The Journal of Credit Risk (1–26) Volume 7/Number 4, Winter 2011/12

Generalized beta regression models for random loss given default

Xinzheng Huang
Group Risk Management, Rabobank Nederland, Graadt van Roggenweg 400,
3531 AH Utrecht, The Netherlands; email: huang.x.z@gmail.com

Cornelis W. Oosterlee
CWI Center for Mathematics and Computer Science,
Amsterdam, the Netherlands
and
Delft Institute of Applied Mathematics, Delft University of Technology,
The Netherlands; email: c.w.oosterlee@cwi.nl

Please also provide full postal addresses for the CWI Center and Delft Institute affiliations (we require post codes and relevant road names if possible).

We describe a framework for modeling systematic risk in loss given default in the context of credit portfolio losses. The class of models is very flexible and accommodates skewness and heteroskedastic errors well. The inference of models in this framework can be unified. Moreover, it allows efficient numerical procedures, such as the normal approximation and the saddlepoint approximation, to calculate the portfolio loss distribution, value-at-risk and expected shortfall.

1 INTRODUCTION

In the context of credit portfolio losses, loss given default (LGD) is the proportion of the exposure that will be lost if a default occurs. Uncertainty regarding the actual LGD is an important source of credit portfolio risk in addition to default risk. In practice, in both CreditMetrics (Gupton *et al* (1997)) and KMV Portfolio Manager (Gupton and Stein (2002)), for example, the uncertainty in the LGD rates of defaulted obligors is assumed to be a beta random variable independent for each obligor. The beta distribution is well-known to be very flexible, modeling quantities constrained in the interval [0, 1]. Depending on the choice of parameters, the probability density function can be unimodal, U-shaped, J-shaped or uniform.

However, extensive empirical evidence (see, for example, Hu and Perraudin (2002) and Altman *et al* (2005)) has shown this simple approach to be insufficient. It is now well-understood that LGD is positively correlated with the default rate. In other words, the LGD is high when the default rate is high, which suggests that systematic risk

Xinzheng Huang would like to thank Dr. Mâcé Mesters from Rabobank for helpful discussions.

exists in LGD just as it does in default rates. A heuristic justification is that the LGD is determined by the collateral value, which is sensitive to the state of the economy.

Based on the results of a nonparametric estimation procedure, Hu and Perraudin (2002) further showed that, without taking the correlation between probability of default (PD) and LGD into account, the economic capital, or value-at-risk (VaR), of a loan portfolio can be significantly underestimated. This has important consequences for risk management practice. In the Basel II Accord this issue is addressed by the notion of "downturn LGD".

The interest in LGD being subject to systematic risk dates back to Frye (2000), where LGD is modeled using a normal distribution. An obvious problem with this model is that it allows the LGD to be negative, which cannot be the case. To ensure the nonnegativity of LGD, Pykhtin (2003) employs a truncated lognormal distribution for the LGD. Andersen and Sidenius (2004) propose the use of a probit transform of the LGD such that the transformed LGD is normally distributed. The probit transformation guarantees that the LGD stays in the interval $[0, 1]$. In a similar manner, Düllmann and Trapp (2004) and Rösch and Scheule (2005) employ a logit transform of the LGD. In contrast with the above approaches, Giese (2006) and Bruche and González-Aguado (2008) extend the static beta distribution assumption in CreditMetrics and the KMV Portfolio Manager by modeling the LGD as a mixture of beta distributions that depend on the systematic risk.

In this paper we describe the concept of a generalized beta regression (GBR) framework for modeling LGD. This framework generalizes the beta regression model proposed by Ferrari and Cribari-Neto (2004) and is very similar to a class of models derived from generalized linear models (GLMs). Our models are called generalized beta regression models since the LGD is always assumed to be (conditionally) beta distributed. The models by Giese (2006) and Bruche and González-Aguado (2008) can be regarded as special examples in this GBR framework. The entities appearing have a simple interpretation as the quantity and quality of the LGD. In contrast with the transformed LGD models, GBR models do not require normality and homoskedasticity. Inference in this framework can be unified for models with a variety of link functions and different degrees of complexity using the least-squares method and maximum likelihood estimation (MLE), making model selection a straightforward task. Moreover, the GBR framework allows both the normal approximation and the saddlepoint approximation to efficiently calculate the portfolio loss distribution. This is the first time that numerical approximation methods have been used successfully to calculate portfolio loss distribution in the presence of random LGD.

The rest of the paper is organized as follows. In Section 2 we introduce Vasicek's Gaussian one-factor model as the default model and give a brief summary of existing random LGD models. Section 3 elaborates on the GBR framework, including the basic beta regression model and two extensions. In Section 4 we discuss methods for parameter estimation and provide a calibration example. Section 5 explains techniques for efficient loss distribution approximation in the GBR framework.

2 CREDIT PORTFOLIO LOSS

Consider a credit portfolio consisting of n obligors, each with exposure at default w_i and PD p_i. Obligor i is subject to default after a fixed time horizon and the default can be modeled as a Bernoulli random variable D_i such that:

$$D_i = \begin{cases} 1 & \text{with probability } p_i \\ 0 & \text{with probability } 1 - p_i \end{cases}$$

Let LGD, the proportion of the exposure that will be lost if a default occurs, be denoted by Λ. Then the loss incurred due to the default of obligor i is given by $L_i = w_i \Lambda_i D_i$. It follows that the portfolio loss is given by:

$$L = \sum_{i=1}^{n} L_i = \sum_{i=1}^{n} w_i \Lambda_i D_i$$

To evaluate the distribution of L, it is necessary to model the various dependence effects, including the dependence between defaults, the dependence between LGDs and the dependence between PD and LGD. A convenient approach is to utilize a latent factor model and introduce systematic risk in both PD and LGD.

2.1 Default model

We consider the Vasicek (2002) Gaussian one-factor model as our default model. Although the Vasicek model is often criticized for being oversimplistic in relying on the Gaussian distribution, the extension of this model to the generic one-factor Lévy model, as outlined in Albrecher *et al* (2007), is straightforward. The Lévy models are able to produce more heavy-tailed loss distributions and provide a better fit to the financial market data.

Based on Merton's firm-value model, the Vasicek model evaluates the default of an obligor in terms of the evolution of its asset value. For obligor i, default occurs when the standardized log of the gross return X_i is less than some prespecified threshold γ_i, where X_i is normally distributed and $\mathbb{P}(X_i < \gamma_i) = p_i$. X_i is decomposed into a systematic part Y, representing the state of the economy, and an idiosyncratic part Z_i, such that:

$$X_i = \sqrt{\rho} Y + \sqrt{1-\rho} Z_i \tag{2.1}$$

where Y and all Z_i are independent and identically distributed (iid) standard normal random variables and ρ is the common pairwise correlation. It is now easily deduced that X_i and X_j are conditionally independent given the realization of Y. This implies that L_i and L_j are also conditionally independent given Y.

Define $p_i(y) = \mathbb{P}[L_i = 1 \mid Y = y]$, ie, the PD conditional on the common factor $Y = y$. Then:

$$\begin{aligned} p_i(y) &= \mathbb{P}[L_i = 1 \mid Y = y] \\ &= \mathbb{P}[X_i < \gamma_i \mid Y = y] \\ &= \Phi\left(\frac{\Phi^{-1}(p_i) - \sqrt{\rho}y}{\sqrt{1-\rho}}\right) \end{aligned} \tag{2.2}$$

where Φ is the cumulative distribution function of the standard normal distribution.

