JOURNAL OF STATISTICAL THEORY AND PRACTICE
2016, VOL. 10, NO. 2, 342–356
http://dx.doi.org/10.1080/15598608.2016.1141733

Modeling stochastic recovery rates and dependence between default rates and recovery rates within a generalized credit portfolio framework

Matthias Fischer[a], Christoph Köstler[a], and Kevin Jakob[b]

[a]Department of Statistics & Econometrics, University of Erlangen-Nürnberg, Nürnberg, Germany; [b]Department of Statistics, University of Augsburg, Augsburg, Germany

ABSTRACT
Without any doubt, the CreditRisk$^+$ model that was launched by Credit Suisse Financial Products in 1997 is one of the most popular credit portfolio models in the banking industry. In order to accommodate more flexible dependence structures, Fischer and Dietz in 2012 introduced a generalized CreditRisk$^+$ framework. Focusing on the extension of Fischer and Dietz, the contribution of this article is twofold: First, we derive an analytic framework that allows for stochastic recovery rates, and for which the corresponding risk figures can be obtained via saddlepoint approximation. Second, we propose a straightforward approach for how to take dependencies between recovery rates and default rates into account. The corresponding loss distribution has to be derived using Monte Carlo simulations. We illustrate the effects of both stochastic recovery rates and dependence between recovery rates and default rates on the level of risk figures for a specific benchmark portfolio.

ARTICLE HISTORY
Received 1 July 2015
Accepted 11 January 2016

KEYWORDS
Credit portfolio model; CreditRisk$^+$; sector correlation; stochastic LGDs; saddlepoint approximation; Monte Carlo; PD-LGD correlation

AMS SUBJECT CLASSIFICATION
62P05

1. Introduction

The focus of this contribution is on credit risk, one of the most important risk types in the classical banking industry. Typically, the amount of economic capital to be reserved for credit risk is determined with a credit portfolio model.

Two of the most widespread models are CreditMetrics, launched by J.P. Morgan in 1997 and CreditRisk$^+$ (briefly: CR$^+$), an actuarial approach proposed by Credit Suisse Financial Products (1997) in 1997. In both cases one faces the challenge to model the dependence between the counterparties in the underlying portfolio. Because in practice the number of counterparties is simply too high to model their dependencies directly, the counterparty dependence structure is commonly reduced to the dependency structure between suitable sector variables.[1]

In the standard CR$^+$ model the sectors are assumed to be independent. In practice, there is no reason to believe that this assumption holds, particularly if we think of industry or country sectors. Consequently, several proposals appeared in order to introduce

CONTACT Matthias Fischer Matthias.Fischer@fau.de Department of Statistics & Econometrics, University of Erlangen-Nürnberg, Lange Gasse 20, D-90402 Nürnberg, Germany.

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/ujsp.

[1]In general, dependence between counterpart arises because they belong to the same sector or because their corresponding sectors itself are dependent.

correlated sectors. At first, Bürgisser et al. (1999) and Giese (2004) came up with two models that can be embedded within the standard CR+ setting but still failed to rebuild different patterns of variance–covariance structures. Han and Kang (2008) published a more sophisticated model, where each sector variable is supplemented by a common background factor. Fischer and Dietz (2011/2012) introduced the common background vector (CR+-CBV) model as a multivariate extension of the common factor model of Han and Kang (2008) from one factor to several background factors. However, within the CR+-CBV framework of Fischer and Dietz, relevant risk parameters such as recovery rates[2] are assumed to be deterministic, and dependence between recovery rates and default rates is neglected.

Against this background, the outline of this article is as follows: section 2 recaps the generalized CR+ model of Fischer and Dietz (2011/2012), whereas section 3 pays attention to the integration of uncertainty in the measurement of recovery rates within an analytic framework, generalizing Gordy (2004), who brought up a sophisticated approach to integrate severity risk into the standard CR+ setting. Section 4 is dedicated to the inclusion of dependence between default rates and recovery rates into the underlying framework[3]. Finally, section 5 illustrates the effects of both stochastic recovery rates and dependence between recovery rates and default rates on the level of risk figures for the benchmark portfolio. Section 6 concludes.

2. The CR+-CBV model: A short primer

Assume that our loan portfolio consists of n counterparties. Let further $\mathcal{A} = \{A, B, \ldots\}$ denote the set of all counterparties, $p_A \in [0,1]$ the probability of default (PD), $\nu_A > 0$ the arbitrary but deterministic and discretized exposure at default (EAD), and $\epsilon_A \in [0,1]$ the loss given default (LGD) of counterparty $A \in \mathcal{A}$. With default indicator $\mathbf{1}_A$ and $\mathbb{P}(1_{\mathbb{A}} = 1) = p_{\mathbb{A}}$, the general portfolio loss (variable) is given by

$$X = \sum_{A \in \mathcal{A}} X_A = \sum_{A \in \mathcal{A}} \nu_A \epsilon_A 1_A.$$

The product $\nu_A \epsilon_A$ quantifies the *potential loss* of counterparty A (in case of a default), whereas the corresponding *expected loss* of counterparty A reads as $\mathbb{E}(\mathbb{X}_{\mathbb{A}}) = \nu_{\mathbb{A}} \epsilon_{\mathbb{A}} p_{\mathbb{A}}$. Within the standard CR+ framework, the default indicator $\mathbf{1}_A$ is replaced by a Poisson variable[4] D_A with specific intensity

$$\lambda_A^S = p_A(w_{A0} + w_{A1} S_1 + \ldots + w_{AK} S_K), \tag{1}$$

where $S_1, \ldots, S_K$ are certain independent and Gamma-distributed sector variables representing the state of the economic sector(s). More generally, Fischer and Dietz (2011/2012), replace Eq. (1) by

[2]The loss given default (LGD) equals the percentage amount of the outstanding exposure at default that cannot be recovered. Equivalently, 1 – *LGD* is commonly termed as recovery rate, briefly RR.
[3]For instance, Frye (2000) or Altman et al. (2005) found empirical evidence for dependence between default rates and recovery rates.
[4]This can be regarded as reasonable approximation, at least for small probabilities of default.

