A Credit Portfolio Framework under Dependent Risk Parameters PD, LGD and EAD

Johanna Eckert[♭], Kevin Jakob[♯] and Matthias Fischer[†]

April 22, 2015

[♭]Department of Insurance Economics and Risk Management, Friedrich-Alexander University Erlangen-Nürnberg (FAU)
[♯] Department of Statistics, University of Augsburg
[†]Department of Statistics & Econometrics, Friedrich-Alexander University Erlangen-Nürnberg (FAU)
mailto: *Johanna.Eckert@fau.de*

This paper introduces a credit portfolio framework which allows for dependencies between default probabilities, secured and unsecured recovery rates and exposures at default. The overall approach is an extension of the factor-models of Pykhtin (2003) and Miu and Ozdemir (2006), with respect to differentiated recovery rates and the inclusion of dependent exposures. As there is empirical evidence for dependence between these risk parameters and observations for the exposure at default, the secured and unsecured recovery rates are available only in case of a default, we propose a multivariate extension of Heckman's (1979) selection model in order to estimate the unknown parameters within a Maximum-Likelihood framework. Finally, we empirically demonstrate the effects of the dependence structure on the portfolio loss distribution and its risk measure for a hypothetical loan portfolio.

Keywords: dependent risk parameters; Factor model; Credit Metrics; Merton model; secured and unsecured recovery rate; CCF

1 Introduction, Literature Review and Motivation

Without any doubt, CreditMetrics^{TM} is one the most popular credit portfolio models in the banking industry. It arises as a natural extension of the well-known Vasicek model, see Vasicek [2002] which in turn forms the basis of the Basel II internal-ratings-based (IRB) capital requirements for credit risk. As a representative of the class of structural models, CreditMetrics$^{\text{TM}}$ is based on the firm value model of Merton [1973] and assumes that defaults or changes in the creditworthiness depend on the firm's asset value, which is driven by systematic as well as idiosyncratic factors. In order to calculate the portfolio loss distribution, relevant input parameters for the underlying borrowers or contracts are the exposure at default (EAD), the loss given default (LGD) and the borrowers' default indicator ($\mathcal{D}$), where a default occurs ($\mathcal{D} = 1$) with a given probability of default (PD). In the classical setting, exposures and LGDs are assumed to be deterministic whereas the default indicators are dependent (Bernoulli) random variables. Default dependence arises, for instance, from a common industry background, the belonging to the same or related countries or regions, or by legal or economical relationships.

Several extensions of this model framework can be found in the recent literature. Focusing on the above-mentioned input parameters, most of the subsequent contributions treat the LGDs as random variables and, above that, allow for dependence between them and the default indicators (commonly called as "PD-LGD correlation") within a factor model approach, see Frye [2000a], Frye [2000b], Pykhtin [2003], Tasche [2004], Miu and Ozdemir [2006], Witzany [2011], Bade et al. [2011] or Rösch and Scheule [2014].

Within this work we introduce a novel model which extends the single-factor model of Pykhtin [2003] and the two-factor model of Miu and Ozdemir [2006] in the following sense: Instead of capturing only dependence between the default indicator and the LGD, we *additionally* take into account the remaining risk parameter EAD by means of the utilization rate at default. Furthermore, we allow for a natural segmentation of the LGD by distinguishing between secured and unsecured recovery rates. Hence, our generalized multi-factor framework allows for dependence between all four risk parameters.

For this purpose, the outline of this article is as follows: Section 2 introduces the underlying definitions and notions. In section 3 we introduce the above-mentioned multi-factor model in order to capture dependence between the four relevant risk parameters. In addition, some remarks are provided how to estimate the unknown parameter within a multivariate sample selection framework. Finally, section 4 illustrates the effects on the risk figures of the portfolio loss distribution for a hypothetical loan portfolio and under different hypothetical but realistic parameter sets which where derived from various empirical studies which at least deal with partial dependence structures. Section 5 concludes.

2 Probabilities of Default, Loss Rates, Exposure at Default and Loss Distributions

We consider a portfolio of $N \in \mathbb{N}$ loans, which is aggregated at the borrower level. A default occurs with probability $\text{PD}_i \in (0, 1)$ and will be represented by a random variable $\mathcal{D}_i \sim \text{Bern}(\text{PD}_i)$ describing if the default of i occurs ($\mathcal{D}_i = 1$) or not ($\mathcal{D}_i = 0$). Typically, PD_i is estimated for a one year horizon. In the event of default the loss arises from the current amount of exposure $\text{EAD}_i \in [0, \infty)$ and the loss rate $\text{LGD}_i \in [0, 1]$. The LGD equals the percentage amount of the EAD that cannot be recovered. This contains both the liquidation of collaterals and the insolvency quota. In general, the risk parameters

EAD and LGD are a-priori unknown. Therefore, we treat them as random variables within our model. The loss of the overall credit portfolio caused by borrower's defaults[1] equals:

$$L := \sum_{i=1}^{N} L_i = \sum_{i=1}^{N} \mathrm{EAD}_i \cdot \mathrm{LGD}_i \cdot \mathcal{D}_i. \tag{1}$$

In order to describe the shape of the portfolio loss distribution we observe the expected loss $\mathrm{EL} := \mathbb{E}\left(L\right)$, the standard deviation $\mathrm{SD} := \sqrt{\mathrm{Var}\left(L\right)}$ and the value at risk at 99.9% $\mathrm{VaR}_{99.9\%} := F_L^*\left(99.9\%\right)$, whereas F_L^* stands for the quantile function of L. Furthermore we derive the economic capital at 99.9%

$$\mathrm{EC}_{99.9\%} := \mathrm{VaR}_{99.9\%} - \mathrm{EL},$$

as well as the expected shortfall at 99.9%

$$\mathrm{ES}_{99.9\%} := \mathbb{E}\left[L | L \geq \mathrm{VaR}_{99.9\%}\right].$$

2.1 Exposure at Default

In general, a bank makes a credit commitment at the commencement of the contract that the borrower is allowed to draw funds up to the specified limits. Therefore, a credit position consists of two parts, the currently drawn (on balance) and the undrawn (off balance) amount [Taplin et al., 2007, BCBS, 2004]. Hence, the amount of the exposure in the event of default is unknown before the time of default. To determine the EAD the utilization in the event of default has to be estimated. The off balance part that has not been drawn yet, but will be utilized at the time of default is calculated by multiplying the current unused amount with the credit conversion factor $\mathrm{CCF}_i \in [0,1]$. Then the EAD is given by

$$\mathrm{EAD}_i = \mathrm{Drawn\ \ Amount}_i + \mathrm{CCF}_i \cdot \mathrm{Undrawn\ \ Amount}_i, \tag{2}$$

which is also illustrated in figure 1. The CCF is also called loan equivalent or usage given at default. Some authors define the CCF differently, for example as the conversion factor applied to the the total commitment or to the current utilization. Following the interpretation of the Basel II capital requirements [Taplin et al., 2007, BCBS, 2004] we stick to the definition of the CCF as the factor applied to the undrawn amount. The factor, applied to the whole commitment (Com_i), is denoted as the utilization rate at default $\mathrm{URD}_i \in [0,1]$. In this case the EAD can be expressed by:

$$\mathrm{EAD}_i = \mathrm{URD}_i \cdot \mathrm{Com}_i. \tag{3}$$

Our model uses the factor URD to specify the EAD. It should be noted that the estimation methods of the various factors can be transformed into each other and are therefore exchangeable.