2.2 LGD models

A variety of models in which LGD is subject to systematic risk can be found in the literature. Within a one-factor framework, Frye (2000) proposed a model in which the LGD is normally distributed and influenced by the same systematic factor Y that drives the PD, so that:

$$\Lambda = \mu + \sigma\xi, \quad \xi = -\sqrt{\tilde{\rho}}Y + \sqrt{1-\tilde{\rho}}\epsilon$$

where ξ and ϵ are both standard normally distributed. The minus sign in front of $\sqrt{\tilde{\rho}}$ reflects the empirical findings that LGD tends to be higher when the economy is weak and lower when the economy is strong. This way the dependence between LGDs and the dependence between PD and LGD are modeled simultaneously. The parameters μ and σ can be understood as the expected LGD and the LGD volatility, respectively. Unfortunately, the LGD is unbounded in $\mathbb{R}$ and can thus be negative. To ensure the nonnegativity of LGD, Pykhtin (2003) employs a lognormal distribution for the LGD:

$$\Lambda = (1 - \mathrm{e}^{\mu+\sigma\xi})^+$$

Other extensions include Andersen and Sidenius (2004), choosing a probit transformation:

$$\Lambda = \Phi(\mu + \sigma\xi)$$

where Φ is again the cumulative distribution function of the standard normal distribution. Düllmann and Trapp (2004) and Rösch and Scheule (2005) employ a logit transformation:

$$\Lambda = \frac{1}{1 + \mathrm{e}^{\mu+\sigma\xi}}$$

All three transformations for the LGD above guarantee that the LGD lies in the interval $[0, 1]$. However the parameters μ and σ do not have a convenient economic interpretation as in Frye's model.

The above models are basically all linear models of the transformed LGD in the form:

$$g(\Lambda(Y)) = \mu - \sigma\sqrt{\tilde{\rho}}Y + \sigma\sqrt{1-\tilde{\rho}}\epsilon \tag{2.3}$$

so that $g(\Lambda(Y))$ is normally distributed with mean $\mu+\sigma\sqrt{\tilde{\rho}}Y$ and variance $\sigma^2(1-\tilde{\rho})$. Hence, the transformed LGD $g(\Lambda(Y))$ is required to be symmetric and homoskedastic, ie, its variance must not vary with the mean. This contradicts the empirical study by Düllmann and Trapp (2004), at least for the Pykhtin (2003) model, in which the Shapiro–Wilk test for normality to $\log(1-\Lambda)$ gives a p-value of 0.05.

A more flexible approach extends the static beta distribution assumption that is made in CreditMetrics and KMV Portfolio Manager. Giese (2006) and Bruche and González-Aguado (2008) model the LGD using a mixture of beta distributions:

$$\Lambda \sim \text{Beta}(\alpha, \beta)$$

where both α and β are functions of common factor Y. However, α and β are both shape parameters and an economic interpretation of such models is nontrivial.

Here we describe a GBR framework for random LGD. The GBR framework includes Giese (2006) and Bruche and González-Aguado (2008) as special examples but calls for a different parameterization of the beta distribution. The class of models is flexible and the quantities in the models have an easy interpretation as the quantity and quality of the LGD. Inference of models in this framework can be unified. Compared with the transformed LGD models given by (2.3), these GBR models better accommodate skewness and heteroskedastic errors.

3 GENERALIZED BETA REGRESSION MODELS

3.1 Parameterization of a beta distribution

Recall that the probability density function of a beta distribution with parameters $\alpha > 0$, $\beta > 0$ reads:

$$f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)} = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}$$

where $B(\cdot,\cdot)$ denotes the beta function and $\Gamma(\cdot)$ denotes the gamma function.

The beta distribution is known for being very flexible, modeling quantities constrained in the interval [0, 1]. Depending on the choice of parameters, the probability density function can be unimodal, U-shaped, J-shaped or uniform. The expectation and variance of a beta distributed variable X are given by:

$$\mu = E[X] = \frac{\alpha}{\alpha+\beta} \tag{3.1}$$

$$\sigma^2 = \text{var}[X] = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} = \frac{\mu(1-\mu)}{\alpha+\beta+1} \tag{3.2}$$

Note here that the variance is permitted to vary with its mean. Let $\varphi = \alpha + \beta$. Then φ can be regarded as a dispersion parameter in the sense that, for a given μ, the variance is determined by the size of φ.

The parameters α and β can be formulated in terms of the mean and dispersion in the following way:

$$\alpha = \mu\varphi, \qquad \beta = (1-\mu)\varphi \tag{3.3}$$

Therefore, a beta distribution can also be uniquely determined by its mean and dispersion.

3.2 Beta regression model

The generalized beta regression framework is characterized by the following elements:

(1) the LGD is assumed to be beta distributed, conditional on some covariates;

(2) the beta distribution is parameterized by its mean and dispersion, rather than its natural parameters (α, β); the mean and dispersion parameters carry the interpretation as the quantity and quality of the LGD, respectively.

This framework generalizes the beta regression model proposed by Ferrari and Cribari-Neto (2004) for modeling rates and proportions. The models from the GBR framework are similar to a class of models derived from GLMs.

Generalized linear models have been developed as an extension to classical linear regression models since the seminal paper by Nelder and Wedderburn (1972). In a GLM, the density function of a response variable X is in the form of:

$$f(x;\theta,\xi) = \exp(a(\xi)[x\theta - b(\theta) + c(x)] + d(\xi, x)) \tag{3.4}$$

where $a(\xi) > 0$, so that, for fixed ξ, we have an exponential family. The parameter ξ could stand for a certain type of nuisance parameter, such as the variance of a normal distribution. For a comprehensive exposition of GLMs we refer the interested reader to McCullagh and Nelder (1989).

We start the explanation of the GBR framework with the beta regression model proposed in Ferrari and Cribari-Neto (2004). This approach mainly models the mean μ and treats the dispersion parameter φ as a nuisance parameter. The mean model in the GBR framework has the following two components.

- A linear predictor η:

$$\eta = a\zeta \tag{3.5}$$

where ζ is a vector of explanatory variables and a is a vector of the corresponding regression coefficients (by convention the first element of ζ is set to be 1, so that the first element of a is an intercept term).

- A monotonic, differentiable link function g:

$$g(\mu) = \eta, \quad \text{where } \mu = E[\Lambda] \tag{3.6}$$

Potential covariates in the linear predictor can be seniority, collateral, type of industry and timing of business cycle. According to Schuermann (2004), these factors drive significant differences on LGD. Meanwhile, the inverse of the link function $g^{-1}(\cdot)$ should form a mapping from $\mathbb{R}$ to [0, 1], which is exactly the range of μ. This can be achieved by a variety of link functions, such as the logit link:

$$\mu = \frac{e^{\eta}}{1 + e^{\eta}}, \qquad \eta = \log\left(\frac{\mu}{1-\mu}\right) \tag{3.7}$$

or the probit link:

$$\mu = \Phi(\eta), \qquad \eta = \Phi^{-1}(\mu) \tag{3.8}$$

Both the logit and probit link functions have a symmetric form about $\mu = \frac{1}{2}$. If, however, it is believed that symmetric links are not justified, asymmetric link functions like the scaled probit link and the complementary log–log link can be used instead.

It should be noted that the model described here can be very different from the transformed models characterized by (2.3) as we take $g(E[\Lambda])$, rather than $E[g(\Lambda)]$, to be linear to the covariates.

A first model for LGD subject to systematic risk is a one-factor model with $\zeta = [1, Y]^{\mathrm{T}}$, where Y is the common factor that also drives the default process. An example of such a model is given in Giese (2005), where the mean is modeled by:

$$\mu = 1 - a_0(1 - p_i(Y)^{a_1})^{a_2} \tag{3.9}$$

and φ is considered a nuisance parameter.

Another special case is the static beta distribution model adopted by CreditMetrics and KMV, which is a degenerated version of the beta regression model, in which the coefficient in front of Y equals zero.

3.3 Extensions

The beta regression model above can be readily extended in various ways. One extension is to model the mean and dispersion jointly, rather than treating the dispersion parameter φ as a nuisance parameter, which is either fixed or known. This is similar to the joint generalized linear model (JGLM) from the GLM framework (see, for example, Nelder and Lee (1991) and Lee and Nelder (1998)).

