Modeling and estimating dependent loss given default

Martin Hillebrand *

September 2006

Abstract

We propose a portfolio credit risk model with dependent loss given default (LGD) which allows for a reasonable economic interpretation and can easily be applied to real data. We build up a precise mathematical framework and stress some general important issues when modeling dependent LGD. Finally, we calibrate the model based on American bond data from 1982 to 2001 and compare the results with recently published alternative models.

1 Introduction

Most credit risk models assume as loss given default (LGD) a constant proportion of any credit loss and ignore the fact that LGD is itself an important driver of the portfolio credit risk because of its possible dependence on economic cycles. The Basel Committee for Banking Supervision (2004) acknowledged this importance by starting a discussion with the banking industry aiming at the investigation of this issue.

Empirical evidence for dependent LGD is indeed provided by data presented in Altman et al. (2003) and Moody's (2003). Approaches of modeling dependent LGD have been suggested during the past five years, but none of them seem to have had an impact on present practice. In fact, it is very hard to obtain a model that

- has a reasonable economic interpretation
- can be calibrated by available data
- is based on a proper statistical setting.

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

In particular, a proper statistical model should incorporate the observation that expected LGD and default probability are, conditional on the economic cycle, dependent but not comonotonous. Hence, a stochastic dependence and not a deterministic functional relation should be modeled. In this article we discuss the up to date proposed models in this context and suggest a new model, which addresses these issues.

We start with a homogeneous portfolio of m credits, for simplicity each with exposure 1. We are interested in the loss L of the portfolio within one year. A statistician's approach would be to take past losses of this portfolio, and calculate quantities, which allow for risk assessment and prediction. To estimate a Value-at-Risk, methods from extreme value theory can be applied to estimate quantiles outside the range of observations. Such purely data and simulation driven methods, however, require a certain data sample size which has not been available for LGD yet. Also, one might miss important economic mechanisms, which influence future losses, and may not yet be visible in past data.

To analyze such an economic mechanism of loss we consider the loss net of recovery L_i of a single credit i. The observed L_i contains two features:

- the default event $D_i := \mathbb{1}_{\{L_i>0\}}$, where $\mathrm{PD}_i = \mathbb{P}(D_i{=}1)$ is the probability of default, and
- the loss given default $\mathrm{LGD}_i := L_i|D_i{=}1$.

These quantities characterize the loss

$$L_i = \mathrm{LGD}_i\, \mathbb{1}_{\{D_i=1\}}.$$

For the expected loss we obtain

$$\mathbb{E}L_i \;=\; \mathbb{P}(D_i{=}1)\cdot\mathbb{E}(L_i|D_i{=}1) \;=\; \mathrm{PD}_i\cdot\mathbb{E}(\mathrm{LGD}_i).$$

As $\mathrm{PD}_i = \mathrm{PD}$ and $\mathrm{LGD}_i\mathbb{1}_{\{D_i=1\}} \stackrel{d}{=} \mathrm{LGD}\mathbb{1}_{\{D=1\}}$ for all $i = 1,\dots,m$, we obtain, for a standardized homogeneous portfolio $L^{(m)} = \frac{1}{m}\sum_{i=1}^m L_i$, as the expected loss:

$$\mathbb{E}L^{(m)} = \mathrm{PD}\frac{1}{m}\sum_{i=1}^{m}\mathbb{E}(L_i|D_i{=}1) = \mathrm{PD}\cdot\mathbb{E}(\mathrm{LGD}).$$

For estimating the risk of the portfolio in terms of the Value-at-Risk (VaR) dependence comes into play. While it is common practice only to model dependence between the D_i, one can imagine that incorporating dependence between the LGD_i may increase the portfolio risk. How to model this dependence is the topic of the next section.

Our paper is organized as follows. In Section 2 we present first approaches to model PD and LGD as jointly dependent on economic factors. Section 3 is devoted to our model, for a homogeneous portfolio we derive its almost sure (a.s.) limit. We show that it can be embedded in a regression model, which opens the way to statistical calibration. We also investigate the Value-at-Risk based on our model and compare it to a model with constant LGD. In Section 4 we compare our model with Tasche's and Giese's model. In Section 5 we explain, how our model extends to non-homogeneous portfolios and draw some conclusions in Section 6.

