HONG KONG INSTITUTE FOR MONETARY RESEARCH

THE EMPIRICAL RELATION BETWEEN CREDIT QUALITY, RECOVERY, AND CORRELATION

Daniel Rösch and Harald Scheule

HKIMR

The Empirical Relation between Credit Quality, Recovery, and Correlation

Daniel Rösch[1]

Leibniz Universität Hannover

and

Harald Scheule[2]

Hong Kong Institute for Monetary Research

The University of Melbourne

Abstract

The majority of industry credit portfolio risk models, as well as recent scientific results, are based on isolated modules for default probabilities and recoveries in the event of default. This paper shows that these common methods lead to various econometric drawbacks when the parameters are interpreted and aggregated for risk capital allocation and pricing purposes.

This paper provides a top down approach in which individual credit risk parameters are derived analytically from a single model. This model allows for a i) dynamic, ii) consistent, and iii) unbiased modeling of credit portfolio risks. An empirical analysis provides evidence for the inferred relationship between credit quality, recovery and correlation.

---

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

[2] Hong Kong Institute for Monetary Research and Department of Finance, Faculty of Economics and Commerce, The University of Melbourne, Victoria 3010, Australia, Phone: +61-3-8344-9078, Fax: +61-3-8344-6914, mailto: hscheule@unimelb.edu.au

1.   Introduction

Financial institutions were surprised that during the current financial crisis, individual risk parameters deteriorated jointly. As a result, credit portfolio losses dramatically exceeded the predictions provided by internal risk models. Measuring credit portfolio losses is also of great concern to fixed income investors. A large growth of investments in credit portfolios rather than single name credits has occurred via mechanisms such as collateralized debt obligations. 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 credit portfolio risks requires the understanding of individual risk drivers as well as their dependence structure.

Credit portfolio risk is measured by various parameters such as default probabilities, loss rates given default, exposures at default and dependence parameters such as correlations and more general copulas. It is common practice to model these parameters independently and to introduce the dependence structure thereafter. This practice is supported by the implementation of isolated models provided by external vendors.

Various authors address the default likelihood. Important contributions are Merton (1974), Leland (1994), Jarrow and Turnbull (1995), Longstaff and Schwartz (1995), Madan and Unal (1995), Leland and Toft (1996), Jarrow *et al.* (1997), Duffie and Singleton (1999), Shumway (2001), McNeil and Wendin (2007) and Duffie *et al.* (2007).

Credit ratings are often used as aggregated explanations of financial risk. Ratings measure the financial risk of corporate bond issuers, corporate bond issues and structured finance securities. Fundamental issues relating to the general extent to which credit rating changes convey new information has a rich pedigree that is the subject of ongoing academic debate and investigation. For example, Radelet and Sachs (1998) find that rating changes are pro-cyclical which would suggest that they provide only a limited amount of new information to the market. Ederington and Goh (1993), Dichev and Piotroski (2001) and Purda (2007) find that corporate credit rating downgrades do provide news to the market, although most studies find that rating upgrades do not. Jorion *et al.* (2005) show that after Regulation Fair Disclosure, the market impact of both downgrades and upgrades is significant and of greater magnitude compared to that observed in the pre-Regulation Fair Disclosure period. The relative roles of different CRAs have also been studied. For example, Miu and Ozdemir (2002) examine the effect of divergent Moody's and S&P ratings of banks.

Research on recoveries and loss rates given default are quite recent. Pan and Singleton (2008) derive the implicit risk structure of recoveries from sovereign CDS spreads. Contributions which focus on recoveries from defaulted issuers include Carey (1998), Pykhtin (2003). Acharya *et al.* (2007), Qi and Yang (2009)

and Grunert and Weber (2009) develop an empirical models for recoveries using explanatory co-variables which are economically motivated. Most of the empirical results in the recent literature using defaulted issuers are from common linear regression models which will be shown to be problematic later on in our paper.

Research on dependencies between risk parameters can be split into two categories. Firstly, dependencies between default events and asset value returns are modeled. Dietsch and Petey (2004) present a non-parametric approach and McNeil and Wendin (2007) apply a generalized mixed model approach using Maximum-Likelihood. Secondly, (compare Hu and Perraudin 2002, Tasche 2004, Altman *et al.* 2005) derive dependencies between default events and loss rates given default.

In relation to the current literature, various shortcomings can be identified. Firstly, default probabilities, recovery rates and correlations are often modeled as constant over time. Secondly, credit risk parameters are modeled independently and possibly inconsistently. The omission of elements of the association structure generally involves an underestimation of credit portfolio risk (compare Altman *et al.* 2005). The most prominent example is the recent proposals by the Basel Committee on Banking Supervision (2006) which are also known as Basel II. Thirdly, conditional parameters such as recoveries which are conditional upon the occurrence of default are modeled by (ordinary least square) regression models which do not take the conditionality into account and lead to a bias of the estimated parameters. One exception is Pykhtin (2003) who accounts for this mortality bias and derives closed-form expressions for the Expected Loss and the Value-at-Risk. However, the paper does not provide empirical solutions for parameter estimation. Interestingly, Pykhtin (2003) even acknowledges that in his paper ("[The average LGD] is impossible to estimate").

The present paper extends the previous literature by empirically parameterizing a PD-recovery model. It includes observable idiosyncratic as well as unobservable systematic information. The following contributions are made:

(1)   *Presentation of an original and relevant model framework:* In a first-in-kind model, credit portfolio risks are modeled by an econometric Tobit framework which involves a limited number of parameters (including correlations) and is therefore subject to a low degree of model risk. This econometric approach involves a top down approach which models the asset value returns of underlying borrowers and subsequently derives the risk parameters. This approach allows for a i) dynamic, ii) consistent, and iii) unbiased modeling of credit portfolio risks and avoids drawbacks of common linear regression models.

(2)   *Calibration to available data:* Credit portfolio risk is explained by observed historic recovery rates. In contrast to market values for debt (or spreads) 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 consists of mortgage loans for which rich recovery histories are available.

(3)   *Provision of empirical reference values:* 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. Using the unbiased estimation technique, the information content of credit ratings is tested for default probabilities as well as recoveries. Credit ratings have been highly criticized in the current financial crisis due to their failure to predict corporate credit default risk.