$$\lambda_A^S = p_A(w_{A0} + w_{A1}\bar{S}_1 + \ldots + w_{AK}\bar{S}_K) \quad \text{with} \quad \bar{S}_k \equiv \delta_k S_k + \sum_{l=1}^{L} \gamma_{lk} T_l \tag{2}$$

with sector variables $S_k \sim \Gamma(\theta_k, 1)$ for $k = 1, \ldots, K$ as in the standard setting and additional background factors $T_l \sim \Gamma\left(\hat{\theta}_l, 1\right)$ for $l = 1, \ldots, L$, which are all assumed to be mutually independent. Hereby L denotes the number of background variables or, equivalently, the dimension of the background vector. For $L = 0$, the classical CR$^+$ setting is recovered, whereas $L = 1$ includes the already-mentioned proposal of Han and Kang (2008). Per construction, $\bar{S}_1, \ldots, \bar{S}_K$ are now dependent sector variables, where dependence arises through the additional common background variables $T_1, \ldots, T_L$. It is one of the major advantages of the CR$^+$-CBV model that Eq. (2) can be reformulated as a standard CR$^+$ model with $K + L$ sectors, setting $\tilde{S}_i = \delta_i S_i$ for $i \leq K$ and $\tilde{S}_i = T_i$ otherwise:

$$\lambda_A^S = p_A\left(w_{A0} + w_{A1}\tilde{S}_1 + \ldots + w_{AK}\tilde{S}_K + w_{A,K+1}\tilde{S}_{k+1} + \ldots + w_{A,K+L}\tilde{S}_{K+L}\right) \tag{3}$$

with artificial weights $w_{A,K+l} = \sum_{k=1}^{K} w_{Ak}\gamma_{lk}$ for $l = 1, \ldots, L$. According to the assumptions already described, the first and second moments derive as

$$\mathbb{E}(\bar{S}_k) = \delta_k\theta_k + \sum_{l=1}^{L} \gamma_{lk}\hat{\theta}_l,\ Var(\bar{S}_k) = \delta_k^2\theta_k + \sum_{l=1}^{L} \gamma_{lk}^2\hat{\theta}_l,\ Cov\left(\bar{S}_i, \bar{S}_j\right) = \sum_{l=1}^{L} \gamma_{li}\gamma_{lj}\hat{\theta}_l, \quad i \neq j.$$

In line with the work of Han and Kang (2008), the unknown parameters of the CR$^+$-CBV model are chosen such that the (normalized) Euclidian distance between the empirical and theoretical covariance matrix is minimal, respecting the side condition that θ_k, $\widehat{\theta}_l$, $\delta_k \geq 0$ and

$$\mathbb{E}(\bar{S}_k) = \delta_k\theta_k + \sum_{l=1}^{L} \gamma_{lk}\widehat{\theta}_l = 1, \quad w_{A,K+l} \geq 0 \ (A \in \mathcal{A}, l = 1, \ldots, L).$$

The last assumption ensures that $\mathbb{E}\left(\lambda_A^S\right) = p_A$.

3. Including stochastic recovery rates

In case of a counterparty default, not all of the outstanding exposure will be lost. Typically, banks receive revenues from the liquidation of collaterals and/or from the insolvency quota. The loss given default (LGD) equals the percentage amount of the outstanding exposure at default that cannot be recovered. Equivalently, 1 – *LGD* is commonly termed as recovery rate (RR). Typically, LGDs (or RRs) are estimated from historical loss data for homogenuous segments, and are applied to the living portfolio as a best a priori estimator. Keeping the finite number of observations and possible idiosyncratic effects or macroeconomic influences in mind, one might imagine that all estimators are fraught with uncertainty that should be taken into account when we derive forecasts for the loss distribution.

3.1. *General framework*

Gordy (2004) found a general solution to the shortcoming of missing stochastic LGDs in the standard CR+ framework. He replaced the standard recursion algorithms to calculate the loss distribution like Panjer recursion (see Gundlach and Lehrbass 2004) and nested evaluation (see Haaf et al. 2004) by the saddlepoint approximation, which is described in section 3.3. Thus, discretization of losses is no longer necessary.

The only additional requirement to apply this approach is the independence of the LGDs ϵ_A, $A \in \mathcal{A}$, among each other and from the intensities λ_A^S. This is necessary to retain an analytic solution for the portfolio loss distribution. In the case of dependence between ϵ_A and λ_A^S (so-called PD-LGD dependence) one has to make use of Monte Carlo simulations, as we demonstrate in section 4.

Given mean and standard deviation of the LGDs on a counterparty level, the distribution of the LGDs can be specified. The only requirement is the existence of the moment generating function M_{ϵ_A} of the stochastic LGD $\epsilon_A(A \in \mathcal{A})$. Therewith, the moment generating function of the portfolio loss X under the assumption of stochastic LGDs can be deduced with some additional calculation steps as follows.

We incorporate severity risk into the CR+-CBV model by defining the potential loss (PL) $\bar{\nu}_A = \nu_A \cdot \epsilon_A$ as a random variable dependent on the deterministic exposure at default ν_A and the loss given default ϵ_A, which is assumed to be a random variable from here on. The conditional moment generating function of x_A is given by

$$\begin{aligned} M_{X_A|\tilde{S}}(z) &= \mathbb{E}_{\epsilon_A}\left[M_{X_A|\tilde{S},\epsilon_A}(z)|\tilde{S}\right] = \mathbb{E}_{\epsilon_A}\left[1-\lambda_A^S+\lambda_A^S\exp(z\nu_A\epsilon_A)|\tilde{S}\right] \\ &= 1-\lambda_A^S+\lambda_A^S\mathbb{E}_{\epsilon_A}[\exp(z\nu_A\epsilon_A)], \end{aligned}$$

where $\mathbb{E}_{\epsilon_A}[\exp(z\nu_A\epsilon_A)]$ is the moment generating function of ϵ_A evaluated at $z\nu_A$, implying

$$\begin{aligned} M_{X_A|\tilde{S}}(z) &= 1-\lambda_A^S+\lambda_A^S M_{\epsilon_A}(z\nu_A) = 1+\lambda_A^S(M_{\epsilon_A}(z\nu_A)-1) \\ &\approx \exp\left(\lambda_A^S(M_{\epsilon_A}(z\nu_A)-1)\right). \end{aligned}$$