2 Modeling Dependent Loss Given Default

Economic reasoning suggests that LGD and PD may both depend on the economic cycle: macroeconomic development such as a decrease of consumption or investment which leads to a downturn period and hence increased PD may also cause a decrease of market value of collateral resulting in a higher LGD. The general idea in most of the proposed models is a common dependence of LGD and PD on a systematic factor Y and assuming independence of the conditional random variables $L_i|Y$ for all $i = 1, \dots, m$. Then, with the homogeneity assumption that $\mathrm{LGD}_i|Y \overset{d}{=} \mathrm{LGD}|Y$ and $\mathrm{PD}_i|Y \overset{d}{=} \mathrm{PD}|Y$ for all $i = 1, \dots, m$), a strong law of large numbers holds for $L^{(m)} = \frac{1}{m}\sum_{i=1}^{m} L_i$; for a proof see e.g. Proposition 2.5.4 in Bluhm et al. (2003)). For $L = L_1$ we formulate this as

$$\mathbb{P}\left(\lim_{m\to\infty}\left[L^{(m)} - \mathbb{E}(L|Y)\right] = 0\right) = 1. \tag{2.1}$$

For all $y \in \mathbb{R}$ we have

$$\begin{aligned}\mathbb{E}(L|Y{=}y) &= \mathbb{P}(D{=}1|Y{=}y) \cdot \mathbb{E}(L|D{=}1, Y{=}y) \\ &= \mathrm{PD}(y) \cdot \mathbb{E}(\mathrm{LGD}|Y{=}y),\end{aligned}$$

leading to

$$\mathbb{E}(L|Y) = \mathrm{PD}(Y) \cdot \mathbb{E}(\mathrm{LGD}|Y). \tag{2.2}$$

Hence we can reduce the problem of modeling the portfolio loss to the problem of finding functions $\mathrm{PD}(Y)$ and $G(Y) := \mathbb{E}(\mathrm{LGD}|Y) = \mathbb{E}(L|D{=}1, Y)$.

Following Merton's model (compare e.g. Bluhm et al. (2002)), company i defaults in time period $[0, T]$, if the log return of the asset value A_i falls below some threshold s_i:

$$D_i = \mathbb{1}_{\{A_i < s_i\}}$$

Choosing the time period $[0, T]$ being one year and assuming the log returns being standard normally distributed, we obtain

$$s_i = \Phi^{-1}(\mathrm{PD}_i).$$

We model A_i by a normal factor model, then for some macroeconomic factor Y and some idiosyncratic factor (or noise term) ε_i, both independent and standard normally distributed, for $0 \leq \beta_i \leq 1$:

$$A_i = \beta_i Y + \sqrt{1-\beta_i^2}\,\varepsilon_i \tag{2.3}$$

We obtain

$$\begin{aligned} \mathrm{PD}(Y) &= P(D_i = 1|Y) \\ &= \mathbb{P}(\beta Y + \sqrt{1-\beta^2}\varepsilon_i \leq s_i) \\ &= \Phi\left(\frac{s_i - \beta Y}{\sqrt{1-\beta^2}}\right). \end{aligned}$$

For the sake of simplicity we take $s_i = s$ for all $i = 1, \ldots, m$, and set $c = s/\sqrt{1-\beta^2}$ and $e = \beta/\sqrt{1-\beta^2}$. Then

$$\mathrm{PD}(Y) = \Phi(c - eY). \tag{2.4}$$

Remark. This model includes a multifactor setting: PD may be dependent on a linear combination of systematic factors $Y_1, \ldots, Y_K$ in the sense that

$$Y = \sum_{k=1}^{K} w_k Y_k \tag{2.5}$$

with $\sum_{k=1}^{K} w_k^2 = 1$. Then (2.4) still holds. □

So we concentrate on the function G. We normalize it by requiring

$$\lim_{y\to-\infty} G(y) = 1 \quad \text{and} \quad \lim_{y\to\infty} G(y) = 0.$$

This means that a worst case economic scenario would imply total loss, while in a perfect economy default would be rare and the LGD negligible.

There have been several approaches of modeling dependent LGD based on (2.2). Early linear modeling by Frye (2000a and 2000b) and extensions by Pykhtin (2003) suggest models, which involve knowledge of the systematic factor Y, which is – as it is a latent variable – rarely available.