(4)   *Ability to stress-test and assess model risk:* The model captures the dependence on the business cycle of credit portfolio risk. Future credit risk losses can be based on stochastic and deterministic economic downturn scenarios. Due to the top down approach, all derived credit risk parameters are stressed consistently for such scenarios.

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 a model with deterministic recoveries and the best practice US industry approach. 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 Credit Default

We derive the default probability and the recovery rate in an asset value model. Let V denote the value of a firm's assets (compare Merton, 1974). 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. The change in the logarithmic firm value ln *V* between time 0 and *T* can be written as

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

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

$$\lambda = P(D = 1 | \upsilon(0)) = P(V(T) < k | \upsilon(0)) = P(S(T) < \ln k - \ln \upsilon(0))$$

$$= P\left(R(T) < \frac{\ln \dfrac{k}{\upsilon(0)} - (\delta - 0.5\sigma^2) \cdot T}{\sigma \cdot \sqrt{T}} \right)$$

$$= \mathbf{\Phi}\left(-d(T) \right) \tag{4}$$

where $\mathbf{\Phi}(\cdot)$ is the standard normal cumulative density function, $R(T) = \dfrac{S(T) - (\delta - 0.5\sigma^2) \cdot T}{\sigma \cdot \sqrt{T}}$ is the

normalized asset return and $d(T) = -\dfrac{\ln \dfrac{k}{\upsilon(0)} - (\delta - 0.5\sigma^2) \cdot T}{\sigma \cdot \sqrt{T}}$ 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

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

Defining the default point *c* by

4

$$c = \ln k - \ln \upsilon_0 \tag{6}$$

gives the transformation

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

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.

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

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

$i = 1, \ldots, n$. This representation assumes that the observed variables Y_i, i.e., the log-repayment ratios, satisfy

$$Y_i = \ln(RR_i) = \min\{Y_i^*, \, 0\} \tag{9}$$

(compare Tobit, 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

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

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

$$\mathbb{E}\left(Y_i \middle| Y_i < 0, \boldsymbol{x}_i\right) = \frac{1}{1 - \boldsymbol{\Phi}\left(\boldsymbol{\beta}'\boldsymbol{x}_i/\sigma\right)} \int_{-\infty}^{0} z f(z) dz$$

$$= \boldsymbol{\beta}'\boldsymbol{x}_i - \sigma \frac{\phi\left(\boldsymbol{\beta}'\boldsymbol{x}_i/\sigma\right)}{1 - \boldsymbol{\Phi}\left(\boldsymbol{\beta}'\boldsymbol{x}_i/\sigma\right)} \tag{12}$$

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

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

The standardized linear predictor $\boldsymbol{\beta}'\boldsymbol{x}_i/\sigma$ equals the Distance-to-Default. Please note that PD_i relates to the econometric model while λ relates to the theoretical 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.

The conditional expectation of Y_i given x_i is

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

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 empirically estimated recovery rates (compare Section 1).

The variance of the conditional expectation of Y_i is given by

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

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

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

that is, it is defined only if the borrower defaults. Then, the expected recovery rate given default is

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

The derivation of the third equation is given in the Appendix. The expected loss rate given default (ELGD) is then defined as

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

Figure 2 shows the relation between PD, expected loss rate given default (ELGD), and the volatility σ. Given the volatility, the relationship between PD and ELGD is monotone: ELGD increases with the PD. 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.

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

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

The likelihood of the log-recovery is

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

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

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

It may be more convenient to calculate the log-likelihood

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

which is then maximized with regard to the parameters $\boldsymbol{\beta}$ and σ. Maximum likelihood estimation implies that 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. This data set includes 446,287 observations with 1,293 default and recovery events. A recovery rate is defined 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.

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.

Generally speaking, default rates decrease and recoveries increase with improving credit quality. Two recessions of the US economy can be identified: a 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 is displayed in Figure 4.

Figure 5 and Figure 6 show histograms for the absolute recoveries and recoveries which are transformed by the natural logarithm. The distribution of the log-recoveries confirms the assumption of a truncated standard normal distribution of Y_i^* in Equation (10).

3.2 Market-Wide Analysis

The base case model is estimated for all observations of the 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., Φ (-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 ÷ 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. The lag enables the use of the models for forecasting applications. Other lags may be chosen. 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 for all observations). The inclusion of rating information into the model reduces the volatility to 2.8097. This demonstrates that credit ratings capture valuable information regarding the idiosyncratic error in the process of the asset returns. Moreover, the last row shows Akaike's information criterion (AIC) declining from 20,875 to 13,295 which indicates a substantial improvement of the goodness-of-fit of the model.

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.

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 similar risk characteristics. Asset values and default thresholds are expected to 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.

4.   Extension of the Model to Asset Return Correlations

4.1   Factor Model

The framework which has been presented thus 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.

The random error U_i of Equation (10) is decomposed into

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

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 $i, i = 1, \ldots, n$. All errors are standard normally distributed and independent from each other. ω and $\tilde{\sigma}$ are parameters which express the exposure to the systematic and idiosyncratic factors. Note that the total variance is $\mathbb{V}(U_i) = \sigma^2 = \omega^2 + \tilde{\sigma}^2$. Thus, the correlation between two latent variables Y_i^* and Y_j^* of asset i and j is given by

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

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^* extends to

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

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

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

where $i \in n_t$, $t = 1, \ldots T$. 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

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

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

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

$$\cdot \prod_{i \in \{y_{it}<0\}} \left(\frac{\phi\left((y_{it} - \boldsymbol{\beta}'\boldsymbol{x}_{it} - \omega \cdot f_t)/\tilde{\sigma}\right)}{\tilde{\sigma}}\right) \phi\left(f_t\right) df_t \tag{31}$$

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

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

(32)

which is then maximized with regard to the parameters β, ω and $\tilde{\sigma}$. This operation can be solved numerically using adaptive Gauss-Hermite quadrature. 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.

