The Empirical Relation between Credit Quality, Recoveries, and Correlation in a Simple Credit Risk Model

Daniel Rösch [a] Harald Scheule [b,1]

[a] *Institute of Banking & Finance, Faculty of Economics and Management, Leibniz Universität Hannover, Königsworther Platz 1, 30167 Hannover, Germany, Phone: +49-511-762-4668, Fax: +49-511-762-4670. mailto: Daniel.Roesch@finance.uni-hannover.de*

[b] *Department of Finance, Faculty of Economics and Commerce, University of Melbourne, Victoria 3010, Australia, Phone: +61-3-8344-9078, Fax : +61-3-8344-6914, mailto: hscheule@unimelb.edu.au*

Preprint submitted to Elsevier *26 November 2008*

The Empirical Relation between Credit Quality, Recoveries, and Correlation in a Simple Credit Risk Model

Abstract

The majority of industry credit portfolio risk models is based on various isolated modules for default probabilities and recoveries in the event of default. However, empirical evidence suggests a negative relation between both. This paper provides an approach for estimating this relation in a simple Merton type credit model framework with stochastic recoveries. An empirical analysis links bond recoveries with credit ratings and subordination levels and provides evidence for the relationship between credit quality, recovery rate and asset correlation. We find that the economic and regulatory capital may be severely underestimated by common regression models and deterministic recoveries.

Key words: Asset Value, Correlation, Credit Portfolio, Loss Given Default, Merton Model, Probability of Default, Recovery, Volatility

JEL classification: G20; G28; C51

[1] The authors gratefully acknowledge the financial support of the Melbourne Centre for Financial Studies and the Australian Prudential Regulation Authority.

1 Introduction

The current US-subprime mortgage crisis highlights that during economic downturns, likelihood and severity of multiple borrowers may deteriorate jointly. This is a particular concern in today's credit markets where via mechanisms such as collateralized debt obligations, credit portfolios rather than single name credits are exposed to losses within contractual boundaries. According to a recent study by the British Bankers' Association (2006), 54 percent of the global $20 trillion credit derivatives market consists of portfolio products. The evaluation of these requires the understanding of individual risk drivers as well as their dependence structure. Simplifying assumptions in existing models often lead to an underestimation of risks, particularly during economic downturns.

For instance, credit portfolio models aggregate parameters for the likelihood, severity and dependence structure underlying a credit portfolio and forecast the distribution of future credit losses. It is common practice to model these parameters independently and to introduce the dependence structure thereafter. Different contributions try to incorporate the dependence structure between the default events (compare e.g., Lucas 1995, Dietsch & Petey 2004, Hamerle et al. 2006, Frey & McNeil 2002, 2003, Rösch & Scheule 2004, Rösch 2005, McNeil & Wendin 2007) and between the default events and related recoveries (compare e.g., Frye 2000, Pykhtin 2003, Tasche 2004, Düllmann & Trapp 2005, Rösch & Scheule 2005).

Examples for well known credit portfolio models are CreditRisk+ (Credit Suisse Financial Products 1997), CreditMetrics (Gupton et al. 1997) and CreditPortfolioManager (Gupton et al. 1997). Newer applications in relation to collateralized debt obligations are VECTOR from Fitch rating agency (see Fitch Ratings 2006), CDOROM from Moody's rating agency (see Moody's 2006) and CDO Evaluator from Standard and Poor's rating agency (see Standard & Poor's 2005).

The literature on the estimation of severity parameters such as recoveries or losses given default includes a limited number of contributions (compare e.g., Carey 1998, Altman et al. 2006, Cantor & Varma 2005, Schuermann 2005, Acharya et al. 2007). Unfortunately, these contributions rely on the estimation of unconditional OLS regression models and do not take into account that recoveries can only be observed when a default event occurs. Three exceptions are Altman et al. (2001), Pykhtin (2003), and Hamerle et al. (2007). Altman et al. (2001) show the relation between default probability and expected recovery rate in a Merton model approach and the sensitivity w.r.t asset volatility. Thereby their model accounts for the fact that recoveries are only observable if a default occurs. Pykhtin (2003) also accounts for this mortality bias and derives closed-form expressions for the Expected Loss and the Value-at-Risk. However, these papers do not provide empirical solutions for parameter estimation. Pykhtin (2003) even acknowledges that "[The average LGD] is impossible to estimate". Hamerle et al. (2007) show by simulation that the estimates of isolated recovery models are biased but do not provide an analytical justification for that. Furthermore, their model is based on two systematic risk factors and thus does not allow a simple analytical solution for the Value-at-Risk in an asymptotic portfolio as for instance in Basel II, see Gordy (2003).

The present paper extends the previous literature, in particular work by Altman et al. (2001) and Pykhtin (2003) by empirically parameterizing a PD-recovery model. It includes observable idiosyncratic as well as unobservable systematic information. The following contributions are made:

(1) In a first-in-kind model, credit portfolio risk is explained by observed historic recovery rates. In contrast to market values for debt or equity, recovery rates are generally observable for past borrower defaults which is particularly useful for retail loans. The majority of commercial banks' loan portfolios consist of mortgage loans for which rich recovery histories are available.
(2) Contrary to most existing models multiple dimensions of credit portfolio risk are simultaneously modeled. Common credit risk parameters such as probabilities of default, expected losses given default or asset correlations may be derived. Models which estimate the parameters with separate models are subject to a mis-specification. This involves generally an underestimation of credit portfolio risk, which results from the omission of elements of the association structure (compare Altman et al. 2006). The most prominent example is the recent proposals by the Basel Committee on Banking Supervision (2006) which are also known as Basel II.
(3) While the mathematical model set-up builds upon Altman et al. (2001) and Pykhtin (2003), additional perspectives of the model are offered. It is shown how the model derived by Altman et al. (2001) can be embedded into an empirical, econometric Tobit framework for which a parameter estimation algorithm is developed. Due to the simplicity of the model, the estimation involves a limited number of parameters and is therefore subject to a low degree of model risk.
(4) Our approach extends the classical Tobit model to an econometric approach which also accounts correlations between the asset returns. Asset return correlations are among the most crucial parameters in current credit portfolio models.
(5) We provide a simple method for stressing credit risk parameters on economic downturns. The Basel II proposal requires banks to build forecasts for losses given default based on economic downturns without providing suggestions for these scenarios. In our model the losses given default de-

pend on a systematic factor which allows for analogous stress scenarios for losses given default as for the default probabilities.

(6) The models are applied to a database provided by the rating agency Moody's. The dynamic behavior of recovery implied asset return volatilities, correlations and their determinants are analyzed.

The rest of the paper proceeds as follows. Section 2 defines a structural default process based on an obligor's asset value and an empirical version of the model. Section 3 describes the data and presents the empirical results. The model is extended to asset return correlations in Section 4 and empirical results for this model are provided. In Section 5, the resulting Basel II capital is compared to commonly used regression models with deterministic recoveries. Closed-end formulas for the Expected Loss, Value-at-Risk and Downturn Loss Given Default are presented. Section 6 concludes with a summary and a discussion of the model and the findings.

2 The Basic Models

2.1 Asset Value Dynamics and Likelihood of a Bond Default

We follow the model outline in Altman et al. (2001) and derive the default probability and the recovery rate in an asset value model. Due to the early work by Merton (1974), let V denote the value of a firm's assets. V is assumed to follow a stochastic process which can be described by

$$dV = \delta \cdot V \cdot dt + \sigma \cdot V \cdot dW, \tag{1}$$

where $\delta \in \Re$ is an exogenous parameter and $\sigma > 0$ is an exogenous volatility parameter. dt represents the passage of time and dW is a Brownian motion. Then the change in the logarithmic firm value $\ln V$ between time 0 and T can be written as

$$S(T) = \ln V(T) - \ln v(0) = (\delta - 0.5\sigma^2)T + \sigma\sqrt{T} \cdot \varepsilon \tag{2}$$

where ε is a standard normally distributed random variable.

The firm is assumed to be financed by debt and equity. Debt consists of a zero coupon bond with nominal k and maturity T. At maturity the bondholders receive either a payment k or the value of the firm's assets, whichever is lower.

In the case $V(T) < k$ the bond issue defaults and bondholders receive a fraction of the notional which is also known as recovery. The default indicator is denoted by the random variable

$$D = \begin{cases} 1 & \text{borrower defaults} \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

Hence, the probability of default is

$$\begin{aligned} \lambda &= P(D = 1|v(0)) = P(V(T) < k|v(0)) = P(S(T) < \ln k - \ln v(0)) \\ &= P\left(R(T) < \frac{\ln \frac{k}{v(0)} - (\delta - 0.5\sigma^2) \cdot T}{\sigma \cdot \sqrt{T}}\right) \\ &= \Phi(-d(T)) \end{aligned} \tag{4}$$

where $\Phi(\cdot)$ is the standard normal cumulative density function, $R(T) = \frac{S(T)-(\delta-0.5\sigma^2)\cdot T}{\sigma\cdot\sqrt{T}}$ is the normalized asset return and $d(T) = -\frac{\ln \frac{k}{v(0)}-(\delta-0.5\sigma^2)\cdot T}{\sigma\cdot\sqrt{T}}$ is the normalized default threshold which is also known as Distance-to-Default.