Due to the conditional independence of $(X_A)_{A\in\mathcal{A}}$ we get

$$\begin{aligned} M_{X|\tilde{S}}(z) &= \prod_{A\in\mathcal{A}} M_{X_A|\tilde{S}}(z) = \exp\left(\sum_{A\in\mathcal{A}}\lambda_A^S(M_{\epsilon_A}(z\nu_A)-1)\right) \\ &= \exp\left(\sum_{A\in\mathcal{A}}p_A\sum_{k=0}^{K+L}w_{Ak}\tilde{S}_k(M_{\epsilon_A}(z\nu_A)-1)\right) = \exp\left(\sum_{k=0}^{K+L}\tilde{S}_kP_k(z)\right) \end{aligned} \tag{4}$$

with sector polynomials $P_k(z) = \sum_{A\in\mathcal{A}} p_A w_{Ak}(M_{\epsilon_A}(z\nu_A)-1)$, $k = 0,\ldots,K+L$, which provide a condensed representation of the obligor-specific risk in sector *k;* see Gundlach and Lehrbass (2004, chap. 6). Thus, the unconditional moment generating function of X has the form

$$M_X(z) = \mathbb{E}_{\tilde{S}}\left[M_{X|\tilde{S}}(z)\right] = \exp(P_0(z))\prod_{k=1}^{K+L}M_{\tilde{S}_k}(P_k(z)).$$

Hence, the moment generating function of X for the CR+-CBV model with stochastic LGDs is given by

$$M_X(z) = \exp(P_0(z)) \prod_{k=1}^{K} (1 - P_k(z)\delta_k)^{-\theta_k} \prod_{l=1}^{L} (1 - P_{K+l}(z))^{-\widehat{\theta_l}}. \tag{5}$$

Thus, we end up with a general representation of the portfolio loss distribution in terms of the moment generating functions M_{ϵ_A} of the LGDs ϵ_A, $A \in \mathcal{A}$. In a second step the distribution of LGDs has to be chosen.

3.2. *The distribution of LGDs*

The only restriction to the distribution of ϵ_A is the existence of its moment generating function. From an economic point of view, the possible values of ϵ_A should be restricted to the interval [0, 1]. A distribution fulfilling these requirements is the beta distribution. Because of its support on [0, 1] and its very adjustable structure, it is used as the standard LGD distribution in the banking practice. The density of the Beta distribution is given by

$$f(x) = \begin{cases} \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1} & \text{for } x \in (0,1), \\ 0 & \text{otherwise,} \end{cases}$$

with $a, b > 0$ and $B(a, b)$ as the Beta function. The corresponding moment generating function does not admit a closed-form representation. Indeed, it can be written as a power series that converges on the whole real axis:

$$M_{\epsilon_A}(t) = 1 + \sum_{n=1}^{\infty} \left(\prod_{r=0}^{n-1} \frac{a_A + r}{a_A + b_A + r} \right) \frac{t^n}{n!}, \quad t \in \mathbb{R}. \tag{6}$$

This power series has to be approximated by a finite sum, when it is implemented on the computer. To save computation time other distributions with closed form moment generating functions can be used. For this purpose, Gordy (2004) suggested using the Gamma distribution as LGD distribution. From an economic point of view, however, this assumption is not appropriate because of the corresponding support of the Gamma distribution on $[0, \infty]$. The parameters of the LGD distribution have to be chosen in such a way that the expectation and the standard deviation of ϵ_A is equal to the corresponding estimated empirical mean $\widehat{\mu}$ and standard deviation $\widehat{\sigma}$.

3.3. *Derivation of the loss distribution*

In order to derive the loss distribution or derivations thereof we use the formula of Lugannani and Rice (1980), which was introduced by Gordy (2004) to the standard CR^+ framework. This formula provides an approximation of the distribution function in the tail area, but it also can be applied to the whole support of the distribution[5]. With regard to some other necessary assumptions, which are satisfied in the CR^+ setting (see Gordy 2004), the main requirement is the knowledge of the cumulant generating function (CGF) and its derivatives. Then the survival function can be approximated by

[5] In the case of very skew loss distributions the Lugannani–Rice formula does not work very well. Wood et al. (1993) generalized this formula and allowed arbitrary continuous base distributions with existing cumulant generating functions.

$$\mathbb{P}(X > y) \approx 1 - \Phi(w) + \phi(w)\left(\frac{1}{u} - \frac{1}{w}\right), \tag{7}$$
$$w = \text{sign}(\hat{z})\sqrt{2(\hat{z}y - K_X(\hat{z}))} \quad u = \hat{z}\sqrt{K_X''(\hat{z})}$$

with K_X as the CGF of the portfolio loss, Φ and respectively ϕ as the distribution and density function of the standard normal distribution, and the saddlepoint $\hat{z}$ as the unique solution of the saddlepoint equation $K_X'(z) = y$. In the CR$^+$-CBV model,

$$K_X(z) = \ln(M_X(z)) = P_0(z) - \sum_{k=1}^{K} \theta_k \cdot \ln(1 - P_k(z)\delta_k) - \sum_{l=1}^{L} \hat{\theta}_l \cdot \ln(1 - P_{K+l}(z)) \tag{8}$$

with $P_k(z)$ as defined earlier. Gordy (2004) showed how to calculate the derivatives of K_X in the standard CR$^+$ setting. Knowing that the CR$^+$-CBV setup can be translated back to the standard CR$^+$ approach, the derivatives in the CR$^+$-CBV setting can be deduced analogously:

$$K_X(z) = P_0(z) + \sum_{k=1}^{K} \psi_k(z) + \sum_{l=1}^{L} \widehat{\psi}_l(z) \tag{9}$$