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 $\sqrt{\omega^2 + \tilde{\sigma}^2} = \sqrt{1.0242^2 + 2.6215^2} = 2.8145$ 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 $\rho = \dfrac{\omega^2}{\omega^2 + \tilde{\sigma}^2} = \dfrac{1.0242^2}{1.0242^2 + 2.6215^2} = 0.1324$. 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 Utility are much larger. As a result, the overall industries model may reflect diversification benefits across industries. The AICs show an improvement of the goodness-of-fit for each industry sector compared to the model without the systematic factor.

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. 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.

Thus, the probability distribution of the future loss of a 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 $\eta_i = \dfrac{a_i}{a}$.

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

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

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

$$\mathbb{L}_i = \mathbb{E}(L_i) = \mathbb{E}(D_i | \boldsymbol{x}_i) - \mathbb{E}(RGD_i \cdot D_i | \boldsymbol{x}_i)$$

$$= \Phi\left(-\boldsymbol{\beta}' \boldsymbol{x}_i / \sigma\right) - \frac{1}{1 - \Phi\left(\boldsymbol{\beta}' \boldsymbol{x}_i / \sigma\right)} \cdot \exp\!\left(\boldsymbol{\beta}' \boldsymbol{x}_i + 0.5\sigma^2\right) \cdot \Phi\left(-\frac{\boldsymbol{\beta}' \boldsymbol{x}_i + \sigma^2}{\sigma}\right) \cdot \Phi\left(-\boldsymbol{\beta}' \boldsymbol{x}_i / \sigma\right)$$

$$= \Phi\left(-\boldsymbol{\beta}' \boldsymbol{x}_i / \sigma\right) - \exp\!\left(\boldsymbol{\beta}' \boldsymbol{x}_i + 0.5\sigma^2\right) \cdot \Phi\left(-\frac{\boldsymbol{\beta}' \boldsymbol{x}_i + \sigma^2}{\sigma}\right)$$

$$= PD_i - ERGD_i \cdot PD_i$$

$$= PD_i \cdot ELGD_i \tag{34}$$

where the second line follows from the fact that the recovery is different from zero only if the borrower defaults and $PD_i = P\left(D_i = 1 \mid \boldsymbol{x}_i\right)$ 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

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

The expected portfolio loss is obtained as

$$\mathbb{L} = \mathbb{E}(\sum_{i=1}^n \eta_i L_i) = \sum_{i=1}^n \eta_i \mathbb{E}(L_i) =$$

$$= \sum_{i=1}^{n} \eta_i \cdot \left[PD_i - ERGD_i \cdot PD_i \right]$$

$$= \sum_{i=1}^{n} \eta_i \cdot PD_i \cdot ELGD_i \tag{36}$$

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 Equation (35) cannot 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 a single stochastic risk factor (which is the case in our model) and an infinitely granular portfolio. The expected loss rate for borrower i is expressed conditional on the systematic risk factor:

$$\mathbb{L}_i(F) = \mathbb{E}(L_i|F) = \mathbb{E}(D_i|\boldsymbol{x}_i, F) - \mathbb{E}(RR_i \cdot D_i|\boldsymbol{x}_i, F)$$

$$= \Phi\left(-\left(\boldsymbol{\beta'x}_i + \omega \cdot F\right)/\tilde{\sigma}\right)$$

$$- \frac{1}{1 - \Phi\left((\boldsymbol{\beta'x}_i + \omega \cdot F)/\tilde{\sigma}\right)} \cdot \exp\left(\boldsymbol{\beta'x}_i + \omega \cdot F + 0.5\tilde{\sigma}^2\right)$$

$$\cdot \Phi\left(-\frac{\boldsymbol{\beta'x}_i + \omega \cdot F + \tilde{\sigma}^2}{\tilde{\sigma}}\right) \cdot \Phi\left(-\left(\boldsymbol{\beta'x}_i + \omega \cdot F\right)/\tilde{\sigma}\right)$$

$$= \Phi\left(-\left(\boldsymbol{\beta'x}_i + \omega \cdot F\right)/\tilde{\sigma}\right) - \exp\left(\boldsymbol{\beta'x}_i + \omega \cdot F + 0.5\tilde{\sigma}^2\right) \cdot \Phi\left(-\frac{\boldsymbol{\beta'x}_i + \omega \cdot F + \tilde{\sigma}^2}{\tilde{\sigma}}\right)$$

$$= CPD_i(F) - CERGD_i(F) \cdot CPD_i(F)$$

$$= CPD_i(F) \cdot CELGD_i(F) \tag{37}$$

where

$$CPD_i(F) = \Phi\left(-\left(\boldsymbol{\beta'x}_i + \omega \cdot F\right)/\tilde{\sigma}\right) \tag{38}$$

is the conditional default probability, while

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

$$\cdot \Phi\left(-\frac{\boldsymbol{\beta'x}_i + \omega \cdot F + \tilde{\sigma}^2}{\tilde{\sigma}}\right) \tag{39}$$

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. The random loss of a granular portfolio is given by

$$L^\infty = \sum_i^n \eta_i \mathbb{L}_i(F) \tag{40}$$

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 instead, is assumed to be deterministic. In Equation (26), the asset correlation was defined as $\rho = \dfrac{\omega^2}{\sigma^2}$ with $\sigma^2 = \omega^2 + \tilde{\sigma}^2$.

Noting that $1 - \rho = \dfrac{\tilde{\sigma}^2}{\sigma^2}$ and rewriting the conditional probability of default results in

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

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 (Equation 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 Equations (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 Equation (39)

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

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

In the granular portfolio the Downturn LGD is then given as in Equation (41) where α can be set to 0.999 as proposed by Basel II. In other words, the Downturn LGD is then based on the same economic stress as the probability of default.

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.

5.2   Application: Basel II Regulatory Capital

Table 7 shows the key risk parameters for the calculation of bank capital for the various rating classes. The risk parameters include the unstressed parameters Basel asset correlation, probability of default (PD) and loss rate given default (ELGD) as well as the stressed parameters conditional probability of default (CPD) and downturn loss rate given default. Two approaches are compared for the latter: the downturn LGD may firstly be calculated according to the empirical derivation presented by Equation (43): CELGD or secondly by a proposal by the Department of the Treasury, Federal Reserve System and Federal Insurance Corporation (2006): US CELGD which applies a linear relationship of the downturns LGD on ELGD:

$$US\ CELGD = 8\% + 92\% * ELGD \tag{44}$$

Please note that under the Basel IRB approach (compare Basel Committee on Banking Supervision 2006), the regulatory capital is equal 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). The Value-at-Risk is based on the 99.9th percentile of the random systematic risk factor and pre-specified asset correlations. The two last rows of Table 7 show that both a deterministic recovery rate as well as the US proposal lead to a severe underestimation of the regulatory capital which increases with the credit risk in a rating category. This underestimation is up to 20.72 per cent in the instance of deterministic recoveries (rating C) and 16.14 per cent for Equation 44 (rating C).

