Dependent credit rating transitions and empirical estimates for event correlations

D.V. Boreiko, School of Economics and Management,
Free University of Bozen-Bolzano,
Via Sernesi 1, 39100 Bolzano, Italy; e-mail: DBoreiko@unibz.it,
Y.M. Kaniovski, School of Economics and Management,
Free University of Bozen-Bolzano,
Via Sernesi 1, 39100 Bolzano, Italy; e-mail: YKaniovskyi@unibz.it
and
G.Ch.Pflug, Department of Statistics and Decision Support Systems,
University of Vienna,
Universitätstraße 5, 1090 Wien, Austria; e-mail: georg.pflug@univie.ac.at

Abstract. Developing the coupled Markov Chain approach by Kaniovski and Pflug (2007) for credit portfolios, we suggest estimates for the parameters of the model and three heuristics for approximately calculating the total loss distribution. The model is based on common risk factors called the *tendencies*. Two of the heuristics deal with the conditional distributions given the tendency. The third one does not require the tendency variables, but employs instead the whole correlation structure of the true distribution. All three approximations match all moments up to the second order of the true distributions substituting them by normal ones. Since computational complexity of each of the heuristics does not depend upon the number of items in the portfolio, but only on the number of cells (i.e. combinations of rating classes and industry sectors), the approximation methods are much more efficient than the full simulation of the rating processes of all debtors in the portfolio.

Based on a data set from Standard and Poor's covering rating transitions of 10413 companies in 30 OECD countries within the period of 1990 – 2006, we estimate a 2×3 Markov transition matrix, default correlations and the parameters of the corresponding coupled Markov chain model with six industry sectors and two credit ratings. Using the estimates, we model a portfolio with 100 debtors for each combination of a non-default credit class and an industry sector. Its evolution for 3, 5 and 7 years is simulated by an exact technique and the three heuristics. The distribution of defaults exhibits much heavier tails than independent defaults would produce.

Key words: credit risk, matching moments, dependent credit rating transitions, event correlation, tendency variables, normal distribution, multinomial distribution.

JEL classification: G31, G11, C15.

1 Introduction

The coupled Markov chain model by Kaniovski and Pflug (2007), referred to as KP07, takes into account the dependence of credit rating transitions in the simplest and the most natural way - using the corresponding event correlations.

Applying the KP07 model for assessing the riskiness of a portfolio requires estimates for the parameters. While Markovian transition matrices have been estimated by major rating agencies, a reliable source for event correlations is missing. We develop here such estimates suited for the model at hand. The approach is tested on estimating default correlations obtained from a particular data set by Standard and Poor's (S&P's). It covers the period from 1990 till 2006. We classify 10413 companies from 30 OECD countries in six industry sectors and two credit classes: investment grade and non-investment grade debtors.

In KP07, the true distributions are given only by their generating functions. They look quite complicated and cannot be used easily for finding the required loss characteristics of the portfolio. However, some moments may be evaluated without big difficulties. This allows to substitute the actual distributions by simpler ones, having the same moments up to a certain order. We implement this idea by matching all moments up to the order two and normal distributions respectively. Dealing with correlations, as a means for taking into account the interdependencies involved, this "second order approximation" seems to be a reasonable approach. One particular heuristics suggests linear combinations of independent normal random variables for approximating the new cell counts. The corresponding coefficients depend upon the tendency variables – the common factors driving the portfolio evolution. Simulating the evolution numerically, this variant of the mean-variance approach allows us to reduce the complexity of the calculation: instead of tracing the evolution of every debtor we may deal with each cell as a whole. Thus, irrespective of the how many debtors are involved, the complexity is determined by the number of cells - that is, the number of credit ratings times the number of industry sectors. The numerical simulations given below allow to compare these heuristics with the true sample distributions.

Our approach bears the features of several known models of dependent credit rating transitions. Like Hull and White (2004), we capture the dependence by a common factor. In our case it is a Bernoulli random variable indicating the tendency of the market: up or down. Similar to what is done by Devis and Lo (2001), the strength of the dependence is measured by the probability of success of another Bernoulli random variable. In a structural model based on Merton's (1974) representation of the cash position of a firm, the coefficient of the linear combination of the individual factors and the common one plays an analogous role. However, we model transitions between all credit classes rather than restricting our attention only to default events. In formal terms, we deal with a copula formed by many identical Markov transition matrices defined by a particular rating agency or estimated by the modeler. The strength of the dependence vary across the industry sectors as well as across the credit classes. Hence we achieve the same conceptual richness as affine Markov chain model by Hurd and Kuznetsov (2007). Our use of coefficients of correlation as the means for transmitting dependence is oriented at a numerical implementation that requires solving a quadratic programming algorithm, turns out to be much simple than correlating jump intensity processes adopted by Hurd and Kuznetsov (2007) who development of the approach by Jarrow at al (1997). As in both cases the coefficients of correlation are positive, the corresponding evolutions are governed by a common tendency. As we deal with discrete time, some technical assumptions, like diagonalizability of the intensity matrix, are not needed here. On the other hand, the intensity based continuous time approach has a long history in financial modeling. See among others Jarrow and Turnbull (1995) or Li (2002) .

