) = 100$. In this case a bank would need a risk capital of 100 to satisfy a regulator

working with VaR at the 95% level. Of course, economic intuition suggests that portfolio B is less risky than portfolio A, showing that measuring risk with VaR can lead to non-sensical results. This intuition is supported by Lemma 2.7 below, which shows that portfolio B is indeed less risky than portfolio A if we measure risk with any coherent risk measure that depends only on the distribution of the loss.

These nonsensical results are directly linked to the lack of subadditivity of VaR. In fact, we get for any coherent risk measure ρ, which depends only on the distribution of L

$$\rho\left(\sum_{i=1}^{50} 2L_i\right) \leq \sum_{i=1}^{50} 2\rho(L_i) = 100\rho(L_1) = \rho(100L_1),$$

hence the fact that $\mathrm{VaR}_{0.95}$ is lower for portfolio A than for portfolio B shows that VaR is in general not subadditive.

2.4   Dangers of a Mean-VaR Portfolio Optimization

Practitioners used to work with value-at-risk sometimes tend to regard the lack of subadditivity of VaR as relatively minor drawback, which is not very relevant in practice. We disagree; while it is admittedly not very likely that we will observe the worst features of VaR for some randomly chosen portfolio, the picture changes, if investors optimize the (expected) return on their portfolios under some constraint on VaR, as the portfolios resulting from such an optimization procedure do exploit the conceptual weaknesses of VaR.

For a concrete example we place ourselves in the context of Section 2.3. Consider a portfolio manager, who has an amount of capital V which he can invest in a riskless asset and the $m = 50$ defaultable bonds. For simplicity we assume that he is not able to borrow additional money or to take short positions in the defaultable bonds. We assume that he determines his portfolio using a mean – value-at-risk optimality criterion. Denote by

$$\Theta := \{\boldsymbol{\theta} = (\theta_0, \theta_1, \ldots, \theta_m)' \,:\, \theta_i \geq 0 \text{ for } i = 0, \ldots, m\}$$

the set of all portfolios, and by Θ_V those portfolios in Θ, whose value at time t equals V. Here θ_0 denotes the amount invested in the riskless security and, for $i \geq 1$, θ_i represents the amount invested in the defaultable bond i. The loss of some portfolio $\boldsymbol{\theta} \in \Theta$ will be denoted by $L(\boldsymbol{\theta})$; the expected profit of a portfolio is clearly given by $E(-L(\boldsymbol{\theta}))$. Given some risk aversion coefficient $\lambda > 0$ our investor chooses a portfolio $\boldsymbol{\theta}^*$ in order to maximize

$$E(-L(\boldsymbol{\theta})) - \lambda \mathrm{VaR}_\alpha(L(\boldsymbol{\theta})) \tag{4}$$

over all $\boldsymbol{\theta} \in \Theta_V$. For concreteness we put $\alpha = 0.95$.

Remark 2.6. Portfolio optimization problems of the form (4) are frequently considered in practice. Moreover, optimization problems closely related to (4) do arise implicitly in the context of *risk-adjusted performance measurement*. Often the performance of portfolio managers or even business units is measured by the ratio of (expected) profits and the amount of risk capital needed to sustain the portfolio; see, for example, Jorion (2001). If the risk capital is determined using VaR, portfolio managers have similar incentives in choosing their portfolios as if operating directly under the simple criterion (4). Our example can therefore be viewed as a warning against the use of VaR in risk-adjusted performance measurement. This is no critique against risk-adjusted performance measurement as such: if combined with a coherent risk measure risk-adjusted performance measurement is a very sensible concept.

To determine the optimal portfolio $\boldsymbol{\theta}^*$ we fix θ_0, the amount invested in the riskless asset. Since the L_i are identically distributed, every portfolio $\tilde{\boldsymbol{\theta}} \in \Theta_V$ with $\tilde{\theta}_0 = \theta_0$ has the

5

same expected loss. Hence maximizing (4) over all portfolios in Θ_V with fixed investment in the riskless asset amounts to minimizing $\text{VaR}_\alpha(L)$ over all these portfolios. For $\alpha = 0.95$ this is clearly achieved by investing all funds into one bond, for instance the first, as was shown in Example 2.3. The optimal portfolio $\boldsymbol{\theta}^*$ is now easily determined: if the "expected risk adjusted return" of holding a defaultable bond, $(E(-L_1) - \text{VaR}_{0.95}(L_1))/95$, exceeds the return r on the riskless asset, the whole capital V is invested in the first bond; otherwise everything is invested in the riskless asset.

In our symmetric situation we would expect the optimal portfolio to consist of a mixture of an investment in the riskless asset and a portfolio consisting of an equal amount of each of the risky bonds. The following result shows that this is indeed the case, if we replace value-at-risk by a coherent risk measure which depends only on the distribution of losses such as (generalized) expected shortfall.