6.   Discussion

The industry has fundamentally changed its risk measurement and management approaches in recent years by a set of isolated modules which are often provided by external vendors. This practice results in independent and often constant recovery rates. Due to the model independence, financial institutions were surprised that during the current financial crisis, individual risk parameters deteriorated jointly. Recently developed internal risk models were unable to predict this.

The current risk measurement approach has multiple drawbacks. Firstly, default probabilities, recovery rates and correlations are often modeled as constant over time. Secondly, credit risk parameters are modeled independently and possibly inconsistently. Thus, dependencies between parameters are not included. Thirdly, conditional parameters such as recoveries which are conditional upon the occurrence of default are modeled by (ordinary least square) regression models, which do not take the conditionality into account and lead to a severe bias of the estimated parameters.

In response to these shortcomings, this paper provides a top down approach in which individual credit risk parameters are derived in a closed formula from a single model. This model allows for a i) dynamic, ii) consistent, and iii) unbiased modeling of credit portfolio risks. 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.

An empirical analysis provides evidence for the inferred relationship between credit quality, recoveries and correlation. This approach allows financial institutions to have a consistent approach across different credit risk measures used to derive provisions, economic and regulatory capital as well as other applications such as credit pricing. The empirical analysis identified an underestimation of the regulatory capital if downturn loss rates given defaults are estimated applying the current best practice approaches.

In relation to the current financial crisis, the paper may facilitate an increase of transparency of credit portfolio risk models. Transparency may increase if dynamic, consistent, and unbiased models are applied. In addition, the formation of a market standard as well as ability to stress-test and evaluate model risk are essential.

References

Acharya, V. V., S. T. Bharath and A. Srinivasan (2007), "Does Industry-Wide Distress Affect Defaulted Firms? - Evidence from Creditor Recoveries," *Journal of Financial Economics*, 85: 787-821.

Altman, E., B. Brady, A. Resti and A. Sironi (2005), "The Link between Default and Recovery Rates: Theory, Empirical Evidence and Implications," *Journal of Business*, 78: 2203-27.

Basel Committee on Banking Supervision (2006), "International Convergence of Capital Measurement and Capital Standards: A Revised Framework, Comprehensive Version,"

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

Carey, M. (1998), "Credit Risk in Private Debt Portfolios," *Journal of Finance*, 53: 1363-87.

Department of the Treasury, Federal Reserve System and Federal Insurance Corporation (2006), "Basel 2 Capital Accord - Notice of Proposed Rulemaking,"

Dichev, I. and J. Piotroski (2001), "The Long-Run Stock Returns Following Bond Ratings Changes," *Journal of Finance*, 56: 173-203.

Dietsch, M. and J. Petey (2004), "Should SME Exposures be Treated as Retail or Corporate Exposures? A Comparative Analysis of Default Probabilities and Asset Correlations in French and German SMEs," *Journal of Banking and Finance*, 28: 773-88.

Duffie, D., L. Saita and K. Wang (2007), "Multi-Period Corporate Default Prediction with Stochastic Covariates," *Journal of Financial Economics*, 83: 635-65.

Duffie, D. and K. Singleton (1999), "Modeling Term Structures of Defaultable Bonds," *Review of Financial Studies*, 12: 687-720.

Ederington, L. H. and J. C. Goh (1993), "Is a Bond Rating Downgrade Bad News, Good News, or No News for Stockholders?" *Journal of Finance*, 48: 2001-8.

Gordy, M. (2003), "A Risk-Factor Model Foundation for Ratings-Based Bank Capital Rules," *Journal of Financial Intermediation*, 12: 199-232.

Grunert, J. and M. Weber (2009), "Recovery Rates of Commercial Lending: Empirical Evidence for German Companies," *Journal of Banking and Finance*, 33: 505-13.

Hu, Y. and W. Perraudin (2002), "The Dependence of Recovery Rates and Defaults," mimeo.

Jarrow, R., D. Lando and S. Turnbull (1997), "A Markov Model for the Term Structure of Credit Risk Spreads," *Review of Financial Studies*, 10: 481-523.

Jarrow, R. and S. Turnbull (1995), "Pricing Derivatives on Financial Securities Subject to Credit Risk," *Journal of Finance*, 50: 53-85.

Jorion, P., Z. Liu and C. Shi (2005), "Informational Effects of Regulation FD: Evidence from Rating Agencies," *Journal of Financial Economics*, 76: 309-30.

Leland, H. (1994), "Corporate Debt Value, Bond Covenants and Optimal Capital Structure," *Journal of Finance*, 49: 1213-52.

Leland, H. and K. Toft (1996), "Optimal Capital Structure, Endogeneous Bankruptcy, and the Term Structure of Credit Spreads," *Journal of Finance*, 51: 987-1019.

Longstaff, F. and E. Schwartz (1995), "A Simple Approach to Valuing Risky Fixed and Floating Rate Debt," *Journal of Finance*, 50: 789-819.

Madan, D. and H. Unal (1995), "Pricing the Risk of Recovery in Default with APR Violation," *Journal of Banking and Finance*, 27: 1001-218.

McNeil, A. and J. Wendin (2007), "Bayesian Inference for Generalized Linear Mixed Models of Portfolio Credit Risk," *Journal of Empirical Finance*, 14: 131-49.

Merton, R. C. (1974), "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates," *Journal of Finance*, 29: 449-70.

Miu, P. and B. Ozdemir (2002), "Rating Banks, Risk and Uncertainty in an Opaque Industry," *American Economic Review*, 92: 874-88.

Pan, J. and K. Singleton (2008), "Default and Recovery Implicit in the Term Structure of Sovereign CDS Spreads," *Journal of Finance*, 68: 2345-84.