[1]Since our focus is on the LGD and EAD, which are only relevant in case of a default, we restrict our analysis to the default mode. Therefore, we do not consider any kind of migration risk.

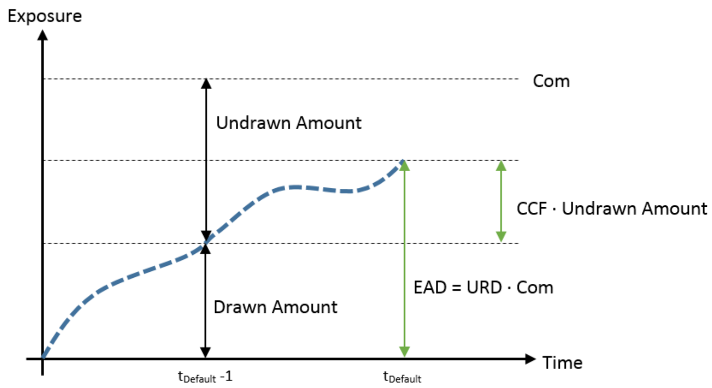

Figure 1: EAD in the context of the CCF and the URD

2.2 Loss Given Default in the Context of a Secured and an Unsecured Recovery Rate

Typically, the loss rate at the time of default is described by the risk parameter LGD as the ratio between the loss in case of a default and the EAD. A more detailed view can be achieved by differentiating between the loss rates of secured and unsecured exposures [BCBS, 2004]. Thereby, we can distinguish between the income from the realization of collaterals and the additional income during the liquidation process. The secured recovery rate $\mathrm{SRR}_i \in [0,1]$ describes the ratio between the payments from the realization of collaterals and the recently stated market value of them denoted by $\mathcal{C}_i$. On the contrary the unsecured recovery rate $\mathrm{URR}_i \in [0,1]$ indicates how much of the residual debt is settled by other payments reduced by the chargeable costs (for example handling and legal costs) [Eller et al., 2010]. Therefore, the loss of a credit position is given by:

$$L_i = (\mathrm{EAD}_i - \mathcal{C}_i \cdot \mathrm{SRR}_i) \cdot (1 - \mathrm{URR}_i) \,. \tag{4}$$

Using a deterministic collateralization quota $q := \frac{\mathcal{C}_i}{\mathrm{EAD}_i} \in [0,1]$, the LGD can be calculated from SRR and URR using the relation

$$\mathrm{LGD}_i = (1 - q \cdot \mathrm{SRR}_i) \cdot (1 - \mathrm{URR}_i) \,. \tag{5}$$

The stepwise approach to determine the loss in the event of a default is illustrated in figure 2.

For the definition of the risk parameters URD, SRR and URR, we follow an approach similar to one of the "potential LGD" as described by Pykhtin [2003]. The potential LGD is a function of the collateral value and is therefore defined irrespectively of the event of default. By contrast, the potential LGD ist defined for all borrowers, whereas the conventional LGD is defined only for defaulted borrowers. This distinction leads to the fact that, assuming that PD and LGD depend on each other, the expected conventional LGD is not only a function of the collateral value but also of the dependence upon PD.

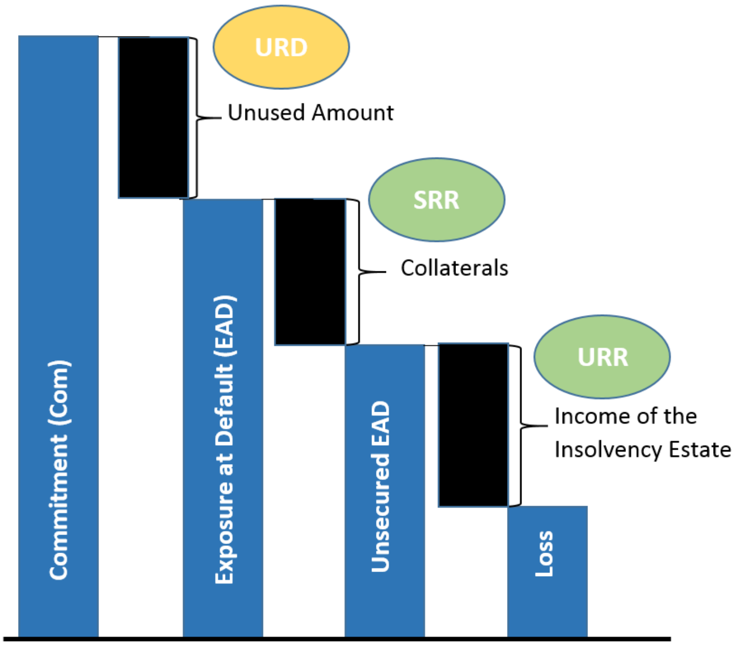

Figure 2: Step by step determination of a borrower's loss

This dependence can be captured by the sample selection model of Heckman [1979] (see chapter 3.2). Neglecting the issue of sample selection might lead to an underestimation of the conventional LGD. Our model framework takes this into account not only for the LGD, but for all risk parameters affecting the borrowers loss in case of a default. These are defined irrespectively of the event of default, but are only observable for defaulted borrowers.

3 Coupling Risk Parameters within a Generalized Multi-Factor Framework

3.1 Model Set-Up

Considering different points in time denoted by $t = 1, ..., T$, the creditworthiness of each firm is driven by two factors: a systematic factor S_t^A, which also influences the other firms simultaneously, and an idiosyncratic factor ϵ_{it}^A, which only affects this specific firm. The risk driver of default A_{it} of borrower i at time t is defined as:

$$A_{it} := \Phi^{-1}\left(\text{PD}_{it}\right) - \left(\alpha S_t^A + \sqrt{1-\alpha^2}\epsilon_{it}^A\right), \tag{6}$$

where the random variables S_t^A and ϵ_{it}^A are assumed to be independent and to follow a standard normal distribution: $S_t^A, \epsilon_{it}^A \overset{iid}{\sim} N(0,1)$. $\Phi^{-1}\left(\text{PD}_{it}\right)$ represents the deterministic default threshold for a borrower with probability of default PD_{it}. The parameter

$\alpha \in (-1, 1)$ governs the impact of the systematic factor S_t^A on the risk driver of default A_{it}. If α is close to zero, the borrower has a minor connection to the overall state of the economy. In this case the idiosyncratic risk is more important[2]. Analogous to CreditMetrics™ we interpret the part $\alpha S_t^A + \sqrt{1-\alpha^2}\epsilon_{it}^A$ as the standardized form of the logarithmic asset return. Particularly for unlisted companies the asset value is not observable on a daily basis. Hence the risk driver of the default A_{it} has to be chosen as an unobservable, latent variable. In our model a firm defaults if A_{it} exceeds 0. Let $\mathcal{D}_{it}$ denote the default event of the firm i at time t, i.e.:

$$\mathcal{D}_{it} = \begin{cases} 1, & \text{if } A_{it} \geq 0 \\ 0, & \text{else.} \end{cases} \tag{7}$$