with $\psi_k(z) = -\theta_k \cdot \ln(1 - P_k(z)\delta_k)$ for $k = 1, \ldots, K$ and $\widehat{\psi}_l(z) = -\widehat{\theta}_l \cdot \ln(1 - P_{K+l}(z))$ for $l = 1, \ldots, L$. By derivating $\psi_k(z)$ and $\widehat{\psi}_l(z)$ we receive the derivatives of K_X. It can be shown that

$$\psi_k'(z) = \theta_k \delta_k \frac{P_k'(z)}{1 - \delta_k P_k(z)}, \quad \widehat{\psi}_l'(z) = \widehat{\theta}_l \frac{P_{K+l}'(z)}{1 - P_{K+l}(z)},$$
$$\psi_k''(z) = \theta_k \delta_k \left(\frac{P_k''(z)}{1 - \delta_k P_k(z)} + \delta_k \left(\frac{P_k''(z)}{1 - \delta_k P_k(z)}\right)^2\right),$$
$$\widehat{\psi_l''}(z) = \widehat{\theta}_l \left(\frac{P_{K+l}''(z)}{1 - P_{K+l}(z)} + \left(\frac{P_{K+l}'(z)}{1 - P_{K+l}(z)}\right)^2\right).$$

With reference to section 3.1, the functions P_k, $k = 0, \ldots, K + L$, have the following representation:

$$P_k(z) = \sum_{A \in \mathcal{A}} p_A w_{Ak} \sum_{n=1}^{\infty} \left(\prod_{r=0}^{n-1} \frac{a_A + r}{a_A + b_A + r}\right) \frac{(z\nu_A)^n}{n!}. \tag{10}$$

We get the derivatives of P_k by derivating each addend:

$$P_k'(z) = \sum_{A \in \mathcal{A}} p_A w_{Ak} \nu_A \sum_{n=0}^{\infty} \left(\prod_{r=0}^{n} \frac{a_A + r}{a_A + b_A + r}\right) \frac{(z\nu_A)^n}{n!},$$
$$P_k''(z) = \sum_{A \in \mathcal{A}} p_A w_{Ak} \nu_A^2 \sum_{n=0}^{\infty} \left(\prod_{r=0}^{n+1} \frac{a_A + r}{a_A + b_A + r}\right) \frac{(z\nu_A)^n}{n!}.$$

Now we have deduced an algorithm to calculate the portfolio loss CGF and its first two derivatives. After that we apply the Lugannani–Rice formula to approximate the loss distribution within the CR$^+$-CBV setting with by stochastic LGDs. In short:

Algorithm 1. Calculating the saddlepoint approximated distribution function between expected loss und maximum loss of a portfolio in the CR+-CBV model enhanced by stochastic LGDs.

Step 1: Calculate z_{max} by solving $K'_X(z) = y_{max}$ with $y_{max} = \sum_A \nu_A$ as the maximum portfolio loss.

Step 2: Calculate z_{min} by solving $K'_X(z) = y_{min}$, where $y_{min} = \sum_A \mathbb{E}(L_A)$ as the expected portfolio loss.

Step 3: Calculate the cumulant generating function and its derivatives for some points $t_1, \ldots, t_n$ between z_{min} and z_{max} with $z_{min} = t_1 < t_2 < \ldots < t_{n-1} < t_n = z_{max}$. Note that the corresponding portfolio loss values y_i are determined by the saddlepoint equation: $y_i = K'_X(t_i)$ for $i = 1, \ldots, n$.

Step 4: Insert the results of step 3 and the corresponding saddlepoints into the formula of Lugannani and Rice to determine the approximated values of the distribution function of the portfolio loss variable X:

$F_X(y_i) = \Phi(w_i) + \phi(w_i)(1/w_i - 1/u_i)$ with
$w_i = \text{sign}(t_i)\sqrt{2(t_i y_i - K_X(t_i))}$ and $u_i = t_i\sqrt{K''_X(t_i)}$.

Following algorithm 1, the saddlepoint equation has to be solved two times only (steps 1 and 2) in order to receive the lower and the upper bound of the relevant saddlepoint interval. If there is an index $i \in \{1, \ldots, n\}$ with $\tilde{F}_X(y_{i-1}) < \alpha < \tilde{F}_X(y_i)$, where $\tilde{F}$ is the approximated distribution function of the portfolio loss, then the following estimate of the approximative value at risk holds: $y_{i-1} < \widetilde{\text{VaR}}_\alpha(X) < y_i$. For $|y_i - y_{i-1}| \to 0$ the approximative value at risk can be estimated with arbitrary precision. If one is interested in the value at risk of one particular confidence level, this value can be directly determined by solving the respective saddlepoint formula for y.

4. Including dependence between default and recovery rates/LGDs

4.1. *General framework*

Uncertainty of stochastic LGDs is only one aspect. There is also empirical evidence for dependence between defaults and recovery rates, see Altman et al. (2005) or Miu and Ozdemir (2006), for instance. One way to integrate this kind of dependence into the CR+-CBV framework is to use again a combination of sector specific and idiosyncratic factors similar to the default mechanism. In order to simplify the construction, we restrict ourselves to the case where every counterparty is assigned to exactly one sector variable $\bar{S}_k$, that is, for every $A \in \mathcal{A}$ there exists exactly one $k*(A) = 1, \ldots, K$ such that $w_{A,k*(A)} = 1$ and $w_{Ak} = 0$ for all $k \neq k*(A)$. For $\rho_A \in [0, 1]$ we define a latent variable driving the LGD of counterparty A as

$$\mathcal{L}_A := \rho_A L_{k*(A)} + \sqrt{1 - \rho_A^2}\, U_A, \tag{11}$$

where L_k (for $k = 1, \ldots, K$) and U_A (for $A \in \mathcal{A}$) follow a standard normal distribution and are independent from each other. By construction, $\mathcal{L}_A$ again is standard normally distributed. Here, L_k represents the sector specific and U_A the idiosyncratic influence, whereas the parameter ρ_A controls their influence on the LGD. In order to introduce dependence between

PD and LGD, the systematic factor L_k is defined as $L_k := \phi^{-1}(F_k(\bar{S}_k))$, where F_k denotes the distribution function[6] of $\bar{S}_k$ and ϕ^{-1} is the quantile function of the standard normal distribution. Therefore, we now have an one-to-one dependence between L_k and $\bar{S}_k$ (i.e., they are comonotone). The resulting dependence between PD and LGD is governed by the parameter ρ_A. In order to ensure that LGD $\in [0, 1]$, $\mathcal{L}_{\mathcal{A}}$ is finally transformed into a Beta-distributed random variable via $\epsilon_A := \beta^{-1}_{a,b}(\phi(\mathcal{L}_A))$, where $\beta^{-1}_{a,b}$ denotes the quantile function of a Beta distribution with parameters a and b according to section 3.2.