2 The basic model

Consider a diversified portfolio consisting of debt obligations. The debtors are non-homogeneous in their credit ratings and they belong to different industry sectors. Let there be M non-default rating classes. The ratings are numbered in a descending order so that 1 corresponds to the safest class, while M is the next to the default. Currently $M = 7$ for all the most respected credit ranking agencies. For example, in terms of S&P's, $1 \longleftrightarrow AAA$, $2 \longleftrightarrow AA$, $3 \longleftrightarrow A$, $4 \longleftrightarrow BBB$, $5 \longleftrightarrow BB$, $6 \longleftrightarrow B$ and $7 \longleftrightarrow CCC$. On the other hand, Nagpal and Bahar (2001), due to scarcity of data concerning defaults, distinguish only two categories – investment grade and non-investment grade. The investment grade rating corresponds to an S&P's class from AAA to BBB, while the non-investment grade rating covers BB or lower S&P's classes. Let there be K industry sectors. For example, Nagpal and Bahar (2001) analyze US debtors classifying them into $K = 11$ industry sectors.

Denote by $N_i^k(1)$ the initial number of obligors from industry sector k in rating class i. The total portfolio is described by the matrix $\mathbf{N}(1)$ with entries $N_i^k(1)$, $i = 1, 2, \ldots, M$, $k = 1, 2, \ldots, K$. A cell i, k is formed by a particular combination of a rating class and an industry sector. In total, there are $\mathcal{N}(1) = \sum_{i=1}^{M} \sum_{k=1}^{K} N_i^k(1)$ non-default debtors in the portfolio.

The methodology suggested in KP07 generates a portfolio consisting of debtors with different credit ratings and industry sectors such that:

1. the migrations of debtors dependent;
2. the degree of dependence varies between industry sectors and rating classes;
3. every individual migration is governed by a Markovian matrix, the same for the whole portfolio.

The fundamental assumption is that each debtor belonging to a credit class follows a rating process with the same distribution, but these processes are coupled in such a way that the joint probability does not necessarily equal the product of the marginals.

Let $p_{i,j}$ be the probability of transition within one year from the i-th credit rating to the j-th. In particular, $p_{i,M+1}$ is the probability that a debtor who is having i-th credit rating at the beginning of a year defaults by the end of this year. The $M \times (M + 1)$ transition matrix $P = (p_{i,j})$ is estimated and reported by rating agencies. In particular, the S&P's transition matrix reads

$$P = \begin{pmatrix} 0.9081 & 0.0833 & 0.0068 & 0.0006 & 0.0012 & 0.0000 & 0.0000 & 0.0000 \\ 0.0070 & 0.9065 & 0.0790 & 0.0064 & 0.0006 & 0.0014 & 0.0002 & 0.0000 \\ 0.0009 & 0.0227 & 0.9105 & 0.0552 & 0.0074 & 0.0026 & 0.0001 & 0.0006 \\ 0.0002 & 0.0033 & 0.0595 & 0.8693 & 0.0530 & 0.0117 & 0.0012 & 0.0018 \\ 0.0003 & 0.0014 & 0.0067 & 0.0773 & 0.8053 & 0.0884 & 0.0100 & 0.0106 \\ 0.0000 & 0.0011 & 0.0024 & 0.0043 & 0.0648 & 0.8346 & 0.0407 & 0.0520 \\ 0.0022 & 0.0000 & 0.0022 & 0.0130 & 0.0238 & 0.1124 & 0.6486 & 0.1979 \end{pmatrix}.$$

See Credit Metrics (1997) p. 69.

Consider $X(t)$, $t \geq 1$, a discrete-time Markov chain evolving in the state space $\{1, 2, \ldots, M+1\}$ governed by the probabilities $p_{i,j}$ and whose absorbing state is $M+1$. That is, the first M rows, corresponding to the transient states, of its Markov transition matrix are the same as in P, while the last one contains 0 at all positions, except for the $(M+1)$-th, which is occupied by 1.

