Unexpected recovery risk

For credit portfolio managers, the priority is to properly incorporate recovery rates into existing models. Here, Michael Pykhtin improves upon earlier approaches, allowing recovery rates to depend on the idiosyncratic part of a borrower's asset return, in addition to the systematic factor. Using a lognormal distribution of collateral value, ensuring that it always remains positive, he derives closed-form expressions for expected loss and economic capital

Models based on a conditional independence framework are central to portfolio credit risk management. Most such models[1] treat individual recoveries in the event of default either as constants or as independent stochastic variables.[2] However, empirical studies (for example, Moody's, 2002) show a significant correlation between annual default and recovery rates: bad years not only bring higher default rates, but also lower recovery rates. Failure to account for this correlation in a model results in underestimation of risk. Frye (2000a) argued that the macroeconomic systematic factors that drive defaults also drive recoveries. To describe this effect quantitatively, he proposed a Merton-type one-factor model. The recovery on a defaulted loan in his model is determined by normally distributed collateral value, which linearly depends on the systematic factor. This dependence results in higher economic capital than that calculated with a Merton-type model with independent recoveries. To simplify model calibration from Moody's default and recovery data, Frye (2000b) presented a different version of his model, where he modelled recoveries directly. A comprehensive survey of recent research in this area, including empirical studies, is given in Altman *et al* (2003).

In this article, we discuss the properties of loss distribution generated by a generalised version of Frye's model. It differs from Frye (2000a) in two important respects. First, apart from its dependence on the systematic factor, the collateral value is also allowed to depend on the idiosyncratic part of the borrower's asset return. The general motivation for this dependence is that borrowers in financial distress often cut back on collateral maintenance and control. This behaviour tends to reduce the value of the collateral for borrowers with lower asset returns. Since asset returns have both systematic and idiosyncratic components, collateral should generally depend on both. There is also additional motivation for this dependence for certain types of loans (such as project finance) where defaults are driven by the value of the collateral itself. Second, we use the lognormal distribution to describe collateral values instead of the normal distribution used by Frye. One of the benefits of the use of the lognormal distribution is that the collateral value is guaranteed to be positive. Another benefit is greater consistency of the model. We derive closed-form expressions for expected loss and, for the case of a fine-grained portfolio, for economic capital. We use these expressions to show how model outputs depend on model inputs.

Model

We consider a portfolio of loans to M distinct borrowers. To avoid cumbersome notations, we assume that each borrower has one loan with principal A_i. We also define the weight of a loan in the portfolio as the ratio of its principal to the total principal of the portfolio, $w_i = A_i/\Sigma_{j=1}^{M} A_j$. Borrower i will default within a chosen time horizon (typically, one year) with probability p_i. Default happens when a continuous variable X_i describing the financial well-being of borrower i over the horizon falls below a threshold. We assume that the variables $\{X_i\}$ have standard normal distribution. This choice brings the model into the Merton framework if we interpret X_i as the standardised continuously compounded return on the ith borrower's assets. The default threshold for borrower i is given by $N^{-1}(p_i)$, where $N^{-1}(\cdot)$ is the inverse of the cumulative normal distribution function.

We assume that asset returns can be written as:

$$X_i = \alpha_i Y + \sqrt{1-\alpha_i^2}\,\xi_i \tag{1}$$

where random variables Y and ξ_i are all independent and have standard normal distribution. Y is the only systematic risk factor in the model describing the state of the economy, while ξ_i is the idiosyncratic variable describing the individual fortunes of borrower i. If borrower i defaults, the amount of loss is determined by the value of the collateral. As in Frye (2000a), we write loss in the event of default Q_i as a unique function of the collateral value C_i (both are expressed in units of the loan principal, A_i) according to:

$$Q_i = \max\{1-C_i, 0\} = \left[1-C_i\right]^+ \tag{2}$$

However, in defining the distributional properties of the collateral, we will diverge from Frye's model. While Frye's collateral has normal distribution, we assume that C_i is lognormally distributed and can be written as:

$$C_i = \exp\left[\mu_i + \sigma_i R_i\right] \tag{3}$$

where random variable R_i has standard normal distribution.

Frye (2000a) argued that collateral, like any other asset, is sensitive to the state of the economy. When the economy is poor, collateral values tend to be lower. To capture this effect, we assume that the standardised collateral return R_i linearly depends on the systematic risk factor Y. There is, however, another effect that Frye's model does not completely account for. Typically, a borrower in financial distress does not invest enough in collateral maintenance, and the collateral loses value as a result. This behaviour is governed by the financial well-being (or asset returns) X_i, and, therefore, the standardised collateral return R_i should depend on X_i in addition to its dependence upon Y. Since X_i is a linear combination of the systematic risk factor Y and the idiosyncratic component ξ_i (equation (1)), we can write the standardised collateral return R_i as:

$$R_i = \beta_i Y + \gamma_i \xi_i + \sqrt{1-\beta_i^2-\gamma_i^2}\,\eta_i \tag{4}$$

where $\{\eta_i\}$ are standard normal variables independent of each other as well as from the other random variables in the model. One should keep in mind that equation (4) is only economically meaningful with non-negative parameters β_i and γ_i satisfying the relations $\beta_i^2 + \gamma_i^2 \le 1$ and $\gamma_i \le (\alpha_i/\sqrt{1-\alpha_i^2})\beta_i$. The latter relation is necessary for interpreting the idiosyncratic component ξ_i in equation (4) as part of the X_i dependence.

