The Journal of Credit Risk (3–20) Volume 8/Number 2, Summer 2012

Two models of stochastic loss given default

Simone Farinelli
Kraus Partner Investment Solutions AG, Militärstrasse 76, 8004 Zurich, Switzerland; email: farinelli@krauspartner.com

Mykhaylo Shkolnikov
Mathematical Sciences Research Institute, 17 Gauss Way, Berkeley, CA 94720, USA; email: mshkolni@gmail.com

We propose two structural models for stochastic loss given default that allow the credit losses of a portfolio of defaultable financial instruments to be modeled. The credit losses are integrated into a structural model of default events accounting for correlations between the default events and the associated losses. We show how the models can be calibrated and analyze the impact of correlations between the occurrences of defaults and recoveries by testing our models for a representative sample portfolio.

1 INTRODUCTION

Many credit risk models assume that loss given default (LGD) is a deterministic proportion of the exposures subject to impairment and ignore the fact that LGD can fluctuate according to the economic cycle. For example, Altman *et al* (2001, 2005) show that default rates and recovery rates are strongly negatively correlated and measure a correlation of 0.75 between yearly average default rate and loss rates in the United States. They provide strong correlation evidence between macro growth variables (such as gross domestic product) and recovery rates and test the impact of correlated defaults and LGD, inferring an understatement of forecasted portfolio losses by up to 30%.

Greening *et al* (2009) show the strong dependence of default rates and recovery rates on the economic cycle for the time period 2000–2009 and detect strong negative correlations between default and recovery rates in various US industries (eg, banking and finance, broadcasting and media) between 2005 and 2009.

The opinions expressed in this document are our own and do not represent the opinions of Kraus Partner Investment Solutions AG. The risk control principles presented are not necessarily used by Kraus Partner Investment Solutions AG. This document does not provide a comprehensive description of concepts and methods used in risk control at Kraus Partner Investment Solutions AG.

An appropriate LGD model should have a reasonable economic interpretation. It should also allow for a calibration by available data and it should be based on a proper statistical setting. In particular, the dependence structure between LGD and default indicators should not arise from a deterministic functional relation.

Frye (2000a,b) proposes a structural model with a systematic risk factor representing the state of the economy and driving both defaults and LGD. The dependence of the default indicator and LGD on the common risk factor gives rise to a strong correlation between the two, which is in line with the empirical evidence. Another single-factor model has been proposed by Tasche (2004) and extended by Pykthin (2003), who unifies the approaches of Frye (2000b) and Tasche (2004). Hillebrand (2006) introduces dependent LGD modeling into a multifactor latent variable framework, providing a good fit to corporate bond data. Marginal distributions for indicator functions and LGD can be specified.

Hamerle *et al* (2007) model default probabilities and LGD jointly using generalized linear mixed-effect models with probit link function and inverse logit function, respectively. All factors are observable, where some represent the general macroeconomic environment and others take obligors' specificities into account.

Inspired by past research in this field, this paper proposes two models for stochastic LGD which are correlated across firms and correlated with occurrences of default events. Both models extend a structural model for the default events to a joint structural model for both defaults and LGD. The models described have the following key features.

- The LGDs are stochastic, correlated with each other and are the occurrences of default events.
- The LGDs follow beta distributions with means estimated from historical data.
- The shapes of the beta distributions vary across firms in such a way that the density function is concave if the corresponding credit instrument is backed by a collateral, and convex otherwise.

The main differences between the two models are as follows.

- In the first model, the parameters of the LGD distributions are random depending on the expected LGD and the risk factors driving the losses in the case of default, whereas, in the second model, the former are deterministic functions of the expected LGD.
- In the first model, the complete joint distribution of the LGD and default indicators can be estimated, whereas, in the second model, the correlations between the risk factors driving defaults and LGD can be fitted exactly. At the same

time, the number of model parameters that have to be estimated coincides for the two models.

We believe that both models are statistically sound and have a reasonable economic interpretation. Moreover, we provide a calibration methodology, which we apply to the available historical data, and we test the models on a representative sample portfolio.

2 A STRUCTURAL MODEL FOR CORRELATED DEFAULTS

Following Pitts (2004), in this section we present a model for the joint equity dynamics of counterparties appearing in a portfolio of financial instruments. Since the default events are triggered by the value of equity crossing a default barrier, this will lead naturally to a structural model for the joint dynamics of defaults. The two models for LGD can then be viewed as attachments to this model, making a joint simulation of default events and LGD possible.

2.1 Joint equity dynamics

Let us consider N firms and introduce the firm index $f = 1, \dots, N$. Furthermore, we assume that these firms are spread over I different industry categories $i = 1, \dots, I$ and R different regions $r = 1, \dots, R$. We denote by i_f and r_f the region category and the industry category, respectively, which the firm f belongs to. Firm f belongs, therefore, to the industry–region cell denoted by:

$$(i_f, r_f) \in \{1, \dots, I\} \times \{1, \dots, R\} =: \mathrm{IR}$$

We assume now that the equity process for all firms $E_t := [E_t^1, \dots, E_t^N]^\dagger$ obeys, in continuous time, the geometric Brownian motion described by the stochastic differential equation:

$$\frac{\mathrm{d}E_t}{E_t} = \mu_t \, \mathrm{d}t + \sigma_t \, \mathrm{d}B_t + \tau_t \, \mathrm{d}W_t \tag{2.1}$$

whereby we have the following.