Lemma 2.7. *Given a coherent risk measure ρ, whose domain includes the set $\{L(\boldsymbol{\theta}) : \boldsymbol{\theta} \in \Theta\}$. Suppose that ρ depends only on the loss distribution. Define the set $\Theta_{0,V}$ by $\Theta_{0,V} := \{\boldsymbol{\theta} \in \Theta_V : \theta_0 = 0\}$, and the portfolio $\boldsymbol{\theta}_0^* \in \Theta_V$ by $\boldsymbol{\theta}_0^* = (0, V/m, \dots, V/m)'$. Then*

$$\rho(L(\boldsymbol{\theta}_0^*)) = \min\{\rho(L(\boldsymbol{\theta})) : \boldsymbol{\theta} \in \Theta_{0,V}\}.$$

A proof of this lemma may be found in Appendix A.

3 Bernoulli Mixture Models for Credit Portfolios

We now turn to the second theme of the paper and show how the risk measures we have discussed in the first part of this paper may be calculated in practice. The method we use is founded on the observation that all standard industry models, such as CreditRisk[+], CreditMetrics, CreditPortfolioView and the model proposed by the KMV corporation, can be represented by Bernoulli mixture models.

In the case of CreditRisk[+] and CreditPortfolioView this is easily seen, since these models are constructed using a mixing philosophy. In the case of CreditMetrics and KMV this is less obvious, since these models are usually presented as firm-value models. The idea of mapping CreditMetrics/KMV type models so that they resemble mixture models can be found in Koyluoglu and Hickman (1998) and Gordy (2000); the theory underlying the mapping procedure is discussed in some detail in Frey and McNeil (2001).

3.1 Notation and definition

Consider a portfolio of m counterparties and fix some time period $[t, t + \Delta t]$ where Δt is typically one year. Assume that at time t all counterparties are in some non-default state. For $1 \le i \le m$, let the random variable Y_i be the *default indicator* for obligor i at time $t + \Delta t$, taking values in $\{0, 1\}$. We interpret the value 1 as default and 0 as non-default. For the arguments of this paper it will suffice to consider losses as arising from defaults only, and to ignore the issue of losses arising from rating class downgrades. The random vector $\mathbf{Y} = (Y_1, \dots, Y_m)'$ is a vector of default indicators for the portfolio over the time horizon of interest.

In a mixture model the default probability of an obligor is assumed to depend on a (typically small) set of *common factors*, which are interpreted as *macroeconomic variables*; given these common factors defaults of different obligors are independent. Dependence between defaults stems from the mutual dependence of the default probabilities on the set of common factors. We now give a formal definition.

Definition 3.1 (Bernoulli Mixture Model). Given some $p < m$ and a p-dimensional random vector $\boldsymbol{\Psi} = (\Psi_1, \dots, \Psi_p)$, the random vector $\mathbf{Y} = (Y_1, \dots, Y_m)'$ follows a Bernoulli

mixture model with factor vector $\mathbf{\Psi}$, if there are functions $Q_i : \mathbb{R}^p \to [0,1]$, $1 \leq i \leq m$, such that conditional on $\mathbf{\Psi}$ the default indicator $\mathbf{Y}$ is a vector of independent Bernoulli random variables with $P(Y_i = 1|\mathbf{\Psi}) = Q_i(\mathbf{\Psi})$.

3.2   CreditRisk$^+$ as Bernoulli Mixture Model

CreditRisk$^+$ may be represented as a Bernoulli mixture model where the distribution of the default indicators is given by

$$
\begin{aligned}
P(Y_i = 1 \mid \mathbf{\Psi}) &= Q_i(\mathbf{\Psi}), \quad \text{and} \\
Q_i(\mathbf{\Psi}) &= 1 - \exp(-\mathbf{w}_i'\mathbf{\Psi}),
\end{aligned}
\tag{5}
$$

where $\mathbf{\Psi} = (\Psi_1, \ldots, \Psi_p)'$ is a vector of independent gamma distributed macroeconomic factors with $p < m$ and $\mathbf{w}_i = (w_{i,1}, \ldots, w_{i,p})'$ is a vector of constant factor weights. Clearly this model has the form specified in Definition 3.1.

This representation of CreditRisk$^+$ facilitates its comparison with other industry models. We note however that CreditRisk$^+$ is usually presented as a *Poisson* mixture model. In this more common presentation it is assumed that, conditional on $\mathbf{\Psi}$, the default of counterparty i in $[t, t + \Delta t]$ occurs independently of other counterparties with a Poisson intensity given by

$$
\Lambda_i(\mathbf{\Psi}) = \mathbf{w}_i'\mathbf{\Psi}.
\tag{6}
$$

Although this assumption makes it possible to default more than once, a realistic model calibration generally ensures that the probability of this happening is very small. Assuming for simplicity that $\Delta t = 1$, the conditional probability given $\mathbf{\Psi}$ that a counterparty defaults over the time period of interest (whether once or more than once) is given by

$$
1 - \exp(-\Lambda_i(\mathbf{\Psi})) = 1 - \exp(-\mathbf{w}_i'\mathbf{\Psi}),
$$

so that the assumption (6) implies the Bernoulli mixture model in (5).