Tasche (2004) extends Merton's model above by assuming that LGD is driven by the value of the undershoot of the asset value log return below the liability s. This yields the model

$$G(y) = (\mathrm{PD}(Y))^{-1} \int_{-\infty}^{c-ey} F^*_{\mu,\sigma}\left(\frac{\mathrm{PD} - \Phi(\beta y + \sqrt{1-\beta^2}x)}{\mathrm{PD}}\right) \varphi(x)\,dx, \quad y \in \mathbb{R}.$$

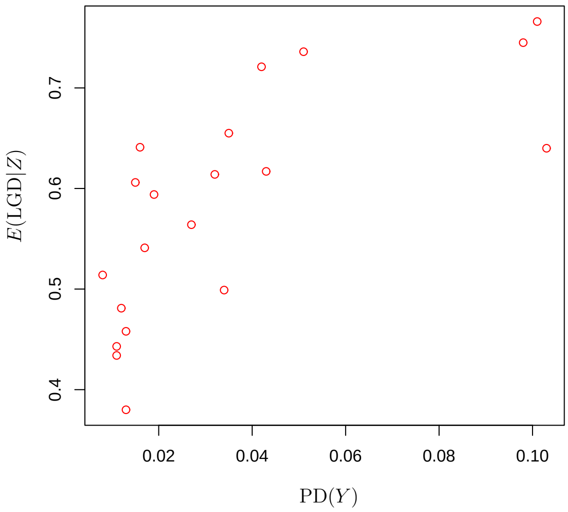

Figure 1: Annual default frequencies (horizontal axis) versus average LGD (vertical axis) of American bonds from 1982 to 2001.

where $F^*_{\mu,\sigma}$ is the inverse of a distribution function on $(0,1)$; he suggests the beta distribution with mean μ and variance σ^2. Fitting this model means then estimating expectation and variance of LGD, but there is no free parameter to adjust the dependence of LGD on Y.

Giese (2005) addresses this problem directly in modeling a non-linear dependence structure. Considering the data set of Figure 1, he suggests as a link function

$$\mathbb{E}(\mathrm{LGD}|Y) = 1 - a_0 \left(1 - \mathrm{PD}(Y)^{a_1}\right)^{a_2}$$

for parameters $a_0, a_1, a_2 \in \mathbb{R}_+$, leading in the Merton framework to

$$G(y) = 1 - a_0 \left(1 - \Phi\left(\frac{\Phi^{-1}(\mathrm{PD}) - \beta y}{\sqrt{1-\beta^2}}\right)^{a_1}\right)^{a_2}$$

This model has two weaknesses. First, the parameter estimation requires a nontrivial optimization procedure. Second, the link function seems rather arbitrary from an economic as well as from a statistical point of view. In particular, it does not explain the data spread which is observed in the data. Instead, as in Tasche's model, a comonotonous relation is assumed.

3 A New Model

Guided by our requirements in the introduction we suggest a new model based on the following assumptions.

- Analogously to (2.5), we allow LGD to be dependent on a linear combination of standard normally distributed systematic factors $Y_1, \ldots, Y_K$ by means of $Z := \sum_{k=1}^{K} v_k Y_k$, where $\sum_{k=1}^{K} v_k^2 = 1$. Hence, Z is standard normally distributed.

- We also assume that the conditional random variables $L_i|Y, Z$ are independent for all $i = 1, \ldots, m$.

Remark. Recently, several data analyses estimating the dependence of LGD on a set of macroeconomic factors such as in Altman et al. (2003) have been published. But these studies require large data sets which are often not available for calibrating region or industry specific portfolios. We introduce dependent LGD modeling into a multifactor latent variable framework because it requires the estimation of fewer parameters and hence also works with smaller data sets. □

This setting allows for a rather flexible dependence structure between PD and $\mathbb{E}$(LGD): Besides a comonotonous influence of factors on both variables, some factors may have converse influences (weights may be negative), or factors may have influence only on one of both variables (weights may be zero). E.g. a (regional) real estate index may influence LGD but not PD of an export oriented small company.

Similarly as in (2.1) a strong law of large numbers holds.

Theorem 1. *Consider a homogeneous portfolio* $L^{(m)} = \frac{1}{m}\sum_{i=1}^{m} L_i$. *Then*

$$\mathbb{P}\left(\lim_{m\to\infty}\left[L^{(m)} - \mathbb{E}(L|Y, Z)\right] = 0\right) = 1. \tag{3.1}$$

As in (2.2) we obtain

$$\mathbb{E}(L|Y, Z) = \mathrm{PD}(Y) \cdot G(Z). \tag{3.2}$$

Based on the Merton model we model PD(Y) by (2.4). Note that $G(Z) = \mathbb{E}(L|D = 1, Z) = \mathbb{E}(\mathrm{LGD}|Z)$. Again, the question is how to model G. If we transform the data by the probit function Φ^{-1} we obtain a linear structure in the data; see Figure 2.