- The processes $B_t := [B_t^1, \dots, B_t^{\mathrm{IR}}]^\dagger$ and $W_t := [W_t^1, \dots, W_t^N]^\dagger$ are two independent standard multivariate Brownian motions in $\boldsymbol{R}^{\mathrm{IR}\times 1}$ and $\boldsymbol{R}^{N\times 1}$, respectively.
- The functions $t \in [0, +\infty[\ \mapsto \mu_t := [\mu_t^1, \dots, \mu_t^N]^\dagger$, $\sigma_t := [\sigma_t^1, \dots, \sigma_t^N]^\dagger$ and $\tau_t := [\tau_t^1, \dots, \tau_t^N]^\dagger$ are called drift, homoskedastic volatility and heteroskedastic volatility, respectively. More precisely, by homoskedasticity, we mean that, for all times t and all firms f, we have $\sigma_t^f = \sigma_t^{(i_f, r_f)}$, while, in the heteroskedastic case, the volatility components explicitly depend on the firm, that is, $\tau_t = \tau_t^f$ cannot be written as $\tau_t = \tau_t^{(i_f, r_f)}$.

- Products and fractions are understood componentwise.

The time discretization of the stochastic differential equation leads to the panel model:

$$y_t = \alpha_t + \beta_t + \varepsilon_t \tag{2.2}$$

in which the following components are represented by the corresponding (possibly random) vectors in $\boldsymbol{R}^{N\times 1}$.

Log returns for the firm equities: $y_t := \log E_t - \log E_{t-1}$.

Deterministic firm-specific components: $\alpha_t := \mu_t - \frac{1}{2}(\sigma_t^2 + \tau_t^2)$.

Industry–region systematic components: $\beta_t := \sigma_t(B_t - B_{t-1}) \sim \mathcal{N}(0, \operatorname{diag}(\sigma_t^2))$.

Firm-specific idiosyncratic components: $\varepsilon_t := \tau_t(W_t - W_{t-1}) \sim \mathcal{N}(0, \operatorname{diag}(\tau_t^2))$.

The log return of the asset value y_t is therefore decomposed into a deterministic part α_t, a stochastic homoskedastic part β_t and an independent heteroskedastic part ε_t. Note that the multiplication and squaring operations applied to vectors are meant componentwise.

We concentrate now on the systematic industry–region components which will be linked with the LGD below. The former can be written as:

$$\beta_t = b_t \gamma_t + v_t \tag{2.3}$$

by setting:

$$\gamma_t \stackrel{\mathrm{d}}{=} \mathcal{N}(0, I) \tag{2.4}$$

$$v_t \stackrel{\mathrm{d}}{=} \mathcal{N}(0, \chi_t^2) \tag{2.5}$$

independently of each other, independently and identically distributed over time, and letting $b_t^{(i,r)} \in \boldsymbol{R}$ be the industry–region "beta" coefficient. The process γ_t^r can be interpreted as the region performance of log-asset returns. The random variables in the components of the vector γ_t can be defined as principal components of β_t, and v_t is a residual quantity. The industry-specific effects are homoskedastic within a region, that is, they have the same variance $(\chi_t^r)^2$. Denoting by:

$$\rho_t^{r_1,r_2} := \operatorname{cov}_{t-1}^{\mathrm{stat}}(\gamma_t^{r_1}, \gamma_t^{r_2}) \tag{2.6}$$

the historical covariance available at time t, we set the covariance between the industry–region effects to the statistical covariance of the latter up to time $t-1$:

$$\operatorname{cov}_{t-1}^{\mathrm{stat}}(\beta_t^{(i_1,r_1)}, \beta_t^{(i_2,r_2)}) = \rho_t^{i_1,i_2} b_t^{(i_1,r_1)} b_t^{(i_2,r_2)} + \chi_t^{r_1}\chi_t^{r_2}\delta^{i_1,i_2}\delta^{r_1,r_2} \tag{2.7}$$

2.2 Joint default dynamics

Assuming that the joint equity dynamics is normalized in such a way that a default of firm f occurs if the value of equity goes below 1, and denoting by X_t^f the default indicator process of firm f, the marginal conditional default probabilities read:

$$\begin{aligned}\mathbb{E}_t[X_{t+1}^f \mid X_t^f = 0] &= \mathbb{P}_t[\log E_{t+1}^f \leqslant 0 \mid \log E_t^f > 0] \\ &= \mathbb{P}_t[G_{t+1}^f \leqslant g_t^f \mid \log E_t^f > 0] \\ &= \Phi(g_t^f)\end{aligned}$$

where we have defined:

$$G_t^f = \frac{\beta_t^{(i_f,r_f)} + \varepsilon_t^f}{\sqrt{(\sigma_t^{(i_f,r_f)})^2 + (\tau_t^f)^2}} \sim \mathcal{N}(0,1) \tag{2.8}$$

$$g_t^f = \frac{-\log E_t^f - \mu_{t+1}^f + \frac{1}{2}((\sigma_{t+1}^{(i_f,r_f)})^2 + (\tau_{t+1}^f)^2)}{\sqrt{(\sigma_{t+1}^{(i_f,r_f)})^2 + (\tau_{t+1}^f)^2}} \tag{2.9}$$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \mathrm{d}u\, \mathrm{e}^{-u^2/2} \tag{2.10}$$