The Poisson formulation of CreditRisk$^+$ has the pleasant analytical feature that the distribution of the number of defaults in the portfolio is equal to the distribution of a sum of independent negative binomial random variables, as is shown in Gordy (2000) and Frey and McNeil (2001). The computational attractions of gamma mixtures of Poisson random variables are well known in the actuarial literature; see for instance Grandell (1997).

Calibration of CreditRisk$^+$ means choosing the factors $\mathbf{\Psi}$ and setting the weight vectors $\mathbf{w}_i$. This is done by considering the likely contribution of the various factors to default risk of each obligor under the constraint that default rates for individuals in the same rating category should be constant. This means effectively that $E(\Lambda_i(\mathbf{\Psi})) = \overline{q}_{g(i)}$ where $\overline{q}_{g(i)}$ represents an estimate of the default rate for all obligors in the group $g(i)$ to which obligor i belongs.

3.3   CreditMetrics and KMV as Bernoulli Mixture Models

Both KMV and CreditMetrics may be considered to descend from the firm-value model of Merton (1974), where default is modelled as occurring when the asset value of a company falls below its liabilities. In statistical texts, such as Joe (1997), such models fall under the general heading of latent variable models.

In both KMV and CreditMetrics we consider a random vector $\mathbf{X} = (X_1, \ldots, X_m)'$ with a multivariate normal distribution, where X_i is an underlying latent variable for company i at time T. We further assume that the vector $\mathbf{X}$ depends on macroeconomic factors according to a classical *linear factor model*. Suppose, without any loss of generality, that $\mathbf{X}$ has mean zero; then $\mathbf{X}$ follows a linear factor model with dimension $p < m$ if $\mathbf{X}$ can be written as

$$
X_i = \mathbf{a}_i'\mathbf{\Theta} + \sigma_i\varepsilon_i,
\tag{7}
$$

for a p-dimensional Gaussian random vector $\boldsymbol{\Theta} \sim N_p(\mathbf{0}, \Omega)$, independent standard normally distributed rv's $\varepsilon_1 \ldots, \varepsilon_m$, which are also independent of $\boldsymbol{\Theta}$, and constant terms given by σ_i and $\mathbf{a}_i = (a_{i,1}, \ldots, a_{i,p})'$. Effectively this model can be thought of as imposing a simplifying structure on the correlation matrix of $\mathbf{X}$.

We define $(D_1, \ldots, D_m)$ to be a vector of deterministic cut-off levels or thresholds. Default of obligor i is modelled as occurring if X_i is less than D_i, so that

$$Y_i = 1 \Longleftrightarrow X_i \leq D_i. \tag{8}$$

Both KMV and CreditMetrics fit into this simple framework and the differences between the two are largely differences of interpretation or calibration, rather than structural differences.

In the KMV model the latent variables X_i are interpreted as relative changes in the firm's asset value (so-called asset returns). For determining the thresholds D_i an option pricing technique based on historical firm value data is used to calculate a *distance-to-default*. CreditMetrics is usually presented as a multi-state model. Instead of considering a single cut-off the range of the X_i is partitioned more finely to represent a series of rating classes of decreasing creditworthiness, culminating in default. The cut-off levels which define these classes are chosen so that default and rating state transition probabilities agree with historical data. In both models the linear factor structure of $\mathbf{X}$ is calibrated by considering common macroeconomic variables that impact firm value.

Simple calculations confirm that this construction defines a Bernoulli mixture model. Setting $\boldsymbol{\Psi} = \boldsymbol{\Theta}$ and using Equations (8) and (7), we see that, conditional on $\boldsymbol{\Psi}$, the Y_i are independent since the ε_i are iid. Moreover

$$P\left(Y_i = 1 \mid \boldsymbol{\Psi}\right) = P\left(X_i \leq D_i \mid \boldsymbol{\Psi}\right) = P\left(\varepsilon_i \leq \left(D_i - \mathbf{a}_i'\boldsymbol{\Psi}\right)/\sigma_i \mid \boldsymbol{\Psi}\right)$$
$$= \Phi\left(\left(D_i - \mathbf{a}_i'\boldsymbol{\Psi}\right)/\sigma_i\right),$$

where Φ is the df of standard normal. Introducing new notation for the constants in this model we may write the stochastic default probability as

$$Q_i(\boldsymbol{\Psi}) = \Phi\left(c_i - \mathbf{w}_i'\boldsymbol{\Psi}\right) \tag{9}$$

which is again of the form specified in Definition 3.1.