Finally, the portfolio loss rate L can be written as the weighted average of individual loss rates L_i:

$$L = \sum_{i=1}^{M} w_i L_i = \sum_{i=1}^{M} w_i 1_{\{X_i < N^{-1}(p_i)\}}\left[1-C_i\right]^+ \tag{5}$$

where $1_{\{\cdot\}}$ is the indicator function. The indicator in equation (5) takes the value one if the borrower defaults and zero otherwise. Equation (5) describes the distribution of the portfolio losses over the horizon and thus completes the model.

[1] *See Bluhm, Overbeck & Wagner (2002) for an excellent review of portfolio credit risk models*
[2] *For large enough portfolios, both assumptions lead to identical loss distributions regardless of the assumed distribution of individual recoveries*

Although one can use Monte Carlo simulations to generate the distribution of losses according to equation (5), some of the loss properties can be calculated analytically. Moreover, for the case of a fine-grained portfolio, there is a closed-form result for any percentile of the loss distribution. The rest of the article is devoted to the derivation of these results and an analysis of their properties.

Expected loss

There is more than one way to derive the expected loss from equation (5) using the law of iterated expectations. We use conditioning on the standardised collateral return R_i and write the expected loss rate for borrower i as:

$$\begin{aligned} E\left[L_i\right] &= E\left\{E\left[L_i \middle| R_i\right]\right\} \\ &= \int_{-\infty}^{-\mu_i/\sigma_i} dr\, n(r) \Pr\left[X_i < N^{-1}(p_i) \middle| R_i = r\right]\left[1 - \exp(\mu_i + \sigma_i r)\right] \end{aligned} \tag{6}$$

The next step is to derive the probability of default conditional on the realisation of R_i. First, let us note from equations (1) and (4) that X_i and R_i are linearly dependent with correlation ρ_i^{XR} given by:

$$\rho_i^{XR} = \alpha_i \beta_i + \sqrt{1 - \alpha_i^2}\,\gamma_i \tag{7}$$

Therefore, X_i can be written as a linear combination of R_i and some standard normal random variable ζ_i independent of R_i:

$$X_i = \rho_i^{XR} R_i + \sqrt{1 - \left(\rho_i^{XR}\right)^2}\,\zeta_i \tag{8}$$

The probability of default conditional on R_i is therefore:

$$\Pr\left[X_i < N^{-1}(p_i) \middle| R_i\right] = N\left[\frac{N^{-1}(p_i) - \rho_i^{XR} R_i}{\sqrt{1 - \left(\rho_i^{XR}\right)^2}}\right] \tag{9}$$

Substituting this probability into equation (6) and evaluating the integral, we obtain the expected loss rate:

$$\begin{aligned} E\left[L_i\right] &= N_2\left[N^{-1}(p_i), -\frac{\mu_i}{\sigma_i}, \rho_i^{XR}\right] \\ &- \exp\left(\mu_i + \frac{\sigma_i^2}{2}\right) N_2\left[N^{-1}(p_i) - \sigma_i \rho_i^{XR}, -\frac{\mu_i}{\sigma_i} - \sigma_i, \rho_i^{XR}\right] \end{aligned} \tag{10}$$

where $N_2(\cdot,\cdot,\cdot)$ is the bivariate normal cumulative distribution function. The algorithms for evaluating this function are discussed in great detail in Vasicek (1998).

An important property of equation (10) is that it does not depend on factor loadings α_i, β_i and γ_i separately. Instead, the expected loss rate depends on their combination ρ_i^{XR} given by equation (7).

LGD subtleties

Typically, the stand-alone risk of a loan is described by two parameters: the probability of default (PD) and the expected loss-given default (LGD). While the former appears as the exogenous parameter p_i, the latter is not explicitly specified in the model.

One might argue that stochastic LGD is the Q_i defined in equation (2), and expected LGD can be obtained by calculating its expectation as:

$$E\left[Q_i\right] = \int_{-\infty}^{-\mu_i/\sigma_i} dr\; n(r)\left[1 - \exp(\mu_i + \sigma_i r)\right]$$

where $n(\cdot)$ is the probability density of the standard normal distribution. Evaluating the last integral is straightforward and yields:

$$E\left[Q_i\right] = N\left(-\frac{\mu_i}{\sigma_i}\right) - \exp\left(\mu_i + \frac{\sigma_i^2}{2}\right) N\left(-\frac{\mu_i}{\sigma_i} - \sigma_i\right) \tag{11}$$

where $N(\cdot)$ is the standard normal cumulative distribution function.