Purda, L. (2007), "Stock Market Reactions to Anticipated Versus Surprise Rating Changes," *Journal of Financial Research*, 30: 301-20.

Pykhtin, M. (2003), "Unexpected Recovery Risk," *Risk*, 16(8): 74-8.

Qi, M. and X. Yang (2009), "Loss Given Default of High Loan-to-Value Residential Mortgages," *Journal of Banking and Finance*, 33: 788-99.

Radelet, S. and J. Sachs (1998), "The East Asian Financial Crisis: Diagnosis, Remedies, Prospects," *Brookings Papers*, 28: 1-90.

Shumway, T. (2001), "Forecasting Bankruptcy More Accurately: a Simple Hazard-Rate Model," *Journal of Business*, 74: 101-24.

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.

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)
AIC	20,875	13,295	19,745	13,297

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. AIC is Akaike's Information Criterion which measures the goodness-of-fit.

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)
AIC	1,195	1,200	2,754	472	3,526	3,843
Obs.	10,288	301,942	37,749	19,107	45,925	31,276

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. AIC is Akaike's Information Criterion which measures the goodness-of-fit.

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)
$\tilde{\sigma}$	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
AIC	12,582	1,166	1,079	2,628	385	3,309	3,728
Obs.	446,287	10,288	301,942	37,749	19,107	45,925	31,276

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. AIC is Akaike's Information Criterion which measures the goodness-of-fit.

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 ($\alpha=0.999$)	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 (6)

	Rating IG	Rating BA	Rating B	Rating C
Basel asset correlation	0.2382	0.2091	0.1423	0.1200
PD	0.0003	0.0060	0.0336	0.1942
CPD	0.0137	0.1077	0.2366	0.5877
ELGD	0.4246	0.4858	0.5419	0.6396
CELGD	0.4709	0.5567	0.6365	0.7700
US CELGD	0.4707	0.5270	0.5786	0.6684
Credit Value-at-Risk (ELGD)	0.0057	0.0498	0.1140	0.2934
Credit Value-at-Risk (CELGD)	0.0063	0.0574	0.1363	0.3701
Credit Value-at-Risk (US CELGD)	0.0063	0.0542	0.1226	0.3104
Underestimation (ELGD)	0.1003	0.1329	0.1642	0.2072
Underestimation (US CELGD)	0.0006	0.0558	0.1006	0.1614

Notes: The (Basel) asset correlation is calculated by inserting the estimated probability of default into the Internal Ratings-based Approach formula for the asset correlation. PD is calculated according to Equation (13). (Basel) CPD is calculated according to Equation (42) and the Basel asset correlation. ELGD is calculated according to Equation (18). CELGD is calculated according to Equation (43). US CELGD is calculated according to a proposal by US regulators (compare Department of the Treasury, Federal Reserve System and Federal Insurance Corporation 2006) according to Equation (44). This proposal applies a linear formula for an economic downturns for the LGD: 8% + 92%*ELGD. Credit Value-at-Risk is equal 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).

Figure 1. Relation between Linear Predictor $\beta'x$ **, Volatility** σ **, and Probability of Default (PD)**
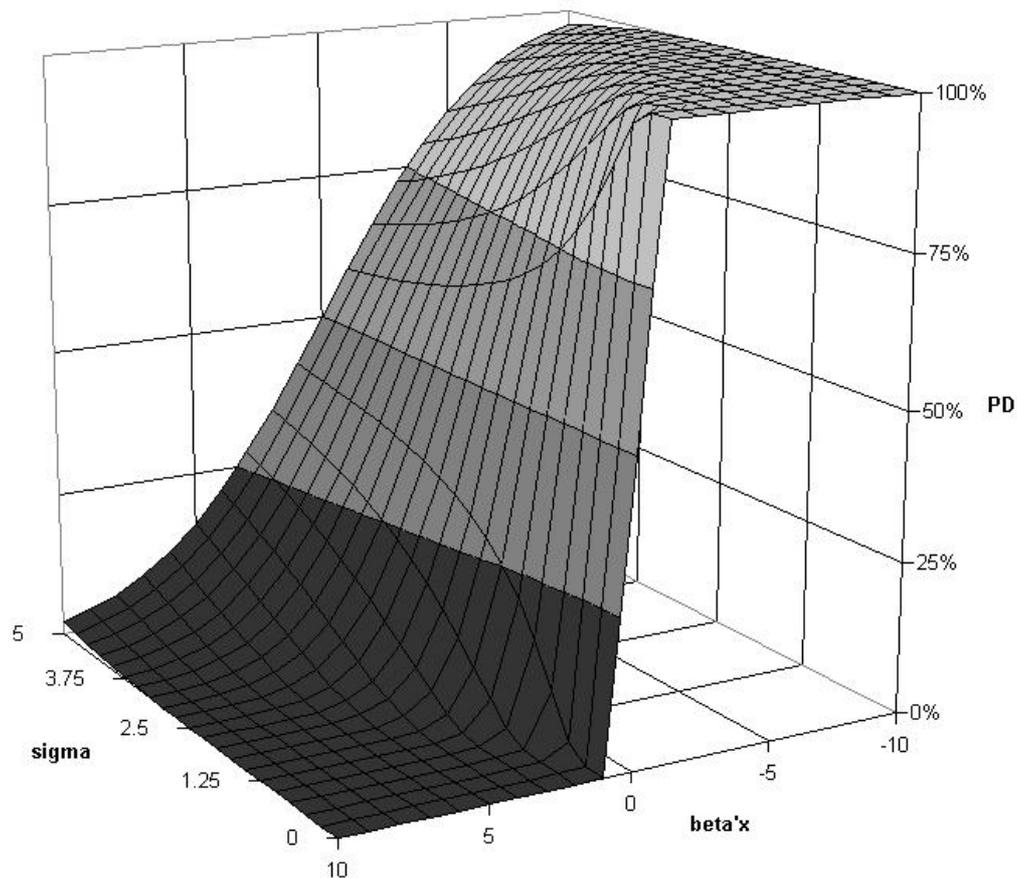
Notes: Probabilities of default are calculated based on σ and $\beta'x$ according to Equation (13). For high σ, the relationship between $\beta'x$ and PD is linear and for low σ, a firm defaults with a high likelihood (i.e., the PD is high) if $\beta'x < 0$ and a firm does not default with a high likelihood (i.e., the PD is low) if $\beta'x > 0$.