The dispersion φ can be modeled by a separate GLM:

$$h(\varphi) = b\zeta$$

where h is also a link function. A simple way to ensure $\varphi > 0$ is to use a log link so that:

$$\varphi = \mathrm{e}^{b\zeta} \tag{3.10}$$

A model of this type, but using a different version of the dispersion parameter, was suggested by Bruche and González-Aguado (2008). They employ the following log-linear model for the two parameters α and β:

$$\alpha = \mathrm{e}^{c\zeta}, \qquad \beta = \mathrm{e}^{d\zeta} \tag{3.11}$$

where, as usual, ζ is a vector of covariates and c and d are vector coefficients. This specification is chosen to ensure positivity of both shape parameters α and β. We note that, by substituting (3.11) into (3.1), we obtain:

$$\mu = \frac{\alpha}{\alpha + \beta} = \frac{\mathrm{e}^{(c-d)\zeta}}{1 + \mathrm{e}^{(c-d)\zeta}}$$

which is a logit model with vector coefficient $c - d$. The variance is then given by:

$$\sigma^2 = \frac{\mu(1-\mu)}{\alpha + \beta + 1} = \frac{\mu^2(1-\mu)}{\alpha + \mu}$$

so that the following dispersion parameter is adopted:

$$\varphi = \alpha = \mathrm{e}^{c\zeta}$$

A second extension is that the mean parameter μ can be modeled by a generalized linear mixed model (GLMM). The GLMM extends the GLM by adding normally distributed random effects in the linear predictor η. A first mixed model is the random intercept model:

$$g(\mu) = \eta = a\zeta + \nu \tag{3.12}$$

where, in addition to the fixed effect $a\zeta$, η also has a single component of random effect ν that follows a univariate normal distribution $N(0, \sigma_\nu^2)$. Here, ν can be thought of as a latent common factor for the LGD independent of the fixed effects and default as well.

Such a GLMM, along with the probit link (3.8), is employed to model the mean LGD in Hillebrand (2006). Other applications of the GLMM for portfolio credit default and migration risk can be found in McNeil and Wendin (2006, 2007).

Note that the two extensions above can be readily combined to form another model that jointly models the mean and dispersion by means of GLMMs, ie, fixed and random effects can be included in the modeling of both mean and dispersion. Further extensions are possible, such as replacing the linear predictor by a generalized additive model (see Hastie and Tibshirani (1990)) or adding multilevel random effects in the GLMM.

4 ESTIMATION

In this section, we discuss the parameter estimation in the GBR framework using

(1) the least-squares method,

(2) MLE.

The former only requires the knowledge of yearly mean LGD and LGD volatility and can be used as the first approximation to the MLE.

Suppose we have a time series of LGD data for $T \in \mathbb{N}$ years. Let K_t be the number of defaulted obligors in year t and let $\lambda_{t,k}$ be the observed LGD for defaulted obligor k, $t = 1, \dots, T$, $k = 1, \dots, K_t$. Each year, a realization of the common factor Y_t can be inferred from the default model and historical default data. The value of Y_t, $t = 1, \dots, T$, should be considered a known fixed effect in the LGD model.

From now on we call the three models in the GBR framework GBR–GLM, GBR–JGLM and GBR–GLMM, respectively. The parameters to be estimated are $\{a, \varphi\}$ in GBR–GLM, $\{a, b\}$ in GBR–JGLM and $\{a, \varphi, \sigma_\nu\}$ in GBR–GLMM, where a denotes the vector coefficients in the linear predictor (3.5), b denotes the vector coefficients in the linear predictor (3.10), φ is the dispersion parameter and σ_ν^2 is the variance of the random effect ν in (3.12).

4.1 Least squares

The method of least squares that we propose here only requires the knowledge of the yearly mean LGD and LGD volatility for parameter estimation. The estimates of the yearly mean LGD and LGD volatility for $t = 1, \dots, T$ can be obtained by matching the first and second moments of the LGD realizations $\lambda_{t,k}$ such that:

$$m_t = \frac{1}{K_t} \sum_{k=1}^{K_t} \lambda_{t,k}, \qquad \sigma_t^2 = \frac{1}{K_t} \sum_{k=1}^{K_t} \lambda_{t,k}^2 - m_t^2$$

Estimation of a and μ

The estimate for parameter a can be obtained by employing a linear regression of the transformed mean LGD $g(m_t)$ on Y_t and other covariates:

$$g(m_t) = \hat{a}\zeta_t + \nu_t \tag{4.1}$$

where ν_t is the residual term. In the GBR–GLM and GBR–JGLM:

$$\hat{\mu}_t = g^{-1}(\hat{a}\zeta_t) \tag{4.2}$$

and, in the GBR–GLMM, ν_t is taken to be the realized random effect in year t so that:

$$\hat{\mu}_t = g^{-1}(\hat{a}\zeta_t + \nu_t) = m_t \tag{4.3}$$

Estimation of b or φ

The estimation of the parameters b and φ takes the prediction of $\hat{\mu}_t$, produced by (4.2) or (4.3), as an input. From (3.2), we obtain:

$$\varphi_t = \frac{\hat{\mu}_t(1-\hat{\mu}_t)}{\sigma_t^2} - 1$$

In both the GBR–GLM and GBR–GLMM the dispersion parameter φ is treated as a nuisance parameter. Its method-of-moments estimator is simply:

$$\hat{\varphi} = \frac{1}{T}\sum_{t=1}^{\mathrm{T}} \varphi_t$$

In the GBR–JGLM, the coefficient b can be calculated by linear regression of the transformed dispersion $h(\varphi_t)$ on covariate vector ζ such that:

$$h(\varphi_t) = \hat{b}\zeta_t + \epsilon_t$$

Estimation of σ_ν in the GBR–GLMM

The moment-based estimate for σ_ν^2 is given by:

$$\hat{\sigma}_\nu^2 = \frac{1}{T}\sum_{t=1}^{\mathrm{T}} \nu_t^2$$

where ν_t is the residual term in (4.1).

4.2 Maximum likelihood estimation

Parameter estimation using the MLE method is also straightforward in the GBR framework. In the models without random effects, ie, the GBR–GLM and GBR–JGLM, the log-likelihood function to be maximized reads:

$$\begin{aligned}\ell(\mu,\varphi) = \sum_{t=1}^{\mathrm{T}}\sum_{k=1}^{K_t}\{&(\mu_t\varphi_t - 1)\log(\lambda_{t,k}) + [(1-\mu_t)\varphi_t - 1]\log(1-\lambda_{t,k}) \\ &+ \log\Gamma(\varphi_t) - \log\Gamma(\mu_t\varphi_t) - \log\Gamma[(1-\mu_t)\varphi_t]\}\end{aligned} \tag{4.4}$$

The score function, the gradient of the log-likelihood function and the Fisher information matrix, ie, the variance of the score, can be formulated explicitly in terms of polygamma functions. These are given in Appendix A. Asymptotic standard errors of the maximum likelihood estimates of the parameters can be computed from the Fisher information matrix.

Since the corresponding estimating equations do not admit a closed-form solution, numerical maximization of the log-likelihood is necessary. Estimates using the method of least squares may be used as the initial approximations to the solutions of the likelihood equations.

We remark that the MLE in Ferrari and Cribari-Neto (2004) beta regression model is already implemented in the statistical computing software R in package "betareg" so that it can be used immediately.[1]

Marginal likelihood in the GBR–GLMM

With the presence of random effects, the samples are no longer independent. In the random intercept model (3.12), the LGDs in year t are only independent conditional on the random effect ν_t. Since we are interested in inference of the variance of the random component ν, but not in its realizations, the random effect needs to be integrated out. Therefore we maximize the marginal log-likelihood:

$$\ell_m(a, \varphi, \sigma_\nu) = \sum_{t=1}^{\mathrm{T}} \log\left(\int \prod_{k=1}^{K_t} L(a, \varphi, \zeta_t, \nu_t; \lambda_{t,k})\, p_{\sigma_\nu}(\nu_t)\, \mathrm{d}\nu_t \right)$$

where $p_{\sigma_\nu}(\cdot)$ is the probability density function of a normal distribution with mean zero and variance σ_ν^2, and $L(\lambda_{t,k})$ is the likelihood of $\{\mathrm{LGD} = \lambda_{t,k}\}$ given ν_t. The integral can be efficiently evaluated by Gaussian quadrature. Alternatively, the marginal likelihood can be approximated analytically by the use of the Laplace approximation to the integral, such as the penalized quasi-likelihood (QL) estimation (Breslow and Clayton (1993)) and the h-likelihood (Lee and Nelder (2001)), thereby avoiding numerical integration.