2.2 Severity of a Bond Default

In this setting, the repayment ratio RR is the minimum of the asset value to debt ratio and one

$$RR = \min\{\frac{V(T)}{k}, 1\} \tag{5}$$

Defining the default point c by

$$c = \ln k - \ln v_0 \tag{6}$$

gives the transformation

$$\begin{aligned} \ln RR &= \min\{\ln V(T) - \ln k, 0\} \\ &= \min\{\ln V(T) - \ln v(0) - (\ln k - \ln v(0)), 0\} \\ &= \min\{S(T) - c, 0\} \end{aligned} \tag{7}$$

Equation (7) shows that the natural logarithm (log) of the repayment ratio is normally distributed but truncated by zero with non-zero values if a default event occurs, see Altman et al. (2001).

2.3 The Empirical Factor Model

The subscript i is introduced for the respective borrower and the number of borrowers is denoted by n. A time-horizon of one-year is considered. Thus, the transformed log-repayment ratio can be written as

$$\ln RR_i = \min\{S_i(1) - c_i, 0\} \tag{8}$$

$i = 1, ..., n$. This representation is known as a Tobit model which assumes that the observed variables Y_i, i.e. the log-repayment ratios, satisfy

$$Y_i = \min\{Y_i^*, 0\} \tag{9}$$

(compare Tobin 1958). Y^* is a latent variable generated by a classical regression model

$$Y_i^* = \boldsymbol{\beta}'\boldsymbol{x}_i + \sigma \cdot U_i \tag{10}$$

where $\boldsymbol{\beta}$ represents a vector of parameters, $\boldsymbol{x}_i$ a vector of covariates, which may include an intercept, and U_i a random error. Note that $y_i < 0$ implies an obligor default event. The errors are assumed to be independent and identically standard normally distributed.

The conditional density of the log-repayment ratio, i.e. the density of the log-recovery rate given default, is given by Bierens (2004):

$$h(y_i|Y_i < 0, \boldsymbol{x}_i) = \frac{\phi(-(y_i - \boldsymbol{\beta}'\boldsymbol{x}_i)/\sigma)}{\sigma \cdot (1 - \Phi(\boldsymbol{\beta}'\boldsymbol{x}_i/\sigma))} \tag{11}$$

for $y_i < 0$, where $\phi(\cdot)$ is the density function of the standard normal distribution. Then a closed-form expression for the conditional expectation of the log-recoveries Y_i given $\boldsymbol{x}_i$ and $Y_i < 0$ can be derived as

$$\begin{aligned}\mathbb{E}(Y_i|Y_i < 0, \boldsymbol{x}_i) &= \frac{1}{1-\Phi(\boldsymbol{\beta}'\boldsymbol{x}_i/\sigma)}\int_{-\infty}^{0} z f(z)\mathrm{d}z \\ &= \boldsymbol{\beta}'\boldsymbol{x}_i - \sigma\frac{\phi(\boldsymbol{\beta}'\boldsymbol{x}_i/\sigma)}{1-\Phi(\boldsymbol{\beta}'\boldsymbol{x}_i/\sigma)} \end{aligned} \tag{12}$$

where $f(\cdot)$ is the density of a normal distributed random variable with mean $\boldsymbol{\beta}'\boldsymbol{x}_i$ and variance σ^2. Note that the probability of default is

$$PD_i = P(D_i = 1|\boldsymbol{x}_i) = \Phi(-\boldsymbol{\beta}'\boldsymbol{x}_i/\sigma) \tag{13}$$

which implies that the default probability and the expectation of the log-recovery are negatively related and that the standardized linear predictor $\boldsymbol{\beta}'\boldsymbol{x}_i/\sigma$ equals the Distance-to-Default. Please note that PD_i relates to the Tobit model while λ relates to the original asset value model.

Figure 1 shows a graphical interpretation of the relation between the linear predictor $\boldsymbol{\beta}'\boldsymbol{x}_i$, the probability of default (PD), and the volatility σ. Equation (13) shows that the PD is a non-linear decreasing function of the linear predictor and a non-linear increasing (decreasing) function of the volatility for low (high) linear predictors.

[Figure 1 about here]

The conditional expectation of Y_i given $\boldsymbol{x}_i$ is

$$\mathbb{E}(Y_i|\boldsymbol{x}_i) = (\boldsymbol{\beta}'\boldsymbol{x}_i)(1 - \Phi(\boldsymbol{\beta}'\boldsymbol{x}_i/\sigma)) - \sigma\phi(\boldsymbol{\beta}'\boldsymbol{x}_i/\sigma) \tag{14}$$

Equations (12) and (14) have important consequences for the estimation of determinants for the recoveries using regression models. In both instances, the expectation of Y_i does not equal the linear predictor $\boldsymbol{\beta}'\boldsymbol{x}_i$. Thus, the estimates for $\boldsymbol{\beta}$ are biased and inconsistent if they are i) estimated using non-zero observations of the Y_i, or ii) by treating the values of Y_i which are zero as regular dependent variables as in common linear regression models. Note that this is the case in most recent contributions which involve the empirical estimation of recovery rates.

The variance of the conditional expectation of Y_i is given by

$$\mathbb{V}(Y_i|Y_i < 0, \boldsymbol{x}_i) = \sigma^2 - \sigma^2 \cdot \left(-\boldsymbol{\beta}'\boldsymbol{x}_i/\sigma + \frac{\phi(\boldsymbol{\beta}'\boldsymbol{x}_i/\sigma)}{1-\Phi(\boldsymbol{\beta}'\boldsymbol{x}_i/\sigma)} \right) \cdot \frac{\phi(\boldsymbol{\beta}'\boldsymbol{x}_i/\sigma)}{1-\Phi(\boldsymbol{\beta}'\boldsymbol{x}_i/\sigma)} \tag{15}$$

Finally, the expectation of the recovery rate given the firm's default is derived. First, we define the recovery rate given default as

$$RGD_i = \exp[Y_i^-] \tag{16}$$

that is, it is defined only if the borrower defaults, and the loss (rate) given default as $LGD_i = 1 - RGD_i$. Then the expected recovery rate given default is

$$\begin{aligned} ERGD_i = \mathbb{E}(RGD_i) = \mathbb{E}(RR_i|D_i = 1, \boldsymbol{x}_i) = \int_{-\infty}^{0} \exp(y_i) \cdot h(y_i|Y_i < 0, \boldsymbol{x}_i)\mathrm{d}y_i \\ = \int_{-\infty}^{0} \exp(y_i) \cdot \frac{\phi(-(y_i - \boldsymbol{\beta}'\boldsymbol{x}_i)/\sigma)}{\sigma \cdot (1-\Phi(\boldsymbol{\beta}'\boldsymbol{x}_i/\sigma))}\mathrm{d}y_i \\ = \frac{1}{1-\Phi(\boldsymbol{\beta}'\boldsymbol{x}_i/\sigma)} \cdot \exp(\boldsymbol{\beta}'\boldsymbol{x}_i + 0.5\sigma^2) \cdot \Phi\left(-\frac{\boldsymbol{\beta}'\boldsymbol{x}_i + \sigma^2}{\sigma}\right) \end{aligned} \tag{17}$$

The derivation of the third equation is given in the Appendix. Please note that this result is analogous to Equation (12) in Pykhtin (2003). The expected loss rate given default (ELGD) is then defined as

$$ELGD_i = 1 - \mathbb{E}(RR_i|D_i = 1, \boldsymbol{x}_i) = 1 - ERGD_i \tag{18}$$

Figure 2 shows the relation between PD, expected loss rate given default (ELGD), and the volatility σ. The relationship between PD and ELGD is monotone: ELGD increases with the PD. Moreover, the slope of the PD-ELGD-curve depends on the volatility resulting in an approximately linear relation for higher values of the volatility. In other words, the positive correlation between the likelihood and severity of credit risk is driven by the random asset value and therefore embedded in a causal model. Note that actual defaults and recoveries (or losses) given default are realizations of random variables (3) and (16) and will take on values different from their expectations shown in Figure 2. We will discuss their properties in more detail in section 5.

[Figure 2 about here]

2.4 Model Estimation

The Tobit model parameters are estimated conditional on default using the Maximum-Likelihood method. The likelihood that obligor i has not defaulted conditional on x_i is

$$1 - PD_i = \Phi(\boldsymbol{\beta}'\boldsymbol{x}_i/\sigma) \tag{19}$$

The likelihood of the log-recovery is

$$h(y_i|\boldsymbol{x}_i) \cdot (1 - \Phi(\boldsymbol{\beta}'\boldsymbol{x}_i/\sigma)) = \frac{\phi((y_i - \boldsymbol{\beta}'\boldsymbol{x}_i)/\sigma)}{\sigma} \tag{20}$$

and therefore the likelihood for an observed pattern of non-defaults and log-recoveries is

$$\mathfrak{L} = \prod_{i \in \{y_i = 0\}} (\Phi(\boldsymbol{\beta}'\boldsymbol{x}_i/\sigma)) \cdot \prod_{i \in \{y_i < 0\}} \left(\frac{\phi((y_i - \boldsymbol{\beta}'\boldsymbol{x}_i)/\sigma)}{\sigma} \right) \tag{21}$$

It may be more convenient to calculate the log-likelihood

$$\ell = \sum_{i \in \{y_i = 0\}} \ln(\Phi(\boldsymbol{\beta}'\boldsymbol{x}_i/\sigma) + \sum_{i \in \{y_i < 0\}} \ln \left(\frac{\phi((y_i - \boldsymbol{\beta}'\boldsymbol{x}_i)/\sigma)}{\sigma} \right) \tag{22}$$

which is then maximized with regard to the parameters $\boldsymbol{\beta}$ and σ. The estimates exist asymptotically, are consistent and asymptotically normally distributed.