4.2. *Derivation of the loss distribution*

As mentioned in section 3, independence of LGDs is necessary in order to retain an analytical solution. Therefore, we have to utilize a Monte Carlo simulation of the CR⁺-CBV model in order to derive the portfolio loss distribution if we switch from independent to dependent LGDs. The simulation framework based on section 4.1 is given in Algorithm 2.

Algorithm 2. Monte Carlo algorithm for simulative CR⁺-CBV model.
For $n = 1, \ldots, N$ #(simulation loop)
Determine sector realizations:

Draw $S^{(n)}_{k=1,\ldots,K} \sim \Gamma(\theta_k, 1)$ and $T^{(n)}_{l=1,\ldots,L} \sim \Gamma\left(\hat{\theta}_l, 1\right)$ independently

Calculate $\bar{S}^{(n)}_k = \delta_k S^{(n)}_k + \sum_{l=1}^{L} \gamma_{l,k} T^{(n)}_l$

In case of PD-LGD dependence: $L^{(n)}_k = \phi^{-1}\left(F_k\left(\bar{S}^{(n)}_k\right)\right)$

For all $A \in \mathcal{A}$ #(counterparty loop)
Determine counterparty default:

$\lambda^{(n)}_A = p_A \cdot \left(w_{A0} + \sum_{k=1}^{K} w_{Ak} \bar{S}_k\right)$

$D^{(n)}_A \sim \text{Pois}\left(\lambda^{(n)}_A\right)$

Determine loss given default if $D^{(n)}_A = 1$:

In case of deterministic LGDs,

$$\epsilon^{(n)}_A = \mathbb{E}(\epsilon_A)$$

In case of stochastic but independent LGDs,

$\epsilon^{(n)}_A \sim \beta(a_A, b_A)$, with a_A, b_A given in section 3.2.

In case of stochastic and dependent LGDs,

$\epsilon^{(n)}_A \sim \beta^{-1}_{a_A,b_A}\left(\phi\left(\mathcal{L}^{(n)}_A\right)\right)$, with $\mathcal{L}^{(n)}_A$ according to Eq. (11)

$X^{(n)}_A = \nu_A \epsilon^{(n)}_A D^{(n)}_A$

[6]Since F_k is unknown, we approximate F_k by its empirical version $\hat{F}_k$.

if $D_A^{(n)} = 0$

$$X_A^{(n)} = 0$$

Determine portfolio loss:

$$X^{(n)} = \sum_{A \in \mathcal{A}} X_A^{(n)}$$

Based on the realizations $X^{(1)}, \ldots, X^{(N)}$, the portfolio loss distribution and the risk figures can be estimated.

5. Results for the IACPM portfolio

In order to demonstrate the effects of our extensions on risk figures we now focus on a *benchmark portfolio* underlying the study on "convergence of credit capital models" that was performed by the International Association of Credit Portfolio Management (IACPM) and International Swaps and Derivatives Association (ISDA) in 2006. The main idea was to compare the capital measures (at the portfolio level) generated by different credit capital models—that is, expected loss (EL) for the portfolio and the amount of economic capital needed to support the credit risk of the portfolio at various specified confidence levels.

5.1. *Portfolio description*

For details on the underlying benchmark portfolio we refer to IACPM or ISDA[7] and to the distributed files *Portfolio Demographics.pdf* and *Test Portfolio for CR⁺ and Similar Models.xls*. In detail, the US \$100 billion test portfolio is comprised of two term loans to each of 3, 000 obligors across a diverse set of industries and 7 countries dispersed along 8 whole-grade rating buckets and varying LGDs. Exposure amounts varied from 1 to 1, 250 million and tenors ranged from 6 months to 7 years. For our purpose, all tenors are set to 1 year. Moreover, we restricted ourselves to 10 GICS[8] sectors only and neglected the country information. More details on the portfolio structure (risk figure in total, by rating structure and by industry) are given in Tables 1, 2, and 3.

5.2. *Specification of the sector VCV*

In order to specify the sector variances and sector correlation, $T = 15$ yearly time series $x_{1,i}, \ldots, x_{15,i}$ (for $i = 1, \ldots, 10$ sectors) on U.S. corporate defaults from 1995 to

Table 1. Portfolio characteristics: Calculation of EL and PL as stated in section 2.

Name	Value
Number of counterparties	3,000
Portfolio exposure (EAD)	100,000,000,000
Portfolio potential loss (PL)	40,552,920,000
Expected loss (EL)	605,258,124
Average PD (arithmetic mean)	0.0173
Average PD (EAD weighted)	0.0151

[7]See http://www.isda.org or http://www.iacpm.org.
[8]Global industry classification standard, see http://www.msci.com/gics.

Table 2. Risk figures by industry sectors $\mathcal{S}$.

Number	Sector name	Number of CP	EAD	PL	EL
1	Energy	239	12,382,000,000	5,021,260,000	57,988,047
2	Materials	472	17,422,000,000	7,104,600,000	134,098,406
3	Industrials	614	14,186,000,000	5,927,880,000	80,364,896
4	Consumer discretionary	634	20,180,000,000	7,803,100,000	148,957,421
5	Consumer staples	222	3,835,000,000	1,578,860,000	17,808,856
6	Health care	103	1,049,000,000	406,480,000	5,358,676
7	Financials	312	14,277,000,000	5,853,480,000	47,012,264
8	Information technology	150	5,012,000,000	2,071,460,000	68,278,467
9	Telecommunication service	84	7,497,000,000	3,066,040,000	11,312,357
10	Utilities	170	4,160,000,000	1,719,760,000	34,078,735
Σ	3000	100,000,000,000	40,552,920,000	605,258,124	

Table 3. Risk figures by rating grades $\mathcal{R}$.