Finally we note that the likelihood-ratio test based on large sample inference can be employed for model selection. Information criteria such as Akaike's information criterion (AIC) or the Bayesian information criterion (BIC) can also be used.

4.3 A simulation study

In this section we show how the models in the GBR framework can be calibrated and how model selection can be dealt with. Our aim is not to identify possible covariates that influence the LGD, however. Our estimation is based on data from Bruche and González-Aguado (2008) that is extracted from the Altman-NYU Salomon Center Corporate Bond Default Master Database and gives the annual default frequency, number of defaults, mean LGD and LGD volatility for a period of twenty-four years (1982–2005). For completeness the data is reproduced in Table 1 on the next page.

Should the values in columns 2, 4 and 5 that have only one decimal place be shown to two decimal places for consistency? If so, should they be shown 0.90, 14.90, 20.40 and so on?

TABLE 1 LGD statistics by year, 1982–2005.

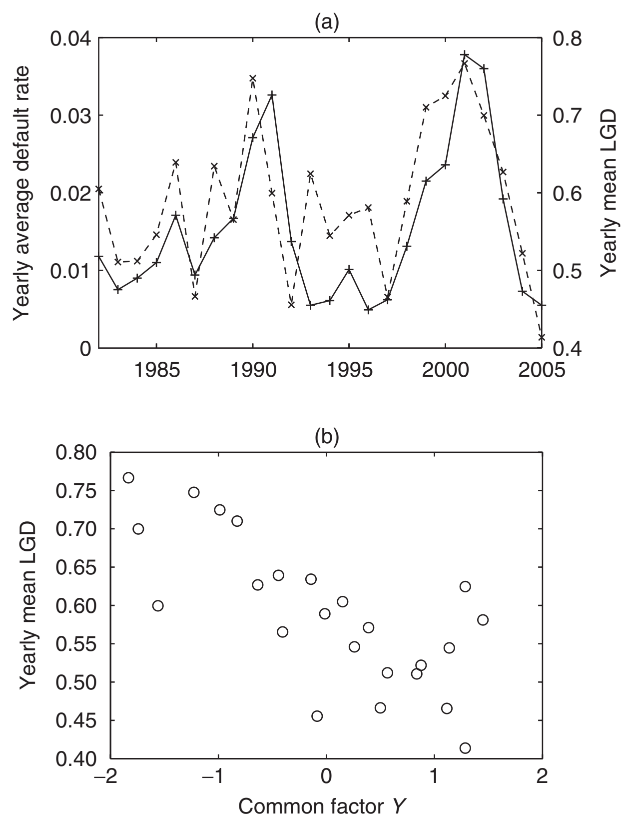
Year	PD (%)	Number of defaults	Mean LGD (%)	LGD volatility (%)
1982	1.18	12	60.49	14.9
1983	0.75	5	51.07	23.53
1984	0.9	11	51.19	17.38
1985	1.1	16	54.59	21.87
1986	1.71	24	63.91	18.82
1987	0.94	20	46.64	26.94
1988	1.42	30	63.43	17.97
1989	1.67	41	56.54	28.78
1990	2.71	76	74.76	22.28
1991	3.26	95	59.95	26.09
1992	1.37	35	45.55	23.38
1993	0.55	21	62.46	20.11
1994	0.61	14	54.46	20.46
1995	1.01	25	57.1	25.25
1996	0.49	19	58.1	24.68
1997	0.62	25	46.54	25.53
1998	1.31	34	58.9	24.56
1999	2.15	102	71.01	20.4
2000	2.36	120	72.49	23.36
2001	3.78	157	76.66	17.87
2002	3.6	112	69.97	17.18
2003	1.92	57	62.67	23.98
2004	0.73	39	52.19	24.1
2005	0.55	33	41.37	23.46

This table is taken from Bruche and González-Aguado (2008), where mean recovery rate (RR) is reported instead of LGD. The column of mean LGD here is calculated to be 1 minus RR, ie, LGD = 1 − RR.

TABLE 2 Estimates given by the method of least squares for different models.

	GBR–GLM	GBR–JGLM	GBR–GLMM
a_1	0.3718	0.3718	0.3718
a_2	−0.3054	−0.3054	−0.3054
φ	4.1914	—	4.0907
b_1	—	1.3505	—
b_2	—	−0.0033	—
σ_ν	—	—	0.2686

FIGURE 1 (a) Yearly average default rate and yearly mean LGD (1982–2005); (b) the common factor Y estimated by Equation (4.6) versus yearly mean LGD (1982–2005).



In part (a), the solid line shows the the yearly average default rate and the dashed line shows the yearly mean LGD.

4.3.1 Estimation results

First, we fit the Vasicek default model. We assume that, across the years, the number of obligors is sufficiently large and that all obligors in the portfolio have the same PD p and asset correlation ρ. Denote by p_t the annual default frequency. We take the maximum likelihood estimates for ρ and p according to Düllmann and Trapp (2004):

$$\rho = \frac{\mathrm{var}[\Phi^{-1}(p_t)]}{1 + \mathrm{var}[\Phi^{-1}(p_t)]}, \qquad p = \Phi\left(\frac{\sum_{t=1}^{\mathrm{T}} \Phi^{-1}(p_t)}{T\sqrt{1 + \mathrm{var}[\Phi^{-1}(p_t)]}}\right)$$

[1] See www.r-project.org.

TABLE 3 Maximum likelihood estimates of various models.

	GBR–GLM	GBR–JGLM	GBR–GLMM
a_1	0.3459 (0.0359)	0.3471	0.3319
a_2	−0.3213 (0.0298)	−0.3246	−0.3307
φ	3.0276 (0.1149)	—	3.3240
b_1	—	1.0879	—
b_2	—	−0.0306	—
σ_v	—	—	0.2943
-2ℓ	−402.74	−403.34	−468.78
AIC	−398.74	−395.34	−460.78
BIC	−381.67	−375.25	−440.69

where:

$$\operatorname{var}[\delta] = \frac{1}{T}\sum_{t=1}^{\mathrm{T}} \delta_t^2 - \left(\frac{1}{T}\sum_{t=1}^{\mathrm{T}} \delta_t\right)^2$$

This yields:

$$\rho = 0.0569, \qquad p = 0.0153 \tag{4.5}$$

The common factor Y_t for year t, assumed to be independent from year to year, can be estimated as follows:

$$Y_t = \frac{\Phi^{-1}(p) - \sqrt{1-\rho}\,\Phi^{-1}(p_t)}{\sqrt{\rho}} \tag{4.6}$$

Before we move on to the LGD model, we run a brief preliminary graphical check. In part (a) of Figure 1 on the preceding page we show the yearly average default rate and yearly mean LGD for the years 1982–2005, from which the correlation between PD and LGD is evident. Part (b) of Figure 1 on the preceding page presents a scatterplot of the common factor Y estimated by Equation (4.6) versus yearly mean LGD. This figure suggests that the common factor Y, which drives the default, may also be an important risk factor for LGD.