In this setting, the same probability of default and the same conditional probability of the default as in the CreditMetrics™ model emerge:

$$\mathbb{P}\left(\mathcal{D}_{it} = 1\right) = \mathrm{PD}_{it}, \tag{8}$$

and:

$$\mathrm{PD}_{it}\left(s_t^A\right) := \mathbb{P}\left(\mathcal{D}_{it} = 1 \mid S_t^A = s_t^A\right) = \Phi\left(\frac{\Phi^{-1}\left(\mathrm{PD}_{it}\right) - \alpha s_t^A}{\sqrt{1-\alpha^2}}\right). \tag{9}$$

The systematic factor of default S_t^A in turn is partitioned into a systematic factor X_t, which influences the systematic factors of all risk parameters, and a specific systematic factor Z_t^A, which influences the systematic factor of default only:

$$S_t^A := \theta_A X_t + \sqrt{1-\theta_A^2} Z_t^A. \tag{10}$$

Like α in case of the risk driver of default A_{it}, the parameter $\theta_A \in (-1, 1)$ measures the sensitivity of the systematic factor of default S_t^A to the mutual systematic factor X_t. The factors $Z_t^A \sim N(0,1)$ and $X_t \sim N(0,1)$ are assumed to be independent and standard normally distributed, such that S_t^A again follows a standard normal distribution.

Analogously, the utilization rate at default URD is assumed to be driven by

$$B_{it} := \beta S_t^B + \sqrt{1-\beta^2} I_{it}^B, \tag{11}$$

where the systematic factor S_t^B and the idiosyncratic factor I_{it}^B are again independent and standard normally distributed and $\beta \in (0, 1)$. The risk driver B_{it} can be interpreted as the transformation of the URD into a standard normally distributed random variable. Letting $F_{\mathrm{URD}_{it}}$ be the cumulative distribution function of the URD of borrower i at time t, the URD of the borrower i at time t can be expressed by:

$$URD_{it} := F^*_{URD_{it}}\left(\Phi\left(B_{it}\right)\right), \tag{12}$$

[2] *Andersen and Sidenius [2005] includes stochastic factor loadings α. This gives the opportunity to model stronger dependencies during economic downturns compared to economic recovery.*

where $F^*_{URD_{it}}(t) := \inf\{v \in (0,1) : F_{URD_{it}}(v) \geq t\}$ denotes the quantile function. Also in this case the systematic factor of the URD S_t^B is divided into two components:

$$S_t^B := \theta_B X_t + \sqrt{1-\theta_B^2} Z_t^B. \tag{13}$$

The parameter $\theta_B \in (-1,1)$ governs the impact of the mutual systematic factor X_t on the systematic factor of the risk driver of the URD S_t^B. The specific systematic factor of the URD $Z_t^B \sim N(0,1)$ and the mutual systematic factor X_t are assumed to be independent, so that S_t^B follows a standard normal distribution. The idiosyncratic factor of the URD I_{it}^B can be also divided into the idiosyncratic factor of default ϵ_{it}^A, which influences the idiosyncratic factors of all risk parameter, and a specific idiosyncratic factor ϵ_{it}^B:

$$I_{it}^B := \rho_B \epsilon_{it}^A + \sqrt{1-\rho_B^2} \epsilon_{it}^B. \tag{14}$$

The parameter $\rho_B \in (-1,1)$ controls how much the mutual idiosyncratic factor ϵ_{it}^A affects the idiosyncratic factor of the URD I_{it}^B. $\epsilon_{it}^B \sim N(0,1)$ is assumed to be independent of ϵ_{it}^A. Therefore, the assumption holds that I_{it}^B is again standard normally distributed.

The secured recovery rate SRR and the unsecured recovery rate URR are modeled in the same way. Additionally, we assume independence of the random variables: $\epsilon_{it}^A, \epsilon_{it}^B, \epsilon_{it}^C, \epsilon_{it}^D, X_t, Z_t^A, Z_t^B, Z_t^C, Z_t^D$, $i = 1, \ldots, N$, $t = 1, \ldots, T$. This implies the following model framework:

	Risk driver	**with:**
$\mathcal{D}$	$A_{it} := \Phi^{-1}(\mathrm{PD}_{it}) - \left(\alpha S_t^A + \sqrt{1-\alpha^2} U_{it}^A\right)$	$S_t^A := \theta_A X_t + \sqrt{1-\theta_A^2} Z_t^A$ $U_{it}^A := \epsilon_{it}^A$
URD	$B_{it} := \beta S_t^B + \sqrt{1-\beta^2} U_{it}^B$	$S_t^B := \theta_B X_t + \sqrt{1-\theta_B^2} Z_t^B$ $U_{it}^B := \rho_B \epsilon_{it}^A + \sqrt{1-\rho_B^2} \epsilon_{it}^B$
SRR	$C_{it} := \gamma S_t^C + \sqrt{1-\gamma^2} U_{it}^C$	$S_t^C := \theta_C X_t + \sqrt{1-\theta_C^2} Z_t^C$ $U_{it}^C := \rho_C \epsilon_{it}^A + \sqrt{1-\rho_C^2} \epsilon_{it}^C$
URR	$D_{it} := \delta S_t^D + \sqrt{1-\delta^2} U_{it}^D$	$S_t^D := \theta_D X_t + \sqrt{1-\theta_D^2} Z_t^D$ $U_{it}^D := \rho_D \epsilon_{it}^A + \sqrt{1-\rho_D^2} \epsilon_{it}^D$

Table 1: Risk drivers of the particular risk parameters, with: $\epsilon_{it}^A, \epsilon_{it}^B, \epsilon_{it}^C, \epsilon_{it}^D, X_t, Z_t^A, Z_t^B, Z_t^C, Z_t^D \overset{iid}{\sim} N(0,1)$, for $i = 1, \ldots, N$, $t = 1, ..., T$.

Remark 3.1. It is straight forward to show that the risk drivers A, B, C and D again follow a multivariate normal distribution with mean $\left(\Phi^{-1}(\mathrm{PD}_{it}), 0, 0, 0\right)^T$ and a covariance given by

$$\mathrm{Cov}(X,Y) = \begin{cases} \varrho_1^X \varrho_1^Y + \varrho_2^X \varrho_2^Y & \text{for } X,Y \in \{A,B,C,D\},\ X \neq Y \\ 1 & \text{for } X,Y \in \{A,B,C,D\},\ X = Y \end{cases}, \tag{15}$$

where $\varrho_{1/2}^{X/Y}$ is defined for risk driver $X, Y \in \{A, B, C, D\}$ as

$$\varrho_1^X := \begin{cases} -\alpha\theta_A & \text{for } X = A \\ \beta\theta_B & \text{for } X = B \\ \vdots & \vdots \end{cases} \quad \text{and} \quad \varrho_2^X := \begin{cases} -\sqrt{1-\alpha^2} & \text{for } X = A \\ \sqrt{1-\beta^2}\rho_B & \text{for } X = B \\ \vdots & \vdots \end{cases}.$$

3.2 Some Remarks on Parameter Estimation

Observations for the utilization rate at default, the secured and unsecured recovery rates are available only in the event of default. Furthermore, there is empirical evidence for dependence between the default variable and the other risk parameters. Hence, the observable sample is necessarily truncated. Neglecting sample selection leads to biased parameter estimators [Heckman, 1979]. The univariate selection model of Heckman [1979], which was used by Bade et al. [2011] in their model for the dependence of default and the recovery rate, provides a solution for this problem. In our multivariate extension the risk drivers B_{it}, C_{it}, D_{it} are only observable (via URD, SRR and URR) if the loan i at time t defaults. Therefore, the selection equation is given by:

$$\begin{pmatrix} B_{it} \\ C_{it} \\ D_{it} \end{pmatrix} = \begin{cases} \text{observable,} & \text{in the event of default, i.e. } A_{it} \geq 0, \\ \text{not observable,} & \text{else, i.e. } A_{it} < 0. \end{cases} \tag{16}$$