The joint distribution of $\{G_t^f\}_{f,t}$ can be simulated sequentially as a multivariate Gaussian distribution with vanishing conditional expectation and conditional covariance matrix having ones on the diagonal and off-diagonal entries:

$$\mathrm{cov}_{t-1}^{\mathrm{stat}}(G_t^{f_1}, G_t^{f_2}) = \mathrm{cov}_{t-1}^{\mathrm{stat}}\left(\frac{\beta_t^{(i_{f_1},r_{f_1})}}{\sqrt{(\sigma_t^{(i_{f_1},r_{f_1})})^2 + (\tau_t^{f_1})^2}}, \frac{\beta_t^{(i_{f_2},r_{f_2})}}{\sqrt{(\sigma_t^{(i_{f_2},r_{f_2})})^2 + (\tau_t^{f_2})^2}}\right)$$

Therefore, to simulate defaults, the default probabilities p_t^f from the macroeconomic model are inverted to $c_t^f = \Phi^{-1}(p_t^f)$ and the vector-valued random variable $\{G_t^f\}_{f,t}$ is simulated. A simulated default event for firm f occurs at the first time for which G_t^f falls below c_t^f.

3 MARKET AND CREDIT EXPOSURES AND LOSS GIVEN DEFAULT

We now present a formal definition for market and credit exposures and LGD. To this end, let us consider a financial instrument at time $t = 0, 1, \dots, T$, and assume that it is a_t-rated, for a rating $a_t \in \mathcal{J} = \{1, 2, \dots, J\}$, which is valid for the interval $]t-1, t]$. The financial instrument, or product, is identified at time t with the stochastic discrete cashflow stream $\{C_s\}_{s \geqslant t}$ that it generates from time t onward. In particular, it accounts for future possible defaults or rating migrations, but not for recovery streams. When we evaluate it to determine its market value at time t, we need to use the a_t-term structure of interest rates. Assuming, for the moment, that there is only one currency, and after having denoted discount factors by $d^{a_t} = d^{a_t}_{t,s}$ for $s \geqslant t$ and short rates by $r^{a_t} = r^{a_t}_t$, we can write the present value of the product at time t as:

$$\begin{aligned}\mathrm{PV}_t(C; d^{a_t}) &= \sum_{s \geqslant t} \mathbb{E}^*_t\biggl[\exp\biggl(-\int_t^s \mathrm{d}u\, r^{a_t}_u\biggr) C_s\biggr] \\ &= \sum_{s \geqslant t} d^{a_t}_{t,s} \mathbb{E}^{*,s}_t[C_s]\end{aligned}$$

Thus, $\mathbb{E}^*_t$ denotes the risk-neutral conditional expectation and $\mathbb{E}^{*,s}_t$ for $s \geqslant t$ denotes the s-forward neutral conditional expectation. We know that risk-neutral measure(s) and forward neutral measure(s) exist by virtue of the fundamental theorem of asset pricing (see Björk (2004, Chapters 10.2, 10.3, 24.4)). The present value of the product is its theoretical price and represents an approximation of its market value in the real-world market.

If a default occurs at time t for the state of nature $\omega \in \Omega$, then the product cashflows are annihilated, that is, $C_s(\omega) = 0$ for all $s \geqslant t$, and there possibly exist recovery cashflows $\{C^{\mathrm{rec}}_s(\omega)\}_{s \geqslant t}$, which will mitigate the loss. These must be evaluated with respect to the government term structure, which is considered to be free of default risk. Therefore, an approximation for the market recovery value is the theoretical price of the recovery cashflow stream:

$$\begin{aligned}\mathrm{PV}_t(C^{\mathrm{rec}}; d^{\mathrm{gov}}) &= \sum_{s \geqslant t} \mathbb{E}^*_t\biggl[\exp\biggl(-\int_t^s \mathrm{d}u\, r^{\mathrm{gov}}_u\biggr) C^{\mathrm{rec}}_s\biggr] \\ &= \sum_{s \geqslant t} d^{\mathrm{gov}}_{t,s} \mathbb{E}^{*,s}_t[C^{\mathrm{rec}}_s]\end{aligned}$$

Exposure at market and credit risk is then defined as:

$$\begin{aligned}\mathrm{EXP}_t(C) &:= \mathrm{PV}_t(C; d^{a_t}) \\ &= \sum_{s \geqslant t} d^{a_t}_{t,s} \mathbb{E}^{*,s}_t[C_s]\end{aligned} \tag{3.1}$$

Note that this exposure definition does not presume that the counterparty is in the default state at time t or later, but covers all possible future states (the default and the nondefault ones); it is just the present value of the cashflow stream for all possible future states. Now we can formally define the LGD at time t for the financial instrument as:

$$\begin{aligned} \mathrm{LGD}_t(C; C^{\mathrm{rec}}) &:= 1 - \frac{\mathrm{PV}_t(C^{\mathrm{rec}}; d^{\mathrm{gov}})}{\mathrm{PV}_t(C; d^{a_t})} \\ &= 1 - \frac{\sum_{s \geqslant t} d_{t,s}^{\mathrm{gov}} \mathbb{E}_t^{*,s}[C_s^{\mathrm{rec}}]}{\sum_{s \geqslant t} d_{t,s}^{a_t} \mathbb{E}_t^{*,s}[C_s]} \end{aligned} \tag{3.2}$$

This definition is consistent with the definition for the recovery of market value (see Lando (2004, Chapter 5.7) and Schönbucher (2003, Chapter 6)).