Next we make inferences about the LGD in the GBR framework with both the least squares and MLE. The LGD models that we consider only include one covariate, which is the common factor Y in the default model. In light of the observation from part (b) of Figure 1 on the preceding page, this may be a reasonable choice. The mean LGD is fitted using a logit link:

$$\mu = \frac{\mathrm{e}^{a_1 + a_2 Y}}{1 + \mathrm{e}^{a_1 + a_2 Y}} \tag{4.7}$$

in the GBR–GLM and GBR–JGLM, and:

$$\mu = \frac{e^{a_1+a_2Y+\nu}}{1 + e^{a_1+a_2Y+\nu}}$$

in the GBR–GLMM. In the GBR–JGLM, the dispersion parameter is modeled as:

$$\varphi = e^{b_1+b_2Y}$$

The estimates given by the method of least squares are presented in Table 2 on page 12. These estimates are used as the first approximation to the MLE. We are already able to get a first impression of the characteristics of the LGD:

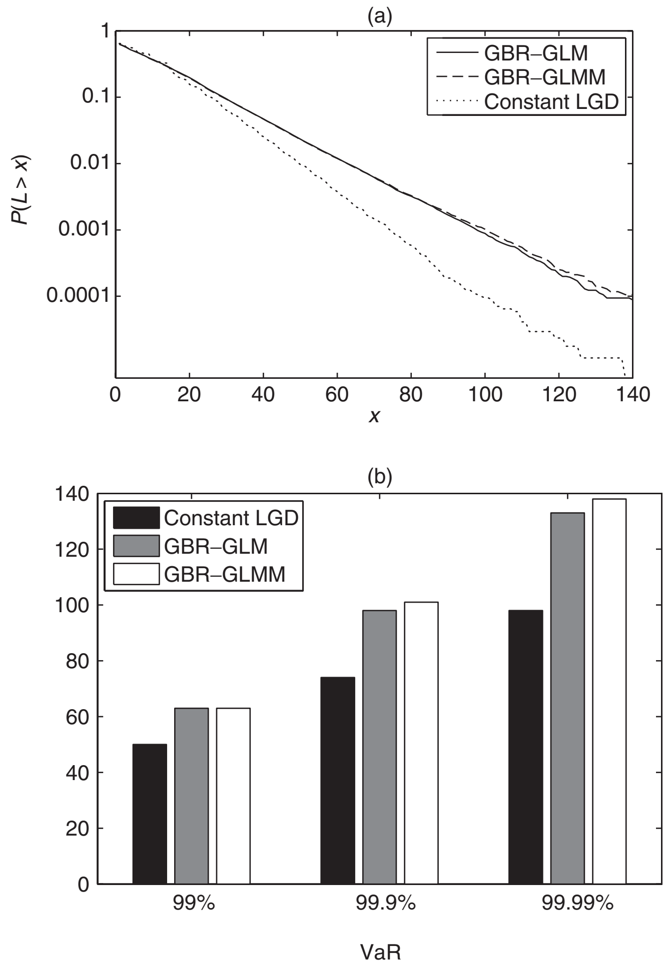
(1) the coefficient a_2 is negative, indicating a negative relation between Y and mean LGD, just as expected;

(2) the coefficient b_2 is very close to zero, suggesting that Y may not be relevant for the estimation of dispersion φ.

It is important to keep in mind that these least-squares estimates are only based on the annual mean LGD and LGD volatility. Consequently, the above observations are not restricted to any particular sample of simulated LGD realizations, as opposed to estimates to be obtained from the MLE.

To carry out the MLE we need a sample of LGD realizations. For each year, a realization of the LGD is simulated for each defaulted obligor from a beta distribution matching the empirical mean and variance. In total, this gives 1123 LGD observations in $T = 24$ years. The maximum likelihood estimates for the various parameters are given in Table 3 on the facing page. For the GBR–GLM, we also report in parentheses the asymptotic standard errors of the estimates. We find that the estimates given by MLE are very similar to those given by least squares. The Wald test confirms that both a_1 and a_2 are statistically significant (both p-values < 0.0001), which justifies the use of Y as a risk factor for the mean LGD. The log-likelihood-ratio statistics of GBR–JGLM and GBR–GLMM to GBR–GLM are $-402.74 - (-403.34) = 0.6$ and $-402.74 - (-468.78) = 66.04$, respectively. They correspond to p-values 0.44 and < 0.0001 for the chi-square distribution with one degree of freedom. It is clear that GBR–GLMM provides a significant improvement on the basic GBR–GLM, whereas GBR–JGLM fails to do so. The AIC and BIC lead to the same conclusion (see Table 3 on the facing page). Additional simulation tests show that the above estimation results are very robust. We note that this does not, however, suggest that GBR–JGLM should be abandoned in general since the idea of jointly modeling mean and dispersion may be meaningful if we include other covariates, for example, seniority and presence and quality of collateral.

FIGURE 2 (a) The portfolio loss distributions; (b) the portfolio VaR at three confidence levels under the three LGD models.



The results are based on Monte Carlo simulation of 200 000 scenarios. For GBR–GLM and GBR–GLMM, the LGD parameters are taken from Table 3 on page 14. In the constant LGD model $\Lambda = 0.58$ for all obligors.

Additionally, we also fit the GBR models to a second sample of LGD realizations, simulated from the probit model where the LGD is given by $\lambda_{t,k} = \Phi(c_t + d_t \epsilon_{t,k})$. The parameters c_t and d_t for all t can be conveniently estimated using the method of moments since:

$$E(\Lambda_t) = \Phi\left(\frac{c_t}{\sqrt{1+d_t^2}}\right), \qquad E(\Lambda_t^2) = \Phi_2\left(\frac{c_t}{\sqrt{1+d_t^2}}, \frac{c_t}{\sqrt{1+d_t^2}}, \frac{d_t^2}{1+d_t^2}\right)$$

where $\Phi_2(\cdot,\cdot,\rho)$ denotes the bivariate cumulative Gaussian distribution function with correlation ρ (for a proof see Andersen and Sidenius (2004)). The MLE procedure gives the estimates $a_1 = 0.3562$, $a_2 = -0.3247$, $\varphi = 3.0589$ for the GBR–GLM model, $a_1 = 0.3571$, $a_2 = -0.3275$, $b_1 = 1.1011$, $b_2 = -0.0260$ for the GBR–JGLM and $a_1 = 0.3393$, $a_2 = -0.3144$, $\varphi = 3.4096$, $\sigma_\nu = 0.3261$ for the GBR–GLMM. These estimates are broadly in agreement with those in Table 2 on page 12 and in Table 3 on page 14, which suggests that the parameters in the GBR models are robust to misspecification of the LGD distribution.

Moreover, for the second sample we also look at the QL (see Wedderburn (1974)) of a model which assumes that the mean and variance of the LGD are given by (4.7) and $\mu(1-\mu)/(1+\varphi)$, respectively, but the distribution of the LGD is unknown. The QL is then given by:

$$\mathrm{QL} = \sum_{t=1}^{\mathrm{T}} \sum_{k=1}^{K_t} [\lambda_{t,k} \log(\mu_t) + (1-\lambda_{t,k}) \log(1-\mu_t)]$$

Maximization of QL gives $a_1 = 0.3799$, $a_2 = -0.3336$, $\varphi = 3.0979$, indicating that the assumption on the distribution of the LGD probably does not matter much. In summary, we conjecture that the distribution assumption of the LGD conditional on the covariates is of low importance as long as the mean and variance of LGD as functions of the covariates are modeled appropriately.

4.3.2 Implication for portfolio risk

It is also interesting to see how much the choice of an LGD model can influence the VaR at the portfolio level. We consider a portfolio of 100 obligors with uniform PD p and correlation ρ as in (4.5) and exposures as follows:

$$w_i = \begin{cases} 1, & k = 1, \dots, 20 \\ 4, & k = 21, \dots, 40 \\ 9, & k = 41, \dots, 60 \\ 16, & k = 61, \dots, 80 \\ 25, & k = 81, \dots, 100 \end{cases}$$

We compare three models for the LGD:

(1) the GBR–GLM;

(2) the GBR–GLMM;

(3) the constant LGD model.