Rating	PD	Number of at CP	EAD	PL	EL
1	0.0001	87	3,332,000,000	1,322,100,000	135,781
2	0.0002	213	6,640,000,000	2,854,340,000	587,071
3	0.0002	568	20,667,000,000	8,389,740,000	1,732,522
4	0.0018	1192	38,862,000,000	15,819,440,000	28,163,209
5	0.0127	637	20,513,000,000	8,219,760,000	104,484,723
6	0.0664	190	7,259,000,000	2,844,360,000	188,851,104
7	0.2550	113	2,727,000,000	1,103,180,000	281,303,715
Σ		3,000	100,000,000,000	40,552,920,000	605,258,124

2009 (from Standard & Poor's Credit Pro) have been used. Let $\bar{x}_i$ denote the average (observed) default rate for sector i; then the classical estimators are given by

$$\widehat{\sigma}_i^2 = \frac{T\sum_{t=1}^{T} x_{t,i}^2 - \left(\sum_{t=1}^{T} x_{t,i}\right)^2}{T(T-1)\bar{x}_i^2}, \qquad \widehat{\Sigma}_{ij} = \frac{\frac{1}{T}\sum_{s,t=1}^{T}\left(x_{t,i}-\bar{x}_i\right)\left(x_{s,j}-\bar{x}_j\right)}{\sqrt{\frac{T\sum_{t=1}^{T} x_{t,i}^2 - \left(\sum_{t=1}^{T} x_{t,i}\right)^2}{T(T-1)}}\sqrt{\frac{T\sum_{t=1}^{T} x_{t,j}^2 - \left(\sum_{t=1}^{T} x_{t,j}\right)^2}{T(T-1)}}},$$

with estimation results given in Tables 4 and 5. We observe sector variances above average for the sectors energy and utilities. Above that, empirical correlations are—with one exception—positive.

5.3. *Results on risk figures: Stochastic LGDs*

Finally, based on the IACPM benchmark portfolio, we calculate the value at risk (VaR as the quantile of the loss distribution) of the portfolio loss for different confidence

Table 4. Estimated sector variances.

Number	Sector name	$\widehat{\sigma}_i^2$
1	Energy	2.0769
2	Materials	0.9047
3	Industrials	1.0270
4	Consumer discretionary	0.9692
5	Consumer staples	0.8019
6	Health care	1.2077
7	Financials	1.1302
8	Information technology	1.2681
9	Telecommunication services	1.6537
10	Utilities	2.5277

Table 5. Empirical sector correlation matrix.

$\widehat{\Sigma}_{ij}$	1	2	3	4	5	6	7	8	9	10
1	1.0000	0.4194	0.3902	0.5235	0.1655	0.4348	0.3972	0.2061	0.2205	0.2023
2		1.0000	0.9141	0.9029	0.5665	0.2629	0.5809	0.5980	0.5681	0.6453
3			1.0000	0.8166	0.7065	0.3674	0.3916	0.7585	0.5831	0.6146
4				1.0000	0.5278	0.2123	0.7658	0.4005	0.2654	0.3604
5					1.0000	0.6334	0.3733	0.7510	0.3725	0.2218
6						1.0000	0.1247	0.6268	0.2619	0.1113
7							1.0000	0.0870	0.0359	-0.0211
8								1.0000	0.7379	0.5889
9									1.0000	0.8497
10										1.0000

Table 6. Quantiles of the loss distribution with deterministic LGDs computed by saddlepoint approximation relative to the standard CR$^+$ model and the corresponding relative deviation (RD) between CR$^+$-CBV/standard CR$^+$ model and saddlepoint approximation.

L	VaR90	RD	VaR95	RD	VaR95	RD	VaR99.95	RD
Standard CR$^+$	1.000	0.07%	1.000	0.08%	1.000	0.09%	1.000	0.11%
CR$^+$-CBV(1)	1.177	1.16%	1.340	0.12%	1.654	0.52%	2.065	0.64%
CR$^+$-CBV(2)	1.199	0.53%	1.349	0.17%	1.625	0.03%	1.978	0.14%
CR$^+$-CBV(5)	1.148	2.95%	1.316	2.04%	1.654	0.65%	2.109	0.15%

levels, applying the approach discussed in section 3. The results are presented in Table 6. With respect to high quantiles, CR$^+$-CBV models put more probability mass in the tail area (see, e.g., Fischer and Dietz 2011/2012) compared to a standard CR$^+$ setting. In addition, in order to quantify the effects of the sector correlation on the risk figures, we also calculated the corresponding quantiles of the CR$^+$-CBV model with co-monotone sectors within the Monte Carlo framework (abbreviation: CR$^+$-MAX): The results are very close to those of CR$^+$-CBV(5): $VaR90 = 1.211$, $VaR95 = 1.364$, $VaR99 = 1.855$, and $VaR99.95 = 2.598$.

Next, we compare the results of the nested evaluation by Haaf et al. (2004) with the saddlepoint approximation of the distribution function. For this purpose, we choose the Lugannani–Rice formula to approximate the value at risk of the portfolio loss. This choice is suitable here because of the low skewness of the portfolio loss distribution in the particular models: The third standardized moment $\mathbb{E}\left[\left(\frac{X-\mathbb{E}[X]}{\mathrm{SD}(X)}\right)^3\right]$ attains the value 1.09 in the standard CR$^+$ model. In the CR$^+$-CBV model the values of the third standardized moment are higher than in the standard CR$^+$ setting, but remain smaller than 3 in the observed cases. As our focus is on extraordinary credit risk, we will look at some quantiles in the tail area. In order to quantify the accuracy of the approximation we use the relative deviation, which we define by

$$\mathrm{RD} := \frac{|\mathrm{VaR} - \widetilde{\mathrm{VaR}}|}{\mathrm{VaR}},$$

where VaR is the exact value at risk calculated by nested evaluation and $\widetilde{\mathrm{VaR}}$ is the approximated value at risk calculated by the Lugannani–Rice formula.