3 Empirical Study

3.1 Data

The empirical analysis is based on recoveries provided by the rating agency Moody's. Moody's measures the recovery of a bond issue upon occurrence of a default event, i.e., if

- Interest and/or principal payments are missed or delayed,
- Chapter 11 or Chapter 7 bankruptcy is filed, or
- Distressed exchange such as a reduction of the financial obligation occurs.

In order to guarantee a homogeneous risk segment, the data set was restricted to regular US bond issues. The observation period includes the years 1982 to 2007. Secured bond issues were excluded from the analysis as their default and recovery characteristics may relate to the collateral value rather than the asset value of a firm and PDs and LGDs. This data set includes 446,287 observations with 1,293 default and recovery events. Moody's defines a recovery rate as the ratio of the price of defaulted debt obligations after 30 days of the occurrence of a default event and the par value.

Table 1 and Table 2 show the number of observations, default rate and mean recovery per year, rating class, industry and seniority/security level. The rating class IG comprises investment grade ratings (i.e., Aaa, Aa, A, Baa) and the rating class C comprises the rating categories Caa, Ca and C.

[**Table 1 about here**]

[**Table 2 about here**]

Figure 3 shows that the ratio of non-investment grade issues to total issues co-moves with the default rate which demonstrates the power of Moody's ratings to predict defaults.

[**Figure 3 about here**]

Generally speaking, default rates decrease and recoveries increase with improving credit quality. Two recessions of the US economy can be identified: the first one in 1991 during the First Gulf War and a second one in 2001 during the downturn in the internet industry and the terrorist attack in the US. This negative relationship between default and recovery rates has been shown in previous studies (see Section 1) and is displayed in Figure 4.

[**Figure 4 about here**]

Figure 5 and Figure 6 show histograms for the recoveries and recoveries which are transformed by the natural logarithm. The distribution of the log-recoveries is truncated at zero which may confirm the assumption of a Tobit model specification.

[**Figure 5 about here**]

[**Figure 6 about here**]

3.2 Market-wide Analysis

The base case model is estimated for all observations of the entire sample period without covariates. Table 3 shows the results of the parameter estimates in the first column which is labeled Model (1). From the first row it can be inferred that the constant (or mean transformed asset return) is 11.4551 and the volatility is 4.1525 as shown in the row labeled σ.

The standard errors are reported in parentheses in each row below the parameter estimates and both estimates are significantly different from zero. This results in a distance to default of 2.7586 (i.e., $11.4551 \div 4.1525$) and an average probability of default of 0.29 per cent (i.e., $\Phi(-2.7586)$. This estimate equals the average default rate of the observation period which is 0.29 per cent (i.e., 1,293 default events divided by 446,287 observations). The expected recovery from Equation (17) is 43.40 per cent. This estimate is also close to the average realized recovery rate from the sample which is 39.9 per cent.

Model (2) to Model (4) extend the base case Model (1) by including covariates. In Model (2), Moody's rating grades which were assigned to each issue at the beginning of a year are included as ex-ante measures for the credit quality of a borrower. Each rating grade is modeled by a dummy variable

$$x_{it}^{j} = \begin{cases} 1 & \text{issue } i \text{ has assigned rating grade } j \text{ at the beginning of year } t \\ 0 & \text{otherwise} \end{cases} \tag{23}$$

for grades j = BA,B,C.

The estimation results are reported in the second column of Table 3. Note that owing to the dummy encoding of the rating grades, grade IG is used as the reference category. Therefore, a borrower with grade IG at the beginning of a year has an estimated constant of 10.2240 (compared to the average of 11.4551 over all observations). The inclusion of rating information into the model reduces the volatility to 2.8097. This shows that credit ratings capture valuable information regarding the idiosyncratic error in the process of the asset returns.

Looking at the results for the other three rating grades it can be seen that all three effects are significantly different from zero indicating significant differences for the three grades. For instance, the constant for a grade Ba borrower is $10.2240 - 2.8013 = 7.4227$ yielding a lower distance to default, a higher PD and a lower expected recovery compared to a grade IG borrower. Similarly, the effects for the other grades can be interpreted where the highest default probability and lowest recovery is assigned to the riskiest grade C.

While the assessment of credit quality made by the rating agency should be an obvious indicator for the default probability and the expected recovery, another indicator should be the seniority. The database allows the differentiation between 'senior unsecured (SU)' and 'subordinated (Sub)' issues. Analogously, the seniority status is coded by a dummy variable

$$x_{i}^{Sub} = \begin{cases} 1 & \text{issue } i \text{ is subordinated} \\ 0 & \text{otherwise} \end{cases} \tag{24}$$

as the reference category SU is used. The results of Model (3) which includes the seniority status are also shown in Table 3. The constant for the reference category SU is 11.4395 and therefore higher than that of the average model, indicating a lower probability of default and a higher recovery rate for senior secured bonds.

Model (4) shows that the subordination does not add statistically significant explanatory power to the model after the credit quality is taken into account. This is plausible as Moody's rating categories reflect expected loss rates and not the likelihood of default.

[Table 3 about here]

3.3 Industry-specific Analysis

In a second step, the data set is split according to the Standard Industry Classification (SIC) code into the industries Commerce (SIC code between 50 and 59), Financial Institutions (FI; SIC code between 60 and 64), Manufacturing (SIC code between 20 and 39), Public Utility (PU; SIC code equal to 49), Services (SIC code between 40 and 48) and Others (SIC code below 19). Industries may constitute homogeneous risk segments in which companies are subject to the similar risk characteristics. Asset values and default thresholds may be similar for the same industry. The descriptive statistics in Table 2 show that in particular Financial Institutions and Public Utility firms have low default rates and high mean recovery rates. Table 4 summarizes the parameter estimates for Model (2) for the individual industries.

[Table 4 about here]

4 Extension of the Model to Asset Return Correlations

4.1 Factor Model

The framework which was presented so far incorporates the residual volatilities but does not take into account that the firms' asset returns may be cross-sectionally correlated. Correlations are an important input into modern credit portfolio risk models. Small changes of the correlation between asset returns may have a high impact on the portfolio loss distribution and related measures (compare Rösch & Scheule (2004) and Rösch & Scheule (2005) among others).

For the empirical estimation of asset correlations, several approaches exist. As long as an obligor's equity returns are observable one can use them as proxies for asset returns subject to a set of assumptions and calculate historical correlations between borrowers' return time series by a multi-factor model (compare Zeng & Zhang 2001). Unfortunately, equity returns are generally unavailable for borrowers. In this case 'implied' asset returns correlations may be estimated from historical default experiences. Here two alternatives exist: Firstly, default correlations can be estimated by non-parametric approaches and then transformed into asset correlations (compare Lucas 1995). Secondly, implied asset correlations can be estimated directly within a class of generalized mixed models by using eg Maximum-Likelihood (compare Frey & McNeil

2003, Rösch & Scheule 2005, McNeil & Wendin 2007).

The present paper basically follows the second way and decomposes the random error U_i of Equation (10) into

$$U_i = \omega \cdot F + \widetilde{\sigma} \cdot V_i \tag{25}$$

where F is a systematic error component which simultaneously affects all assets (which is also known as a systematic random effect), and V_i is an idiosyncratic error affecting only asset i, $i = 1, ..., n$. All errors are standard normally distributed and independent from each other. ω and $\widetilde{\sigma}$ are parameters which express the exposure to the systematic and idiosyncratic factors. Note that the total variance is $\mathbb{V}(U_i) = \sigma^2 = \omega^2 + \widetilde{\sigma}^2$. Thus, the correlation between two latent variables Y_i^* and Y_j^* of asset i and j is given by

$$\rho = \frac{\mathbb{C}(Y_i^*, Y_j^*)}{\sigma \cdot \sigma} = \frac{\omega^2}{\sigma^2} = \frac{\omega^2}{\omega^2 + \widetilde{\sigma}^2} \tag{26}$$

where $\mathbb{C}(\cdot)$ denotes the covariance. This parameter plays a crucial role in most commercial credit risk models as well as Basel II which will be discussed in Section 5.

The latent variable Y_i^* in the Tobit model then extends to

$$Y_i^* = \boldsymbol{\beta}' \boldsymbol{x}_i + \omega \cdot F + \widetilde{\sigma} \cdot V_i \tag{27}$$

F is an annual realization and ω can be estimated using the econometric specification

$$Y_{it}^* = \boldsymbol{\beta}' \boldsymbol{x}_{it} + \omega \cdot F_t + \widetilde{\sigma} \cdot V_{it} \tag{28}$$

where $i \in n_t$, $t = 1, ...T$. T is the number of time series observations available (e.g., the number of years) and n_t is the set of borrowers in period t. Given this notation the parameters can be estimated by the Maximum-Likelihood method as shown below.