The corresponding model for G is given by

$$G(z) = \Phi(a - bz), \quad z \in \mathbb{R}, \tag{3.3}$$

for parameters $a, b \in \mathbb{R}$.

Since we assume–in the tradition of the standard Merton model–the systematic factors $Y_1, \ldots Y_K$ to be latent (i.e. not observable) variables, it suffices to regard the aggregated factors Y and Z. We can decompose Z into Y and a second independent factor X as follows: Define $d := \mathrm{cov}(Y, Z)$ and $X := 1/\sqrt{1-d^2}(Z - dY)$. Then X is by definition standard normally distributed and independent of Y, and

$$Z = dY + \sqrt{1-d^2}X \tag{3.4}$$

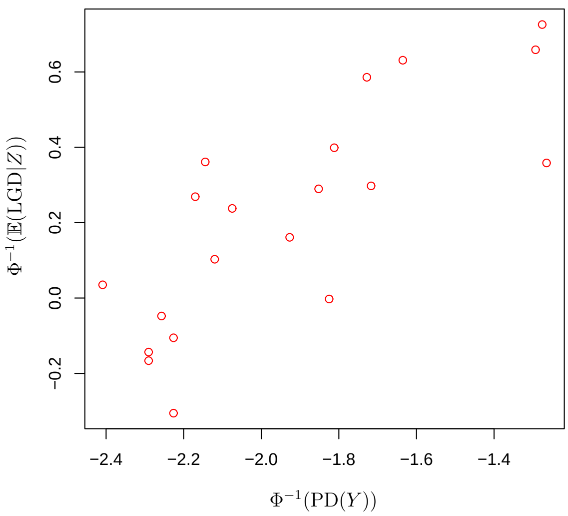

Figure 2: Probit transformed annual default frequencies (horizontal axis) versus average LGD (vertical axis) of American bonds from 1982 to 2001.

holds. X can be regarded as aggregated influence of systematic factors on $\mathbb{E}(\text{LGD}|Z)$ which is independent of the systematic influence on $\text{PD}(Y)$.

Model equations (2.4), (3.3) and (3.4) provide a linear model for the data shown in Fig. 2 as follows:

From (3.3) and (3.4) we obtain

$$\Phi^{-1}(E(\text{LGD}|Z)) = a - bdY - b\sqrt{1-d^2}X. \tag{3.5}$$

Plugging equation (2.4) into (3.5) yields the linear regression equation

$$\Phi^{-1}(E(\text{LGD}|Z)) = a - \frac{b\,d\,c}{e} + \frac{b\,d}{e}\Phi^{-1}(\text{PD}(Y)) - b\,\sqrt{1-d^2}\,X, \tag{3.6}$$

where $\Phi^{-1}(\mathbb{E}(\text{LGD}|Z))$ is the response variable, $\Phi^{-1}(\text{PD}(Y))$ the predictor variable and X the residual.

Model Calibration Issues

For this regression model, we obtain estimates for intercept $a - b\,d\,c/e$, slope $b\,d/e$, and standard deviation $b\sqrt{1-d^2}$ of the residuals. For known values c and e this yields estimates for a, b, and d. For unknown values c and e, we estimate them first, and then get plug-in estimates for a, b, and d.

With these estimates we can estimate the downturn LGD at level $\alpha \in (0,1)$ from (3.3) by

$$\widehat{\mathbb{E}(\text{LGD}|Z = z_\alpha)} = \Phi(\hat{a} - \hat{b} \cdot z_\alpha), \tag{3.7}$$

where z_α is the standard normal α-quantile.

For a proper risk assessment, we need the asymptotic distribution of the portfolio loss. The following result gives an analytic form for the relevant ranges of parameters.

Theorem 2. *With the model assumptions as above let* $L = \Phi(c - eY)\Phi(a - bZ)$, *where* $Z = dY + \sqrt{1-d^2}X$, *for* $e > 0$ *and* $bd > 0$. *Then*

$$P\left(\lim_{m\to\infty} \left|L^{(m)} - L\right| = 0\right) = 1, \tag{3.8}$$

and the distribution function of L *has the form*

$$\mathbb{P}(L \le l) = 1 - \int_{-\infty}^{A(l)} \Phi(B(l,y))\varphi(y)\,dy, \quad l \in \mathbb{R},$$

where

$$A(l) := \frac{c - \Phi^{-1}(l)}{e} \quad \text{and} \quad B(l,y) := \frac{1}{b\sqrt{1-d^2}}\left(a - bdy - \Phi^{-1}\left(\frac{l}{\Phi(c-ey)}\right)\right).$$