In the next step we integrate stochastic LGDs into our model and compare the results with them above. The Beta distribution is chosen for the distribution of LGDs because of its support on [0, 1] and sufficient flexibility (see also discussion in section 3.2).

Table 7. Quantiles of the loss distribution with deterministic and stochastic LGDs relative to the case of deterministic LGDs.

L	VaR90	VaR95	VaR99	VaR99.9	VaR99.95	VaR99.99
Standard scenario: σ_{high}						
Standard CR$^+$	1.090	1.119	1.162	1.196	1.204	1.217
CR$^+$-CBV(1)	1.045	1.049	1.052	1.055	1.055	1.056
CR$^+$-CBV(2)	1.044	1.050	1.056	1.059	1.060	1.062
CR$^+$-CBV(5)	1.046	1.050	1.053	1.057	1.057	1.059
CR$^+$-MAX	1.058	1.055	1.053	1.072	1.074	1.086
Scenario B: σ_{low}						
Standard CR$^+$	1.024	1.033	1.046	1.056	1.059	1.063
CR$^+$-CBV(1)	1.012	1.013	1.013	1.014	1.014	1.014
CR$^+$-CBV(2)	1.011	1.013	1.014	1.015	1.015	1.016
CR$^+$-CBV(5)	1.012	1.013	1.014	1.015	1.015	1.015
CR$^+$-MAX	1.015	1.014	1.016	1.018	1.017	1.021
Scenario C: σ_{low}, μ_{low}						
Standard CR$^+$	1.089	1.125	1.181	1.227	1.237	1.256
CR$^+$-CBV(1)	1.046	1.051	1.055	1.058	1.058	1.059
CR$^+$-CBV(2)	1.044	1.051	1.059	1.063	1.064	1.065
CR$^+$-CBV(5)	1.047	1.052	1.056	1.060	1.061	1.062
CR$^+$-MAX	1.052	1.056	1.057	1.069	1.077	1.078

In Table 7 the significant influence of stochastic LGDs on the risk figures is discernible. The real data LGD volatilities of the IACPM portfolio are denoted as (σ_{high}). In order to compare the effects of LGD volatilities, Table 7 also contains the results for halved real data LGD volatilities $\sigma_{low} = 0.5 \cdot \sigma_{high}$ (Scenario B) and, in addition, for the halved LGD mean, that is, simulating higher collateralization quota (Scenario C). In addition, we again calculated the figures for the co-monotone dependence structure in order to check whether the results are depending primarily on the chosen sector VCV. Our results indicate that this is not the case.

There are four main effects that can be recognized from Table 7:

At first, the increase of value at risk depends on the correlation structure used in the model: The relative markup for (independent) stochastic LGDs, which can be understood as some sort of idiosyncratic risk, is higher if there is no sector dependence compared to the case if there is a positive dependence.

This is reasonable because higher systematic risk like in the case of the CR$^+$-CBV model leads to higher risk figures (even without stochastic LGD) and consequently to a higher value at risk in the preceding calculations (in comparison to the standard CR$^+$ model) if PD and LGD are independent.

A second observation is related to the quantile level. In our setup the impact of stochastic LGDs is larger for the higher quantile level. This, however, is not a general rule. The relative markups on different quantiles due to stochastic LGDs depend strongly on portfolio structure and absolute value at risk figures in relation to exposure values and LGD parameters (mean and standard deviation).

The third effect arises from the standard deviation of LGDs. As can be seen in Table 7, a doubling of the standard deviation leads to a substantial increase of value at risk (in our setup the markup rises nearly by factor 4).

Finally, Scenario C indicates that the relative markups on different quantiles due to stochastic LGDs increases when the collateralization quota increases.

Thus, severity risk should not be neglected in CR^+ because otherwise credit risk will be underestimated.

5.4. Results on risk figures: PD-LGD dependence

In a first step, we show that the risk figures of the simulation model (based on 5 million repetitions) correspond to those of the analytical one. In this way, we can prove that the simulation framework works well and that simulation errors are not substantial. Secondly, we illustrate how the dependence between PD and LGD affects the risk figures.

Table 8 shows the value at risk for two loss levels of the analytical frameworks (i.e., CR^+-CBV(5) model and saddlepoint approximation) and the simulation models as multiples of the CR^+-CBV(5) model. As one can see, the risk figures of the simulation model equals those of the CR^+-CBV(5) model in case of deterministic LGDs with a level of precision of three decimal places. Repeating the simulation again 10 times always gives us nearly the same results with a maximum deviation below 0.05%. Therefore, we can state that simulation errors are not substantial. Regarding the saddlepoint approximation we can state that the deviation to the ordinary CR^+-CBV model is clearly below 1%. Although precision of Monte Carlo simulation is higher compared to the saddlepoint approximation, the runtime of saddlepoint approximation is lower, on the other hand.

Finally, we analyze how the dependence between PD and LGD affects the loss distribution. Table 9 shows two quantiles of loss distributions of the CR^+-CBV(5) model, the simulation model with stochastic but independent LGDs, and the model with PD-LGD dependence for two different values of ρ as well as the expected loss. In contrast to the model with stochastic but independent LGDs, the expected loss increases by 7% (21%) if PD and LGD depend on each other. Similarly, the value at risk increases by 24% up to 79% depending on the loss level (higher levels increase faster) and the degree of dependence. The risk figures show that the dependence between risk parameters is a clearly higher source of risk compared to their purely random behavior.

The increase of risk figures is also illustrated in Figure 1, where the upper tail of the loss distributions in case of deterministic, stochastic, and dependent LGDs is illustrated. The

Table 8. Quantiles of analytical CR^+-CBV model (deterministic and stochastic LGDs with saddlepoint approximation, SA) and Monte Carlo simulation (MCS) framework relative to the CR^+-CBV(5) model.