The regression function for the sub-sample of available data results in the conditional expectation of the observable sub-sample:

$$\mathbb{E}\left[\begin{pmatrix} B_{it} \\ C_{it} \\ D_{it} \end{pmatrix} \middle| X_t, Z_t^A, Z_t^B, Z_t^C, Z_t^D, \text{ selection} \right] \tag{17}$$

$$= \begin{pmatrix} \beta S_t^B \\ \gamma S_t^C \\ \delta S_t^D \end{pmatrix} + \mathbb{E}\left[\begin{pmatrix} \sqrt{1-\beta^2}U_{it}^B \\ \sqrt{1-\gamma^2}U_{it}^C \\ \sqrt{1-\delta^2}U_{it}^D \end{pmatrix} \middle| X_t, Z_t^A, Z_t^B, Z_t^C, Z_t^D, \text{ selection} \right] \tag{18}$$

$$= \begin{pmatrix} \beta S_t^B \\ \gamma S_t^C \\ \delta S_t^D \end{pmatrix} - \begin{pmatrix} \rho_B\sqrt{1-\beta^2} \\ \rho_C\sqrt{1-\gamma^2} \\ \rho_D\sqrt{1-\delta^2} \end{pmatrix} \cdot \lambda_{it},$$

with $\lambda_{it} := \frac{\varphi(\kappa_{it})}{1-\Phi(\kappa_{it})}$ and $\kappa_{it} := \frac{-\Phi^{-1}(\mathrm{PD}_{it})+\alpha S_t^A}{\sqrt{1-\alpha^2}}$.

Proof. The derivation of the conditional expectation of the observed sub-sample uses Theorem 19.5 of Greene [2012]. □

Dependence between the default rate and the loss rate respectively the utilization rate has been attested by the empirical literature, see Frye [2000b], Altman et al. [2005], Miu and Ozdemir [2006], Asarnow and Marker [1995] and Jiménez et al. [2009]. Therefore, we have to assume that, in general, the error terms $U_{it}^A, U_{it}^B, U_{it}^C, U_{it}^D$ are not independent and that the regression parameters of the variable λ_{it} are unequal to 0. Regression

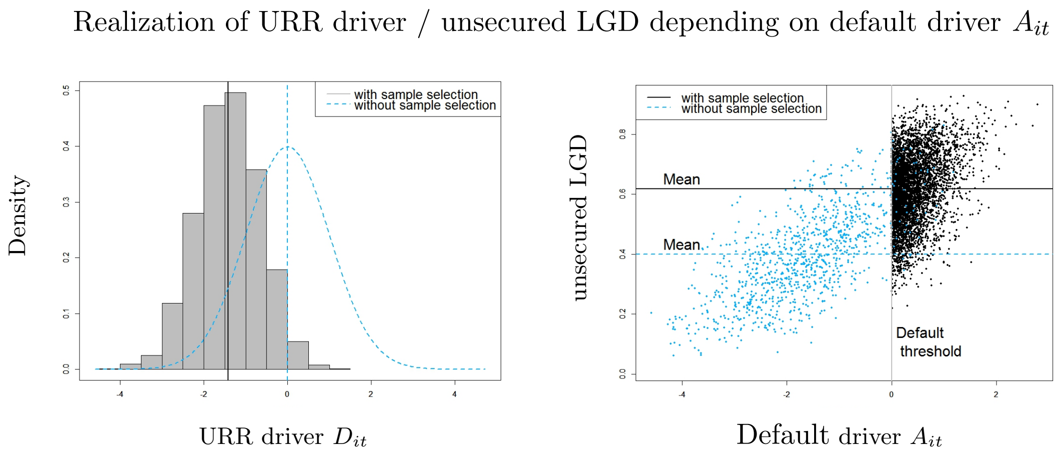

Figure 3: Sample selection in case of URR / unsecured LGD

estimators of the risk parameters on the selected sample that do not take into account the sample selection omit the final term of equation (18). This leads to a bias arising from the well-known problem of omitted variables. The bias of the expectation of a risk parameter directs toward the sign of the correlation between the error term U_{it}^B, U_{it}^C and U_{it}^D, respectively, and the error term of the selection equation U_{it}^A.

The issue is illustrated in figure 3, where the histogram shows the distribution of the URR driver D_{it} with and without sample selection. Similarly, realizations of the unsecured LGD (=1-URR) are pictured on the right hand side of the figure. For the marginal distribution of the unsecured LGD we choose a Beta distribution with mean 0.4 and standard deviation 0.15. The mean in case with (without) sample selection is illustrated via a black (red) line. In the unconditional case $D_{it} \sim N(0,1)$ and the LGD mean equals the mean of the corresponding Beta distribution, whereas in the conditional case $\mathbb{E}\left(D_{it} \,|\, A_{it} \geq 0\right) = \text{Cor}\left(A, D\right) \frac{\phi\left(-\Phi^{-1}(\text{PD}_{it})\right)}{\text{PD}_{it}} \approx -1.44$.[3] Therefore, the observable mean of the unsecured LGD is shifted from 0.4 to approximately 0.62 in our example.

In order to address the problem of sample selection, Heckman suggests two estimation procedures: a two-step estimator and a maximum likelihood estimator. Like Bade et al. [2011] we suggest the second option. The maximum likelihood estimator goes back to Heckman [1974] and is asymptotically efficient if the assumption of normal distributed error terms is correct. Additionally, our multivariate extension of the Heckman model takes into account several unobservable random variables, namely the systematic factors $X_t, Z_t^A, Z_t^B, Z_t^C, Z_t^D$. The likelihood function for time $t = 1, \ldots, T$ conditional on $X_t, Z_t^A, Z_t^B, Z_t^C, Z_t^D$ with n_t denoting the number of observations at time t can be

[3] Using Theorem 19.5 of Greene [2012] and remark 3.1 it is easy to show that the conditional expectation of B_{it} under sample selection is given by $\text{Cor}\left(A, D\right) \frac{\phi\left(-\Phi^{-1}(\text{PD}_{it})\right)}{\text{PD}_{it}}$, which in our parameter setting (0.4 for all parameters $\alpha, \beta, \gamma, \delta, \theta, \rho$) is around -1.44.