REMARK 3.1 (Industry standards for traditional credit risk management) The definitions for LGD and exposure presented here differ from the industry credit risk standards, where the following apply.

- Exposure is typically termed exposure-at-default and is understood as conditional expectation of the theoretical value in the case of default.
- Loss given default is typically understood as a fraction.
- Exposures are always positive. Typically, for structured products, this is enforced by taking only the positive part of the distribution of the theoretical values.
- For a loan or a bond, the exposure is typically defined in nominal terms (in other words, as the issued amount).
- For Lombard credit, the exposure does not require netting of collateral, which needs to be treated separately.

Of course, industry standards have their fundament. They service the book-value accounting perspective, which segregates credit portfolio profits (cash inflows) from losses (cash outflows). In the traditional credit risk management approach, only losses are considered, and these in nominal terms. For a book of plain-vanilla loans, bonds or mortgages, these approaches can be reconciled with a mark-to-market valuation by modeling the cash inflows. But for structured products like those from investment banking, the requirement that exposure must be positive impedes diversification. It leads to a conservative overestimation of credit risk, which, for a risk manager, is reassuring on the one hand but annoying on the other, since it binds more risk capital than is effectively necessary. We believe that our definition is the appropriate one to provide a fair valuation for all products and to allow for the integration of market and credit risk. As a matter of fact, using the notation introduced above, the value of a portfolio of N financial instruments at time t can be written as:

$$V_t = \sum_{j=1}^{N} (1 - X_t^j \, \mathrm{LGD}_t(C^j; C^{j,\mathrm{rec}})) \, \mathrm{EXP}_t(C^j)$$

whereas X_t^j denotes the default indicator at time t for the jth financial instrument. This formula shows how the portfolio value depends on both market risk and credit risk factors. Exposures depend on market risk factors, default indicators depend on credit risk factors, and LGDs depend on both.

4 LOSS GIVEN DEFAULT MODELS

In this section we explain two models for LGD that allow for a joint simulation of default events and LGD. Thus, not only are the default events correlated, but the correlations between default indicators and LGD are included as well. For both models, we first present the theoretic framework and then explain how the model can be calibrated and used for simulations of the credit loss of a portfolio.

4.1 Model A

4.1.1 Theoretic framework

We propose a structural model for LGD which is connected to the default times model through correlations between the risk factors in the two models. Following the recent literature, we make the assumption that the LGD of the financial instrument f in the industry–region cell (i, r) at time t follows a beta distribution $\mathrm{Beta}(\mu_t^f, \nu_t^f)$. The corresponding parameters are modeled as functions of the risk factors driving the losses occurring at defaults. These risk factors fall into two categories:

- systematic risk factors $Y_t^{(i,r)}$ characteristic for an industry–region cell (i, r) at time t;
- macroeconomic risk factor $\bar{Y}_t$ varying with the economic cycle and common to all industry–region cells.

Moreover, we assume that the distribution of LGDs of counterparties of industry i in region r depends on the value of the combined risk factor:

$$Z_t^{(i,r)} = \eta_t^{(i,r)} Y_t^{(i,r)} + \sqrt{1 - (\eta_t^{(i,r)})^2}\,\bar{Y}_t \tag{4.1}$$

and work directly with the latter for the purposes of parameter estimation and simulation. The risk factors $Z_t^{(i,r)}$ are assumed to be jointly normally distributed with the credit risk factors specified in Section 2, with the only nontrivial covariances being:

$$\operatorname{cov}(Z_t^{(i,r)}, Z_t^{(i',r')}) = \theta^{i,r,i',r'} \tag{4.2}$$

$$\operatorname{cov}(Z_t^{(i,r)}, \beta_t^{(i',r')}) = \psi^{i,r,i',r'} \tag{4.3}$$

It remains to specify the dependence of the parameters μ_t^f, ν_t^f of the beta distribution on the risk factor $Z_t^{(i,r)}$. We note first that, since the mean of the beta distribution

$\text{Beta}(\mu, \nu)$ is given by $\mu/(\mu+\nu)$, it suffices to specify μ_t^f as a function of $Z_t^{(i,r)}$, since the value of ν_t^f is then automatically determined after a value for the expected LGD is prescribed. Since $Z_t^{(i,r)}$ reflects the regional- and industry-type specificities as well as the economic cycle, we are quite free in our choice of a functional dependence between the stochastic parameter μ_t^f and the stochastic factor $Z_t^{(i,r)}$, provided this dependence is given by a bijective function. Later, when calibrating the model, different choices of the function will lead to different covariance parameters θ and ψ. Since $Z_t^{(i,r)}$ is Gaussian and μ_t^f always assumes strictly positive values, we can choose:

$$\mu_t^f = \mathrm{e}^{Z_t^{(i,r)}} \tag{4.4}$$

since the exponential is a bijective function from the real axis to the positive reals. The mean of the LGD distribution m_t^f is already given as the result of estimation from historical LGD data for the industry–region cell (i_f, r_f) or as the result of expert judgment based on information about the corresponding counterparty. Finally, to ensure that the mean of the LGD distribution is matched, we set:

$$\nu_t^f = \mu_t^f \frac{1 - m_t^f}{m_t^f} \tag{4.5}$$

4.1.2 Parameter estimation and simulations

We assume that the model for default times is already implemented and that the time series of the risk factors therein are already estimated. In this section we propose a simple and quick-to-implement estimation procedure that allows the joint modeling of the default times and the LGD according to the previous sections. The estimation procedure is composed of the following three steps.