However, equation (11) does not provide the correct answer. This is because the random variable Q_i is defined irrespective of whether borrower i defaults, while conventional LGD is defined only for defaulted borrowers. The variable Q_i is the potential LGD introduced in Pykhtin & Dev (2002). It determines what would be the loss if default happened regardless of whether default has happened. Therefore, the expected LGD (which we will denote by λ_i) should be calculated as the expectation of the potential LGD Q_i conditional on the default of borrower i according to:

$$\lambda_i = E\left[Q_i \middle| X_i < N^{-1}(p_i)\right] = E\left[L_i \middle| X_i < N^{-1}(p_i)\right] = \frac{E\left[L_i\right]}{p_i} \tag{12}$$

In conventional portfolio credit risk models, where variables $\{Q_i\}$ are independent of defaults, this conditioning is immaterial and both approaches (given by equations (11) and (12)) yield the same result. However, if there is correlation between the default and recovery drivers, expected LGD given by equation (12) ceases to be just a function of the collateral – it picks up the dependence upon PD. This dependence does not allow for the traditional interpretation of LGD as a measure of the collateral.

One can think of this PD dependence in terms of the sample selection bias studied by Heckman (1979). He considered two sets of observations corresponding to two distinct random variables. An observation was recorded in the second set only when the corresponding observation in the first set satisfied certain conditions. If the two random variables are independent, an estimate of the expectation (or another moment) of the second variable is unbiased. If they are interdependent, the condition on the first variable creates a selection bias for the second. An estimate of the expectation of the second random variable based on its data set will be biased. In the case of LGDs, the selection rule is the event of default because the only LGDs that we actually observe are for defaulted companies. Expected LGD can be thought of as a biased estimate of the expectation of potential LGD.

Let us look at the origin of the PD dependence by considering two borrowers: borrower A, with a low PD, and borrower B, with a high PD. Let us assume that these borrowers were given loans on identical terms, including identical collaterals. Contrary to the 'traditional' intuition based on the independence between defaults and recoveries, expected LGDs for these two loans will be different. Indeed, borrower A's asset return must go much further down to force its default than borrower B's asset return. Since the collateral value of each borrower is an increasing function of its asset return, the collateral posted by borrower A will be more depressed in the event of default than the collateral posted by borrower B. Therefore, on average, recoveries from the default of borrower A will be lower than recoveries from the default of borrower B. Thus, given identical collaterals, expected LGD will be higher for borrowers with higher credit quality. The strength of the expected LGD dependence upon PD is determined by the correlation ρ_i^{XR} between the 'asset return' of the borrower X_i and the 'collateral return' of the loan R_i.

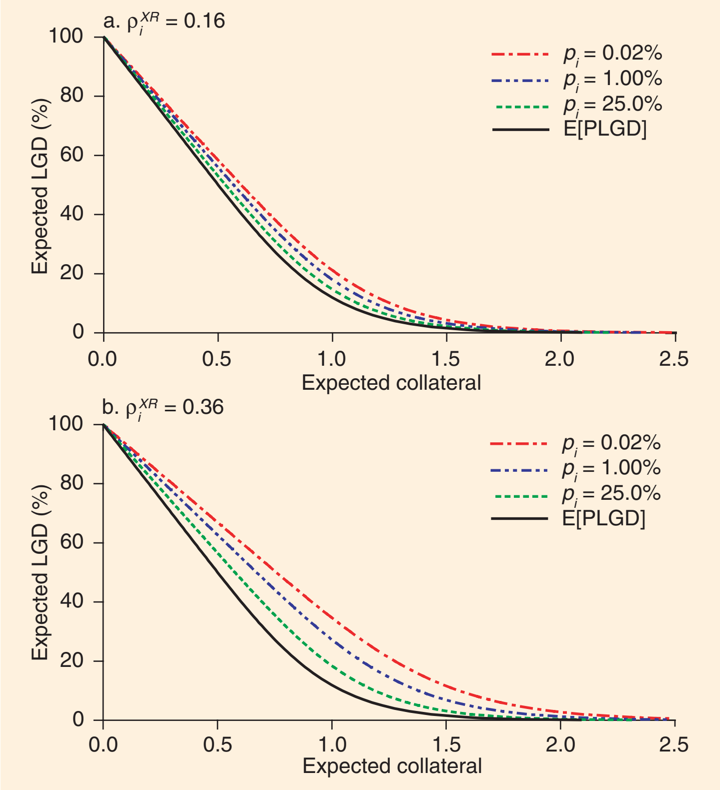
In figure 1, we show expected LGD λ_i as a function of expected collateral $E[C_i] = \exp(\mu_i + \sigma_i^2/2)$ at two levels of the asset-collateral correlation: $\rho_i^{XR} = 0.16$ in panel a and $\rho_i^{XR} = 0.36$ in panel b. We calculated expected LGD according to equation (12), varying μ_i and keeping σ_i fixed at 0.30. The three coloured dashed curves correspond to three levels of PD (0.02%, 1.00% and 25.0%). For comparison, the black solid curve shows expected potential LGD calculated according to equation (11) (assuming the same σ_i). As one moves to higher expected collaterals and lower expected LGDs, the dependence of λ_i upon PD becomes stronger and equation (11) becomes less appropriate to approximate expected LGD.