calculated by

$$\begin{aligned}
&\mathbb{L}_t\left(X_t, Z_t^A, Z_t^B, Z_t^C, Z_t^D\right) \qquad (19)\\
=&\prod_{i=1}^{n_t}\left[\mathbb{P}\left(\mathcal{D}_{it}=0 \mid X_t, Z_t^A, Z_t^B, Z_t^C, Z_t^D\right)\right]^{1-\mathcal{D}_{it}} \cdot \left[\mathbb{P}\left(\mathcal{D}_{it}=1 \mid X_t, Z_t^A, Z_t^B, Z_t^C, Z_t^D\right)\right.\\
&\left.\cdot f_{(B_{it},C_{it},D_{it})}\left(b_{it}, c_{it}, d_{it} \mid X_t, Z_t^A, Z_t^B, Z_t^C, Z_t^D, \mathcal{D}_{it}=1\right)\right]^{\mathcal{D}_{it}}\\
=&\prod_{i=1}^{n_t}\left[1-\Phi\left(\frac{\Phi^{-1}\left(\mathrm{PD}_{it}\right)-\alpha s_t^A}{\sqrt{1-\alpha^2}}\right)\right]^{1-D_{it}} \cdot \left[\Phi\left(\frac{\Phi^{-1}\left(\mathrm{PD}_{it}\right)-\alpha s_t^A}{\sqrt{1-\alpha^2}}\right)\right.\\
&\left.\cdot f_{(B_{it},C_{it},D_{it})}\left(b_{it}, c_{it}, d_{it} \mid X_t, Z_t^A, Z_t^B, Z_t^C, Z_t^D, \mathcal{D}_{it}=1\right)\right]^{\mathcal{D}_{it}},
\end{aligned}$$

where $f_{(B_{it},C_{it},D_{it})}\left(b_{it}, c_{it}, d_{it} \mid X_t, Z_t^A, Z_t^B, Z_t^C, Z_t^D, \mathcal{D}_{it}=1\right)$ denotes the conditional density of $(B_{it}, C_{it}, D_{it})^t$ given $X_t, Z_t^A, Z_t^B, Z_t^C, Z_t^D, \mathcal{D}_{it}=1$. It holds that:

$$\begin{aligned}
&f_{(B_{it},C_{it},D_{it})}\left(b_{it}, c_{it}, d_{it} \mid X_t, Z_t^A, Z_t^B, Z_t^C, Z_t^D, \mathcal{D}_{it}=1\right) \qquad (20)\\
=&\frac{1}{2\pi\Phi\left(\frac{\Phi^{-1}(\mathrm{PD}_{it})-\alpha S_t^A}{\sqrt{1-\alpha^2}}\right)\sqrt{\left(1-\beta^2\right)\left(1-\gamma^2\right)\left(1-\delta^2\right)\left(1-\rho_B^2\right)\left(1-\rho_C^2\right)\left(1-\rho_D^2\right)}}\\
&\cdot \phi_1\left(\tilde{\mu};\, 0, \tilde{\sigma}^2\right)\Phi\left(\frac{\frac{\Phi^{-1}(\mathrm{PD}_{it})-\alpha S_t^A}{\sqrt{1-\alpha^2}}+\hat{\mu}}{\hat{\sigma}}\right).
\end{aligned}$$

with:

$$\tilde{\mu} := \left(\sum_{i=1}^{4}\sum_{j>i}\left(\frac{\mu_i-\mu_j}{\sigma_i\sigma_j}\right)^2\right)^{\frac{1}{2}}, \qquad \hat{\mu} := \frac{\sum_{i=1}^{4}\sigma_i^{-2}\mu_i}{\sum_{i=1}^{4}\sigma_i^{-2}},$$

$$\tilde{\sigma} := \left(\sum_{i=1}^{4}\sigma_i^{-2}\right)^{\frac{1}{2}}, \qquad \hat{\sigma} := \tilde{\sigma}^{-1},$$

and

$$\begin{aligned}
&\mu_1 := \tfrac{-b_{it}+\beta s_t^B}{\rho_B\sqrt{1-\beta^2}}, \quad \mu_2 := \tfrac{-c_{it}+\gamma s_t^C}{\rho_C\sqrt{1-\gamma^2}}, \quad \mu_3 := \tfrac{-d_{it}+\delta s_t^D}{\rho_D\sqrt{1-\delta^2}}, \quad \mu_4 := 0,\\
&\sigma_1 := \sqrt{\tfrac{1-\rho_B^2}{\rho_B^2}}, \quad \sigma_2 := \sqrt{\tfrac{1-\rho_C^2}{\rho_C^2}}, \quad \sigma_3 := \sqrt{\tfrac{1-\rho_D^2}{\rho_D^2}}, \quad \sigma_4 := 1.
\end{aligned}$$

Proof. The derivation of the conditional density is available upon request.. □

Since the realizations of $X_t, Z_t^A, Z_t^B, Z_t^C, Z_t^D$ are not observable, one has to calculate the respective expectation of the likelihood function unconditional of these parameters in order to determine the likelihood function for time t:

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{L}_t\left(X_t, Z_t^A, Z_t^B, Z_t^C, Z_t^D\right)\right] = \qquad (21)\\
&\int\!\!\cdots\!\!\int \mathbb{L}_t\left(x_t, z_t^A, z_t^B, z_t^C, z_t^D\right)\varphi\left(x_t\right)\varphi\left(z_t^A\right)\varphi\left(z_t^B\right)\varphi\left(z_t^C\right)\varphi\left(z_t^D\right)dx_t dz_t^A dz_t^B dz_t^C dz_t^D.
\end{aligned}$$

The log-likelihood function for all times $t = 1, \ldots, T$ is then given by:

$$\ln \mathbb{L} = \ln \left(\prod_{t=1}^{T} \mathbb{E}\left[\mathbb{L}_t\left(X_t, Z_t^A, Z_t^B, Z_t^C, Z_t^D\right)\right] \right) = \sum_{t=1}^{T} \ln \mathbb{E}\left[\mathbb{L}_t\left(X_t, Z_t^A, Z_t^B, Z_t^C, Z_t^D\right)\right]. \tag{22}$$

Maximizing the log-likelihood function by numerical optimization with respect to the unknown parameters leads to the corresponding estimates.

4 Empirical Application

This section contains the numerical analysis regarding the impact of the dependence between the different risk parameters on the risk figures. Results are based on several Monte Carlo simulations. Besides a basic setting with hypothetical but reasonable assumptions regarding the dependence parameters, sensitivity analysis are conducted. After presenting the considered credit portfolio and deducing the sign of the dependence parameters, we present the simulation algorithm and our results regarding the uncertainties of different model specifications in terms of changing risk figures.

4.1 Portfolio characteristics and distributional assumptions

We consider a hypothetical credit portfolio with a total (drawn and undrawn) amount of 1.000.000, which is distributed on 5000 borrowers heterogeneously. The probabilities of default are expressed via a rating scale of seven categories, see table 2. The major part of the overall commitment belongs to the better rating categories. Taking into account the structure of our model, we do not distinguish between different economic sectors. However, it is straight forward to generalize the framework in order to a deal with multiple sectors.