4.2 Model Estimation

Consider a given realization of the systematic factor $F_t = f_t$. Conditional on f_t the Likelihood for each period is

$$\mathfrak{L}_t = \prod_{i\in\{y_{it}=0\}} (\Phi((\boldsymbol{\beta}'\boldsymbol{x}_{it}+\omega\cdot f_t)/\widetilde{\sigma}))\cdot \prod_{i\in\{y_{it}<0\}} \left(\frac{\phi((y_{it}-\boldsymbol{\beta}'\boldsymbol{x}_{it}-\omega\cdot f_t)/\widetilde{\sigma})}{\widetilde{\sigma}}\right) \quad (29)$$

Please note that f_t is not observable and that the expectation is calculated with respect to F_t

$$\mathbb{E}(\mathfrak{L}_t) = \int_{-\infty}^{\infty} \prod_{i\in\{y_{it}=0\}} (\Phi((\boldsymbol{\beta}'\boldsymbol{x}_{it}+\omega\cdot f_t)/\widetilde{\sigma})) \quad (30)$$

$$\cdot \prod_{i\in\{y_{it}<0\}} \left(\frac{\phi((y_{it}-\boldsymbol{\beta}'\boldsymbol{x}_{it}-\omega\cdot f_t)/\widetilde{\sigma})}{\widetilde{\sigma}}\right) \phi(f_t)\,\mathrm{d}f_t \quad (31)$$

Finally, using a time series of T observations, the Log-Likelihood is

$$\ell = \ln \mathfrak{L} = \ln \left(\prod_{t=1}^{T} \mathbb{E}(\mathfrak{L}_t)\right) = \sum_{t=1}^{T} \ln \mathbb{E}(\mathfrak{L}_t) \quad (32)$$

which is then maximized with regard to the parameters β, ω and $\widetilde{\sigma}$. This operation can be solved numerically using adaptive Gauss-Hermite-quadrature (compare Pinheiro & Bates 1995, Rabe-Hesketh et al. 2002). [2]

4.3 Empirical Results

Table 5 shows the estimation results for the entire data base in the first column and the industries in the remaining columns.

The model includes the rating factors which are comparable in relation to the parameters and significance of the models without asset correlation. For the overall database we can calculate the total volatility as $\sqrt{\omega^2+\widetilde{\sigma}^2} = \sqrt{1.0242^2+2.6215^2} = 2.8145$ which is very close to the volatility from the model without a systematic risk component. The asset correlation given in the last row is then calculated as $\rho = \frac{\omega^2}{\omega^2+\widetilde{\sigma}^2} = \frac{1.0242^2}{1.0242^2+2.6215^2} = 0.1324$. For the industry sectors we find large differences for volatilities and correlations. Correlations in Commerce, Manufacturing, Services and Others are similar to the overall correlation. The correlations for Financial Institutions and Public

[2] A simulation study was conducted to ensure the performance of the estimators. For space reasons the results are not reported here. Details are available upon request from the authors.

Utility are much larger. As a result, the overall industries model may reflect diversification benefits across industries. The high correlation in the financial institutions industry may be interpreted as a warning sign for the current financial markets and banking crisis.

[Table 5 about here]

4.4 A Note on PD and LGD Correlations

In the proposed model a mathematical relationship between PD and expected LGD is derived. Empirically, as evidence for association between these parameters sometimes the sample correlation between realized default rates and realized LGD (or recoveries respectively) is calculated. For our database and the whole sample the correlation between default rates and recovery rates is -0.34, see also Figure 4 which shows an obvious negative association. However, the mathematical relationship from the model on the one hand and the sample correlation on the other hand are different in nature as the latter refers to a dependence measure between random variables and the first is the relation between two parameters. In other words, the realized default rate and the recovery rate may or may not be correlated in the sample whatsoever the relationship between the underlying parameters is. To see this, we conduct a simple simulation exercise. We consider a simple artificial portfolio with 5,000 borrowers from which time series of defaults and recoveries are randomly sampled by the proposed Tobit type model. The time series length is $T = 20$ years and the parameters are set to $\boldsymbol{\beta}'\boldsymbol{x}_{it} = 11$ and $\sigma = 4$, respectively. For each sampled time series the sample correlation coefficient between default rates and recovery rates is computed. The sampling is then repeated 1,000 times and the empirical distribution of the correlation coefficients is calculated. Figure 7 shows the frequency distribution for 1,000 simulation trials. The distribution centers near zero and is highly dispersed indicating that empirically positive and negative correlations can occur in the underlying model even if the asset correlation is zero. In Figure 8 the distribution is generated when the asset correlation of the underlying process is 50%, i.e. $\omega = \widetilde{\sigma} = \sqrt{8}$. The distribution is shifted to the left and negative values of the correlation are much more likely. This is because the systematic risk factor affects both the default rates and the recovery rates. Hence, in a ’bad year’ the default rates are higher while the recoveries are lower et vice versa in a ’good year’. The exercise shows that the mathematical relation between the parameters and the statistical correlation between the random variables should be differently interpreted.

[Figure 7 about here]

[Figure 8 about here]

5 Implications for Portfolio Credit Risk

5.1 Measurement of Portfolio Credit Risk

Finally, we determine the economic and regulatory capital under the Basel II rules. Due to a controversy in literature, the behavior of the loss distributions for various ways of modeling the recovery rate is of particular interest. Please note that general and specific provisions by the financial institutions should be sufficient to cover the expected losses while the Tier I and Tier II capital should be sufficient to cover the difference between the 99.9th percentile of the future loss and the Expected Loss which is also known as the Credit-Value-at-Risk, (see e.g., C. Bluhm 2002).

Thus, the probability distribution of the future loss of the whole credit portfolio and risk figures derived thereof, such as the Expected Loss or the Value-at-Risk are of a central concern to financial institutions. This generally requires the forecast of the loss distribution for a future time period, e.g., one year. In the following, the time subscript is dropped for efficiency of exposition. We denote the exposure of loan i in the portfolio by a_i which is assumed to be known. Then, the total exposure of the portfolio is $a = \sum_i^n a_i$ and the proportion of loan exposure i in the entire portfolio is defined as $\eta_i = \frac{a_i}{a}$.

The random loss of borrower $i, i = 1, ..., n$ as a fraction of its total exposure is denoted by

$$L_i = (1 - RGD_i) \cdot D_i \tag{33}$$

where RGD_i is the recovery rate given default.

The expected loss of borrower i as a fraction of its total exposure can be calculated as

$$
\begin{aligned}
\mathbb{L}_i &= \mathbb{E}(L_i) = \mathbb{E}(D_i|\boldsymbol{x}_i) - \mathbb{E}(RGD_i \cdot D_i|\boldsymbol{x}_i) \\
&= \Phi(-\boldsymbol{\beta}'\boldsymbol{x}_i/\sigma) - \frac{1}{1-\Phi(\boldsymbol{\beta}'\boldsymbol{x}_i/\sigma)} \cdot \exp(\boldsymbol{\beta}'\boldsymbol{x}_i + 0.5\sigma^2) \cdot \Phi\left(-\frac{\boldsymbol{\beta}'\boldsymbol{x}_i + \sigma^2}{\sigma}\right) \cdot \Phi(-\boldsymbol{\beta}'\boldsymbol{x}_i/\sigma) \\
&= \Phi(-\boldsymbol{\beta}'\boldsymbol{x}_i/\sigma) - \exp(\boldsymbol{\beta}'\boldsymbol{x}_i + 0.5\sigma^2) \cdot \Phi\left(-\frac{\boldsymbol{\beta}'\boldsymbol{x}_i + \sigma^2}{\sigma}\right) \\
&= PD_i - ERGD_i \cdot PD_i \\
&= PD_i \cdot ELGD_i
\end{aligned} \tag{34}
$$

where the second line follows from the fact that the recovery is different from zero only if the borrower defaults and $PD_i = P(D_i = 1|\boldsymbol{x}_i)$ is the probability of default from Equation (13).

The loss rate of a portfolio of loans is the weighted average of the individual loan loss rates given by

$$
L = \sum_i^n \eta_i (1 - RR_i) \cdot D_i \tag{35}
$$

The expected portfolio loss is obtained as

$$
\begin{aligned}
\mathbb{L} &= \mathbb{E}(\sum_{i=1}^n \eta_i L_i) = \sum_{i=1}^n \eta_i \mathbb{E}(L_i) = \\
&= \sum_{i=1}^n \eta_i \cdot [PD_i - ERGD_i \cdot PD_i] \\
&= \sum_{i=1}^n \eta_i \cdot PD_i \cdot ELGD_i
\end{aligned} \tag{36}
$$

For the probability distribution of the portfolio loss and risk measures such as the Value-at-Risk the dependency structure of the loans is crucial. Generally speaking, the density of (35) can not be expressed analytically but can be obtained by Monte-Carlo simulation. Following Gordy (2003) and Pykhtin (2003) an analytical solution for the percentiles of the distribution can be given in the special case of an infinitely granular portfolio. First, the expected loss rate for borrower i is expressed conditional on the systematic risk factor. Analogously to (34) one obtains

$$
\begin{aligned}
\mathbb{L}_i(F) &= \mathbb{E}(L_i|F) = \mathbb{E}(D_i|\boldsymbol{x}_i, F) - \mathbb{E}(RR_i \cdot D_i|\boldsymbol{x}_i, F) \\
&= \Phi(-(\boldsymbol{\beta}'\boldsymbol{x}_i + \omega \cdot F)/\widetilde{\sigma}) \\
&- \frac{1}{1-\Phi((\boldsymbol{\beta}'\boldsymbol{x}_i + \omega \cdot F)/\widetilde{\sigma})} \cdot \exp(\boldsymbol{\beta}'\boldsymbol{x}_i + \omega \cdot F + 0.5\widetilde{\sigma}^2) \\
&\cdot \Phi\left(-\frac{\boldsymbol{\beta}'\boldsymbol{x}_i + \omega \cdot F + \widetilde{\sigma}^2}{\widetilde{\sigma}}\right) \cdot \Phi(-(\boldsymbol{\beta}'\boldsymbol{x}_i + \omega \cdot F)/\widetilde{\sigma}) \\
&= \Phi(-(\boldsymbol{\beta}'\boldsymbol{x}_i + \omega \cdot F)/\widetilde{\sigma}) - \exp(\boldsymbol{\beta}'\boldsymbol{x}_i + \omega \cdot F + 0.5\widetilde{\sigma}^2) \cdot \Phi\left(-\frac{\boldsymbol{\beta}'\boldsymbol{x}_i + \omega \cdot F + \widetilde{\sigma}^2}{\widetilde{\sigma}}\right) \\
&= CPD_i(F) - CERGD_i(F) \cdot CPD_i(F) \\
&= CPD_i(F) \cdot CELGD_i(F) \qquad (37)
\end{aligned}
$$