Another interesting observation from figure 1 is the asymptotic linear dependence of the expected LGD upon the expected collateral at high λ_i. This linear dependence appears naturally when there is no restriction on possible recoveries in the model. If one defines potential LGD without limiting it to non-negative values (that is, as $Q_i = 1 - C_i$), expected LGD is given by:

$$\lambda_i = 1 - \frac{1}{p_i} N\left[N^{-1}(p_i) - \sigma_i \rho_i^{XR}\right] E\left[C_i\right] \tag{13}$$

which is linear in the expected collateral. When recoveries and defaults

1. Expected LGD as a function of expected collateral



are independent ($\rho_i^{XR} = 0$), equation (13) simplifies to $\lambda_i = 1 - E[C_i]$. Thus, all the non-linearity in the dependence of expected LGD upon expected collateral results from restricting maximum possible recovery to the loan principal (or, equivalently, restricting potential LGD to non-negative values). When expected collateral is small (expected LGD is large), the chance of the collateral exceeding one is negligible. Therefore, for small $E[C_i]$, there is no essential difference between the model specifications with or without the restriction on recovery, and equation (13) provides large-LGD asymptotics for the model with the restriction.

In most practical applications, the expected LGD λ_i is used as an input rather than as a derived quantity. While this choice is convenient empirically and makes the calculation of the expected loss rate as trivial as $E[L_i] = p_i\lambda_i$, it also makes the expected log-collateral μ_i a PD-dependent derived quantity. If we assume that the collateral volatility σ_i and the asset-collateral correlation ρ_i^{XR} are known, the calculation of μ_i reduces to numerically inverting λ_i as a function of μ_i. The function is given by equation (12), and its derivative can be obtained by straightforward differentiation:

$$\frac{d\lambda_i}{d\mu_i} = -\exp\left(\mu_i + \frac{\sigma_i^2}{2}\right)\frac{1}{p_i}N_2\left[N^{-1}(p_i) - \sigma_i\rho_i^{XR}, -\frac{\mu_i}{\sigma_i} - \sigma_i, \rho_i^{XR}\right]$$

This derivative is strictly negative, so λ_i is a monotonically decreasing function of μ_i and the inversion presents no problems. Moreover, since the derivative can be evaluated as easily as the function itself, the fast-converging Newton-Raphson method can be applied.

Loss distribution for a fine-grained portfolio

Unfortunately, the distribution of losses is not analytically tractable for an arbitrary portfolio. However, the case of a large number of borrowers M allows for tremendous simplification of the portfolio loss distribution if there are no individual large exposures (such a portfolio is called fine-grained). The intuition behind this simplification is diversification, which essentially eliminates all idiosyncratic risk in a fine-grained portfolio. Gordy (2002) has shown that the portfolio loss rate distribution has a limiting form as $M \to \infty$ provided that w_M goes to zero faster than $1/\sqrt{M}$. This limiting loss rate is given by the expected loss rate conditional on the systematic factors.[3] Thus, by taking the conditional on Y expectation of the right-hand side of equation (5), we obtain the limiting portfolio loss rate L^∞:

$$L^\infty = E\left[L|Y\right] = \sum_{i=1}^{M} w_i E\left[L_i|Y\right] \tag{14}$$

The contribution of borrower i to the portfolio loss can be shown to be (see Appendix):

$$\begin{aligned} E\left[L_i|Y\right] = {} & N_2\left[\bar{X}_i(Y), \bar{\eta}_i(Y), \frac{\gamma_i}{\sqrt{1-\beta_i^2}}\right] \\ & - \exp\left[\mu_i + \sigma_i\beta_i Y + \frac{\sigma_i^2}{2}\left(1-\beta_i^2\right)\right] \\ & N_2\left[\bar{X}_i(Y) - \sigma_i\gamma_i, \bar{\eta}_i(Y) - \sigma_i\sqrt{1-\beta_i^2}, \frac{\gamma_i}{\sqrt{1-\beta_i^2}}\right] \end{aligned} \tag{15}$$