Figure 2. Relation between Probability of Default (PD), Expected Loss given Default (ELGD), and Volatility σ
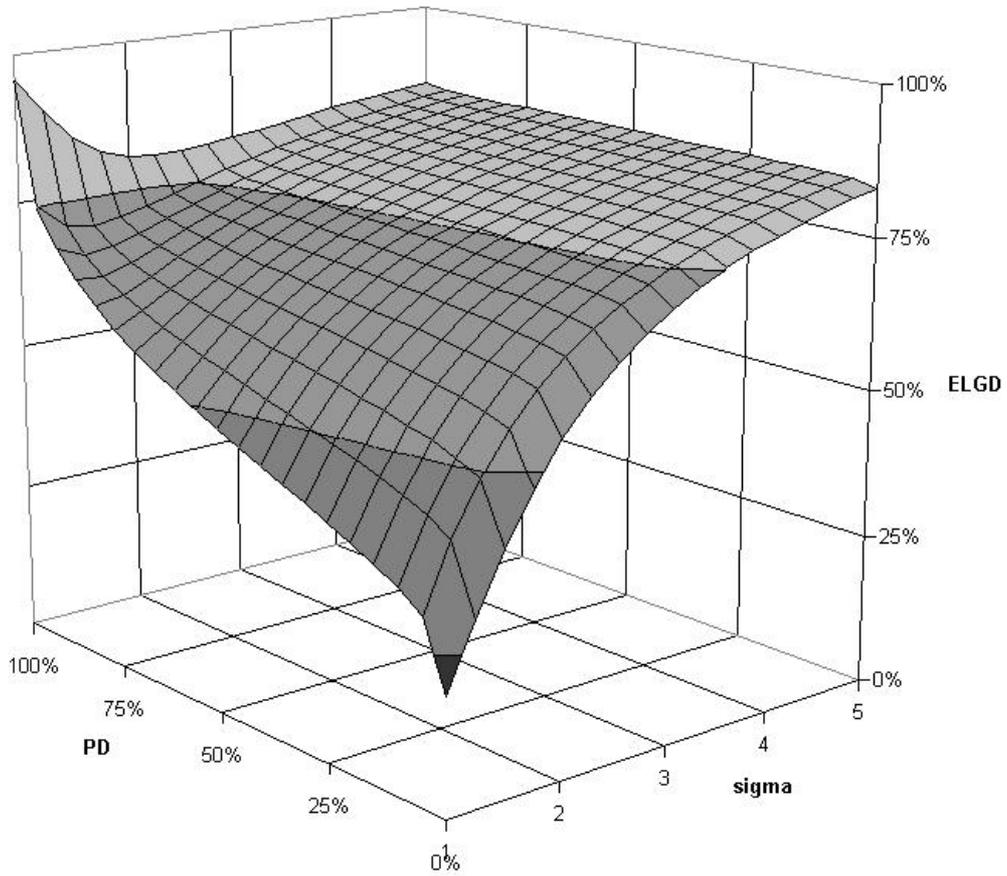
Notes: ELGD is calculated based on PD and σ according to Equation (17) and Equation (18).

Figure 3. Default Rate and Non-Investment Grade Rate
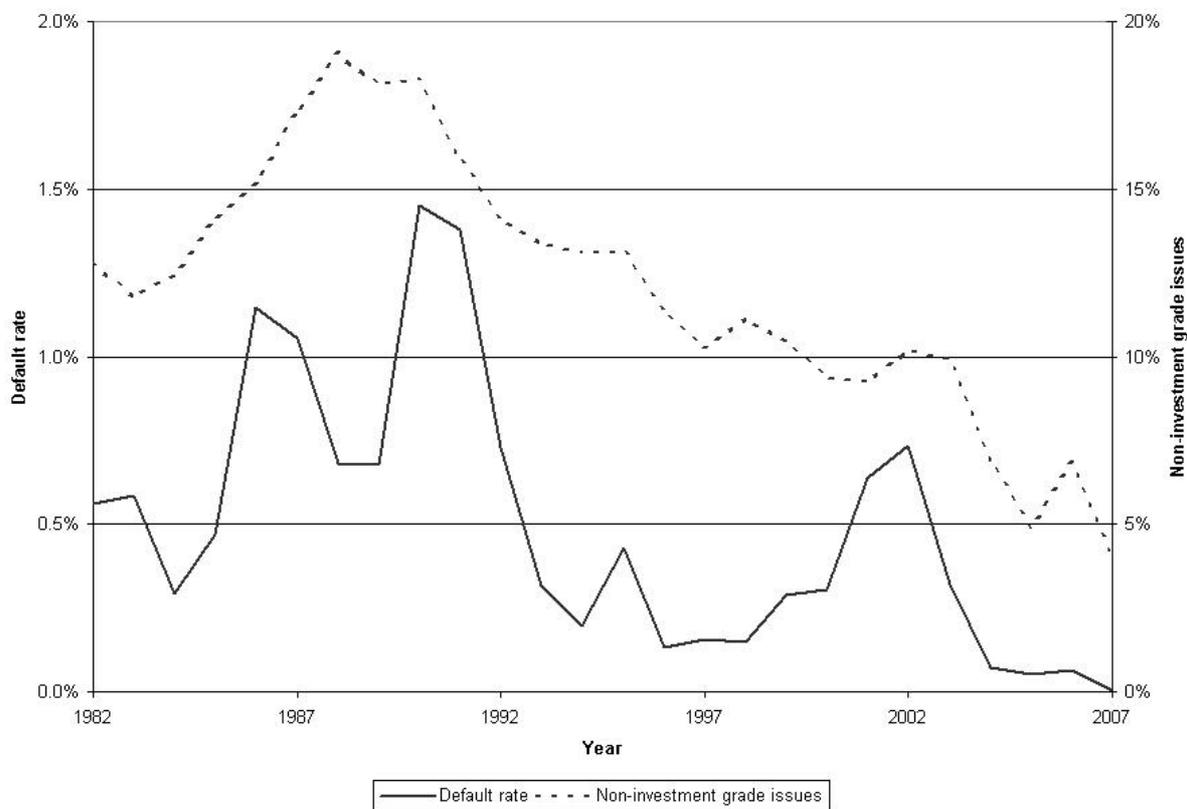
Notes: Default rate is the ratio between the number of defaulted issues and the total number of issues. The non-investment grade rate is the number of non-investment grade issues to the total number of issues.