3.4 Other Possibilities

Frey, McNeil, and Nyfeler (2001) have shown that the idea behind CreditMetrics and KMV may be extended to construct other latent variable models where the distribution of $\mathbf{X}$ is no longer multivariate Gaussian, but rather some model from the family of multivariate normal mixtures, such as a multivariate t distribution or a multivariate hyperbolic. Set $\boldsymbol{\Psi} = (W, \Theta_1, \ldots, \Theta_p)'$ where $\boldsymbol{\Theta} \sim N_p(\mathbf{0}, \Omega)$ as above and W is a positive random variable independent of $\boldsymbol{\Theta}$. The Bernoulli mixture model representation of these extensions has the general form

$$
\begin{aligned}
P\left(Y_i = 1 \mid \boldsymbol{\Psi}\right) &= Q_i(\boldsymbol{\Psi}) \\
Q_i(\boldsymbol{\Psi}) &= \Phi\left(c_i W - \mathbf{w}_i'\boldsymbol{\Theta}\right),
\end{aligned}
$$

for constants c_i and $\mathbf{w}_i$. As a concrete example, if the distribution of W is such that νW^2 has a chi-squared distribution with ν degrees of freedom, we get a mixture model which is equivalent to a CreditMetrics/KMV type model with Student t_ν distributed asset returns following a linear factor model rather than Gaussian asset returns. (A general definition of the linear factor model which may be applied to non-Gaussian random vectors is given in the Appendix.) For a ful discussion of the theory behind these generalised latent variable model see Frey and McNeil (2001).

Another possible Bernoulli mixture model with relevance to practice is obtained by taking

$$P\left(Y_i = 1 \mid \boldsymbol{\Psi}\right) = Q_i(\boldsymbol{\Psi})$$
$$Q_i(\boldsymbol{\Psi}) = \frac{\exp(c_i - \mathbf{w}_i'\boldsymbol{\Theta})}{1 + \exp(c_i - \mathbf{w}_i'\boldsymbol{\Theta})}.$$

In comparison with Equation (9) the normal distribution function Φ has been replaced by a logistic distribution function; in statistical language the probit link between Q_i and $c_i - \mathbf{w}_i'\Theta$ has been replaced by a logit link. CreditPortfolioView is a model of this kind; see Wilson (1997).

3.5   Homogeneous Portfolios

In their most general form the models described above can model fully heterogeneous portfolios with different counterparty default probabilities and different default correlations between counterparties. To gain a better mathematical understanding of the differences between models it is useful to simplify to the case of homogeneous groups. Moreover, fully heterogeneous models are extremely difficult to calibrate reliably and it is quite common in practice to segment large portfolios into a small number of fairly homogeneous groups, corresponding to some external or internal rating class, which may be more realistically calibrated.

The correct way to mathematically formalise this notion of homogeneity is to assume that the default indicator vector $\mathbf{Y}$ is *exchangeable* (distributionally invariant under permutations):

$$(Y_1, \ldots, Y_m) \overset{\mathrm{d}}{=} (Y_{\Pi(1)}, \ldots, Y_{\Pi(m)}),$$

for any permutation $(\Pi(1), \ldots, \Pi(m))$ of $(1, \ldots, m)$.

An exchangeable Bernoulli-mixture model is obtained in Definition (3.1) when the functions Q_i are all identical. In this case it is convenient to introduce the rv $Q := Q_1(\boldsymbol{\Psi})$ and to observe that for $\mathbf{y} = (y_1, \ldots, y_m)'$ in $\{0, 1\}^m$

$$P(\mathbf{Y} = \mathbf{y} \mid Q) = Q^{\sum_{i=1}^m y_i}(1 - Q)^{m - \sum_{i=1}^m y_i},$$

and, in particular, $P(Y_i = 1 \mid Q) = Q$. Thus exchangeable models may essentially be thought of as *one-factor* versions of the more general multi-factor models.

The distributions of Q in the homogeneous-group versions of the models we have described are summarised below.

- CreditRisk$^+$. $Q = 1 - \exp(-Y)$, where $Y \sim Ga(a, b)$ is a single gamma distributed factor with density

$$f(y) = \frac{b^a}{\Gamma(a)} y^{a-1} \exp(-by). \tag{10}$$

  The 2 parameters a and b fully specify the model.

- CreditMetrics/KMV. $Q = \Phi(Z)$, where $Z \sim N(\mu, \sigma^2)$. This model is thus specified by the 2 parameters μ and σ. Q is said to have a probit-normal distribution. In the latent variable language of (7) this model is equivalent to a model where all possible pairs of firm asset values X_i and X_j have a common correlation given by $\rho = \sigma^2/(1 + \sigma^2)$.

- Frey-McNeil extensions to CreditMetrics/KMV (Frey, McNeil, and Nyfeler 2001). $Q = \Phi(cW + Z)$, where now $Z \sim N(0, \sigma^2)$. These extended models have the 2 parameters c and σ, as well as extra parameters stemming from the distribution of W.

- CreditPortfolioView. $Q = \exp(Z)/(1 + \exp(Z))$ with $Z \sim N(\mu, \sigma^2)$. Q is said to have a logit-normal distribution.