For the GBR–GLM and GBR–GLMM, the LGD parameters are taken from Table 3 on page 14. In the constant LGD model, for all obligors we take $\Lambda = 0.58$, matching the expected LGD $E_Y[\mu(Y)]$ in the GBR–GLM, where $E_Y(\cdot)$ denotes the expectation obtained by integrating over Y.

The portfolio loss distributions plotted in part (a) of Figure 2 on page 16 are based on Monte Carlo simulation with 200 000 scenarios. On the one hand, the curves of the GBR–GLMM and the GBR–GLM are almost identical, with the GBR–GLMM producing a slightly heavier tail. This is again an indication of the robustness of the GBR models. On the other hand, the loss distribution under the constant LGD model deviates substantially from the other two models with random LGD.

We then look at the portfolio VaR at three particular confidence levels, 99%, 99.9% and 99.99%, illustrated in part (b) of Figure 2 on page 16. Compared with the constant LGD model, the GBR–GLM (respectively, GBR–GLMM) increases the VaR at the three levels by a factor of 1.26, 1.32 and 1.36 (respectively, 1.26, 1.36 and 1.41). It is apparent that ignoring the systematic risk in the LGD significantly underestimates risk. Moreover, the further along the tail, the higher the degree of underestimation. These results are in line with those reported in Altman *et al* (2005) and Giese (2006).

5 LOSS DISTRIBUTION APPROXIMATIONS

The calculation of portfolio loss distribution with random LGD is mostly based on Monte Carlo simulation in the literature. To our knowledge the only exception is Giese (2006), where the saddlepoint approximation was employed. An important advantage of the generalized beta regression framework for random LGD is that it allows both the normal approximation and the saddlepoint approximation to efficiently calculate the portfolio loss distribution, thereby avoiding the need for time-consuming simulation. Both approximations apply to completely heterogeneous portfolios. For simplicity, we derive the formulas only for the basic GBR–GLM with a single covariate, Y, or, equivalently, a single-factor model, where the tail probability reads:

$$\mathbb{P}(L \geqslant x) = \int \mathbb{P}(L \geqslant x \mid Y)\,\mathrm{d}\Phi(Y)$$

Generalization to more complex models is fairly straightforward.

5.1 Normal approximation

First of all, in the case of a large homogeneous portfolio, the expected loss from obligor i conditional on Y reads:

$$E[L_i(Y)] = w_i E[D_i(Y)] E[\Lambda_i(Y)] = w_i p_i(Y) \mu_i(Y) \tag{5.1}$$

TABLE 4 Approximations to the portfolio VaR at three confidence levels.

	$\mathrm{VaR}_{99\%}$	$\mathrm{VaR}_{99.9\%}$	$\mathrm{VaR}_{99.99\%}$
Monte Carlo	63	98	133
Normal approximation	58	90	123
Saddlepoint approximation	63	97	133

The LGD model adopted here is GBR–GLM. The Monte Carlo results are based on 200 000 simulated scenarios and can be regarded as our benchmark.

A version of the large homogeneous portfolio approximation similar to that in the Vasicek model can also be obtained for random LGD:

$$\frac{L(Y)}{\sum_{i=1}^{n} w_i} \to \frac{\sum_{i=1}^{n} w_i p_i(Y)\mu_i(Y)}{\sum_{i=1}^{n} w_i} \quad \text{almost surely}$$

However, when the portfolio is not sufficiently large or not very homogeneous, unsystematic risk arises. The normal approximation improves on the large homogeneous portfolio approximation by taking into account the variability of portfolio loss L conditional on the common factor Y. The conditional portfolio loss $L(Y)$ can be approximated by a normally distributed random variable with mean $M(Y)$ and variance $V^2(Y)$ such that:

$$\begin{aligned} M(Y) &= \sum_{i=1}^{n} w_i p_i(Y)\mu_i(Y) \\ V^2(Y) &= \sum_{i=1}^{n} E[L_i^2(Y)] - \sum_{i=1}^{n} E[L_i(Y)]^2 \end{aligned}$$

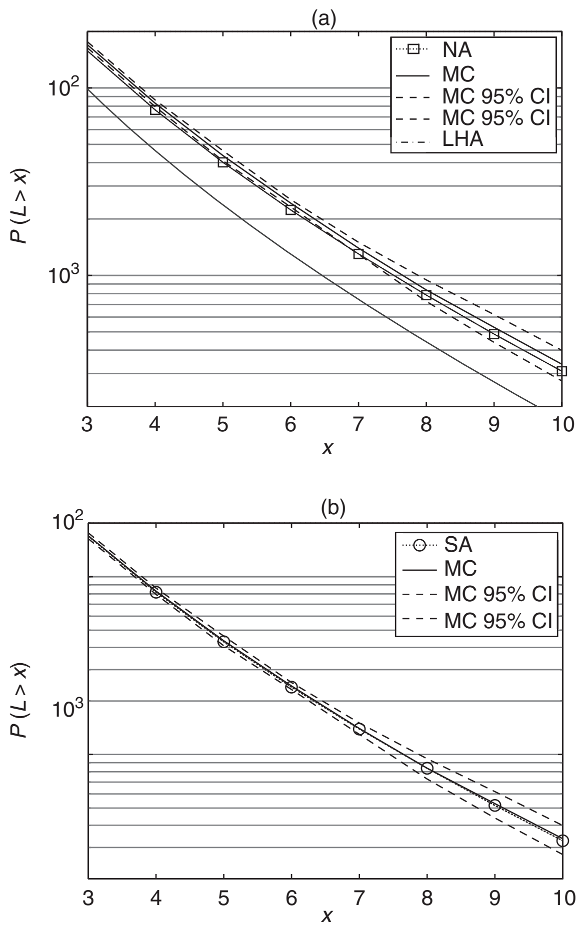
where:

$$\begin{aligned} E[L_i^2(Y)] &= w_i^2 E[D_i(Y)] E[\Lambda_i^2(Y)] \\ &= w_i^2 p_i(Y) E[\Lambda_i^2(Y)] \\ &= w_i^2 p_i(Y)[\mu_i^2(Y) + \mathrm{var}(\Lambda \mid Y)] \\ &= w_i^2 p_i(Y)\left[\mu_i^2(Y) + \mu_i(Y)\frac{1-\mu_i(Y)}{1+\varphi_i}\right] \end{aligned}$$

The conditional tail probability is $\mathbb{P}(L \geqslant x \mid Y) = \Phi((M(Y)-x)/V(Y))$ and it follows that the unconditional tail probability reads:

$$\mathbb{P}(L \geqslant x) = \int \Phi\left(\frac{M(Y)-x}{V(Y)}\right) \mathrm{d}\Phi(Y) = E_Y\left[\Phi\left(\frac{M(Y)-x}{V(Y)}\right)\right] \tag{5.2}$$

FIGURE 3 Loss distribution.



The loss distribution obtained from (a) the large homogeneous approximation (LHA), the normal approximation (NA) and (b) the saddlepoint approximation (SA) compared with results based on Monte Carlo simulation of 200 000 scenarios. The Monte Carlo (MC) 95% confidence interval (CI) is based on the standard deviation calculated using ten simulated subsamples of 20 000 scenarios each.

5.2 Saddlepoint approximation

The use of the saddlepoint approximation in portfolio credit loss was pioneered by Martin *et al* (2001a,b). The only paper to apply the saddlepoint approximations to the calculation of portfolio credit risk in the presence of random LGD is Giese (2006).

Martin *et al* (2001a,b) and Giese (2006) apply the saddlepoint approximation to the unconditional moment generating function (MGF) of portfolio loss L, despite the fact that the L_i are not independent. In Huang *et al* (2007b) the saddlepoint approximation was applied to the conditional MGF of L given the common factor Y, so that $L(Y) = \sum L_i(Y)$ is a weighted sum of independent random variables, which is exactly the situation where the saddlepoint approximation will work well. Here we extend Huang *et al* (2007b) to models with random LGD and show, using numerical examples, that the saddlepoint approximation is able to produce accurate tail probability approximations to all loss levels and handles heterogeneous portfolios with exposure concentration well.