Proof.
The strong law of large numbers in (3.8) follows directly from Theorem 3.1 and (3.2). From (2.4), (3.4), (3.2) and (3.3) we obtain

$$E(L|Y,Z) \;=\; \Phi(c - eY)\cdot\Phi(a - b\,d\,Y - b\sqrt{1-d^2}\,X).$$

Taking advantage of the monotonicity of distribution functions, we have the equivalences

$$\begin{aligned} & \mathbb{E}(L|X,Y) \le l \\ \Longleftrightarrow & \begin{cases} \quad\;\; \Phi(c-eY) \le l \\ \text{or} \;\; \Phi(c-eY) > l \text{ and } \Phi(c-eY)\cdot\Phi(a - b\,d\,Y - b\sqrt{1-d^2}\,X) \le l. \end{cases} \end{aligned}$$

With this we calculate

$$\begin{aligned} \mathbb{P}(L \le l) &= \mathbb{P}(\mathbb{E}(L|X,Y)) \le l) \;=\; \mathbb{E}\mathbb{1}_{\{\mathbb{E}(L|X,Y))\le l\}} \\ &= \mathbb{E}\left(\mathbb{1}_{\{\Phi(c-eY)\le l\}} + \mathbb{1}_{\{\Phi(c-eY)>l\}}\mathbb{1}_{\{\Phi(c-eY)\Phi(a-b\,d\,Y-b\sqrt{1-d^2}\,X)\le l\}}\right) \\ &= 1 - \mathbb{E}\left(\mathbb{1}_{\{\Phi(c-eY)>l\}}\left(1 - \mathbb{1}_{\{\Phi(c-eY)\Phi(a-b\,d\,Y-b\sqrt{1-d^2}\,X)\le l\}}\right)\right) \\ &= 1 - \mathbb{E}\left(\mathbb{1}_{\{\Phi(c-eY)>l\}}\mathbb{1}_{\{\Phi(c-eY)\Phi(a-b\,d\,Y-b\sqrt{1-d^2}\,X)>l\}}\right). \end{aligned}$$

Using the independence of X and Y we calculate for $e > 0$ and $bd > 0$,

$$\mathbb{P}(L \le l) \;=\; 1 - \int_{-\infty}^{A(l)} \Phi(B(l,y))\varphi(y)dy,$$

where $A(l)$ and $B(l, y)$ are the functions in the assertion. □

Observe that, if LGD $\equiv 1$, then

$$\mathbb{P}(L \leq l) = 1 - \Phi(A(l)),$$

which allows for a comparison of a deterministic LGD with the stochastic LGD in our model.

Fitting the Model

We fit the data of Altman et al. (2003) as shown in Figure 1. It consists of $n = 20$ data points of annual default frequencies and average LGD of American bonds from 1982 to 2001. We want to fit them to our model.

First, we fit the Merton model to the default frequency data. As $\text{PD} = \text{PD}_i = \mathbb{E}D_i$, we estimate it by the empirical mean of the annual default frequency. To estimate β we use relation (2.4):

$$\beta = \sqrt{\frac{\text{var}\left(\Phi^{-1}(\text{PD}(Y))\right)}{1 + \text{var}\left(\Phi^{-1}(\text{PD}(Y))\right)}}.$$

Hence β can be estimated as a function of the empirical variance of the probit transformed annual default frequencies. This yields

$$\widehat{\text{PD}} = 0.035 \quad \text{and} \quad \hat{\beta} = 0.336.$$

To estimate the remaining parameters of the model we use the regression model (3.6) with the plug-in parameters c and e. This yields

$$\hat{a} = 0.220, \quad \hat{b} = 0.300 \quad \text{and} \quad \hat{d} = 0.620.$$

Kolmogorov-Smirnov test, Durbin-Watson test, and White test support the assumption of normally distributed homoscedastic and not autocorrelated residuals. It turns out that all estimated parameters yield $e > 0$ and $bd > 0$, hence we can apply Theorem 2. Figure 3 shows the calibrated function $\Phi(\hat{a} - \hat{b} \cdot z_\alpha)$ as a function of α.