where

$$
CPD_i(F) = \Phi(-(\boldsymbol{\beta}'\boldsymbol{x}_i + \omega \cdot F)/\widetilde{\sigma}) \qquad (38)
$$

is the conditional default probability, while

$$
\begin{aligned}
CERGD_i(F) = &\frac{1}{1-\Phi((\boldsymbol{\beta}'\boldsymbol{x}_i + \omega \cdot F)/\widetilde{\sigma})} \cdot \exp(\boldsymbol{\beta}'\boldsymbol{x}_i + \omega \cdot F + 0.5\widetilde{\sigma}^2) \\
&\cdot \Phi\left(-\frac{\boldsymbol{\beta}'\boldsymbol{x}_i + \omega \cdot F + \widetilde{\sigma}^2}{\widetilde{\sigma}}\right) \qquad (39)
\end{aligned}
$$

and $CELGD_i(F) = 1 - CERGD_i(F)$ are the conditional expected recovery rate given default and expected loss given default given the systematic factor. According to Gordy (2003) and Pykhtin (2003) the random loss of a granular portfolio is given by

$$
L^{\infty} = \sum_{i}^{n} \eta_i \mathbb{L}_i(F) \qquad (40)
$$

and is therefore a monotonically increasing function of the systematic factor. Thus, the α-percentile of the future loss, referred to as Value-at-Risk, is obtained as

$$L^\alpha = \sum_i^n \eta_i \mathbb{L}_i(F = \Phi^{-1}(1-\alpha)) \tag{41}$$

for $0 < \alpha < 1$. Note that this expression reduces to the core of IRB Basel II formula after a simple reparameterization if the recovery is not modeled via the asset value model and is assumed to be deterministic instead. In (26) the asset correlation was defined as $\rho = \frac{\omega^2}{\sigma^2}$ with $\sigma^2 = \omega^2 + \widetilde{\sigma}^2$. Noting that $1 - \rho = \frac{\widetilde{\sigma}^2}{\sigma^2}$ and rewriting the conditional probability of default results in

$$\begin{aligned} CPD_i(F) &= \Phi\left(-(\boldsymbol{\beta}'\boldsymbol{x}_i + \omega \cdot F)/\widetilde{\sigma}\right) \\ &= \Phi\left(-\frac{\boldsymbol{\beta}'\boldsymbol{x}_i \cdot \sigma}{\widetilde{\sigma} \cdot \sigma} - \frac{\omega \cdot F \cdot \sigma}{\widetilde{\sigma} \cdot \sigma}\right) \\ &= \Phi\left(-\frac{\boldsymbol{\beta}'\boldsymbol{x}_i}{\sigma} \cdot \frac{\sigma}{\widetilde{\sigma}} - \frac{\omega \cdot F}{\sigma} \cdot \frac{\sigma}{\widetilde{\sigma}}\right) \\ &= \Phi\left(-\frac{\boldsymbol{\beta}'\boldsymbol{x}_i}{\sigma} \cdot \frac{1}{\sqrt{1-\rho}} - \sqrt{\rho} \cdot F \cdot \frac{1}{\sqrt{1-\rho}}\right) \\ &= \Phi\left(\frac{\Phi^{-1}(PD_i) - \sqrt{\rho} \cdot F}{\sqrt{1-\rho}}\right) \end{aligned} \tag{42}$$

which is the conditional default probability in the Basel II IRB approach in terms of asset correlation where the systematic factor is fixed to the 99.9th percentile of a standard normally distributed variable and the asset correlation is expressed as a function of the default probability.

Finally, the model allows for a straightforward definition of so-called 'Downturn Loss Given Default' for the Basel II model. While a downturn probability of default can be defined by the conditional default probability (42) a similar interpretation is possible for the recovery (or the loss given default) and the individual or portfolio loss rate. To see this, note that Equation (37), (39) and (41) depend only on the systematic factor. Therefore a 'downturn recovery' is defined as the conditional expected recovery given an adverse realization of the systematic factor according to (39)

$$
\begin{aligned}
CERGD_i(F = \Phi^{-1}(1-\alpha)) = & \frac{1}{1 - \Phi((\boldsymbol{\beta}'\boldsymbol{x}_i + \omega \cdot \Phi^{-1}(1-\alpha))/\tilde{\sigma})} \\
& \cdot \exp(\boldsymbol{\beta}'\boldsymbol{x}_i + \omega \cdot \Phi^{-1}(1-\alpha) + 0.5\tilde{\sigma}^2) \\
& \cdot \Phi\left(-\frac{\boldsymbol{\beta}'\boldsymbol{x}_i + \omega \cdot \Phi^{-1}(1-\alpha) + \tilde{\sigma}^2}{\tilde{\sigma}}\right)
\end{aligned} \tag{43}
$$

with a downturn loss given default given as $CELGD_i(F = \Phi^{-1}(1-\alpha)) = 1 - CERGD_i(F = \Phi^{-1}(1-\alpha))$.

In the granular portfolio the downturn loss is then given as in Equation (41) where α can be set to 0.999 as proposed by Basel II.

In summary, given the estimation of a single credit risk model all common credit risk measures may be calculated. This is shown exemplary for the random effects model for all industries from Section 4.3. Table 6 shows in the first panel the unstressed measures probability of default, loss given default and Expected Loss for different credit ratings categories. The second panel shows the stressed credit measures conditional probability of default, conditional expected loss given default and Value-at-Risk based on the 99.9th percentile of the random systematic risk factor.

[Table 6 about here]

5.2 Application: Basel II regulatory capital

Table 7 shows the stressed Basel II credit measures conditional probability of default, conditional expected loss given default and Value-at-Risk based on the 99.9th percentile of the random systematic risk factor and pre-specified asset correlations as proposed by Basel Committee on Banking Supervision (2006).

[Table 7 about here]

The regulatory capital is derived from the Credit-Value-at-Risk and may be based on a Constant LGD Model (i.e. ELGD) or a Stochastic LGD Model (i.e. CELGD). It can be seen that the resulting regulatory capital may be underestimated by as much as 23 per cent (rating category C) if the Constant LGD Model is assumed.

6 Discussion

The current US sub-prime mortgage crisis suggest that the likelihood and severity of the risk in credit portfolios may have been underestimated. The industry fundamentally changed its risk measurement and management approaches in recent years by a set of isolated modules often provided by various external vendors resulting in independent and constant recovery rates. An empirical study showed that regulatory capital based on a constant recovery assumption may underestimated by as much as 23 per cent.

The contributions to literature and credit portfolio risk measurement and management are that all relevant credit risk parameters are modeled in one consistent and unbiased framework. This framework is regression based and requires the observation of past recoveries or losses but no market prices. A causal relationship between credit quality, recovery rate, volatility, and correlation is established. In particular, the model accounts for i) stochastic recoveries ii) dependencies between the PDs and recoveries, and iii) the fact that recoveries can only be observed after a default event which renders estimates from common regression models biased.

An empirical analysis linked bond recoveries with credit ratings and subordination levels. This approach allows financial institutions to have a consistent approach across different credit risk measures used to derive provisions, economic and regulatory capital. Other applications may exist.

The presented model focuses on unsecured regular exposures. Credit products may occasionally be more complicated and future research may extend the presented framework by including macroeconomic risk factors or an additional process for the collateral value. In addition, the empirical study may be extended to retail loans such as mortgage, credit card and auto loans.

References

Acharya, V. V., Bharath, S. T. & Srinivasan, A. (2007), 'Does industry-wide distress affect defaulted firms? - evidence from creditor recoveries', *Journal of Financial Economics* **85**, 787–821.

Altman, E., Brady, B., Resti, A. & Sironi, A. (2006), 'The link between default and recovery rates: Theory, empirical evidence and implications', *Journal of Business* **78**, 2203–2227.

Altman, E., Resti, A. & Sironi, A. (2001), 'Default recovery rates in credit risk modeling: A review of the literature and empirical evidence', *Working Paper* .

Basel Committee on Banking Supervision (2006), 'International convergence of capital measurement and capital standards: A revised framework, comprehensive version'.

Bierens, H. (2004), 'The Tobit model', *Working Paper* .

British Bankers' Association (2006), 'BBA Credit Derivatives Report 2006'.

C. Bluhm, L. Overbeck, C. W. (2002), *An Introduction to Credit Risk Modeling*, Chapman & Hall/CRC.

Cantor, R. & Varma, P. (2005), 'Determinants of recovery rates on defaulted bonds and loans for north american corporate issuers: 1983-2003', *Journal of Fixed Income* **14**, 29–44.

Carey, M. (1998), 'Credit risk in private debt portfolios', *Journal of Finance* **53**, 1363–1387.

Credit Suisse Financial Products (1997), 'Creditrisk+ - a credit risk management framework'.

Dietsch, M. & Petey, J. (2004), 'Should SME exposures be treated as retail or corporate expo-sures? a comparative analysis of default probabilities and asset correlations in french and german smes', *Journal of Banking and Finance* **28**, 773–788.

Düllmann, K. & Trapp, M. (2005), 'Systematic risk in recovery rates- an empirical analysis of U.S. corporate credit exposures', *Edward I. Altman, Andrea Resti and Andrea Sironi (eds.) Recovery Risk: The Next Challenge in Credit Risk Management* .

Fitch Ratings (2006), 'Exposure draft: Introducing the Fitch VECTOR Default Model Version 3.0'.