It is interesting to note that our simplified version of CreditRisk$^+$ is essentially identical to modelling Q with a beta distribution with parameters a and b. This is seen by noting that the density $g(q)$ of Q in the one-factor gamma model can be calculated from the gamma density in (10) according to $g(q) = f(-\log(1-q))/(1-q)$; we obtain

$$g(q) = \frac{b^a}{\Gamma(a)} \left(-\log(1-q)\right)^{a-1} (1-q)^{b-1}.$$

In a realistic credit model the parameters a and b will be such that the probability mass of the distribution is concentrated on the left side of the interval $[0, 1]$, default being for most credit classes a rare event. For q small we may use the approximation $-\log(1-q) \approx q$ to observe that the functional form is extremely close to that of the beta distribution, which has density $g(q) = \beta(a, b)^{-1} q^{a-1} (1-q)^{b-1}$, $a, b > 0$. The idea of using the beta distribution as a mixture distribution for a Bernoulli failure probability is a classical one in statistics and the resulting Bernoulli mixture model proves to be particularly analytically tractible; see Joe (1997) for more information.

In exchangeable models we introduce the notation π for the probability that a counterparty defaults; it is easily calculated that $\pi = E(Y_i) = E(Q)$. Moreover, joint default probabilities can be expressed as higher moments of the mixing distribution; we can define and calculate

$$\pi_k \quad := \quad P(Y_1 = 1, \ldots, Y_k = 1) = E\left(E(Y_1 \cdots Y_k \mid Q)\right) = E(Q^k).$$

For $i \neq j$, the covariance between default indicators and the default correlation are thus determined by the first two moments of the mixing distribution.

$$\begin{aligned} \text{cov}(Y_i, Y_j) &= \pi_2 - \pi^2 = \text{var}(Q) \geq 0, \\ \rho_Y := \rho(Y_i, Y_j) &= \frac{\pi_2 - \pi^2}{\pi - \pi^2}. \end{aligned}$$

3.6 A Warning Concerning Correlations

The common *default* correlation ρ_Y, should be carefully distinguished from the common *asset* correlation ρ which appears in the one-factor version of the KMV/CreditMetrics model.

It is common practice to compare one-factor credit models for homogeneous groups by calibrating them so that they have the same default probability π and default correlation ρ_Y or, in other words, so that the first two moments of the distribution of Q agree in all cases. Clearly if we do this for the four standard industry models then we fix the two parameters of the mixing distribution and fully specify the models. Of the examples we have studied, only the Frey-McNeil family of extensions to KMV/CreditMetrics allow the possibility of extra parameters. In Section 3.8 we give an example of this kind of default-correlation-based calibration.

If we restrict our attention to the KMV/Creditmetrics style of latent variable model, and its extensions, then another approach to calibration is possible, although in our opinion misguided. We can calibrate these models so that default probability π and latent variable correlation ρ take prespecified values. However, this approach is subject to considerable *model risk* and the values of risk measures for the portfolio loss can vary greatly depending on our choice of multivariate model for the latent variables $\mathbf{X}$.

In Frey, McNeil, and Nyfeler (2001) it is shown that it is essentially the *copula* of the latent variables that drives the tail of the loss distribution and that it is possible to define

models with identical default probabilities and latent variable correlations, but very different copulas. We give an example here to reinforce our warning against model calibrations based on asset correlation.

Group	π	ρ
1	0.05	0.10
2	0.05	0.20

Table 1: Values of π and ρ for the two groups in Table 2.

We consider two homogeneous groups of 1000 counterparties of poor credit quality as defined in Table 1. Thus in both groups the default probability is 5%; in Group 1 the asset correlation is 10% and in Group 2 it is 20%. To each group we calibrate three latent variable models: a model which assumes multivariate normality of the latent variables as in standard KMV/Creditmetrics; two models from the Frey-McNeil class which assume that the latent variables have a multivariate t distribution with 10 and 5 degrees of freedom respectively. We assume that all losses given default are equal to one unit so that the overall loss is given by the number of defaulting counterparties. We then use Monte Carlo with 1 million replications to estimate the 99% and 99.9% VaRs for the credit loss distribution, which we label $\text{VaR}_{0.99}(L^{(1000)})$ and $\text{VaR}_{0.999}(L^{(1000)})$. Results are presented in Table 2.

Group	Risk Measure	KMV/CreditMetrics	t Model (10 d.f.)	t Model (5 d.f.)
1	$\text{VaR}_{0.99}(L^{(1000)})$	170	255	320
1	$\text{VaR}_{0.999}(L^{(1000)})$	242	384	482
2	$\text{VaR}_{0.99}(L^{(1000)})$	250	327	389
2	$\text{VaR}_{0.999}(L^{(1000)})$	386	512	600

Table 2: Estimates of VaR for two different portfolios of 1000 counterparties. In both cases default probability π and latent variable correlation ρ are held fixed according to the values in Table 1. Estimates are based on 1000000 Monte Carlo simulations. Losses are rounded to nearest unit.