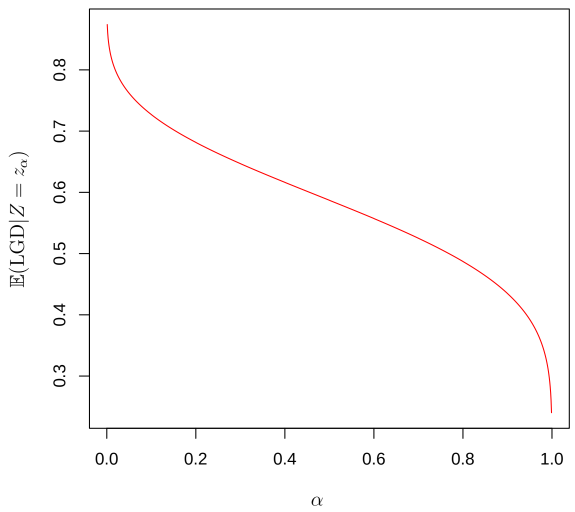

Figure 3: Expected LGD dependent on systematic factors.

Consequences of Dependent LGD

With Theorem 2, we can compute the loss distribution of a large homogeneous portfolio. Figures 4 and 5 show the portfolio loss distribution of our model compared with Merton models with constant loss given default: LGD $\equiv 0.65$ and LGD $\equiv 1$ which can be regarded as lower and upper bounds for LGD in the right tail of the loss distribution.

For confidence level α, we define the portfolio Value-at-Risk (VaR) as usual by by

$$\text{VaR}_\alpha(L) = \inf\{x \geq 0 | \mathbb{P}(L \leq x) \geq \alpha\},$$

where L has representation $L = \Phi(c - eY)\Phi(a - bZ)$ as given in Theorem 2.

There are suggestions to estimate LGD and PD quantiles separately and obtain a portfolio loss VaR by multiplication (see e.g. [2]). This implies the assumption that PD and LGD are comonotonous. In our modeling framework, it means that quantiles of $\Phi(c - eY)$ and $\Phi(a - bZ)$ are calculated separately, assuming comonotonicity of Y and Z. This leads to a systematic overestimation of risk as seen in Table 3.1.

VaR_α	$\alpha = 99\%$	$\alpha = 99.5\%$	$\alpha = 99.9\%$
Dependent LGD	0.107	0.126	0.171
LGD $\equiv 0.65$	0.089	0.102	0.134
LGD $\equiv 1$	0.137	0.157	0.206
Comonotonous LGD	0.112	0.132	0.180

Table 3.1: Portfolio Value-at-Risk for different confidence levels.

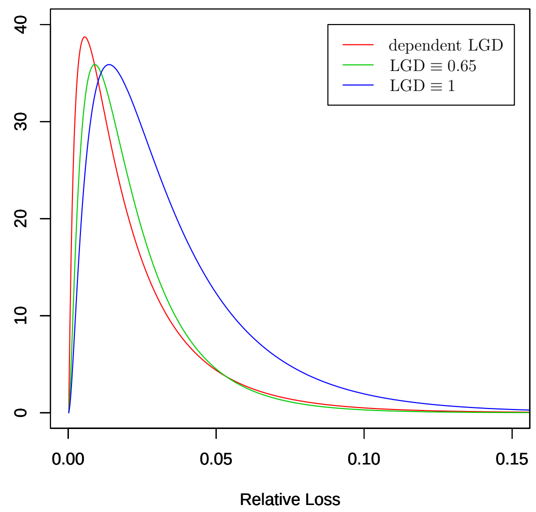

Figure 4: Portfolio loss distribution densities.

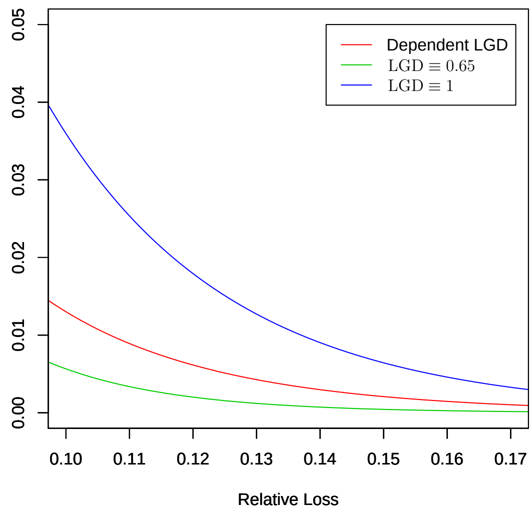

Figure 5: Portfolio loss distribution tails.