Frey, R. & McNeil, A. (2002), 'VaR and Expected Shortfall in portfolios of dependent credit risks: Conceptual and practical insights,', *Journal of Banking and Finance* **26**, 1317–1334.

Frey, R. & McNeil, A. (2003), 'Dependent defaults in models of portfolio credit risk', *Journal of Risk* **6**, 59–92.

Frye, J. (2000), 'Collateral damage', *Risk* **13**, 91–94.

Gordy, M. (2003), 'A risk-factor model foundation for ratings-based bank capital rules', *Journal of Financial Intermediation* **12**, 199–232.

Gupton, G., Finger, C. & Bhatia, M. (1997), *CreditMetrics Technical Document*, J.P. Morgan & Co., New York.

Hamerle, A., Knapp, M. & Wildenauer, N. (2007), 'Default and recovery correlations - a dynamic econoetric approach', *Risk* **20**, 100–105.

Hamerle, A., Liebig, T. & Scheule, H. (2006), 'Forecasting credit event frequency - empirical evidence for West German firms', *Journal of Risk* **9**, 75–98.

Lucas, D. (1995), 'Default correlation and credit analysis', *Journal of Fixed Income* **4**, 76–87.

McNeil, A. & Wendin, J. (2007), 'Bayesian inference for generalized linear mixed models of portfolio credit risk', *Journal of Empirical Finance* **14**, 131–149.

Merton, R. C. (1974), 'On the pricing of corporate debt: The risk structure of interest rates', *Journal of Finance* **29**, 449–470.

Moody's (2006), 'CDOROM v2.3 user guide'.

Pinheiro, J. & Bates, D. (1995), 'Approximations to the log-likelihood function in the nonlinear mixed-effects model', *Journal of Computational and Graphical Statistics* **4**, 12–35.

Pykhtin, M. (2003), 'Unexpected recovery risk', *Risk* **16**(8), 74–78.

Rabe-Hesketh, S., Skrondal, A. & Pickles, A. (2002), 'Estimation of generalized linear mixed models', *Stata Journal* **1**, 1–21.

Rösch, D. (2005), 'An empirical comparison of default risk forecasts from alternative credit rating philosophies', *International Journal of Forecasting* **25**(1), 37–51.

Rösch, D. & Scheule, H. (2004), 'Forecasting retail portfolio credit risk', *Journal of Risk Finance* **5**(2), 16–32.

Rösch, D. & Scheule, H. (2005), 'A multi-factor approach for systematic default and recovery risk', *Journal of Fixed Income* **15**(2), 63–75.

Schuermann, T. (2005), 'What do we know about loss given default?', *Edward I. Altman, Andrea Resti and Andrea Sironi (eds.) Recovery Risk: The Next Challenge in Credit Risk Management* .

Standard & Poor's (2005), 'CDO Evaluator Version 3.0: Technical Document'.

Tasche, D. (2004), 'The single risk factor approach to capital charges in case of correlated loss given default rates', *www.defaultrisk.com* .

Tobin, J. (1958), 'Estimation of relationships for limited dependent variables', *Econometrica* **26**.

Zeng, B. & Zhang, J. (2001), 'An empirical assessment of asset correlation models', *Working Paper Moody's KMV Corporation* .

Appendix: derivation of the Expected Recovery Rate Given Default

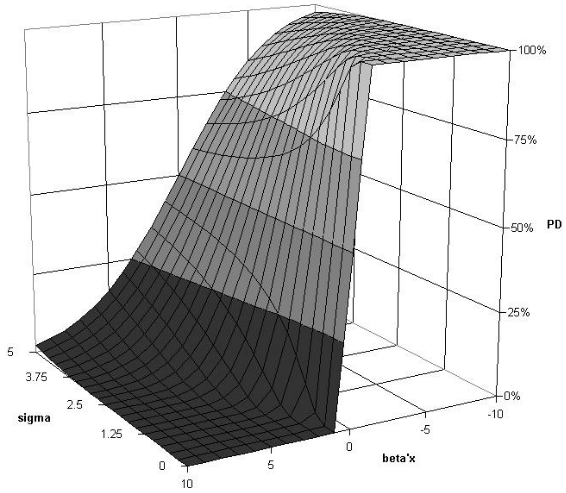
$$ERGD_i = \mathbb{E}(RR_i|D_i = 1, \boldsymbol{x}_i) = \frac{1}{1 - \Phi(\boldsymbol{\beta}'\boldsymbol{x}_i/\sigma)} \cdot \exp(\boldsymbol{\beta}'\boldsymbol{x}_i + 0.5\sigma^2) \cdot \Phi\left(-\frac{\boldsymbol{\beta}'\boldsymbol{x}_i + \sigma^2}{\sigma}\right)$$

Substitute $\mu_i = \boldsymbol{\beta}'\boldsymbol{x}_i$ and $PD_i = 1 - \Phi(\mu_i/\sigma)$ and write

$$\begin{aligned}
ERGD_i &= \int_{-\infty}^{0} \exp(y_i) \cdot h(y_i|Y_i < 0, \boldsymbol{x}_i)\mathrm{d}y_i \\
&= \int_{-\infty}^{0} \exp(y_i) \cdot \frac{\phi(-(y_i - \mu_i)/\sigma)}{\sigma \cdot (1 - \Phi(\mu_i/\sigma))}\mathrm{d}y_i \\
&= \frac{1}{\sigma \cdot PD_i} \int_{-\infty}^{0} \exp(y_i) \cdot \frac{1}{\sqrt{2\pi}} \cdot \exp\left(\frac{-(-y_i + \mu_i)^2}{2\sigma^2}\right)\mathrm{d}y_i \\
&= \frac{1}{\sigma \cdot PD_i} \cdot \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{0} \exp\left(y_i - \frac{\mu_i^2 - 2y_i\mu_i + y_i^2}{2\sigma^2}\right)\mathrm{d}y_i \\
&= \frac{1}{\sigma \cdot PD_i} \cdot \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{0} \exp\left(\frac{2\sigma^2 y_i - \mu_i^2 + 2y_i\mu_i - y_i^2}{2\sigma^2}\right)\mathrm{d}y_i \\
&= \frac{1}{\sigma \cdot PD_i} \cdot \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{0} \exp\left(\frac{-y_i^2 + 2y_i(\mu_i + \sigma^2) - (\mu_i + \sigma^2)^2 - \mu_i + (\mu_i + \sigma^2)^2}{2\sigma^2}\right)\mathrm{d}y_i \\
&= \frac{1}{\sigma \cdot PD_i} \cdot \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{0} \exp\left(\frac{-(y_i - (\mu_i + \sigma^2))^2 + 2\mu_i\sigma^2 + \sigma^4}{2\sigma^2}\right)\mathrm{d}y_i \\
&= \frac{1}{\sigma \cdot PD_i} \cdot \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{0} \exp\left(\frac{-(y_i - (\mu_i + \sigma^2))^2}{2\sigma^2}\right) \cdot \exp(\mu_i + 0.5\sigma^2)\mathrm{d}y_i \\
&= \frac{1}{PD_i} \cdot \exp(\mu_i + 0.5\sigma^2) \cdot \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(y_i - (\mu_i + \sigma^2))^2}{2\sigma^2}\right)\mathrm{d}y_i \\
&= \frac{1}{PD_i} \cdot \exp(\mu_i + 0.5\sigma^2) \cdot \Phi\left(-\frac{\mu_i + \sigma^2}{\sigma}\right)
\end{aligned}$$

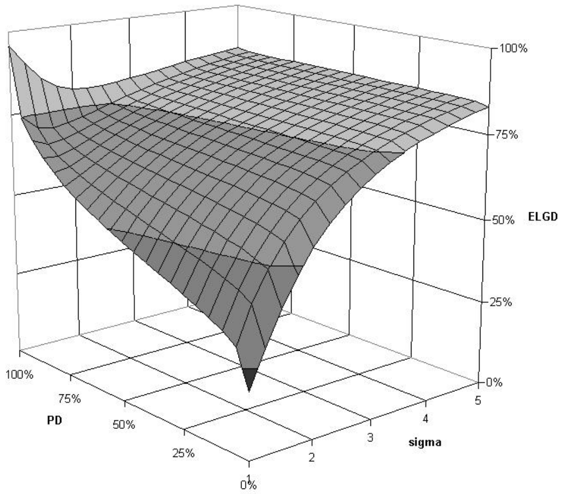
Figures

Fig. 1. Relation between linear predictor $\beta' x$, volatility σ, and probability of default (PD)