Any transition of $X(t)$ may be divided in two moves. During the first one a *tendency* is determined. That is, whether $X(t) \geq X(t+1)$, i.e. a non-deteriorating step takes place, or whether $X(t) < X(t+1)$, i.e. a deteriorating one occurs. Let $\chi(t)$ be the tendency variable, i.e.

$$\chi(t) = \begin{cases} 1 & \text{if } X(t) \geq X(t+1), \\ 0 & \text{if } X(t) < X(t+1). \end{cases}$$

In other words, $\chi(t) = \mathbb{1}_{\{X(t+1) \leq X(t)\}}$ is the indicator of a non-deteriorating move at time t. Here $\mathbb{1}_A$ denotes the indicator of a statement A. That is $\mathbb{1}_A = 1$ when A is true and $\mathbb{1}_A = 0$ when A is false. The tendency variable has a Bernoulli distribution whose probability of a success reads

$$p_i^+ \stackrel{def}{=} \sum_{j=1}^{i} p_{i,j} = \mathbb{P}\mathrm{r}\{\chi(t) = 1 \mid X(t) = i\}.$$

Then, conditioned on the tendency, the next move follows the distribution $p'_{i,j}$ such that

$$\mathbb{P}\mathrm{r}\{X(t+1) = j \mid X(t) = i, \chi(t) = 1\} = \begin{cases} p_{i,j}/p_i^+ & \text{if } j \leq i, \\ 0 & \text{if } j > i; \end{cases}$$

$$\mathbb{P}\mathrm{r}\{X(t+1) = j \mid X(t) = i, \chi(t) = 0\} = \begin{cases} p_{i,j}/(1 - p_i^+) & \text{if } j > i, \\ 0 & \text{if } j \leq i. \end{cases}$$

The evolution of the total portfolio is modeled by a multi-dimensional random process $\overrightarrow{X}(t) = (X_1(t), X_2(t), \ldots, X_{\mathcal{N}(1)}(t))$ whose components evolve like $X(t)$, but are dependent. The starting values are

$$X_n(1) = m(n), \; n = 1, 2, \ldots, \mathcal{N}(1),$$

where $m(n)$ is the credit rating of firm n at time $t = 1$. Denote by $s(n)$ the industry sector of firm n. Since $X_n(t)$ is modeled as time-homogeneous Markov process, it is enough to look at the transition from time $t = 1$ to time $t = 2$ and, consequently, we may drop the index t. The rating may randomly change in time, becoming say $X_n(2)$ at time $t = 2$, while the assignment to the sector $s(n)$ remains constant. To define $X_n(2)$, we need several steps.

First introduce $\mathcal{N}(1)$ independent random variables $\xi_n, n = 1, 2, \ldots, \mathcal{N}(1)$, assuming the values $1, 2, \ldots, M+1$ and such that

$$\mathbb{P}\mathrm{r}\{\xi_n = j\} = p_{m(n),j}.$$

For a given $M \times M$ matrix of correlations $C = (c_{i,I})$, a tendency random vector

$$\overrightarrow{\chi} = (\chi_1, \chi_2, \ldots, \chi_M)$$

is generated in such a way that the tendency variables χ_i satisfy

- $\mathbb{P}\mathrm{r}\{\chi_i = 1\} = p_i^+ = 1 - \mathrm{Pr}\{\chi_i = 0\}$,
- $\mathbb{C}\mathrm{orr}(\chi_i, \chi_I) = c_{i,I}$,
- the χ_i are independent of the ξ_n.

The numerics of this generation using quadratic optimization is described in KP07. When no solution to the corresponding quadratic optimization problem exists, the approach may not be applied, because the corresponding set of parameters is not feasible.

We define further $\mathcal{N}(1)$ random variables η_n, $n = 1, 2, \ldots, \mathcal{N}(1)$, by setting

$$\mathbb{P}\mathrm{r}\{\eta_n = j \mid \chi_{m(n)} = 1\} = p_{m(n),j}/p^+_{m(n)},\ j = 1, 2, \ldots, m(n),$$

$$\mathbb{P}\mathrm{r}\{\eta_n = j \mid \chi_{m(n)} = 0\} = p_{m(n),j}/(1 - p^+_{m(n)}),\ j = m(n)+1, m(n)+2, \ldots, M+1.$$