4 Comparing the Models

The problem with comparing the different models is that we estimate extreme losses lying far outside the sample. For our data of yearly default frequencies and average LGD, we calculate yearly losses within the range of 0.004 to 0.077. A comparison with Table 3.1 shows that no data point corresponds to such a high VaR.

What we can do is, of course, show the fit of the different models within the sample range. The most interesting part when comparing the fit of the models would be at the upper end of the sample.

An excellent method to investigate exactly this part of the sample are quantile plots, also called QQ-plots. In a QQ-plot the quantiles of a theoretical model are plotted against the empirical quantiles. Ideally, the quantiles should be match (at least for large samples), i.e. the plot should lie on the 45 degree line. If the plot is linear with different slope/intercept, then the location/scale family of the model is correctly chosen, but scale/location parameters have to be adjusted. If the plot has negative curvature at the right end, the estimated distribution has heavier tails than the empirical distribution. For more details, see, e.g., Embrechts et al. (1997), Section 6.2, and references therein.

In Figure 6 we observe that our model seems to provide a good estimate, while Giese's model systematically overestimates large quantiles. This is caused by the fact that Giese models $\mathbb{E}(\text{LGD})$ and PD(Y) to be comonotonous. This means that high $\mathbb{E}(\text{LGD})$ and PD always occur at the same time leading to a higher probability of extreme losses. On the other hand, Tasche's model systematically underestimates for this data set the

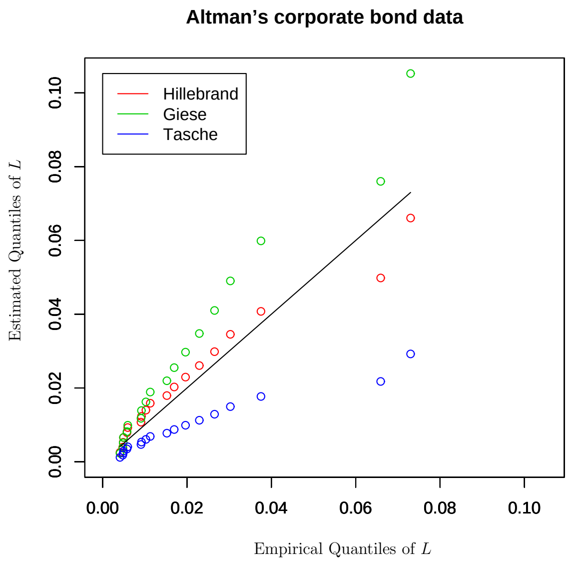

Figure 6: Q-Q-plots of the estimated loss distribution for different models. The plotted 45 degree line indicates the quality of the fit.

scale parameter. The reason for this lies in the model, which implies a functional relation between PD and $\mathbb{E}(\text{LGD})$ that is too weak. There is no additional parameter to remedy this weakness.

To investigate the quality of our model further we also calibrate it to data of defaulted corporate bonds gathered by Moody's (2003). It consists of annual default frequencies and average LGD of bonds and preferred stocks from 1982 to 2003. The result is simply summarized in a QQ-plot; see Figure 7.

5 Nonhomogeneous Portfolios

In the preceeding considerations we made two simplifications. We assumed exposure at default (EAD) of 1 and equal distributions of the single losses L_i. This is a good approximation for large portfolios with exposures of same magnitude. If this is not the case, one has to take into account the individual randomness, also known as idiosyncratic or unsystematic risk.

This can be done analogously to the approaches of e.g. Tasche (2004) or Giese (2005): for each credit class k for $k = 1, \dots, K$ with the same industry sector, global region and company size, the systematic risk is estimated by calibrating $\text{PD}_k(Y)$ and $\mathbb{E}(\text{LGD}_k|Z)$ as shown in the preceeding section. Individual randomness is then modeled by introducing for instance a beta function with scaling parameters σ_k, which yields the model

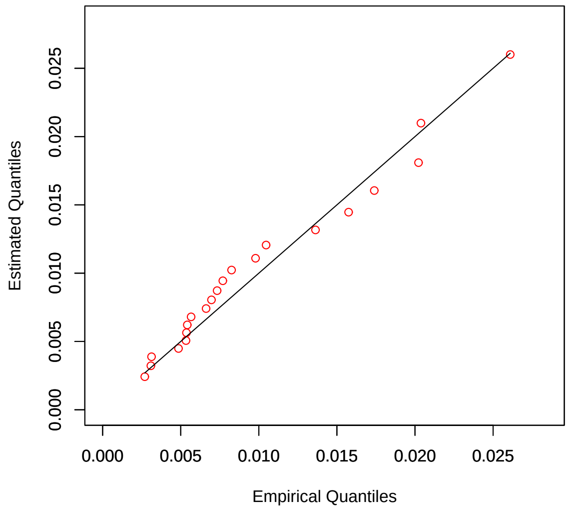

Figure 7: QQ-plot of Moody's data.