(1) For each industry–region cell (i, r) at time t, we compute the maximum likelihood estimate of the parameter μ_t^f (taking the same value for all firms f belonging to the same cell).

(2) We obtain a time series for the combined LGD risk factor for each industry–region cell (i, r) up to time t from the estimates of the parameters μ_t^f by the formula:

$$Z_t^{(i,r)} = \log \mu_t^f \tag{4.6}$$

(3) Lastly, we use the time series for $\beta_t^{(i,r)}$ and $Z_t^{(i,r)}$ to obtain an estimate on the covariances:

$$\operatorname{cov}(Z_t^{(i,r)}, \beta_t^{(i',r')})\colon\ 1 \leqslant i, i' \leqslant I,\ 1 \leqslant r, r' \leqslant R$$
$$\operatorname{cov}(Z_t^{(i,r)}, Z_t^{(i',r')})\colon\ 1 \leqslant i, i' \leqslant I,\ 1 \leqslant r, r' \leqslant R$$

The LGD distributions can be modeled consecutively in the following two steps.

(1) At each point of time t, model the LGD risk factors $Z_t^{(i,r)}$ jointly with the credit risk factors $\beta_t^{(i,r)}$, ε_t^f as a multivariate normal random variable with the previously estimated covariance structure.

(2) Set the parameters of the LGD distribution for the financial instrument f in cell (i, r) at time t by the formulas:

$$\mu_t^f = \mathrm{e}^{Z_t^{(i,r)}} \tag{4.7}$$

$$\nu_t^f = \mu_t^f \frac{1 - m_t^f}{m_t^f} \tag{4.8}$$

where m_t^f is again the estimated or prescribed expected LGD.

4.2 Model B

In this second model, LGDs will be also modeled as random variables assuming that the expectation of the LGD conditioned on a default event affecting the corresponding company:

$$\overline{\mathrm{LGD}}_t^f = \mathbb{E}[\mathrm{LGD}_t^f \mid X_t^f = 1] \tag{4.9}$$

has already been determined by estimation or expert judgment. In previous work (see Lévy (2008) and Schönbucher (2003)), the LGDs (conditional on a default event) are assumed to be beta-distributed:

$$\mathrm{LGD}_t^f \mid (X_t^f = 1) \sim \mathrm{Beta}(\mu_t^f, \nu_t^f) \tag{4.10}$$

where:

$$\mu_t^f = (\kappa - 1)\overline{\mathrm{LGD}}_t^f \tag{4.11}$$

$$\nu_t^f = (\kappa - 1)(1 - \overline{\mathrm{LGD}}_t^f) \tag{4.12}$$

Note that, for any choice of the shape parameter κ, the desired expectation is matched. For the special choice $\kappa := 4$, it was believed (on the basis of apparently persuasive previous research that crystallized to a consensus in the literature) that the density of the LGD distribution has a concave shape. This turns out to be true only if the expected LGD is near 50%. However, the density is no more concave if the latter is close to 100% or to 0%. If the expected LGDs are close to 0% or to 100%, then the corresponding probability density near 1 and 0 is not negligible.

We assume as before that the conditional expectation $\overline{\mathrm{LGD}}_t^f$ of the LGD event is our best guess for the average loss rate in the case of a default event. Moreover, we recall that, in the case of a collateralized loan, the empirical LGD distribution tends to have a concave shape, putting most of the weight on a small interval around the expectation. This is, of course, explained by the nontrivial recovery values. For unsecured loans, the empirical LGD distribution tends to have a convex form, accounting for frequent LGD values close to 1. Therefore, we propose to model the density as a symmetric function with respect to the expectation. To this end, we apply a linear transformation to a beta-distributed random variable $B \sim \text{Beta}(\mu, \nu)$:

$$\text{LGD}_t^f \mid (X_t^f = 1) = (\overline{\text{LGD}}_t^f - \delta_t^f)B + (\overline{\text{LGD}}_t^f + \delta_t^f)(1 - B) \tag{4.13}$$

where:

$$\mu = \nu = 2, \qquad \delta_t^f = 0.2 \min(\overline{\text{LGD}}_t^f, 1 - \overline{\text{LGD}}_t^f) \tag{4.14}$$

in the case where the financial instrument f is backed by collateral, and:

$$\mu = \nu = 0.5, \qquad \delta_t^f = \min(\overline{\text{LGD}}_t^f, 1 - \overline{\text{LGD}}_t^f) \tag{4.15}$$

otherwise. This guarantees the symmetry of the probability density function with an appropriate support in a symmetric interval $[\overline{\text{LGD}}_t^f - \delta_t^f, \overline{\text{LGD}}_t^f + \delta_t^f]$ around the expectation. This choice reflects our knowledge (and ignorance) about the LGD: we have a best guess given by its expectation but no opinion about its skewness. Therefore, a symmetric distribution supported by a neighborhood of the expectation is a legitimate choice.