where $\bar{X}_i(Y)$ is the default threshold conditional on Y given by:

$$\bar{X}_i(Y) = \frac{N^{-1}(p_i) - \alpha_i Y}{\sqrt{1-\alpha_i^2}}$$

and:

$$\bar{\eta}_i(Y) = \frac{-\mu_i/\sigma_i - \beta_i Y}{\sqrt{1-\beta_i^2}}$$

When the collateral does not depend on the idiosyncratic component of the asset return (see below), equation (15) reduces to:

$$\begin{aligned} & E\left[L_i|Y\right] = N\left[\bar{X}_i(Y)\right] \\ & \left(N\left[\bar{\eta}_i(Y)\right] - \exp\left[\mu_i + \sigma_i\beta_i Y + \frac{\sigma_i^2}{2}\left(1-\beta_i^2\right)\right] N\left[\bar{\eta}_i(Y) - \sigma_i\sqrt{1-\beta_i^2}\right]\right) \end{aligned} \tag{16}$$

The conditional expected loss given by equation (16) is the product of the conditional probability of default and the conditional expected LGD. One would expect this result because, conditionally on Y, the default and recovery drivers of borrower i are independent when $\gamma_i = 0$.

Substitution of equation (15) (or equation (16)) into equation (14) results in a closed-form expression for the portfolio loss rate distribution.

Calibration issues

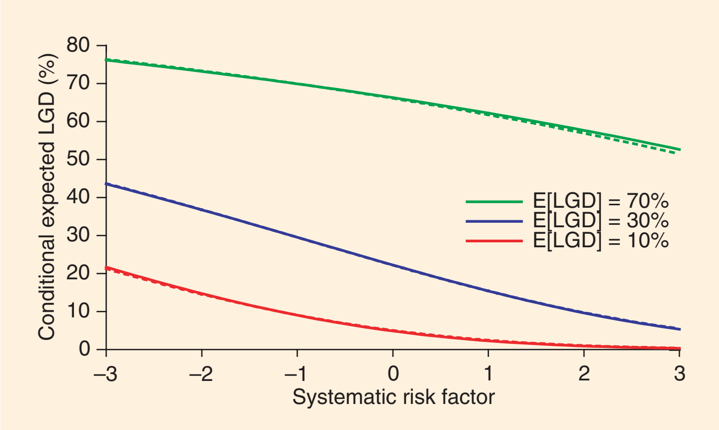
Let us assume that we have a history of default rates and losses in the event of default for loans all having the same PD p and expected LGD λ. Practically, these loans may belong to the same cell of a two-dimensional rating system based on PD and expected LGD. Each time point t of the time series can be characterised by a certain realisation of the systematic risk factor y_t. If the number of borrowers at each t is large enough, both y_t and the factor loading α can be estimated from the history of default rates as shown in Frye (2000b).[4] The expected log-collateral μ can be derived from the given expected LGD, as discussed above. The collateral volatility σ can be estimated from cross-sectional LGD data.

Unfortunately, in most cases the complete parameterisation of the model is not possible. Factor loading γ, which describes the dependence of the collateral upon the idiosyncratic component of the asset return, can only be determined from time series of asset values, which are not observable. However, if loans are bucketed[5] by expected LGD, the inability to determine γ is not a problem.

[3] Limiting loss rate distribution for a homogeneous portfolio was obtained earlier by Vasicek (1991). Vasicek (2002) extended his earlier results to non-homogeneous portfolios with the fine-granularity condition specified as $\Sigma_{i=1}^{M} w_i^2 \to 0$

[4] This method is not suitable for small samples

2. Conditional expected LGD as a function of systematic risk factor for different expected LGD levels and collateral specifications



The observed average LGD at time t should be distributed around the expected LGD conditional on $Y = y_t$, which can be defined as $\bar{\lambda}(y_t) = E[L_i|Y = y_t]/N[\bar{X}(y_t)]$. Although this quantity is a function of both β and γ, it is impossible to estimate these parameters separately. It turns out that there are infinitely many pairs of β and γ that generate almost indistinguishable curves $\bar{\lambda}(y)$. Each such pair (β, γ) will lead to a different asset-collateral correlation ρ^{XR} and, therefore, to a different collateral log-return μ. Assuming PD $p = 1\%$, collateral volatility $\sigma = 0.3$ and factor loading $\alpha = 0.4$, we plotted three pairs of curves in figure 2, each pair (distinguished by colour) representing the function $\bar{\lambda}(y)$ for a certain level of expected LGD. The solid curve in each pair corresponds to the values $\beta = 0.5$, $\gamma = 0.3$, and the dashed curve corresponds to $\beta = 0.4$, $\gamma = 0$. The curves in the pairs are almost indistinguishable.