$$\begin{aligned} D_i &\stackrel{d}{=} \text{Binomial}(\text{PD}_k(Y)) \\ \text{LGD}_i &\stackrel{d}{=} \text{Beta}(\mathbb{E}(\text{LGD}_k|Z), \sigma_k) \end{aligned} \tag{5.1}$$

for all credits $i = 1, \ldots, m$ and corresponding credit classes $k = 1, \ldots, K$ to obtain a simulated portfolio loss

$$L^{(m)} = \sum_{i=1}^{m} \text{EAD}_i \cdot D_i \cdot \text{LGD}_i.$$

The parameter σ_k models the variances of $\text{LGD}_i|Y$ of obligors in credit class k, it can be estimated by the average of empirical annual LGD variance in that credit class, see e.g. Schuerman (2004) for more details. The portfolio loss distribution function is then obtained by a Monte Carlo simulation as follows:

1. Simulate X, Y.

2. For each sector k: calculate $\text{PD}_k(Y)$ and $\mathbb{E}(\text{LGD}|Z)$, where

$$Z = dY + \sqrt{1 - d^2}X.$$

3. For each obligor i: simulate D_i and LGD_i according to (5.1).

4. Calculate the aggregated loss $L^{(m)}$.

5. Repeat steps 1-4 sufficiently often.

Remark. This algorithm also works in a multifactor framework. Then the first two steps look as follows:

1. Simulate $X, Y_1, \dots, Y_l$.

2. For each sector k: Calculate $Y := \sum_{j=1}^{l} w_j Y_j$ and $Z = dY + \sqrt{1-d^2}X$. Then calculate $\mathrm{PD}_k(Y)$ and $\mathbb{E}(\mathrm{LGD}|Z)$.

□

An alternative to Monte Carlo simulations with a reduced computational effort is calculating a granularity adjustment to the VaR obtained from the model of the preceeding section, see e.g. Giese (2005) for a short description and further references.

6 Conclusion

Default probability and loss given default are features of the portfolio loss with a special stochastic relationship, which – when ignored – can easily lead to biased estimators.

A model for the portfolio loss including dependence of PD and LGD on the economic cycle has to take account for this and should also be transparent in terms of model risk.

Our proposed model integrates the possibility of economic interpretation, a proper statistical setting and easy calibration. Compared to other models, it provides an excellent fit of corporate bond data.

Our model also addresses the observation that the probability of default and the expected LGD do not necessarily vary comonotonously over the economic cycle, but require at least two different factors for an appropriate fit.

Some experts suggest to estimate "downturn LGD" and "downturn PD" separately and obtain a portfolio loss VaR by multiplication (see e.g. [2], a document of the Basel Committee (2004)). Based on the insight achieved from the investigation we cannot support this suggestion. As we need at least two factors to explain the data spread, we suggest to use a complete model for the portfolio loss such as the one presented in this paper.

References

[1] Altman E, Brooks B, Resti A and Sironi A (2003). The link between default and recovery rates: theory, empirical evidence and implications. Available at www.defaultrisk.com.

[2] Basel Committee on Banking Supervision (2004). Background note on LGD quantification. Available at www.bis.org/bcbs.

[3] Bluhm C, Overbeck L, Wagner C (2003). *An Introduction to Credit Risk Modeling.* Chapman & Hall/CRC.

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

[5] Frye J, (2000a). Collateral damage. *RISK.* April, p 91-94.

[6] Frye J, (2000b). Depressing recoveries. Risk November, p 108-111.

[7] Giese G (2005). The impact of PD/LGD correlations on credit risk capital. *RISK.* April, p. 79-84.

[8] Moody's Investors Service (2003). Recovery rates on defaulted corporate bonds and preferred stocks, 1982-2003.

[9] Pykhtin M (2003). Unexpected recovery risk. *RISK.* August, p 74-78.

[10] Tasche D (2004). The Single risk factor approach to capital charges in case of correlated loss given default rates. Available at www.defaultrisk.com.

[11] Schuerman T (2004) What do we know about loss given default? Available at www.defaultrisk.com.