The use of the saddlepoint approximation only requires the existence of the MGF, which makes the beta distribution assumption for LGD in the framework described very attractive. Recall that the MGF of a beta distributed random variable with parameters (α, β) is a confluent hypergeometric function as follows:

$$\mathrm{MGF}(t) = {}_1F_1(\alpha, \alpha + \beta; t)$$

By basic differentiation, we obtain the following first and second derivatives of the MGF:

$$\begin{aligned}
\mathrm{MGF}'(t) &= {}_1F_1(\alpha + 1, \alpha + \beta + 1; t)\frac{\alpha}{\alpha + \beta} \\
\mathrm{MGF}''(t) &= {}_1F_1(\alpha + 2, \alpha + \beta + 2; t)\frac{\alpha(\alpha + 1)}{(\alpha + \beta)(\alpha + \beta + 1)}
\end{aligned}$$

In this setting, the obligors are independent conditional on the common factor Y. For obligor i, (α_i, β_i) conditional on Y can be determined by (3.3). The conditional MGF and the cumulant generating function, denoted by κ, of the portfolio loss are then given by:

$$\begin{aligned}
\mathrm{MGF}(t, Y) &= \prod_{i=1}^{n} [1 - p_i + p_i \, {}_1F_1(\alpha_i, \alpha_i + \beta_i; w_i t)] \\
\kappa(t, Y) &= \log(\mathrm{MGF}(t, Y)) \\
&= \sum_{i=1}^{n} \log[1 - p_i + p_i \, {}_1F_1(\alpha_i, \alpha_i + \beta_i; w_i t)]
\end{aligned}$$

For simplicity of notation, we have suppressed the explicit dependence of p_i and (α_i, β_i) on the common factor Y.

The derivatives of the conditional cumulant generating function up to second order are:

$$\kappa'(t, Y) = \sum_{i=1}^{n} \frac{w_i p_i \, {}_1F_1(\alpha_i + 1, \alpha_i + \beta_i + 1; w_i t)}{1 - p_i + p_i \, {}_1F_1(\alpha_i, \alpha_i + \beta_i; w_i t)} \frac{\alpha_i}{\alpha_i + \beta_i}$$

$$\kappa''(t, Y) = \sum_{i=1}^{n} \left\{ \frac{w_i^2 p_i \alpha_i (\alpha_i + 1) \, {}_1F_1(\alpha_i + 2, \alpha_i + \beta_i + 2; w_i t)}{(\alpha_i + \beta_i)(\alpha_i + \beta_i + 1)[1 - p_i + p_i \, {}_1F_1(\alpha_i, \alpha_i + \beta_i; w_i t)]} - \frac{w_i^2 p_i^2 \alpha_i^2 \, {}_1F_1(\alpha_i + 1, \alpha_i + \beta_i + 1; w_i t)^2}{(\alpha_i + \beta_i)^2 [1 - p_i + p_i \, {}_1F_1(\alpha_i, \alpha_i + \beta_i; w_i t)]^2} \right\}$$

After finding the saddlepoint $\tilde{t}$ that solves $\kappa'(\tilde{t}, Y) = x$ for the loss level x, the tail probability conditional on Y can be approximated by the Lugannani and Rice (1980) formula:

$$\mathbb{P}(L \geqslant x \mid Y) = 1 - \Phi(z_l) + \phi(z_l)\left(\frac{1}{z_w} - \frac{1}{z_l}\right) \tag{5.3}$$

where:

$$z_w = \tilde{t}\sqrt{\kappa''(\tilde{t}, Y)}, \qquad z_l = \operatorname{sgn}(\tilde{t})\sqrt{2[x\tilde{t} - \kappa(\tilde{t}, Y)]}$$

and ϕ is the probability density function of the standard normal distribution.

Integrating over Y gives the unconditional tail probability $\mathbb{P}(L \geqslant x)$, from which the portfolio VaR can be derived. Formulas for the calculation of other risk measures like VaR contribution, expected shortfall and expected shortfall contribution can be found in Huang *et al* (2007b).

5.3 Numerical results

We now illustrate the performance of the normal and saddlepoint approximations in loss distribution calculation. We first take a homogeneous portfolio with $n = 100$ obligors, each with:

$$w = 1, \qquad p = 0.005, \qquad \rho = 0.18$$

The parameters in the LGD are:

$$a = [0.37, -0.32], \qquad \varphi = 3.16$$

with a logit link for mean LGD. This leads to the following specification of (conditional) mean LGD:

$$\mu = \frac{1}{1 + \exp(-0.37 + 0.32Y)}$$

We compare the loss distributions obtained from various approximation methods with the results from a Monte Carlo simulation. Our benchmark is the sample mean and the accompanying 95% confidence intervals obtained by ten subsamples of Monte Carlo simulation with 20 000 replications each. The performance of the approximations is demonstrated in parts (a) and (b) of Figure 3 on page 20.

The large homogeneous approximation (LHA) results deviate considerably from our benchmark. This is not surprising as the size of the portfolio is rather small. The normal approximation provides a significant improvement over the LHA and underestimates risk only slightly. Some of its tail probability estimates, however, fall outside the 95% confidence interval. By comparison, the saddlepoint approximation is able to give all tail probability estimates within the 95% confidence interval. The loss distribution given by the saddlepoint approximation is indistinguishable from the benchmark. We remark that the calculation of the loss distribution in MATLAB costs roughly four seconds for the normal approximation and four minutes for the saddlepoint approximation on a Pentium 4 2.8 GHz desktop.

Finally, we calculate the VaR for the portfolio considered in Section 4.3.2, with LGD modeled by the GBR–GLM. The results are given in Table 4 on page 19. The Monte Carlo results are based on 200 000 simulated scenarios and can be regarded as our benchmark. In this example the saddlepoint approximation is again very accurate. The normal approximation is, however, rather unsatisfactory. At all three levels, relative errors are around 8%. This is certainly due to the existence of exposure concentration as the variation in the exposures is not negligible. For more details on how robust the normal approximation and the saddlepoint approximation are in terms of handling exposure concentration, we refer the interested reader to Huang *et al* (2007a).

6 CONCLUSIONS

In this paper we have described a GBR framework for modeling systematic risk in LGD in the context of credit portfolio losses. The GBR framework provides great flexibility in random LGD modeling and accommodates skewness and heteroskedastic errors well. We have shown that parameter estimation and model selection are straightforward in this framework. Moreover, it has been demonstrated that the portfolio loss distribution can be efficiently evaluated using both the normal approximation and the saddlepoint approximation.

APPENDIX A: SCORE FUNCTION AND FISHER INFORMATION MATRIX

In this appendix we give details about the score function and the Fisher information matrix for the parameters appearing in the GBR–GLM and GBR–JGLM. The score function may help to accelerate the convergence in the MLE procedure and the Fisher information matrix leads to the asymptotic standard errors of the maximum likelihood estimates of the parameters in the models. In the GBR–GLMM the corresponding formulas get more complicated and lengthy and they are therefore omitted here. We refer the interested reader to Pan and Thompson (2007) for an example.

The score function, ie, the partial derivative of the log-likelihood function with respect to parameters (μ, φ), reads:

$$\frac{\partial \ell}{\partial \mu} = \varphi \left\{ \log\left(\frac{\lambda}{1-\lambda}\right) - \Psi(\mu\varphi) + \Psi[(1-\mu)\varphi] \right\} \tag{A.1}$$

$$\begin{aligned} \frac{\partial \ell}{\partial \varphi} = {} & \mu \log \lambda + (1-\mu)\log(1-\lambda) + \Psi(\varphi) \\ & - \mu\Psi(\mu\varphi) - (1-\mu)\Psi[(1-\mu)\varphi] \end{aligned} \tag{A.2}$$

where λ is a realization of the LGD and $\Psi(\cdot)$ is the digamma function.