Rating	1	2	3	4	5	6	7
PD	0,002	0,0005	0,0012	0,0041	0,014	0,047	0,15
Number of borrowers	246	92	1091	705	2700	70	96
% of commitment	27.45	12.26	28.43	21.26	7.76	2.25	0.58

Table 2: Rating categories and exposure distribution

The risk parameters utilization rate at default URD, secured recovery rate SRR, and unsecured recovery rate URR are modeled by a Beta-distribution on the unit interval[4]. The density is given by:

$$f_{\mathrm{Beta}(p,q)}(x) = \frac{1}{B(p,q)} x^{p-1} (1-x)^{q-1}, \; 0 < x < 1,$$

[4]This assumption is not irrefutable. For example, the SRR could take values higher than 1. This would be the case, if the payments from the realization of collaterals is higher than the recently stated market value of the collaterals.

where $B(\cdot,\cdot)$ is the Beta-function and $p > 0$, $q > 0$. The parameters p, q are determined by the expectation μ and the variance σ^2 via moment-matching. Following Tasche [2004], we choose a conservative value for the variance by specifying it as a fixed percentage $v = 0.25$ of the maximum possible variance $\mu(1-\mu)$. In this case the two parameters can be expressed by the first moment:

$$p = \mu\frac{1-v}{v}, q = (1-\mu)\frac{1-v}{v}.$$

Referring to the empirical study of Jiménez et al. [2009], the expectation of the utilization rate at default is chosen as $\mathbb{E}(\text{URD}) = 0.6$. Due to the lack of data, we choose hypothetical values for the expectation of secured recovery rate and unsecured recovery rate given by $\mathbb{E}(\text{SRR}) = 0.6$, $\mathbb{E}(\text{URR}) = 0,4$, which imply a similar, but more conservative expectation for the loss given default compared to Miu and Ozdemir [2006].

4.2 Dependence Parameters

In this section, our dependence parameters are specified as follows: In a first step, we deduce the signs of the dependence parameters. Afterwards, we propose hypothetical but plausible parameter values for the systematic and idiosyncratic factor weights in particular for URD, SRR and URR, which form the basis for further analysis.

4.2.1 Signs of the Dependence Parameters

The basis of the following derivation is the assumption regarding the states of the risk parameters in times of a high (low) creditworthiness of the borrower. A high creditworthiness can have two reasons: economic upturn (systematic) and borrower specific success (idiosyncratic). The literature states that the asset return is high during an economic upturn, whereas the loss given default and the utilization rate is low, see for example Frye [2000b] and Jiménez et al. [2009]. Assuming analogous mechanisms for the borrower specific success, the relations in table 3 hold.

Risk Driver of	**in Times of High Creditworthiness**
$\mathcal{D}$	Low
URD	Low
SRR	High
URR	High

Table 3: State of the risk parameters in times of high creditworthiness

Assuming without loss of generality that the systematic factors $X_t, Z_t^A, Z_t^B, Z_t^C, Z_t^D$ exhibit high values during economic upturns and the mutual idiosyncratic factor ϵ_{it}^A is also high during borrower specific success, we can derive the signs of the dependence parameters in our model. We demonstrate exemplarily the procedure for the utilization rate at default. The URD is low in times of high creditworthiness and thus its risk driver:

$$B_{it} := \beta \left(\theta_B X_t + \sqrt{1-\theta_B^2} Z_t^B \right) + \sqrt{1-\beta^2} \left(\rho_B \epsilon_{it}^A + \sqrt{1-\rho_B^2} \epsilon_{it}^B \right).$$

Deriving its correlation with the total systematic factor, the mutual systematic factor and the mutual idiosyncratic factor leads to the following signs of the dependence parameters:

- $\text{Cor}\left(B_{it}, \theta_B X_t + \sqrt{1-\theta_B^2} Z_t^B\right) = \beta < 0 \Rightarrow \beta < 0.$
- $\text{Cor}\left(B_{it}, X_t\right) = \beta\theta_B < 0 \Rightarrow \theta_B > 0.$
- $\text{Cor}\left(B_{it}, \epsilon_{it}^A\right) = \sqrt{1-\beta^2}\rho_B < 0 \Rightarrow \rho_B < 0.$

In doing so, the correlation of B_{it} with its risk parameter specific idiosyncratic factor $\text{Cor}\left(B_{it}, \epsilon_{it}^B\right) = \sqrt{1-\beta^2}\sqrt{1-\rho_B^2} > 0$ implies that the specific idiosyncratic factor ϵ_{it}^B exhibits low values under borrower specific success. The signs of the remaining risk parameters can be derived analogously. The results are summarized in table 4.

Weight of the...	D	**URD**	**SRR**	**URR**
systematic factor	$\alpha > 0$	$\beta < 0$	$\gamma > 0$	$\delta > 0$
mutual systematic factor	$\theta_A > 0$	$\theta_B > 0$	$\theta_C > 0$	$\theta_D > 0$
mutual idiosyncratic factor	/	$\rho_B < 0$	$\rho_C > 0$	$\rho_D > 0$

Table 4: Signs of the dependence parameters

Moreover it should be noted that this constellation leads to high risk parameter specific idiosyncratic drivers $\epsilon_{it}^C, \epsilon_{it}^D$ and low ϵ_{it}^B during borrower specific success, which implies that in this situations the recovery rates (secured and unsecured) tend to be high and the usage rate is low.

4.2.2 Initial Values of the Dependence Parameters

On the basis of the deduced signs, we make hypothetical but realistic assumptions regarding the dependence parameters. Given a suitable set of data they could be estimated by the methods mentioned in section 3.2 instead. Referring to the asset correlation in Basel II, the weight of the systematic factor of the default driver is set to $\alpha = 0.24$. Following Jiménez et al. [2009], a moderate, but significant decrease of the URD is associated with a macroeconomic upturn. Therefore, we choose the weight of the systematic factor of URD to $\beta = -0.2$. Due to the lack of empirical data, we assume that the weights of the systematic factors of SRR and URR are equal to $\gamma = 0.2$ and $\delta = 0.1$. Since the URR represents the income of the insolvency estate, it depends more on jurisdiction than on the macroeconomic state. Therefore, we propose a smaller value for δ compared to the other risk parameters. We assume equality for the parameters of the mutual systematic factor $\theta_A = \theta_B = \theta_C = \theta_D = \theta$ and examine the cases $\theta \in \{0, 0.5, 0.7, 0.9, 1\}$.

For the basic setting we choose $\theta = 0.7$, which implies relatively strong impact of the mutual systematic factor and leads to a similar development of the four risk parameters. The weights of the mutual idiosyncratic factor are examined for the values $(\rho_B, \rho_C, \rho_D) \in \{(-0.1, 0.025, 0.1), (-0.2, 0.05, 0.2), (-0.3, 0.075, 0.3)\}$, whereas the medium setting denotes the initial parameters from the basic setting.

Table 5 summarizes the definitions of the model parameters and their initial values for our numerical analysis (Basic setting).

Weight of the ...	***D***	**URD**	**SRR**	**URR**
systematic factor	$\alpha = 0.24$	$\beta = -0.2$	$\gamma = 0.2$	$\delta = 0.01$
mutual systematic factor	$\theta_A = 0.7$	$\theta_B = 0.7$	$\theta_C = 0.7$	$\theta_D = 0.7$
mutual idiosyncratic factor	/	$\rho_B = -0.2$	$\rho_C = 0.05$	$\rho_D = 0.2$

Table 5: Initial values of the dependence parameters

4.3 Algorithm of the Monte Carlo Simulation

The following algorithm outlines the steps of our Monte Carlo simulation:

Algorithm 1 Simulation of portfolio loss distribution

For each scenario $t = 1, \ldots, T$:

Draw a realization of the systematic factors:
$\left(X_t,\, Z_t^A,\, Z_t^B,\, Z_t^C,\, Z_t^D\right)^t \sim N_5\left(\mathbf{0}_5, I_5\right)$, with $\mathbf{0}_5 = (0,0,0,0,0)^T$ and I_5 being the identity matrix of size 5.