The η_n variables are conditionally independent given the tendency vector $\overrightarrow{\chi}$. Denote the corresponding conditional probabilities by $p'_{i,j}$. The unconditional probabilities are as desired

$$\mathbb{P}\mathrm{r}\{\eta_n = j\} = p_{m(n),j},\ j = 1, 2, \ldots, M+1.$$

Notice that the random variables $\{\eta_n\}$ are dependent. The dependence acts through the corresponding tendency variables. Its strength varies among industry sectors. This variability is modeled as follows. Let $\{\delta_n\}$ be $\mathcal{N}(1)$ independent Bernoulli random variables which are independent of $\{\xi_n\}$ and $\{\eta_n\}$. The probability of a success of δ_n is $q^{s(n)}_{m(n)}$. The constants q_i^k, $i = 1, 2, \ldots, M$, $k = 1, 2, \ldots, K$, are collected in an $M \times K$ matrix Q. We set finally

$$X_n(2) = (1 - \delta_n)\xi_n + \delta_n \eta_n,\ n = 1, 2, \ldots, \mathcal{N}(1).$$

The larger is $q^{s(n)}_{m(n)}$, the stronger will the common tendency affect the evolution of debtor n.

The counts $N_i^k(2)$ at time $t = 2$ are obtained by the following formula

$$N_i^k(2) = \sum_{n=1}^{\mathcal{N}(1)} \mathbb{1}_{\{X_n(2)=i, s(n)=k\}},\ i = 1, 2, \ldots, M+1.$$

The same construction applies to the $M \times K$ matrix $\mathbf{N}(2) = (N_i^k(2))$ to produce $\mathbf{N}(3)$ and so on.

The number of defaults D_2 at time $t = 2$ reads

$$\mathcal{N}(1) - \mathcal{N}(2) = \sum_{n=1}^{\mathcal{N}(1)} \mathbb{1}_{\{X_n(2)=M+1\}}.$$

In order to calibrate the model, we have to estimate $M \times (M+1) + \frac{M(M-1)}{2} + M \times K = M\frac{3M+2K+1}{2}$ parameters using the observed data. Conceptually, the parameters are: the Markovian transition probabilities forming the matrix P, the coefficients of correlations between the tendency variables given in the symmetric non-negative definite matrix C, and the probabilities of success $\{q_i^k\}$ summarized in the matrix Q.

In the following, we describe how to estimate the parameters of our model from real data. For estimating the overall transition matrix P one may either use the given data set or use the a matrix of a rating agency. Typically the second case the value of M is so high, that the existing data do not allow to estimate all needed events correlation with any satisfactory precision.

The model has been formulated assuming that all obligors are of the same size. If they differ, we may split a big debtor into several standard size smaller ones. Introducing a new industry sector formed by them and setting the corresponding $q_{m(n)}^{s(n)}$ equal to 1, we embed the situation into the above framework.

3 A scheme for parameters estimation

We denote the given empirical data set as $x_n(t),\ n = 1, 2, \ldots, \mathcal{N},\ t = 1, 2, \ldots, T$, where $x_n(t)$ is the rating of the n-th firm in year t, T is the time span of the observations and $\mathcal{N}$ is the total number of firms in the data set.

The natural estimate for $p_{i,j}$ is

$$\hat{p}_{i,j} = [\sum_{n=1}^{\mathcal{N}} \sum_{t=2}^{T} \mathbb{1}_{\{x_n(t)=j,\ x_n(t-1)=i\}}]/[\sum_{n=1}^{\mathcal{N}} \sum_{t=2}^{T} \mathbb{1}_{\{x_n(t-1)=i\}}], \tag{1}$$

(If the rating of a firm is missing in some year, the formula (1) has to be adapted accordingly).

Let $\mathcal{I}_{i,j}^k$ be the indicator of the event that a firm in sector k moves from rating i to rating j.

We are especially interested in the event correlations

$$\rho_{i\to j;I\to J}^{k,\ell} = \mathbb{C}\mathrm{orr}(\mathcal{I}_{i,j}^k, \mathcal{I}_{I,J}^\ell) = c_{i,I} d_{i\to j;I\to J} q_i^k q_I^\ell. \tag{2}$$

where

$$d_{i\to j;I\to J} = \begin{cases} \sqrt{\frac{p_{i,j}p_{I,J}(1-p_i^+)(1-p_I^+)}{(1-p_{i,j})(1-p_{I,J})p_i^+ p_I^+}} & \text{for } j \le i,\ J \le I \\ -\sqrt{\frac{p_{i,j}p_{I,J}(1-p_i^+)p_I^+}{(1-p_{i,j})(1-p_{I,J})p_i^+(1-p_I^+)}} & \text{for } j \le i,\ J > I \\ -\sqrt{\frac{p_{i,j}p_{I,J}p_i^+(1-p_I^+)}{(1-p_{i,j})(1-p_{I\ J})(1-p_i^+)p_I^+}} \text{ for } j > i, & J \le I \\ \sqrt{\frac{p_{i,j}p_{I,J}p_i^+ p_I^+}{(1-p_{i,j})(1-p_{I,J})(1-p_i^+)(1-p_I^+)}} & \text{for } j > i,\ J > I. \end{cases}$$