Figure 4. Default Rates for All Issues and Recovery Rates for All Issues
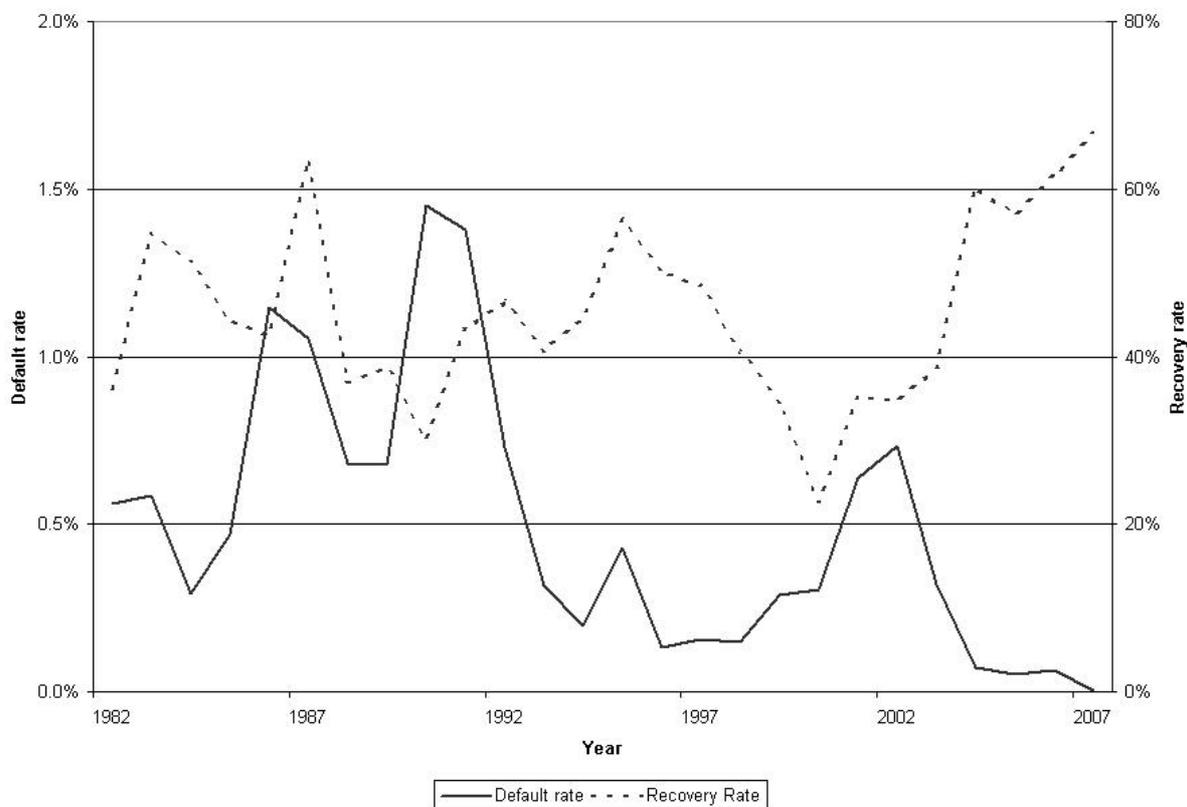
Notes: Default rate is the ratio between the number of defaulted issuers and the total number of issuers. Recovery rate is the ratio of the price of defaulted debt obligations after 30 days of the occurrence of a default event and the par value.

Figure 5. Absolute Frequencies for Recoveries
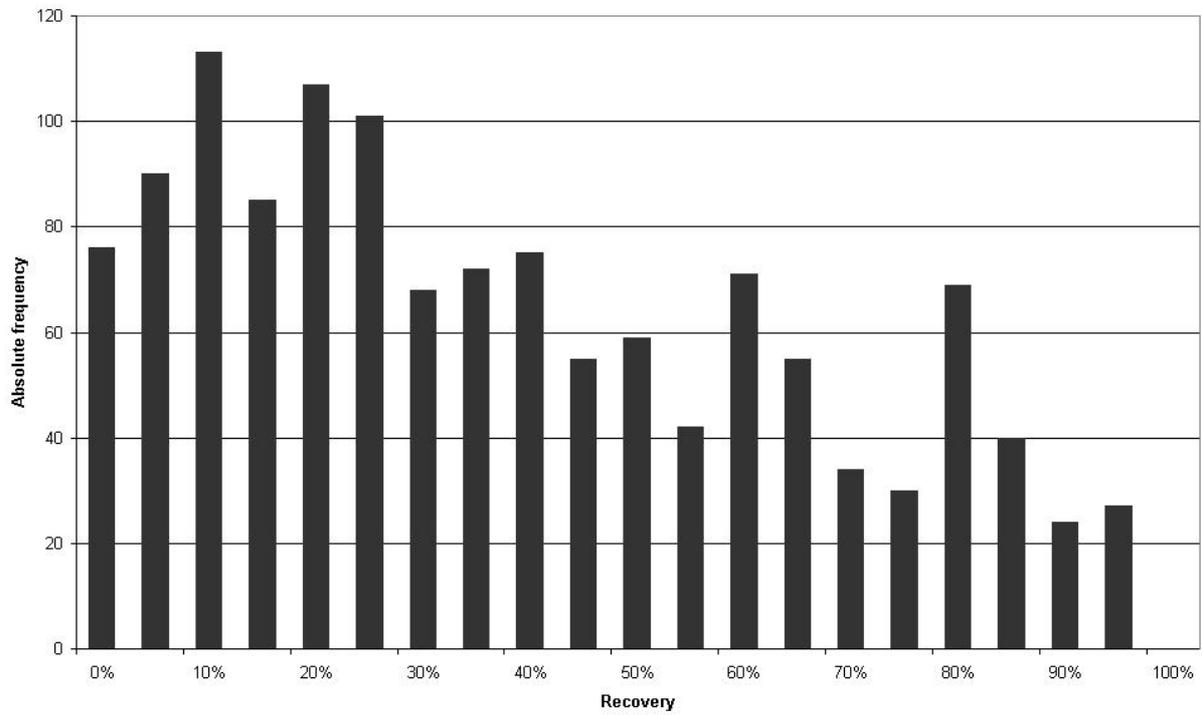
Notes: Recovery rate is the ratio of the price of defaulted debt obligations after 30 days of the occurrence of a default event and the par value.