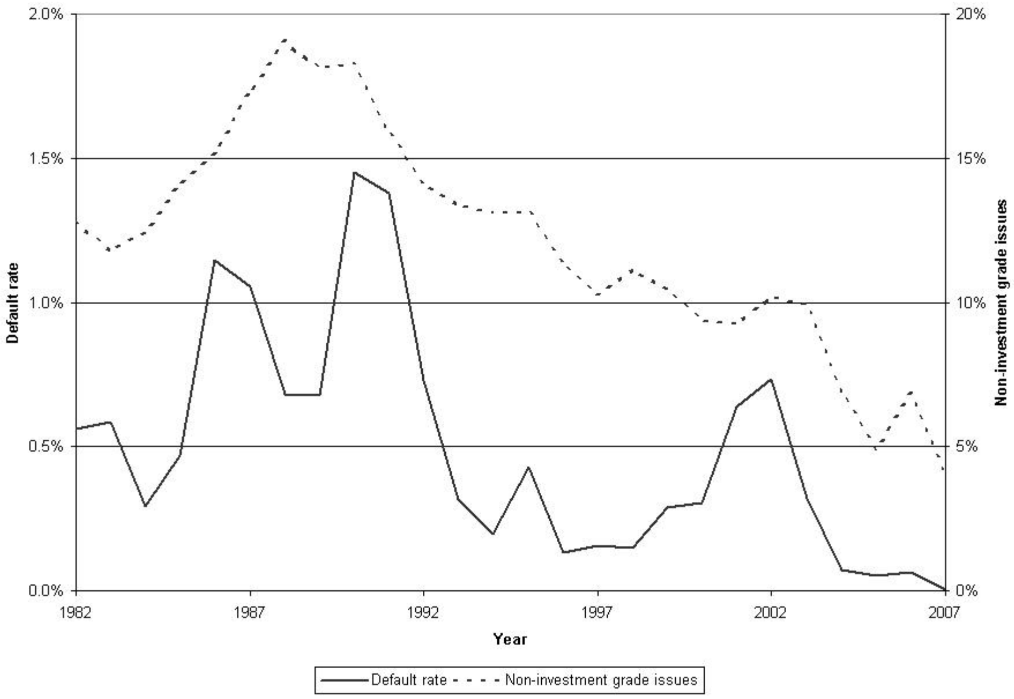
Notes: Probabilities of default are calculated based on σ and $\beta' x$ according to Equation (13). For high σ, the relationship between $\beta' x$ and PD is linear and for low σ, a firm defaults with a high likelihood (i.e., the PD is high) if $\beta' x < 0$ and a firm does not default with a high likelihood (i.e., the PD is low) if $\beta' x > 0$.

Fig. 2. Relation between probability of default (PD), expected loss given default (ELGD), and volatility σ



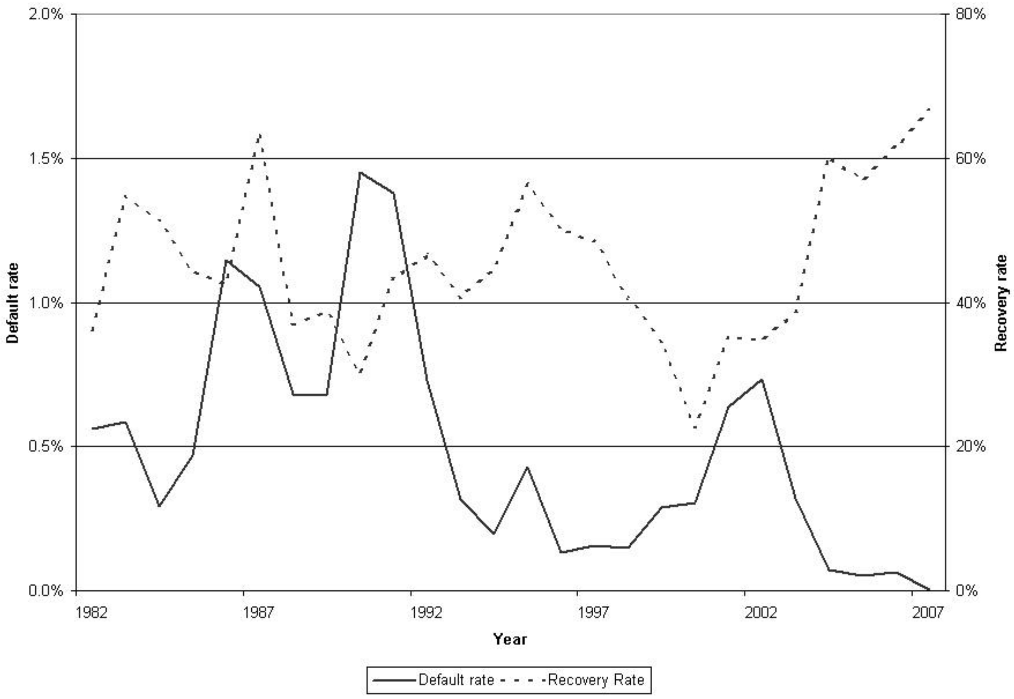
Notes: ELGD is calculated based on PD and σ according to Equation (17) and Equation (18).

Fig. 3. Default rate and non-investment grade rate



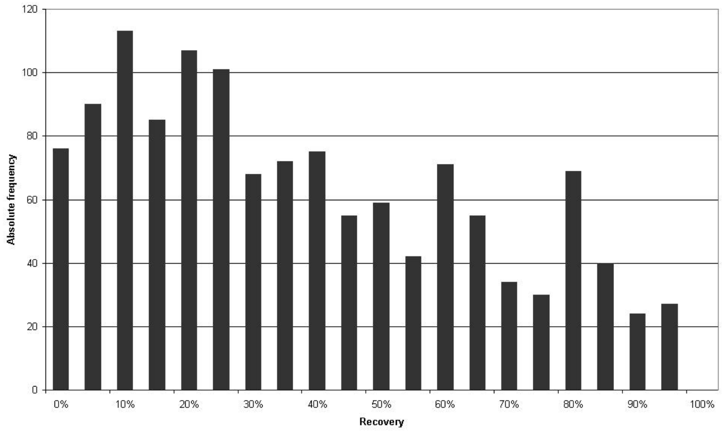
Notes: Default rate is the ratio between the number of defaulted issues and the total number of issues. The non-investment grade rate is the number of non-investment grade issues to the total number of issues.

Fig. 4. Default rates for all issues and recovery rates for all issues



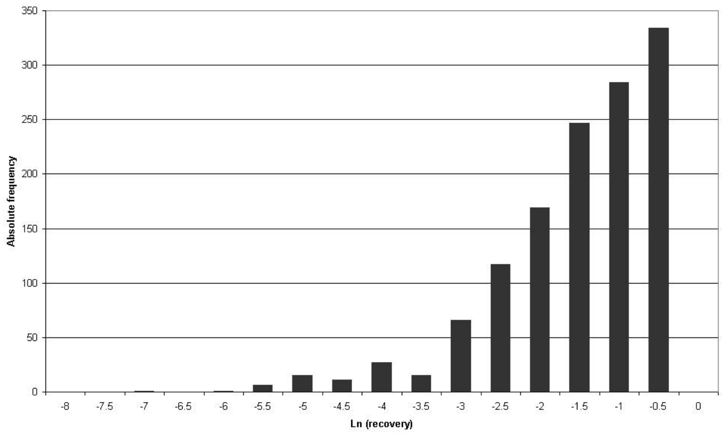
Notes: Default rate is the ratio between the number of defaulted issuers and the total number of issuers. Recovery rate is the ratio of the price of defaulted debt obligations after 30 days of the occurrence of a default event and the par value.

Fig. 5. Absolute frequencies for recoveries



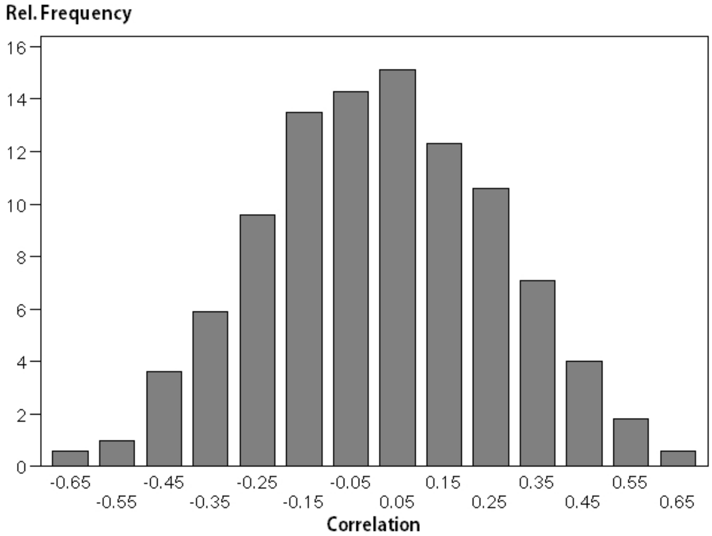
Notes: Recovery rate is the ratio of the price of defaulted debt obligations after 30 days of the occurrence of a default event and the par value.

Fig. 6. Absolute frequencies for log-recoveries



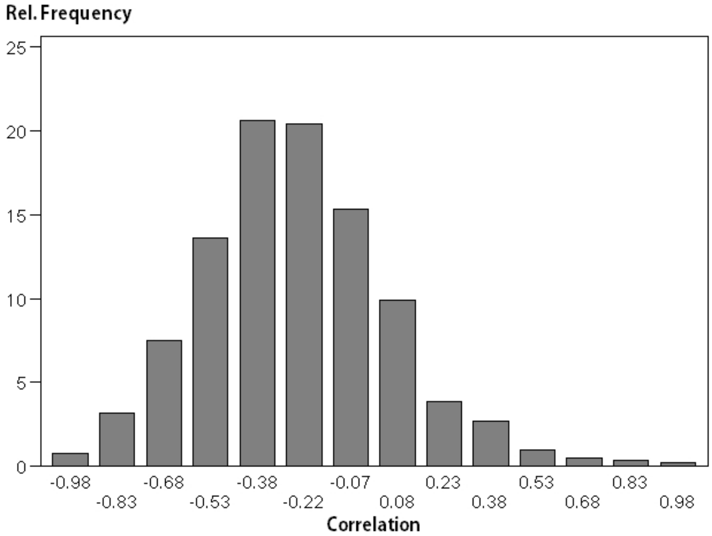
Notes: Recovery rate is the ratio of the price of defaulted debt obligations after 30 days of the occurrence of a default event and the par value.