For every borrower $i = 1, \ldots, N$:

Draw the borrower specific, idiosyncratic factors:
$\left(\epsilon_{it}^A,\, \epsilon_{it}^B,\, \epsilon_{it}^C,\, \epsilon_{it}^D\right)^t \sim N_4\left(\mathbf{0}_4, I_4\right)$.
Compute the default driver A_{it} according to table 1.

If $A_{it} \geq 0$:

Compute the residual drivers B_{it}, C_{it}, and D_{it} according to table 1.
Compute the risk parameters by quantile transformation under the above mentioned Beta distribution with $v = 0,25$:
$\text{URD}_i = F^{-1}_{\text{Beta}(p,q)}\left(\Phi\left(B_{it}\right)\right)$, with $p = \mathbb{E}(\text{URD}_i) \cdot \frac{1-v}{v}$ and $q = (1 - \mathbb{E}(\text{URD}_i)) \cdot \frac{1-v}{v}$
$\text{SRR}_i = F^{-1}_{\text{Beta}(p,q)}\left(\Phi\left(C_{it}\right)\right)$,with $p = \mathbb{E}(\text{SRR}_i) \cdot \frac{1-v}{v}$ and $q = (1 - \mathbb{E}(\text{SRR}_i)) \cdot \frac{1-v}{v}$
$\text{URR}_i = F^{-1}_{\text{Beta}(p,q)}\left(\Phi\left(D_{it}\right)\right)$,with $p = \mathbb{E}(\text{URR}_i) \cdot \frac{1-v}{v}$ and $q = (1 - \mathbb{E}(\text{URR}_i)) \cdot \frac{1-v}{v}$.

The portfolio loss of scenario t is given by:

$$L_t \quad = \sum_{i|A_{it} \geq 0} \max\left(\left(\text{Com}_i \cdot \text{URD}_i - \mathcal{C}_i \cdot \text{SRR}_i\right) \cdot \left(1 - \text{URR}_i\right), 0\right) \qquad (23)$$

Based on the realizations $L_1, \ldots, L_T$, the portfolio loss distribution can be estimated by the empirical distribution function. The maximum operator in (23) is necessary in

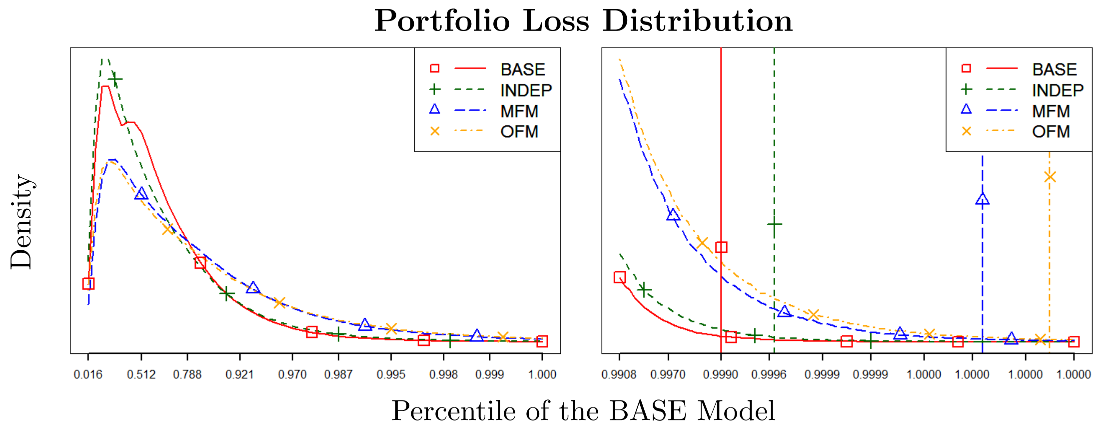

Figure 4: Density function of the loss distribution: base case (BASE), multi factor model under independence (INDEP), multi factor model (MFM), one factor model (OFM) as well as indicators for $\mathrm{VaR}_{0.999}$.

order to ensure that losses are not negative (i.e. profits), in case of $\mathrm{Com}_i \cdot \mathrm{URD}_i < \mathcal{C}_i \cdot \mathrm{SRR}_i$. One should keep in mind that this leads to a truncation of the simulated portfolio loss distribution.

4.4 Simulation Study

In order to better judge the results of our multi factor model, which takes into account the dependencies between all risk parameters, we compare this model with the following model specifications: In the base case (BASE) we use a model that simulates the defaults similar to the CreditMetrics$^{\mathrm{TM}}$ model, and treats the other risk parameters as deterministic given by their expected values. Additional, we investigate our multi factor model under independence (INDEP), by changing the weights of the mutual factors to $\theta_A = \theta_B = \theta_C = \theta_D = 0$ and $\rho_B = \rho_C = \rho_D = 0$, and the one factor model (OFM), i.e. $\theta_A = \theta_B = \theta_C = \theta_D = 1$. We perform a Monte Carlo simulation of each model and estimate the loss distribution based on 5.000.000 repetitions[5]. The resulting portfolio loss distributions are illustrated in figure 4 together with vertical lines, indicating the $\mathrm{VaR}_{0.999}$. The horizontal axis represents the loss percentile in case of the BASE model. Table 6 comprises the corresponding risk measures.

As expected, the risk figures increase together with dependence. Since the BASE and the INDEP model assume independence between the different risk parameters, their values are closer together compared to the OFM / MFM. The small deviation is a consequence of the truncation to prevent negative losses in equation (23) and the fact that, because of stochastic risk parameters, losses are no longer discrete. The second effect also implies that scenarios of high losses (within the tail) now can have different reasons, namely a high number of defaults, high utilization rates and/or low recovery

[5] Across 10 simulations of 5mio repetitions each, the maximum deviation for $\mathrm{EC}_{0.999}$ was below 0.4%. Therefore, we think that simulation errors are negligible.

Model	EL	SD	$\mathbf{VaR_{0.999}}$	$\mathbf{EC_{0.999}}$	$\mathbf{ES_{0.999}}$
BASE	100%	100%	100%	100%	100%
INDEP	101%	114%	117%	121%	121%
MFM	157%	178%	186%	193%	192%
OFM	163%	196%	209%	218%	215%

Table 6: Simulated expected loss, standard deviation, value at risk, economic capital and expected shortfall at 99.9%. The risk measures are denoted as multiples of the values in the base case.

$\boldsymbol{\theta}$	EL	SD	$\mathbf{VaR_{0.999}}$	$\mathbf{EC_{0.999}}$	$\mathbf{ES_{0.999}}$
0	96%	90%	87%	86%	87%
0.5	98%	95%	94%	93%	94%
0.7	100%	100%	100%	100%	100%
0.9	103%	106%	108%	109%	107%
1 (OFM)	104%	110%	112%	113%	112%

Table 7: Simulated expected loss, standard deviation, value at risk, economic capital and expected shortfall at 99.9% with varying θ, as well as in the OFM. The risk measures are denoted by multiples of the values in the original parametrization $\theta = 0.7$.

rates. The possibility of a combination of all these events (by incident because of missing dependence) leads to a small increase of risk figures.

In the other two models, the dependence between the risk parameters splits in two dimensions: the systematic and idiosyncratic component. Whereas both models share identical risks in the idiosyncratic dimension, the two models differ strongly from each other in the systematic dimension. In the OFM the systematic risks are completely determined by the same single factor, and are therefore co-monotone. However, in the MFM the systematic risks are reduced by diversification effects to some extend because every risk parameter is influenced by a specific systematic factor besides the mutual systematic factor. The risk measures of these two models exhibit strong growth, by up to more than 100% in the OFM compared to the base case. Since, a stronger dependence typically implicates a heavier right tail, this is a reasonable observation.