LGD	Deterministic			σ_{low}		σ_{high}	
model	CR^+-CBV(5)	SA	MCS	SA	MCS	SA	MCS
VaR99	1	0.994	1.000	1.007	1.019	1.046	1.072
VaR99.95	1	1.002	1.000	1.016	1.031	1.059	1.102

Table 9. Risk figures within simulation model with PD-LGD dependence relative to deterministic model (i.e. CR^+-CBV(5) and CR^+-MAX model).

		CR^+-CBV(5)			CR^+-MAX
	-	$\rho = 0$	$\rho = 0.1$	$\rho = 0.3$	$\rho = 0.3$
	$\sigma = 0$	σ_{high}	σ_{high}	σ_{high}	σ_{high}
EL	1	1	1.070	1.214	1.214
VaR99	1	1.072	1.242	1.600	1.610
VaR99.95	1	1.101	1.330	1.792	1.840

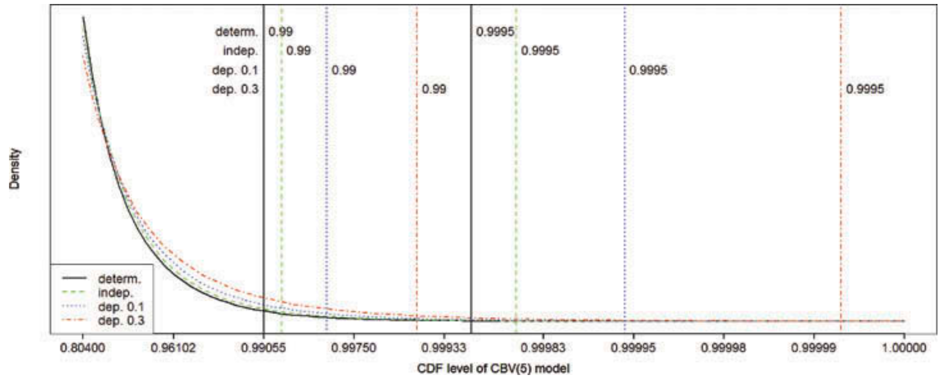

Figure 1. Upper tail of loss distributions of models with deterministic LGDs, stochastic LGDs (σ_{high}), and PD-LGD dependence with parameter $\rho \in \{0.1, 0.3\}$. The horizontal axis denotes the CDF level in case of deterministic LGDs. The vertical lines indicate the value at risk on level 0.99 and 0.9995.

horizontal axis denotes the level of the cumulative distribution function (CDF) in case of deterministic LGDs, whereas the vertical lines indicate the value at risk on different loss levels, clearly demonstrating how the risk increases along with dependence. For example, the value at risk on level 0.99 in case of dependent LGDs with $\rho_A = 0.3$ roughly corresponds to the value at risk in the deterministic model on level 0.999, whereas the 0.9995 value at risk of the dependent LGD model is around the 0.999996 loss level in the deterministic case, implying that losses above that level are approximately 120 times more probable in the framework with PD-LGD dependence compared to the deterministic case. In the case of co-monotone sectors, the effect is even higher.

6. Summary

The main focus of this contribution is to extend the CR$^+$-CBV framework of Fischer and Dietz (2011/2012) to stochastic recovery rates, on the one hand, and to dependence between default rates and recovery rates, on the other hand. Adopting the original proposal of Gordy (2004) for the standard CR$^+$ setting, the integration of stochastic recovery rates can be implemented within an analytic framework using the theory of saddlepoint approximation. In contrast, integrating dependence between default rates and recovery rates has to be done within a Monte Carlo setting, which we introduced in the second part of the article. For a well-known benchmark portfolio, the empirical results illustrate that PD-LGD dependence might increase the risk figures significantly (in particular compared to the increases that might arise from stochastic LGDs or modifications in the underlying sector correlation structure). However, a reliable estimation of the unknown parameters that govern the dependence between PD and LGD is still a crucial issue.

Acknowledgment

The authors thank two anonymous referees for their helpful comments and suggestions, which significantly improved the presentation of the article.

References

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

Bürgisser, P., A. Kurth, A. Wagner, and M. Wolf. 1999. Integrating correlations. *Risk Magazine* 12 (7):57–60.

Fischer, M., and C. Dietz. 2011/2012. Modelling sector correlations with CreditRisk$^+$: The common background vector (CBV) model. *Journal of Credit Risk* 7:23–43.

Frye, J. 2000. Depressing recoveries. *Risk* 13 (4):108–111.

Giese, G. 2004. Dependent risk factors. In *CreditRisk$^+$ in the banking industry*, M. Gundlach and F. Lehrbass, ed. 153–166. Berlin, Germany: Springer.

M. B. Gordy. 2004. Saddlepoint approximation. In *CreditRisk$^+$ in the banking industry*, M. Gundlach and F. Lehrbass, ed. 91–110. Berlin, Germany: Springer.

Gundlach, M., and F. Lehrbass. 2004. *CreditRisk$^+$ in the Banking Industry*. Berlin, Germany: Springer.

Haaf, H., O. Reiß, and J. Schoenmakers. 2004. Numerically stable computation of CreditRisk$^+$. In *CreditRisk$^+$ in the banking industry*, M. Gundlach and F. Lehrbass, ed. 69–78. Berlin, Germany: Springer.

Han, C., and J. Kang. 2008. An extended CreditRisk$^+$ framework for portfolio credit risk management. *Journal of Credit Risk* 4 (4):63–80.

Lugannani, R., and S. O. Rice. 1980. Saddlepoint approximation for the distribution of the sum of independent random variables. *Advances in Applied Probability* 12:1118–1124.

Miu, P., and B. Ozdemir. 2006. Basal requirement of downturn LGD: Modeling and estimating PD & LGD correlations. *Journal of Credit Risk* 2 (2):43–68.

Credit Suisse Financial Products. 1997. CreditRisk$^+$: A credit risk management framework. Technical paper. Credit Suisse. http://www.csfb.com/institutional/research/assets/creditrisk.pdf

Wood, A. T. A., J. G. Booth, and R. W. Butler. 1993. Saddlepoint approximations to the CDF of some statistics with nonnormal limit distribution. *Journal of the American Statistical Association* 88:680–686.