Although this 'collateral degeneracy' prevents us from complete parameterisation of the model, it also works to our advantage. All 'degenerate' collateral specifications μ, β and γ result in almost identical loss distributions. Only one of these specifications describes the real collateral, but it does not need to be chosen if the exposures are bucketed by expected LGD. We can always choose the one that simplifies the model: $\gamma = 0$. Once this choice has been made, the model parameterisation can be completed by fitting the curve $\bar{\lambda}(y)$ to the observed time series of average LGD and the implied time series of y_t with β as the only fitting parameter.

If loans are bucketed by collateral, this presents a real problem that can be resolved only in special cases when γ can be determined from other considerations. One such case is project finance loans, where project cashflows drive both default and recovery. In this case, collateral return is comonotonic with asset return, and collateral factor loadings must satisfy the relations $\beta_i = \alpha_i$ and $\gamma_i = \sqrt{1-\alpha_i^2}$.

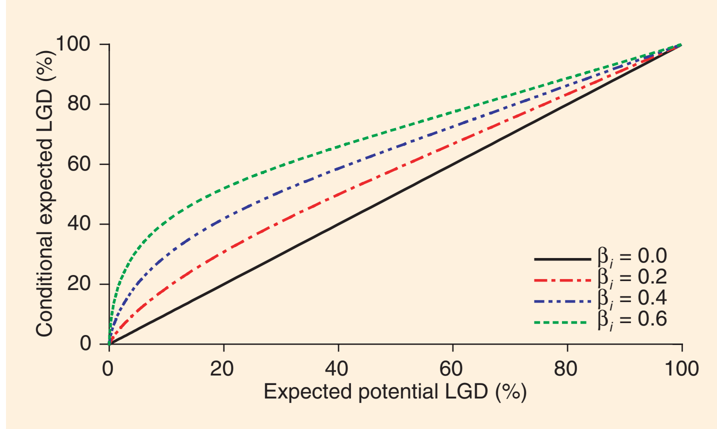
Economic capital

Economic capital is defined as the qth percentile of the portfolio loss distribution. Since the portfolio loss rate given by equation (15) is a deterministic monotonically decreasing function of Y, the economic capital rate can be obtained by substituting $(1 - q)$th percentile of Y (that is, $N^{-1}(1 - q)$) into the function:

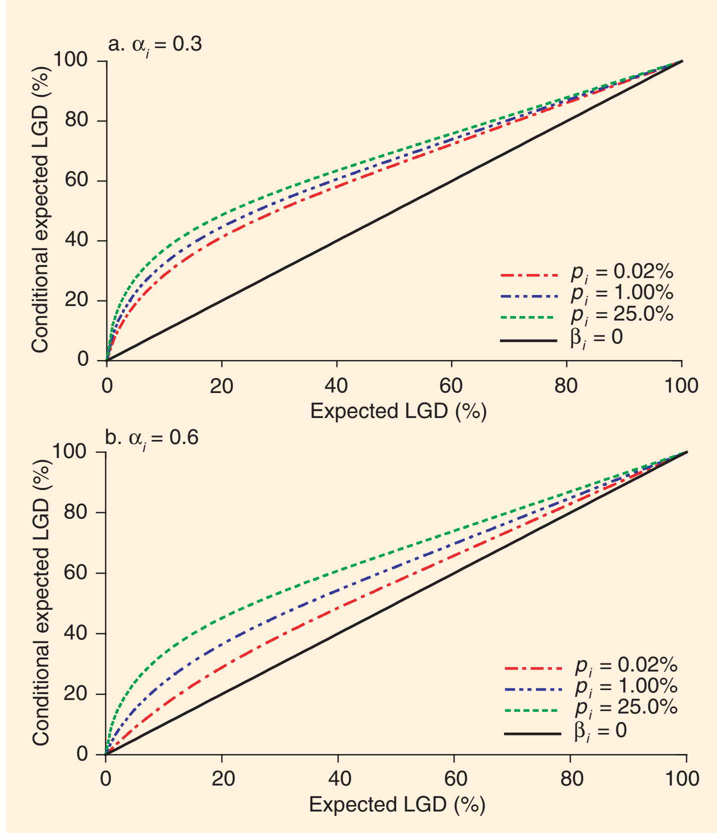
$$K^q = \sum_{i=1}^{M} w_i E\left[L_i \middle| Y = N^{-1}(1-q)\right] \qquad (17)$$

In what follows, we show how the risk contribution from borrower i to the capital depends on the model parameters. Throughout this section, we assume the confidence level $q = 99.9\%$ and collateral volatility $\sigma_i = 0.3$. We will consider only the simplified version of the model, where we assume that collateral does not depend on the idiosyncratic component of the asset return ($\gamma_i = 0$).