The second-order partial derivatives of the log-likelihood function with respect to parameters (μ, φ) are:

$$\frac{\partial^2 \ell}{\partial \mu^2} = -\varphi^2\{\Psi'(\mu\varphi) + \Psi'[(1-\mu)\varphi]\} \tag{A.3}$$

$$\frac{\partial^2 \ell}{\partial \varphi^2} = \Psi'(\varphi) - \mu^2\Psi'(\mu\varphi) - (1-\mu)^2\Psi'[(1-\mu)\varphi] \tag{A.4}$$

$$\frac{\partial^2 \ell}{\partial \mu \partial \varphi} = \frac{1}{\varphi}\frac{\partial \ell}{\partial \mu} - \varphi\{\mu\Psi'(\mu\varphi) - (1-\mu)\Psi'[(1-\mu)\varphi]\} \tag{A.5}$$

where $\Psi'(\cdot)$ is the trigamma function.

In the GBR–GLM, the parameters to be estimated are a and φ. The score function for φ is given by (A.2); the score function with respect to a_i, the ith element of a, is given by:

$$\frac{\partial \ell}{\partial a_i} = \frac{\partial \ell}{\partial \mu}\frac{\partial \mu}{\partial a_i} = \varphi \left\{ \log\left(\frac{\lambda}{1-\lambda}\right) - \Psi(\mu\varphi) + \Psi[(1-\mu)\varphi] \right\} \frac{\zeta_i}{g'(\mu)} \tag{A.6}$$

The Fisher information matrix is the negative of the expectation of the second derivative of the log-likelihood with respect to the parameters. The entries in the Fisher information matrix are:

$$-E\left(\frac{\partial^2 \ell}{\partial \varphi^2}\right) = -\frac{\partial^2 \ell}{\partial \varphi^2} \tag{A.7}$$

$$-E\left(\frac{\partial^2 \ell}{\partial a_i \partial a_j}\right) = -\frac{\partial^2 \ell}{\partial \mu^2}\frac{\zeta_i \zeta_j}{(g'(\mu))^2} \tag{A.8}$$

$$-E\left(\frac{\partial^2 \ell}{\partial a_i \partial \varphi}\right) = \varphi\{\mu\Psi'(\mu\varphi) - (1-\mu)\Psi'[(1-\mu)\varphi]\}\frac{\zeta_i}{g'(\mu)} \tag{A.9}$$

In the GBR–JGLM, the parameters to be estimated are a and b. The score function for the coefficient a is given by (A.6) and that for b_i, the ith element of b, is as follows:

$$\frac{\partial \ell}{\partial b_i} = \frac{\partial \ell}{\partial \varphi}\frac{\partial \varphi}{\partial b_i} = \{\mu \log \lambda + (1-\mu)\log(1-\lambda) + \Psi(\varphi) - \mu\Psi(\mu\varphi) - (1-\mu)\Psi[(1-\mu)\varphi]\}\frac{\zeta_i}{h'(\varphi)} \quad (\text{A.10})$$

The Fisher information matrix contains:

$$-E\left(\frac{\partial^2 \ell}{\partial a_i \partial a_j}\right)$$

given by (A.8) and:

$$-E\left(\frac{\partial^2 \ell}{\partial b_i \partial b_j}\right) = -\frac{\partial^2 \ell}{\partial \varphi^2}\frac{\zeta_i \zeta_j}{(h'(\varphi))^2} \quad (\text{A.11})$$

$$-E\left(\frac{\partial^2 \ell}{\partial a_i \partial b_j}\right) = \varphi\{\mu\Psi'(\mu\varphi) - (1-\mu)\Psi'[(1-\mu)\varphi]\}\frac{\zeta_i \zeta_j}{g'(\mu)h'(\varphi)} \quad (\text{A.12})$$

REFERENCES

Albrecher, H., Ladoucette, S., and Schoutens, W. (2007). A generic one-factor Lévy model for pricing synthetic CDOs. In *Advances in Mathematical Finance*, Elliott, R. J. *et al* (eds). Birkhäuser.

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), 2203–2227.

Andersen, L., and Sidenius, J. (2004). Extensions to the Gaussian copula: random recovery and random factor loadings. *The Journal of Credit Risk* **1**(1), 29–70.

Breslow, N., and Clayton, D. (1993). Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association* **88**, 9–25.

Bruche, M., and González-Aguado, C. (2008). Recovery rates, default probabilities and the credit cycle. Discussion Paper, CEMFI. URL: http://ssrn.com/abstract=934348.

Düllmann, K., and Trapp, M. (2004). Systematic risk in recovery rates: an empirical analysis of US corporate credit exposures. Deutsche Bundesbank Discussion Paper 02/2004, Series 2 (Banking and Financial Supervision).

Ferrari, S., and Cribari-Neto, F. (2004). Beta regression for modeling rates and proportions. *Journal of Applied Statistics* **31**, 799–815.

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

Giese, G. (2005). The impact of PD/LGD correlations on credit risk capital. *Risk* **18**(4), 79–84.

Giese, G. (2006). A saddle for complex credit portfolio models. *Risk* **19**(7), 84–89.

Gupton, G., and Stein, R. (2002). LossCalc™: Moody's model for predicting loss given default (LGD). Moody's Investors Service.

Gupton, G., Finger, C., and Bhatia, M. (1997). CreditMetrics. Technical Document, JP Morgan, New York.

Hastie, T. J., and Tibshirani, R. J. (1990). *Generalized Additive Models*. Chapman & Hall, New York.

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

Hu, Y. T., and Perraudin, W. (2002). The dependence of recovery rates and defaults. Working Paper, Birkbeck College.

Huang, X., Oosterlee, C. W., and Mesters, M. (2007a). Computation of VaR and VaR contribution in the Vasicek portfolio credit loss model: a comparative study. *The Journal of Credit Risk* **3**(3), 75–96.

Huang, X., Oosterlee, C. W., and van der Weide, J. A. M. (2007b). Higher order saddlepoint approximations in the Vasicek portfolio credit loss model. *The Journal of Computational Finance* **11**(1), 93–113.

Lee, Y., and Nelder, J. A. (1998). Generalized linear models for the analysis of quality-improvement experiments. *Canadian Journal of Statistics* **26**(1), 95–105.

Lee, Y., and Nelder, J. A. (2001). Hierarchical generalized linear models: a synthesis of generalized linear models, random effect models and structured dispersions. *Biometrika* **88**, 987–1006.

Lugannani, R., and Rice, S. (1980). Saddlepoint approximations for the distribution of the sum of independent random variables. *Advances in Applied Probability* **12**, 475–490.

Martin, R., Thompson, K., and Browne, C. (2001a). Taking to the saddle. *Risk* **14**(6), 91–94.

Martin, R., Thompson, K., and Browne, C. (2001b). VaR: who contributes and how much? *Risk* **14**(8), 99–102.

McCullagh, P., and Nelder, J. (1989). *Generalized Linear Models*. Chapman & Hall, London.

McNeil, A. J., and Wendin, J. P. (2006). Dependent credit migrations. *The Journal of Credit Risk* **2**(3), 87–114.

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

Nelder, J. A., and Lee, Y. (1991). Generalized linear models for the analysis of Taguchi-type experiments. *Applied Stochastic Models and Data Analysis* **7**, 107–120.

Nelder, J. A., and Wedderburn, R. W. M. (1972). Generalized linear models. *Journal of the Royal Statistical Society* A **135**(3), 370–384.

Pan, J., and Thompson, R. (2007). Quasi-Monte Carlo estimation in generalized linear mixed models. *Computational Statistics and Data Analysis* **51**, 5756–5775.

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

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

Schuermann, T. (2004). What do we know about loss given default? In *Credit Risk Models and Management*, 2nd edn, Shimko, D. (ed), pp. 249–274. Risk Books, London.

Vasicek, O. (2002). Loan portfolio value. *Risk* **15**(12), 160–162.

Wedderburn, R. (1974). Quasi-likelihood functions, generalized linear models, and the Gauss–Newton method. *Biometrika* **61**(3), 439–447.