Clearly moving from a Gaussian assumption to a multivariate t assumption for the latent variables has a massive effect on the VaR, even though the expected number of defaults is alway the same ($1000\pi = 50$). As the "true" multivariate model for latent variables is generally unknown, this example should be interpreted as a warning against calibration approaches which are based exclusively on assumptions concerning latent variable correlations.

3.7 Loss Distributions for Large Homogeneous Portfolios

In this section we state a result which shows that for large homogeneous groups the tail of the portfolio loss distribution and, in particular, the risk measures that describe it, are driven by the tail of the underlying mixture distribution of Q. This insight will help us implement a simple risk measure estimation methodology for homogeneous groups that obviates the need for simulation.

We introduce a sequence of iid exposures $\{E_i\}_{i\in\mathbb{N}}$ with mean exposure μ_E and finite variance. Assuming that the entire exposure is lost in the event of default the loss in a portfolio of m obligors over the time period $[t, t + \Delta t]$ is given by $L^{(m)} = \sum_{i=1}^{m} E_i Y_i$

Proposition 3.2. *Let* $\text{VaR}_\alpha(L^{(m)})$ *and* $\text{ES}_\alpha(L^m)$ *be the value-at-risk and expected shortfall for a portfolio of* m *counterparties. Assume that the quantile function* $\alpha \to q_\alpha(Q)$ *is*

continuous in α, i.e. that

$$G(q_\alpha(Q) + \delta) > \alpha \ \text{for every} \ \delta > 0. \tag{11}$$

Then

$$\lim_{m \to \infty} \frac{1}{m} \text{VaR}_\alpha(L^{(m)}) = \mu_E \, q_\alpha(Q) \tag{12}$$

$$\lim_{m \to \infty} \frac{1}{m} \text{ES}_\alpha(L^{(m)}) = \mu_E \, E(Q \mid Q \geq q_\alpha(Q)). \tag{13}$$

A proof of (12) may be found in Frey and McNeil (2001) and (13) follows easily from (12). A similar result is proved in Gordy (2001) and this work has been very influential in the setting of capital charges in the Basel II proposals on credit risk. An early version of the limit result, in the special case of the probit-normal mixing distribution (i.e. the one-factor version of the KMV-model), can also be found in KMV-Corporation (1997).

In all models we consider the mixing variable Q has a strictly positive density and the condition (11) is satisfied. This motivates our use of the following approximations to our risk measures in large balanced portfolios.

$$\begin{aligned}
\text{VaR}_\alpha(L^{(m)}) &\approx m \, \mu_E \, q_\alpha(Q) \\
\text{ES}_\alpha(L^{(m)}) &\approx m \, \mu_E \, E(Q \mid Q \geq q_\alpha(Q)).
\end{aligned}$$

For the standard industry models it is possible to calculate $q_\alpha(Q)$ and $E(Q \mid Q \geq q_\alpha(Q))$ accurately. To calculate the former simply requires inversion of the distribution function G of Q, which is readily performed for the standard models. To calculate the latter we can also use the formula

$$E(Q \mid Q \geq q_\alpha(Q)) = q_\alpha(Q) + (1 - \alpha)^{-1} \int_{q_\alpha(Q)}^\infty \overline{G}(q) dq,$$

where $\overline{G} = 1 - G$ denotes the tail function of Q; the integral can be evaluated numerically.

Accurate calculation is slightly more difficult in the case of the Frey-McNeil extensions to KMV/Creditmetrics, since the distribution of $\Phi^{-1}(Q)$ is a convolution, and the distribution function G of the mixing variable Q must itself be calculated by numerical integration. We restrict our attention in the following example to the standard industry models.

3.8 An Example

In this concluding example we show how both the risk measures discussed in Section 2 may be calculated for large homogeneous portfolios.

Group	π	π_2	ρ_Y
1	0.05	0.0037	0.0255
2	0.05	0.0052	0.0578

Table 3: Values of π, π_2 and ρ_Y for the two groups in Table 4.

We compare risk measures for the loss distributions in the standard industry models for two homogeneous portfolios of poor credit quality defined in Table 3. Thus for both groups the default probability π and the default correlation ρ_Y, or equivalently the joint default probability π_2, are assumed to be known and fixed. We set $\mu_E = 1$ and consider a portfolio of size $m = 1000$. Note that, in the special case of the KMV/CreditMetrics model, the two groups coincide with the groups in Table 1; that is, the asset correlation values ρ in Table 1 are precisely the values that give the default correlations ρ_Y in Table 3.

Using direct calculation, and numerical methods where necessary, we determine the parameters of the mixing distribution which are necessary to give the required first and second moments. We then calculate risk measures for the mixing distribution and scale them by portfolio size to estimate VaR and expected shortfall at different probability levels. The results are contained in Table 4.