3. Conditional expected LGD as a function of expected potential LGD at $\gamma_i = 0$



4. Expected conditional LGD as a function of expected LGD at $\beta_i = 0.6$ and $\gamma_i = 0$



The risk contribution can be written as the product of the conditional PD $\bar{p}_i^q = N[\bar{X}_i(N^{-1}[1 - q])]$ and the conditional expected LGD $\bar{\lambda}_i^q$, which can be obtained from equation (16):

$$\bar{\lambda}_i^q = N\left[\bar{\eta}_i^q\right] - \exp\left[\mu_i + \sigma_i\beta_i Y + \frac{\sigma_i^2}{2}\left(1-\beta_i^2\right)\right] N\left[\bar{\eta}_i^q - \sigma_i\sqrt{1-\beta_i^2}\right] \qquad (18)$$

where $\bar{\eta}_i^q = \bar{\eta}_i[N^{-1}(1 - q)]$. The conditional PD is independent of the recovery assumptions and has been studied by Vasicek and others. Therefore, we will focus the analysis on the conditional expected LGD. When

[5] *By bucketing, we mean grouping the loans with a certain parameter (in this case, expected LGD) within a pre-specified range into 'buckets', which are presumed to be homogeneous*

recoveries are independent from defaults, conditional expected LGD coincides with the unconditional one, that is, $\bar{\lambda}_i^q = \lambda_i$. When there is non-zero correlation between default and recovery drivers, the deviation of $\bar{\lambda}_i^q$ from λ_i describes the effect of this correlation on the risk contribution – unexpected recovery risk.

Equation (18) has no explicit dependence on either PD or factor loading α_i. If the other parameters are fixed, $\bar{\lambda}_i^q$ clearly increases with β_i. In figure 3, we plot $\bar{\lambda}_i^q$ as a function of the expected potential LGD $E[Q_i]$ for several values of β_i. The case of independent recoveries ($\beta_i = 0$) is represented by the solid black line. Since the expected potential LGD is a unique function of only μ_i and σ_i (it is given by equation (11)), $\bar{\lambda}_i^q\,(E[Q_i])$ is the same for all PDs and α_i-s. From figure 3, one can see that the conditional expected LGD is a concave function of the expected potential LGD with fixed end values $\bar{\lambda}_i^q(0) = 0$ and $\bar{\lambda}_i^q(1) = 1$. For non-zero β_i, the relative difference between $\bar{\lambda}_i^q\,(E[Q_i])$ (the curve) and $E[Q_i]$ (the straight line) decreases with $E[Q_i]$. Therefore, models with independent recoveries are particularly inappropriate at low LGDs.

Now let us consider what happens if one uses conventional expected LGD λ_i as an input. Expected log-collateral μ_i now becomes a decreasing function of PD and an increasing function of the asset-collateral correlation, which now simplifies to $\rho_i^{XR} = \alpha_i \beta_i$. Therefore, even if we fix λ_i, σ_i and β_i, the conditional expected LGD will still depend on PD and α_i. In figure 4, we illustrate this dependence at a fixed $\beta_i = 0.6$ (which corresponds to the top curve in figure 3) by plotting the conditional expected LGD as a function of the unconditional expected LGD, that is $\bar{\lambda}_i^q\,(\lambda_i)$, for two values for factor loadings α_i. The three coloured dashed curves in each panel describe three PD levels, while the black solid line shows the diagonal $\bar{\lambda}_i^q\,(\lambda_i) = \lambda_i$ describing the case of independent recoveries (that is, $\beta_i = 0$). The order of the curves clearly follows from the economic intuition we developed above.

Conclusion

Most portfolio models of credit risk used in the banking industry assume that recoveries are either deterministic constants or random variables independent from default events. However, there are empirical as well as conceptual arguments against such assumptions.

In this article, we have presented a one-factor portfolio model where the recoveries and the default events are correlated. There are two sources for this correlation in the model. One is the explicit dependence of the collateral upon the systematic risk factor. This dependence describes the direct effect of the macroeconomic environment on collateral values (and, therefore, recoveries): apart from bringing more defaults, an economic downturn also tends to reduce collateral values. Another source of correlation comes from the dependence of the collateral upon the default driver – the borrower's financial well-being. This dependence accounts for the negligence in collateral maintenance by borrowers in financial distress. It also allows one to handle cases (such as project finance loans) when the defaults are driven by the collateral itself. We have derived closed-form expressions for the expected loss and, for the case of a fine-grained portfolio, for the economic capital. We have used these expressions to show how the model output depends on the model inputs.