To illustrate the effect of the weight of the mutual systematic factor, we calculate the mentioned risk measures under varying specifications: besides the original specification ($\theta = \theta_A = \theta_B = \theta_C = \theta_D = 0.7$), we consider the values $\theta = 0$, 0.5, 0.9. The remaining parameters are unchanged. The OFM represents the limiting case, i.e. $\theta_A = \theta_B = \theta_C = \theta_D = 1$. Table 7 states that increasing weight of the mutual systematic factor has a significant impact on risk figures.

In particular, the economic capital at 99.9% increases significantly from 86% under systematic independence ($\theta_A = \theta_B = \theta_C = \theta_D = 0$) to 109% under $\theta_A = \theta_B = \theta_C = \theta_D = 0.9$. Since, the OFM represents the limiting case, also the risk figures of the MFM

Scenario	ρ_B	ρ_C	ρ_D	EL	SD	$\text{VaR}_{0.999}$	$\text{EC}_{0.999}$	$\text{ES}_{0.999}$
low	−0.1	0.025	0.1	84%	86%	87%	88%	87%
interm.	−0.2	0.05	0.2	100%	100%	100%	100%	100%
high	−0.3	0.075	0.3	117%	114%	113%	113%	113%

Table 8: Simulated expected loss, standard deviation, value at risk, economic capital and expected shortfall at 99.9% under varying idiosyncratic scenarios of the MFM. The risk measures are denoted as multiples of the values in the original parametrization (interm.).

converge to those of the OFM when θ goes to 1.

The idiosyncratic factor splits up into a risk parameter specific and a mutual component. We consider two additional idiosyncratic scenarios to investigate the influence of the weight of the mutual idiosyncratic factor. The previously described idiosyncratic weights $\rho_B = 0,2$, $\rho_C = 0,05$ and $\rho_D = -0,2$ represent an intermediate idiosyncratic dependence between the risk parameters. For a lower respectively higher dependence we set these parameters to 0.5 respectively 1.5 times the original values. The other parameters coincide with the original specification. Table 8 lists the resulting risk measures for each parameter setting. Increasing the weight of the mutual idiosyncratic factor is associated with an increase of credit risk, ceteris paribus. Relative to the economic capital at 99.9% of the MFM under intermediate idiosyncratic dependence, i.e. the original parametrization, the $\text{EC}_{0.999}$ increases from 88% under low to 113% under high idiosyncratic dependence.

For our portfolio, the mutual factors of the risk parameters in the systematic as well as in the idiosyncratic dimension have both similar strong influences. For example, the economic capital at 99.9% of the artificial portfolio increases by 31% under perfect dependence between the systematic factors compared to independent risk parameters.

5 Conclusion

A generalized credit portfolio framework of the CreditMetrics type is introduced which captures the dependence between the default indicator, the secured and the unsecured recovery rate as well as the utilization rate at default. Each of these variables is driven by an individual risk driver which in turn depends on common (global) and specific systematic factors as well as on idiosyncratic factors. Risk drivers and factors are connected through a linear structure and with individual (and in general unknown) weights. As there is empirical evidence for dependence between the four relevant variables and because observations for the utilization rate at default, the secured and unsecured recovery rates are available only in case of the default, we develop a multivariate extension of Heckman's (1979) selection model in order to estimate the unknown parameters within a Maximum-Likelihood framework. Finally, we present some indicative results on the (strong) sensitivity of the risk figures (expected loss, value at risk, expected shortfall) with respect to the factor structure for a hypothetical credit portfolio. For this portfolio, both specific systematic factors and idiosyncratic factors indicate strong influence on the risk figures. With this in mind, our results suggest that a correct specification of the underlying dependence structure (here: weights of the factor model) in a credit portfolio setting is essential. Because loss data collection is still an on-going process and there are no (sufficiently large) data sets to jointly estimate the unknown parameter, the parameterization of our multi-factor models was derived from different empirical studies which itself only cover a partial aspects of our model.

References

Edward I. Altman, Brooks Brady, Andrea Resti, and Andrea Sironi. The Link between Default and Recovery Rates: Theory, Empirical Evidence, and Implications. *Journal of Business*, 78(6):2203–2228, 2005.

Leif Andersen and Jakob Sidenius. Extensions to the Gaussian copula: Random Recovery and Random Factor Loadings. *Journal of Credit Risk*, 1 (1):29–70, 2005.

Elliot Asarnow and James Marker. Historical Performance of the U.S. Corporate Loan Market: 1988-1993. *Commercial Lending Review*, 10(2):13–32, 1995.

Benjamin Bade, Daniel Rösch, and Harald Scheule. Default and Recovery Risk Dependencies in a Simple Credit Risk Model. *European Financial Management*, 17(1):120–144, 2011.

BCBS. International convergence of capital measurement and capital standards: a revised framework, 2004. www.bis.org.

Roland Eller, Markus Heinrich, Rene Perrot, and Markus Reif. *Kompaktwissen Risikomanagement*. Gabler, Wiesbaden, 2010.

Jon Frye. Collateral Damage: A Source of Systematic Credit Risk. *Risk*, 13(4):91–94, 2000a.

Jon Frye. Depressing Recoveries. *Risk*, 13(11):108–111, 2000b.

William H. Greene. *Econometric Analysis*. Pearson, Boston, 2012.

James J. Heckman. Shadow Prices, Market Wages and Labor Supply. *Econometrica*, 42(4):679–694, 1974.

James J. Heckman. Sample Selection Bias as a Specification Error. *Econometrica*, 47(1):153–161, 1979.

Gabriel Jiménez, Jose A. Lopez, and Jesús Saurina. Empirical Analysis of Corporate Credit Lines. *The Review of Financial Studies*, 22(12):5069–5098, 2009.

Robert C. Merton. On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *The Journal of Finance*, 29(2):449–470, 1973.

Peter Miu and Bogie Ozdemir. Basel Requirement of Downturn LGD: Modeling and Estimating PD & LGD Correlations. *Journal of Credit Risk*, 2(2):43–68, 2006.

Michael Pykhtin. Recovery Rates: Unexpected Recovery Risk. *Risk*, 16(8): 74–78, 2003.

Daniel Rösch and Harald Scheule. Forecasting probabilities of default and loss rates given default in the presence of selection. *Journal of the Operational Research Society*, 65:393–407, 2014.

Ross Taplin, Huong Minh To, and Jarrad Hee. Modeling Exposure at Default, Credit Conversion Factors and the Basel II Accord. *Journal of Credit Risk*, 3(2):75–84, 2007.

Dirk Tasche. The Single Risk Factor Approach to Capital Charges in Case of Correlated Loss Given Default Rates. Working paper, Deutsche Bundesbank, 2004. URL http://arxiv.org/pdf/cond-mat/0402390v2.pdf.

Oldrich Alfons Vasicek. The Distribution of Loan Portfolio Value. *Risk*, 15(12):160–162, 2002.

Jiři Witzany. A Two Factor Model for PD and LGD Correlation. *Bulletin of the Czech Econometric Society*, 18(28):1–19, 2011.