Fig. 7. Simulated Correlations between default rates and recovery rates



Notes: Artificial portfolio consisting of 5,000 homogenous borrowers. Time series of $T = 20$ years for defaults and recoveries are randomly sampled by the Tobit type model and correlation coefficient between realized default rate and recovery rate is computed. Sampling is repeated 5,000 times. Parameters are $\boldsymbol{\beta}'\boldsymbol{x}_{it} = 11$ and $\sigma = 4$.

Fig. 8. Simulated Correlations between default rates and recovery rates



Notes: Artificial portfolio consisting of 5,000 homogenous borrowers. Time series of $T = 20$ years for defaults and recoveries are randomly sampled by the Tobit type model and correlation coefficient between realized default rate and recovery rate is computed. Sampling is repeated 5,000 times. Parameters are $\boldsymbol{\beta}'\boldsymbol{x}_{it} = 11$ and $\omega = \widetilde{\sigma} = \sqrt{8}$.

Tables

Table 1
Number of observations, default rates and mean recoveries per year

Year	Total observations	Default rate	Mean recovery
1982	2,491	0.6%	36.1%
1983	2,908	0.6%	54.7%
1984	3,079	0.3%	51.2%
1985	3,420	0.5%	44.5%
1986	4,183	1.1%	42.5%
1987	4,749	1.1%	63.5%
1988	4,996	0.7%	36.8%
1989	5,474	0.7%	38.6%
1990	5,865	1.4%	30.1%
1991	5,871	1.4%	43.2%
1992	5,880	0.7%	46.5%
1993	6,030	0.3%	40.7%
1994	6,645	0.2%	44.5%
1995	7,730	0.4%	56.2%
1996	9,694	0.1%	50.1%
1997	14,223	0.2%	48.3%
1998	19,650	0.1%	40.6%
1999	25,606	0.3%	34.7%
2000	29,405	0.3%	22.5%
2001	29,586	0.6%	35.1%
2002	27,113	0.7%	34.6%
2003	27,595	0.3%	38.7%
2004	37,622	0.1%	60.2%
2005	48,741	0.1%	57.0%
2006	55,246	0.1%	61.5%
2007	52,485	0.0%	66.5%
Sum/ Average	446,287	0.3%	39.9%

Notes: Default rate is the ratio between the number of defaulted issuers and the total number of issuers. Recovery rate is the ratio of the price of defaulted debt obligations after 30 days of the occurrence of a default event and the par value.

Table 2
Number of observations, default rates and mean recoveries per rating class, industry and seniority/security level

	Total observations	Default rate	Mean recovery
Rating class			
IG	406,497	0.0%	55.8%
Ba	16,672	0.4%	54.8%
B	18,753	2.4%	37.5%
Caa-C	4,365	16.5%	38.7%
Industry			
Commerce	10,288	1.1%	33.5%
FI	301,942	0.0%	60.5%
Manufacturing	37,749	0.7%	34.6%
PU	19,107	0.2%	64.8%
Services	45,925	0.8%	35.6%
Others	31,276	1.3%	40.0%
Seniority/Security class			
Senior unsecured	412,928	0.2%	44.0%
Subordinated	33,359	1.6%	34.0%

Notes: Default rate is the ratio between the number of defaulted issuers and the total number of issuers. Recovery rate is the ratio of the price of defaulted debt obligations after 30 days of the occurrence of a default event and the par value.

Table 3
Parameter estimates for the Tobit models

	Model (1)	Model (2)	Model (3)	Model (4)
const	11.4551***	10.2240***	11.4395***	10.2239***
	(0.3036)	(0.2706)	(0.3015)	(0.2705)
Rating BA		-2.8013***		-2.8016***
		(0.1642)		(0.1645)
Rating B		-4.7005***		-4.7017***
		(0.1575)		(0.1611)
Rating C		-7.4386***		-7.4398***
		(0.2074)		(0.2098)
Sub			-2.9703***	0.0029
			(0.1096)	(0.0783)
σ	4.1525***	2.8097***	3.9423***	2.8097***
	(0.1079)	(0.0691)	(0.1018)	(0.0691)

Notes : Table shows the results of Tobit models for the logarithm of the recovery rate with rating grades and seniority status as explanatory variables; standard deviations are in parentheses; ***indicates significance at the 1%-level, **indicates significance at the 5%-level, *indicates significance at the 10%-level.

Table 4
Parameter estimates for the Tobit Model (2) for different industries

	Commerce	FI	Manufacturing	PU	Services	Others
const	12.1580***	6.915***	10.7076***	4.7703***	10.8667***	9.1466***
	(1.3377)	(0.5200)	(0.6548)	(0.6759)	(0.5840)	(0.5797)
Rating BA	-3.4122***	-1.7607***	-1.7115***		-2.7238***	-3.6083***
	(1.0691)	(0.2370)	(0.5252)		(0.3974)	(0.4736)
Rating B	-5.6839***	-2.4715***	-4.3839***	-0.1331	-4.8920***	-4.7396***
	(1.0107)	(0.2594)	(0.3873)	(0.3350)	(0.3655)	(0.4697)
Rating C	-8.4791***	-5.6007***	-7.2724***	-2.7397***	-7.6831***	-6.9853***
	(1.1260)	(0.4454)	(0.4908)	(0.4143)	(0.4525)	(0.5173)
σ	3.3327***	1.6932***	3.1259***	1.5457***	3.1604***	2.4564***
	(0.2759)	(0.1330)	(0.1716)	(0.2136)	(0.1505)	(0.1072)

Notes : Table shows the results of Tobit models for the logarithm of the recovery rate with rating grades as explanatory variables; standard deviations are in parentheses; ***indicates significance at the 1%-level, **indicates significance at the 5%-level, *indicates significance at the 10%-level. The industry PU (Public Utility) did not experience any default event for the rating category Ba in the observation period. Hence, the respective parameter can not be determined empirically.

Table 5

Parameter estimates for the Random Effects Model for different industries

	All industries	Commerce	FI	Manufacturing	PU	Services	Others
const	9.7353***	12.1599***	6.7840***	10.3729***	5.7391***	10.6772***	9.0600***
	(0.3270)	(1.3388)	(0.6086)	(0.6691)	(0.9312)	(0.6377)	(0.6005)
Rating BA	-2.5604***	-3.5017***	-1.9502***	-1.5998***		-2.6079***	-3.5103***
	(0.1606)	(1.0255)	(0.2766)	(0.5154)		(0.3853)	(0.4714)
Rating B	-4.3857***	-5.5270***	-2.7743***	-4.1707***	-1.5229***	-4.6295***	-4.4934***
	(0.1514)	(0.9651)	(0.2887)	(0.3667)	(0.4652)	(0.3521)	(0.4651)
Rating C	-7.1286***	-8.3112***	-5.5659***	-7.0264***	-3.5482***	-7.4022***	-6.7618***
	(0.1986)	(1.0801)	(0.4489)	(0.4694)	(0.5514)	(0.4353)	(0.5103)
ω	1.0242***	1.3033***	1.1159***	1.1453***	1.4920***	1.3353***	0.8619
	(0.1565)	(0.3450)	(0.2637)	(0.2206)	(0.4411)	(0.2549)	(0.1585)
$\widetilde{\sigma}$	2.6215***	3.1191***	1.5858***	2.8518***	1.3549***	2.8567***	2.3213***
	(0.0560)	(0.2569)	(0.1245)	(0.1549)	(0.1843)	(0.1342)	(0.1007)
σ	2.8145	3.3874	1.9391	3.0732	2.0154	3.1334	2.4761
ρ	0.1324	0.1486	0.3312	0.1389	0.5480	0.1793	0.1212

Notes : Table shows the results of Tobit models for the logarithm of the recovery rate with rating grades as explanatory variables; standard deviations are in parentheses; ***indicates significance at the 1%-level, **indicates significance at the 5%-level, *indicates significance at the 10%-level. The industry PU (Public Utility) did not experience any default event for the rating category Ba in the observation period. Hence, the respective parameter can not be determined empirically.

Table 6
Summary of credit risk measures derived from Random Effects Model (6)

	Rating IG	**Rating BA**	**Rating B**	**Rating C**
PD	0.0003	0.0060	0.0336	0.1942
ELGD	0.4246	0.4858	0.5419	0.6396
Expected Loss	0.0001	0.0029	0.0182	0.1242
Empirical asset correlation	0.1397	0.1397	0.1397	0.1397
CPD	0.0070	0.0713	0.2334	0.6238
CELGD	0.4709	0.5567	0.6365	0.7700
Value-at-Risk (α=0.999)	0.0033	0.0397	0.1486	0.4803

Notes : PD is calculated according to Equation (13), ELGD is calculated according to Equation (18), Expected Loss is calculated according to Equation (34), CPD is calculated according to Equation (38), CELGD is calculated one minus ERGD which is calculated according to Equation (39), Value-at-Risk is calculated according to Equation (37).

Table 7
Summary of Basel II credit risk measures derived from Random Effects Model

	Rating IG	**Rating BA**	**Rating B**	**Rating C**
Basel asset correlation	0.2382	0.2091	0.1423	0.1200
CPD	0.0137	0.1077	0.2366	0.5877
Credit Value-at-Risk (ELGD)	0.0057	0.0494	0.1100	0.2517
Credit Value-at-Risk (CELGD)	0.0063	0.0571	0.1324	0.3284
Underestimation	0.1003	0.1338	0.1691	0.2335

Notes : The Basel CPD is calculated according to Equation (42), the regulatory capital, i.e., Credit Value-at-Risk is equals to the difference between the Value-at-Risk (i.e. the product of Basel CPD and loss given default) and the Expected Loss (i.e. the product of PD and ELGD).