Group	Risk Measure	CreditRisk$^+$	KMV/CreditMetrics	CreditPortfolioView
1	$\mathrm{VaR}_{0.99}(L^{(1000)})$	162	169	175
1	$\mathrm{VaR}_{0.999}(L^{(1000)})$	218	241	265
1	$\mathrm{ES}_{0.99}(L^{(1000)})$	186	200	214
1	$\mathrm{ES}_{0.999}(L^{(1000)})$	241	271	301
2	$\mathrm{VaR}_{0.99}(L^{(1000)})$	237	250	268
2	$\mathrm{VaR}_{0.999}(L^{(1000)})$	340	384	445
2	$\mathrm{ES}_{0.99}(L^{(1000)})$	282	308	344
2	$\mathrm{ES}_{0.999}(L^{(1000)})$	380	439	514

Table 4: Estimates of risk measures for two different portfolios of 1000 counterparties. Losses are rounded to nearest unit.

Although the model risk is not as severe as in Table 2 it is still clear that for these portfolios of poorer quality the choice of model has an effect on the tail of the loss distribution and the associated risk measures. Up to the 99th percentile the differences are not too great, but at the 99.9th percentile there are large differences between the most optimistic model (in this case CreditRisk$^+$) and the most pessimistic model (CreditPortfolioView); the expected shortfall estimate at the 99.9th percentile corresponds to a loss of 38% of portfolio value in the case of CreditRisk$^+$ and 51% for CreditPortfolioView. In Figure 1 we plot the tail functions on a logarithmic y-scale for the three models. The differences clearly emerge beyond the 99th percentile.

Clearly this methodology can provide useful insights into the extremal behaviour of credit models. It is interesting to note that the technique may be applied equally easily to both VaR and expected shortfall; therefore there is no insurmountable barrier to basing our practical risk management on the coherent alternative measure.

References

ACERBI, C., and D. TASCHE (2001): "On the coherence of expected shortfall," working paper, Department of Mathematics, TU-München.

ARTZNER, P., F. DELBAEN, J. EBER, and D. HEATH (1999): "Coherent Measures of Risk," *Mathematical Finance*, 9, 203–228.

CREDIT-SUISSE-FINANCIAL-PRODUCTS (1997): "CreditRisk$^+$ a Credit Risk Management Framework," Technical Document, available from `htpp://www.csfb.com/creditrisk`.

CROUHY, M., D. GALAI, and R. MARK (2000): "A comparative analysis of current credit risk models," *Journal of Banking and Finance*, 24, 59–117.

FREY, R., and A. McNEIL (2001): "Modelling dependent defaults," Preprint, ETH Zürich, available from `http://www.math.ethz.ch/~frey`.

FREY, R., A. McNEIL, and N. NYFELER (2001): "Copulas and Credit Models," *Risk*, 10, 111–114.

FRITELLI, M., and E. GIANIN (2002): "Putting Order in Risk Measures," *Journal of Banking and Finance.*

GORDY, M. (2000): "A comparative anatomy of credit risk models," *Journal of Banking and Finance*, 24, 119–149.

GORDY, M. (2001): "A Risk-Factor Model Foundation for Ratings-Based Bank Capital Rules," Board of Governors of the Federal Reserve System, Working Paper.

GRANDELL, J. (1997): *Mixed Poisson Processes.* Chapman and Hall, London.

JOE, H. (1997): *Multivariate Models and Dependence Concepts.* Chapman & Hall, London.

JORION, P. (2001): *Value at Risk: the New Benchmark for Measuring Financial Risk.* McGraw-Hill, New York.

KMV-CORPORATION (1997): "Modelling Default Risk," Technical Document, available from `http://www.kmv.com`.

KOYLUOGLU, U., and A. HICKMAN (1998): "Reconciling the Differences," *RISK*, 11(10), 56–62.

MERTON, R. (1974): "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates," *Journal of Finance*, 29, 449–470.

RISKMETRICS-GROUP (1997): "CreditMetrics – Technical Document," available from `http://www.riskmetrics.com/research`.

WILSON, T. (1997): "Portfolio Credit Risk I and II," *Risk*, 10(Sept and Oct).

A   Proof of Lemma 2.7

Proof. Denote by $\Delta^{m-1} := \{\boldsymbol{\gamma} \in \mathbb{R}^m : \gamma_i \geq 0, \sum_{i=1}^m \gamma_i = 1\}$ the standard $(m-1)$-simplex. Define a function

$$f_\rho : \quad \Delta^{m-1} \to \mathbb{R} \quad , \boldsymbol{\gamma} \mapsto f_\rho(\boldsymbol{\gamma}) := \rho\left(L((0, \gamma_1 M, \ldots, \gamma_m M)')\right) .$$

Then the lemma is equivalent to the claim

$$f_\rho((1/m, \ldots, 1/m)') = \min\{f_\rho(\boldsymbol{\gamma}) : \ \boldsymbol{\gamma} \in \Delta^{m-1}\} .$$