To specify the joint distribution of default events and LGD, recall that the standardized company equity return conditional on default is truncated Gaussian distributed. More precisely:

$$\begin{aligned} G_t^f \mid (X_t^f = 1) &= G_t^f \mid (G_t^f < \Phi^{-1}(p_t^f)) \\ &\sim F_{G_t^f \mid (G_t^f < \Phi^{-1}(p_t^f))} \end{aligned}$$

where:

$$F_{G_t^f \mid (G_t^f < \Phi^{-1}(p_t^f))}(x) = \frac{\Phi(x)}{p_t^f} \tag{4.16}$$

for all $x < p_t^f$, and 1 otherwise. Therefore, assuming that the simulation of standardized equity returns has already occurred, we simulate LGD in such a manner that it follows the specified marginal distributions and that correlations with the standardized equity returns match historical values. More exactly, we set:

$$\text{LGD}_t^f = F^{-1}_{\text{LGD}_t^f \mid (X_t^f = 1)}(F_{H_t^f}(H_t^f)) \tag{4.17}$$

where:

$$H_t^f = F_{G_t^f \mid (G_t^f < \Phi^{-1}(p_t^f))}(G_t^f \mid (G_t^f < \Phi^{-1}(p_t^f)))(1 + \xi_t^f(V_t^f - \tfrac{1}{2}))$$

Thus,

- $F_{\mathrm{LGD}_t^f \mid (X_t^f = 1)}$ is the cumulative probability distribution function of the marginal specified above,
- V_t^f is a [0, 1]-valued uniformly distributed random variable, independent of all G_t^f,
- ξ_t^f is a parameter which can be chosen in such a manner that the covariance between:

$$F_{\mathrm{LGD}_t^f \mid (X_t^f = 1)}(\mathrm{LGD}_t^f \mid (X_t^f = 1))$$

and:

$$F_{G_t^f \mid (G_t^f < \Phi^{-1}(p_t^f))}(G_t^f \mid (G_t^f < \Phi^{-1}(p_t^f)))$$

attains its historical value.

Thus, H_t^f can be viewed as the standardized company equity return, conditioned on the occurrence of a default event, transformed to the uniform distribution and perturbed by the auxiliary LGD risk factor V_t^f. The distribution function of the random variable H_t^f is given by the following lemma.

LEMMA 4.1 *Let $\xi \in\,]2, \infty[$ and let U and V be two standard uniformly distributed independent random variables. Then:*

$$H = U(1 + \xi(V - \tfrac{1}{2}))$$

has support in $[1 - \frac{1}{2}\xi, 1 + \frac{1}{2}\xi]$ and distribution function:

$$F_H(y) = \begin{cases} \dfrac{1}{2} - \dfrac{1}{\xi} + \dfrac{y}{\xi} + \dfrac{y}{\xi}(\log(\tfrac{1}{2}\xi - 1) - \log|y|), & 1 - \tfrac{1}{2}\xi < y < 0 \\ \dfrac{1}{2} - \dfrac{1}{\xi} + \dfrac{y}{\xi} + \dfrac{y}{\xi}(\log(1 + \tfrac{1}{2}\xi) - \log y), & 0 < y < 1 + \tfrac{1}{2}\xi \end{cases}$$

PROOF Since U and V are standard uniformly distributed, their probability density functions are the indicator functions for the interval [0, 1]. Furthermore, since U and V are independent, their joint probability density function is the indicator function of the square $[0, 1]^2$. It follows that:

$$\begin{aligned} F_H(y) &= P[H \leqslant y] \\ &= \int_{\{u(1+\xi(v-(1/2)))\leqslant y\}} \mathrm{d}u\, \mathrm{d}v\, \mathbf{1}_{[0,1]^2}(u, v) \end{aligned} \tag{4.18}$$

Integration completes the proof. □

The choice of the parameters ξ_t^f can be implemented by matching the historical covariance between:

$$F_{\mathrm{LGD}_t^f \mid (X_t^f = 1)}(\mathrm{LGD}_t^f \mid (X_t^f = 1))$$

and:

$$F_{G_t^f \mid (G_t^f < \Phi^{-1}(p_t^f))}(G_t^f \mid (G_t^f < \Phi^{-1}(p_t^f)))$$

The equation connecting the two quantities is specified by the following lemma.

LEMMA 4.2 *The parameters ξ_t^f fulfill:*

$$\xi_t^f = \frac{10 + 8\sqrt{1 + 54\lambda_t^f}}{288\lambda_t^f - 3}, \quad \xi_t^f > 2 \tag{4.19}$$

where:

$$\lambda_t^f = \mathrm{cov}(F_{G_t^f \mid (G_t^f < \Phi^{-1}(p_t^f))}(G_t^f \mid (G_t^f < \Phi^{-1}(p_t^f))), \\ F_{\mathrm{LGD}_t^f \mid (X_t^f = 1)}(\mathrm{LGD}_t^f \mid (X_t^f = 1)))$$

REMARK 4.3 (Calibration of model B) The model can be calibrated by setting the value of λ_t^f to the historical covariance of the appropriate quantities in the industry–region cell (i_f, r_f) before time t, ie:

$$\mathrm{cov}_{t-1}^{i_f, r_f, \mathrm{stat}}(F_{G_t^f \mid (G_t^f < \Phi^{-1}(p_t^f))}(G_t^f \mid (G_t^f < \Phi^{-1}(p_t^f))), \\ F_{\mathrm{LGD}_t^f \mid (X_t^f = 1)}(\mathrm{LGD}_t^f \mid (X_t^f = 1)))$$

The parameters ξ_t^f which then depend only on the particular industry–region cell (i_f, r_f) can be subsequently obtained by using the equation in the lemma. Note that, by the Cauchy–Schwarz inequality, it holds that:

$$\lambda_t^f \leqslant \tfrac{1}{12} = 0.083 =: \lambda_{\max} \tag{4.20}$$

with equality if and only if the two random variables are perfectly correlated. Hence, we expect that the estimates for λ_t^f lie in $[\frac{1}{7}\lambda_{\max}, \frac{2}{3}\lambda_{\max}]$, in which case $\xi_t^f > 2$ and the equation in the lemma can be applied. If the estimate falls in one of the intervals $[-\lambda_{\max}, \frac{1}{7}\lambda_{\max}[$ or $]\frac{2}{3}\lambda_{\max}, \lambda_{\max}]$, we should replace it by $\frac{1}{7}\lambda_{\max}$ or $\frac{2}{3}\lambda_{\max}$, respectively.