Figure 6. Absolute Frequencies for Log-Recoveries
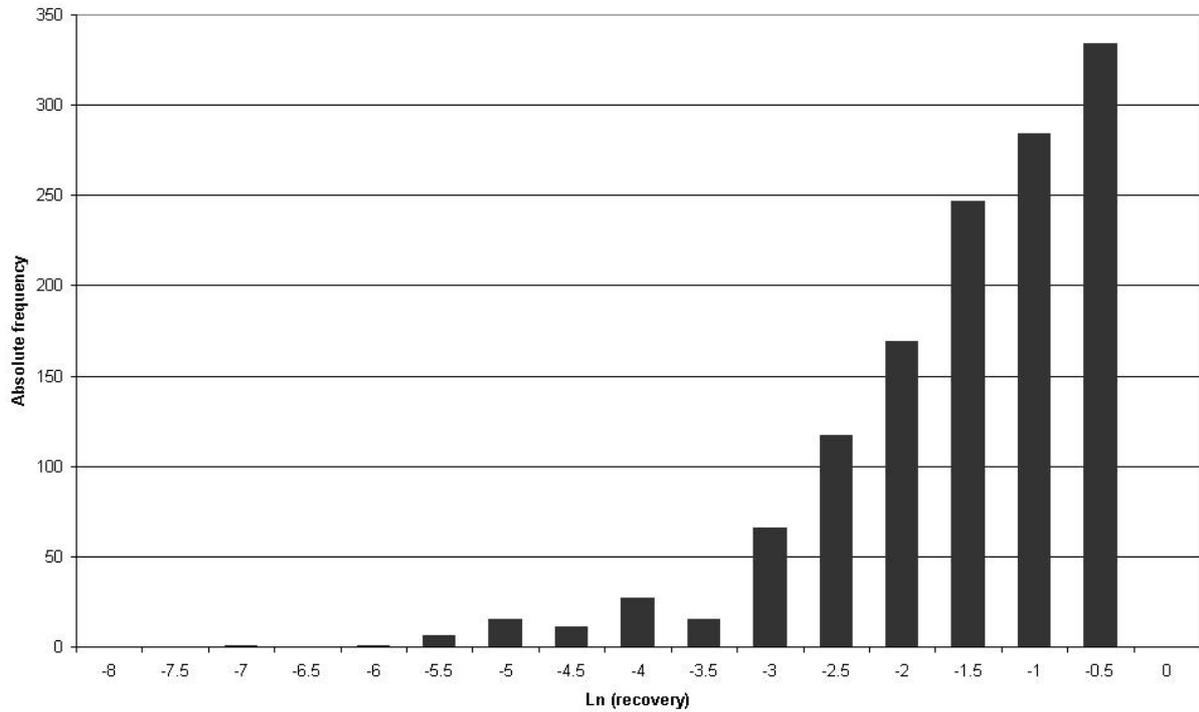
Notes: Recovery rate is the ratio of the price of defaulted debt obligations after 30 days of the occurrence of a default event and the par value.

Appendix. Derivation of the Expected Recovery Rate Given Default

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

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

$$ERGD_i = \int_{-\infty}^{0} \exp(y_i) \cdot h(y_i | Y_i < 0, \boldsymbol{x_i}) dy_i$$

$$= \int_{-\infty}^{0} \exp(y_i) \cdot \frac{\phi(-(y_i - \mu_i)/\sigma)}{\sigma \cdot (1 - \Phi(\mu_i/\sigma))} dy_i$$

$$= \frac{1}{\sigma \cdot PD_i} \int_{-\infty}^{0} \exp(y_i) \cdot \frac{1}{\sqrt{2\pi}} \cdot \exp\left(\frac{-(-y_i + \mu_i)^2}{2\sigma^2}\right) dy_i$$

$$= \frac{1}{\sigma \cdot PD_i} \cdot \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{0} \exp\left(y_i - \frac{\mu_i^2 - 2y_i\mu_i + y_i^2}{2\sigma^2}\right) dy_i$$

$$= \frac{1}{\sigma \cdot PD_i} \cdot \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{0} \exp\left(\frac{2\sigma^2 y_i - \mu_i^2 + 2y_i\mu_i - y_i^2}{2\sigma^2}\right) dy_i$$

$$= \frac{1}{\sigma \cdot PD_i} \cdot \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{0} \exp\left(\frac{-y_i^2 + 2y_i(\mu_i + \sigma^2) - (\mu_i + \sigma^2)^2 - \mu_i + (\mu_i + \sigma^2)^2}{2\sigma^2}\right) dy_i$$

$$= \frac{1}{\sigma \cdot PD_i} \cdot \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{0} \exp\left(\frac{-(y_i - (\mu_i + \sigma^2))^2 + 2\mu_i\sigma^2 + \sigma^4}{2\sigma^2}\right) dy_i$$

$$= \frac{1}{\sigma \cdot PD_i} \cdot \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{0} \exp\left(\frac{-(y_i - (\mu_i + \sigma^2))^2}{2\sigma^2}\right) \cdot \exp(\mu_i + 0.5\sigma^2) dy_i$$

$$= \frac{1}{PD_i} \cdot \exp(\mu_i + 0.5\sigma^2) \cdot \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(y_i - (\mu_i + \sigma^2))^2}{2\sigma^2}\right) dy_i$$

$$= \frac{1}{PD_i} \cdot \exp(\mu_i + 0.5\sigma^2) \cdot \Phi\left(-\frac{\mu_i + \sigma^2}{\sigma}\right)$$