In traditional models with recoveries that are independent of defaults, LGD is a measure of the collateral quality. In any model with non-zero correlation between default and recovery drivers, expected LGD is a mixed measure of borrower and collateral quality. Therefore, if one buckets loans by PD and expected LGD, the collateral quality depends on the PD within the same LGD bucket. Generally, the collateral quality improves as one goes from low PD to high PD within the same expected LGD bucket. We used the closed-form expressions derived for the presented model to demonstrate how this dependence affects the economic capital. ■

Michael Pykhtin is assistant vice-president, capital allocation and quantitative analysis, at KeyCorp in Cleveland, Ohio. **He would like to thank Ashish Dev for valuable discussions, and anonymous referees for their helpful comments.** **E-mail: Michael_V_Pykhtin@KeyBank.com**

Appendix

The contribution of borrower i to the portfolio loss can be written via mutually independent random variables as:

$$E\left[L_i \middle| Y\right] = E\left[1_{\{\xi_i < \bar{X}_i(Y)\}}\left(1 - \exp\left[\mu_i + \sigma_i r_i\left(Y, \xi_i, \eta_i\right)\right]\right)^+ \middle| Y\right] \quad (19)$$

and function $r_i(Y, \xi_i, \eta_i)$ is defined according to equation (4). Both factors in the product under the conditional expectation in equation (19) depend on the idiosyncratic component of the asset return ξ_i. Therefore, the conditional expected loss rate cannot be written as the product of the conditional expectations of the default and the recovery factors. However, the conditional expected loss rate can be calculated in two steps with the help of the law of iterated expectations, with extra conditioning on ξ_i according to:

$$E\left[L_i \middle| Y\right] = E\left\{1_{\{\xi_i < \bar{X}_i(Y)\}} E\left[\left(1 - \exp\left[\mu_i + \sigma_i r_i\left(Y, \xi_i, \eta_i\right)\right]\right)^+ \middle| Y, \xi_i\right] \middle| Y\right\} \quad (20)$$

The evaluation of the inner expectation can be performed as:

$$\begin{aligned} E\left[Q_i \middle| Y, \xi_i\right] &= \int_{-\infty}^{\bar{\bar{\eta}}_i(Y,\xi_i)} d\eta\, n(\eta)\left(1 - \exp\left[\mu_i + \sigma_i r_i\left(Y, \xi_i, \eta\right)\right]\right) \\ &= N\left[\bar{\bar{\eta}}\left(Y, \xi_i\right)\right] - \exp\left[\mu_i + \sigma_i\left(\beta_i Y + \gamma_i \xi_i\right) + \sigma_i^2\left(1 - \beta_i^2 - \gamma_i^2\right)/2\right] \\ &\quad N\left[\bar{\bar{\eta}}\left(Y, \xi_i\right) - \sigma_i\sqrt{1 - \beta_i^2 - \gamma_i^2}\right] \end{aligned} \quad (21)$$

where:

$$\bar{\bar{\eta}}_i\left(Y, \xi_i\right) = \frac{-\mu_i/\sigma_i - \beta_i Y - \gamma_i \xi_i}{\sqrt{1 - \beta_i^2 - \gamma_i^2}} \quad (22)$$

With this intermediate result, equation (20) can be rewritten as:

$$\begin{aligned} E\left[L_i \middle| Y\right] = \int_{-\infty}^{\bar{X}_i(Y)} d\xi\, n(\xi)\Big\{ & N\left[\bar{\bar{\eta}}\left(Y, \xi\right)\right] \\ & - \exp\left[\mu_i + \sigma_i\left(\beta_i Y + \gamma_i \xi\right) + \sigma_i^2\left(1 - \beta_i^2 - \gamma_i^2\right)/2\right] \\ & N\left[\bar{\bar{\eta}}\left(Y, \xi\right) - \sigma_i\sqrt{1 - \beta_i^2 - \gamma_i^2}\right]\Big\} \end{aligned} \quad (23)$$

Evaluation of the integral in the right-hand side of equation (23) yields equation (15). ■

REFERENCES

Altman E, B Brady, A Resti and A Sironi, 2003
The link between default and recovery rates: theory, empirical evidence and implications
Working paper, March

Bluhm C, L Overbeck and C Wagner, 2002
An introduction to credit risk modeling
Chapman & Hall/CRC

Frye J, 2000a
Collateral damage
Risk April, pages 91–94

Frye J, 2000b
Depressing recoveries
Risk November, pages 108–111

Gordy M, 2002
A risk-factor model foundation for ratings-based bank capital rules
Finance and Economics Discussion Series 2002-55, Washington: Board of Governors of the Federal Reserve System. Forthcoming in Journal of Financial Intermediation

Heckman J, 1979
Sample selection bias as a specification error
Econometrica 47, pages 153–161

Moody's Investors Service, 2002
Default and recovery rates of corporate bond issuers
Global Credit Research, February

Pykhtin M and A Dev, 2002
Analytical approach to credit risk modelling
Risk March, pages S26–S32

Vasicek O, 1991
Limiting loan loss probability distribution
KMV Corporation

Vasicek O, 1998
A series expansion for the bivariate normal integral
Journal of Computational Finance 1(4), summer, pages 5–10

Vasicek O, 2002
Loan portfolio value
Risk December, pages 160–162