PROOF OF LEMMA 4.2 For ease of notation we drop all indices. We then have:

$$F_{\mathrm{LGD} \mid (X=1)}(\mathrm{LGD} \mid (X = 1)) = F_H(H) = F_H(U(1 + \xi(V - \tfrac{1}{2})))$$

where we have introduced the random variable:

$$U = F_{G\mid(G<\Phi^{-1}(p))}(G \mid (G < \Phi^{-1}(p))$$

so that U and V satisfy the assumptions of Lemma 4.1. Therefore:

$$\begin{aligned}\lambda &= \operatorname{cov}(F_{\mathrm{LGD}\mid(X=1)}(\mathrm{LGD} \mid (X = 1)), F_{G\mid(G<\Phi^{-1}(p))}(G \mid (G < \Phi^{-1}(p))) \\ &= \operatorname{cov}(F_H(H), U) \\ &= \operatorname{cov}(F_H(U(1 + \xi(V - \tfrac{1}{2}))), U)\end{aligned}$$

which can be explicitly computed, because the joint density of U, V is the indicator function of the unit square, and F_H is known from Lemma 4.1. The result of the computation is the expression for ξ displayed in the lemma statement. □

5 IMPACT ANALYSIS FOR A SAMPLE PORTFOLIO

5.1 Portfolio

We consider an invented portfolio of approximately 17 000 firms distributed across forty rating classes, nine industry types and fourteen world regions. Rating class 1 corresponds to the best possible creditworthiness, while a firm displaying rating 40 is in the default state. The world regions and industry types considered are shown in Table 1 on the facing page and Table 2 on the facing page, respectively.

To describe the portfolio, we show the repartition of expected potential losses at time 0 for the one-year default horizon with respect to rating (see Figure 1 on page 18), industry (see Figure 2 on page 18) and region (see Figure 3 on page 19). The potential loss at time t for a firm with EXP_t and LGD_t is defined as $\mathrm{PTL}_t := \mathrm{LGD}_t\,\mathrm{EXP}_t$.

5.2 Calibration, simulation and numerical results

To calibrate the parameters of the default model, we utilize the Moody's KMV asset values database covering approximately 30 000 companies. To calibrate the parameters of the LGD models, we employed the Moody's KMV recovery database, from which we extracted LGDs of approximately 5000 defaults that had already occurred. We considered the 2001–6 time period with monthly quotes.

The models for both defaults and LGD presented in the preceding sections have been implemented as a Monte Carlo simulation in MATLAB. By construction, the correlation between simulated LGD and default indicators corresponds to the historical values, and LGDs are beta-distributed. The resulting statistics for the yearly portfolio credit loss for 500 000 simulations can be found in Table 3 on page 19. Thus, in the deterministic LGD model, the stochastic LGD of models A and B are replaced by their respective deterministic means. The "EL" column refers to the expected loss of

TABLE 1 World regions.

Region	Countries
1	Australia and New Zealand
2	Hong Kong, Korea, Malaysia, Singapore, Thailand and Taiwan
3	Rest of Asia
4	Japan
5	European Union
6	Switzerland
7	Rest of Europe
8	United Kingdom
9	Offshore jurisdictions
10	Argentina, Brazil and Chile
11	Rest of Latin America
12	Unites States and Canada
13	Oil
14	Least-developed countries

TABLE 2 Industry types.

Industry	Activity
1	Banks, insurance companies and other financials
2	Manufacturing, including energy and mining
3	Services
4	Health care
5	Real estate
6	Technology
7	Utilities
8	Government
9	Private customers

the portfolio, the columns Q_{90}, Q_{95}, Q_{99}, $Q_{99.95}$ and $Q_{99.98}$ display several value-at-risk quantiles of the portfolio, while the columns ETL_{90}, ETL_{95}, ETL_{99}, $ETL_{99.95}$ and $ETL_{99.98}$ show the expected tail losses above the respective quantiles.

We remark that, at least for this example, there are small tangible differences between the output of models A and B on the one hand and deterministic LGD on the other. This effect is probably due to the dominance of well-rated companies in the sample portfolio. More differences in the tails can probably be obtained by mod-

FIGURE 1 Potential loss versus rating (Swiss francs).