Since the L_i (the losses on the individual bonds) are iid, and since ρ depends only on the loss distribution, the function f_ρ is invariant under permutations of its arguments. Coherence of ρ (subadditivity and positive homogeneity) implies that f_ρ is convex. Denote by Π^m the set of all permutations of $\{1, \ldots, m\}$. Then we have for $\Pi \in \Pi^m$ and $\boldsymbol{\gamma} \in \Delta^{m-1}$

$$\begin{aligned}
f_\rho(\boldsymbol{\gamma}) &= f_\rho\left((\gamma_{\Pi(1)}, \ldots, \gamma_{\Pi(m)})'\right) \\
&= \frac{1}{m!} \sum_{\Pi \in \Pi^m} f_\rho\left((\gamma_{\Pi(1)}, \ldots, \gamma_{\Pi(m)})'\right) \\
&\geq f_\rho\left(\frac{1}{m!} \sum_{\Pi \in \Pi^m} \gamma_{\Pi(1)}, \ldots, \frac{1}{m!} \sum_{\Pi \in \Pi^m} \gamma_{\Pi(m)}\right) \\
&= f_\rho((1/m, \ldots, 1/m)') ,
\end{aligned}$$

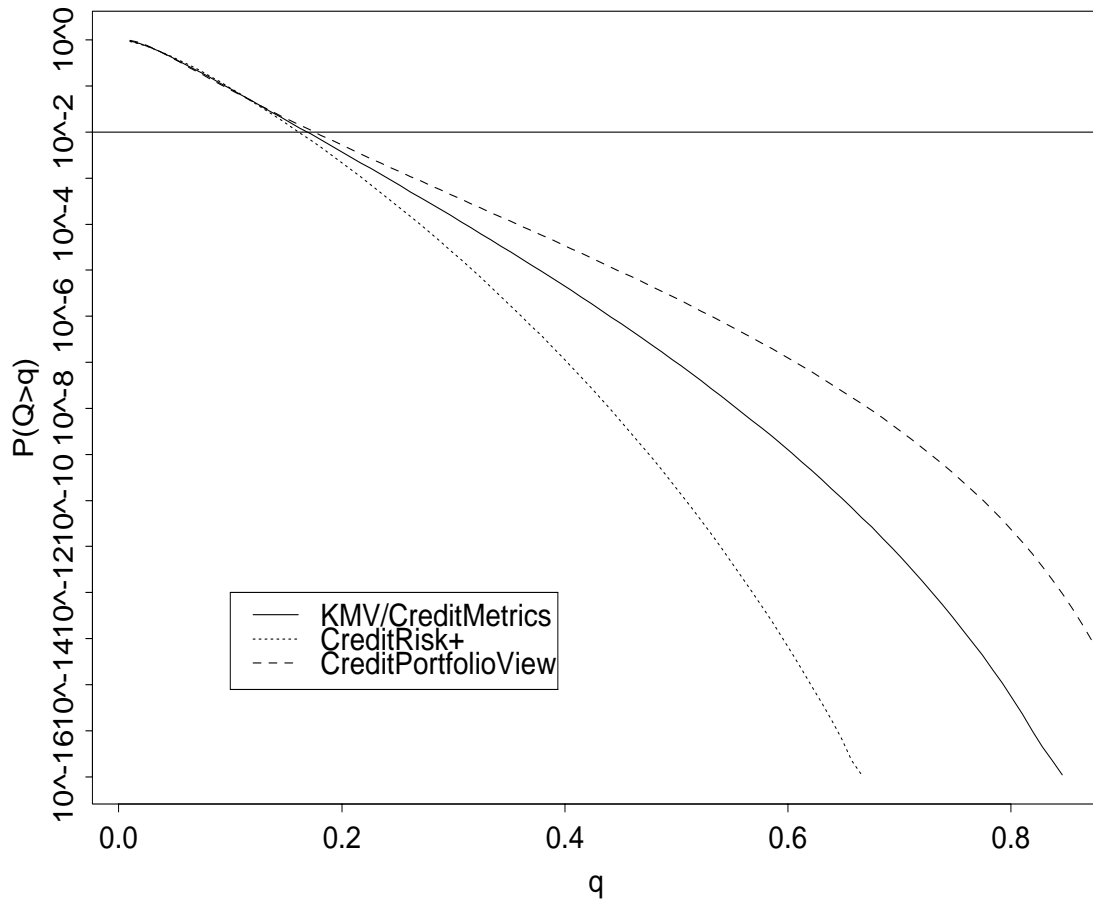where the inequality follows immediately from the convexity of f_ρ. $\qquad\square$

14

Figure 1: Tail of the mixing distribution G of Q in three different exchangeable Bernoulli mixture models: CreditRisk$^+$ (essentially a beta mixing distribution); KMV/CreditMetrics (probit-normal mixture); CreditPortfolioView (logit-normal mixture). In all cases the first two moments π and π_2 have the values for group 1 in Table 1. Horizontal line at 0.01 shows that models only really start to diverge after 99th percentile of mixing distribution.