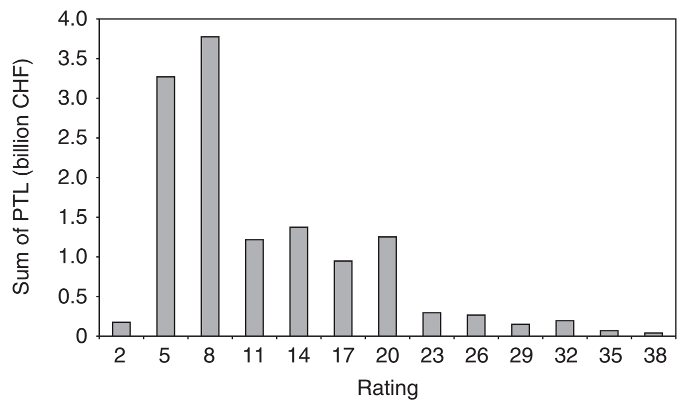
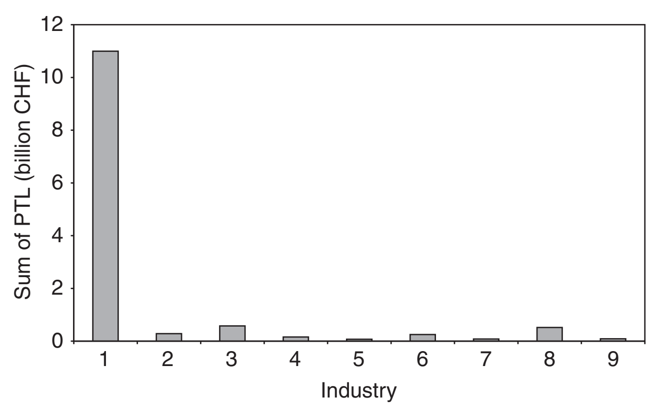


FIGURE 2 Potential loss versus industry (Swiss francs).



eling conditional expectations of LGD (conditional on default events) which strongly vary with respect to the absolute expected LGD. In both model A and model B, the expected conditional LGDs are frozen to their absolute expectations and stochasticity is induced by second-order and higher conditional moments differing from the absolute moments. This will be the subject of future research.

FIGURE 3 Potential loss versus region (Swiss francs).

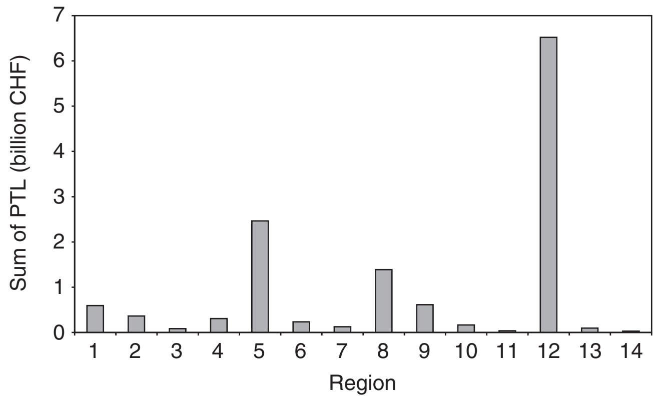


TABLE 3 Loss statistics (million CHF).

(a) Expected loss and value-at-risk quantiles

Model	EL	Q_{90}	Q_{95}	Q_{99}	$Q_{99.95}$	$Q_{99.98}$
Deterministic LGD	81.51	155.81	198.28	315.43	596.67	685.68
Model A	81.46	155.72	198.64	315.91	596.58	678.50
Model B	81.86	156.59	199.42	316.77	596.68	677.32

(b) Expected tail losses

Model	ETL_{90}	ETL_{95}	ETL_{99}	$ETL_{99.95}$	$ETL_{99.98}$
Deterministic LGD	223.47	272.34	402.54	686.08	769.06
Model A	223.43	272.23	401.17	678.56	753.33
Model B	224.19	273.11	402.64	678.80	753.78

REFERENCES

Altman, E. I., Resti, A., and Sironi, A. (2001). Analyzing and explaining default recovery rates. Report submitted to the International Swaps and Derivatives Association (December).

Altman, E. I., Brady, B., Resti, A., and Sironi, A. (2005). The link between default and recovery rates: theory, empirical evidence, and implications. *Journal of Business* **78**(6), 2203–2228.

Björk, T. (2004). *Arbitrage Pricing in Continuous Time*, 2nd edition. Academic Press, New York.

Frye, J. (2000a). Collateral damage: a source of systematic credit risk. *Risk* **13**(4), 91–94.

Frye, J. (2000b). Depressing recoveries. *Risk* **13**(11), 108–111.

Greening, T., Oline, M., Rosenthal, E., and Verde, M. (2009). Defaults surge, recoveries sink in 2009: understanding the fundamental and cyclical drivers of corporate recovery rates. Credit Market Research report, FitchRatings (July).

Hamerle, A., Knapp, M., and Wildenauer, N. (2007). Default and recovery correlations: a dynamic econometric approach. *Risk* **20**(1), 100–105.

Hillebrand, M. (2006). Modeling and estimating dependent loss given default. *Risk* **19**(9), 120–125.

Lando, D. (2004). *Credit Risk Modeling: Theory and Applications*. Princeton Series in Finance. Princeton University Press.

Lévy, A. (2008). An overview of modeling credit portfolios. Moody's KMV. URL: www.moodyskmv.com/research/files/wp/overview_modeling_credit_portfolios.pdf.

Pitts, A. (2004). Correlated defaults: let's go back to the data. *Risk* **17**(6), 75–79.

Pykthin, M. (2003). Unexpected recovery risk. *Risk* **16**(4), 74–78.

Schönbucher, P. H. (2003). *Credit Derivatives Pricing Models: Models, Pricing and Implementation*. Wiley Finance.

Tasche, D. (2004). The single risk factor approach to capital charges in case of correlated loss given default rates. Working Paper, Deutsche Bundesbank.