The values in the left hand side of (2) may be estimated empirically. In fact, a natural estimate for $\rho^{k,\ell}_{i\to j;I\to J}$ is

$$\frac{\sum_{(n,m)} \sum_{t=2}^{T} \mathbb{1}_{\{x_n(t)=j,\ x_n(t-1)=i,\ k(n)=k\}} \mathbb{1}_{\{x_m(t)=J,\ x_m(t-1)=I,\ k(s)=\ell\}}}{\sum_{(n,m)} \sum_{t=2}^{T} \mathbb{1}_{\{x_n(t)=j,\ k(n)=k\}} \mathbb{1}_{\{x_m(t)=I,\ k(m)=\ell\}}} - \hat{p}_{i,j}\hat{p}_{I,J}.$$

Here $n \neq m$, $1 \le n,\ m \le \mathcal{N}(t)$.

The values $d_{i\to j;I\to J}$ are known as long as a matrix P has been given or it has been estimated from the observed data.

From now onwards we regard $\mathbb{C}\mathrm{orr}(\mathcal{I}^k_{i,j}, \mathcal{I}^\ell_{I,J})$ and $d_{i\to j;j\to J}$ as known values. Using them, we want to estimate the unknown matrices C and Q. In what follows next we assume that the required $d_{i\to j;I\to J} \neq 0$. Depending upon the matrix P, this may not always be the case when the calculations are not precise enough. See the numerical example given in KP07. In terms of the problem in hands, $d_{i\to j;I\to J} = 0$ means that the corresponding event correlations are useless for identifying the unknown parameters.

Setting $v^{k,\ell}_{i\to j;I\to J} = \rho^{k,\ell}_{i\to j;I\to J}/d_{i\to I;j\to J}$, we may separate in (2) the given variables and the unknowns. Then for the correlation between the credit rating moves $i \to j$ and $I \to J$ in industry sectors k and ℓ, correspondingly, relation (2) implies that

$$v^{k,\ell}_{i\to j;I\to J} = c_{i,I}\, q_i^k q_I^\ell. \tag{3}$$

Here $i, I = 1, 2, \ldots, M$, $j, J = 1, 2, \ldots, M+1$, $k, \ell = 1, 2, \ldots, K$. (Note that $i \neq j$, as for $i = j$ no move takes place. In a similar manner, $I \neq J$.) For a fixed pair i, j of credit ratings, summing up relations (3) over all possible industry sectors, we have

$$v_{i\to j;I\to J} = c_{i,I} \sum_{k,\ell=1}^{K} q_i^k q_I^\ell, \tag{4}$$

where

$$v_{i\to j;I\to J} = \sum_{k,\ell=1}^{K} v^{k,\ell}_{i\to j;I\to J}.$$

Setting

$$q_i = \sum_{k=1}^{K} q_i^k,$$

expressions (4) become

$$v_{i\to j;I\to J} = c_{i,I} q_i q_I. \tag{5}$$

Since $c_{i,i} = 1$, we conclude that

$$q_i = \sqrt{v_{i\to j;i\to J}},\ i = 1, 2, \dots, M, \tag{6}$$

for all possible $j, J \neq i$. Substituting these values in (5), we obtain that

$$c_{i,I} = v_{i\to j;I\to J} \big/ \sqrt{v_{i\to s;i\to S} v_{I\to l;I\to L}},\ i, I = 1, 2, \dots, M. \tag{7}$$

Note that in these relations only the initial credit ratings matter while the left hand side remains unchanged when the destinations vary.

For a fixed pair i, I of credit ratings and a fixed industry sector k, summing up relations (3) over all possible industry sectors ℓ, we have

$$v^k_{i\to j;I\to J} = c_{i,I} q^k_i \sum_{\ell=1}^{K} q^\ell_j = c_{i,I} q^k_i q_j,$$

where

$$v^k_{i\to j;I\to J} = \sum_{\ell=1}^{K} v^{k,\ell}_{i\to j;I\to J}.$$

Substituting here the expressions for q_I and $c_{i,I}$ from (6) and (7) correspondingly, we obtain that

$$q^k_i = v^k_{i\to j;I\to J} \sqrt{v_{i\to l;i\to L}} / v_{i\to s;I\to S},\ i = 1, 2, \dots, M,\ k = 1, 2, \dots, K. \tag{8}$$

Here again only the initial index i matters.

Relations (7) and (8) allow to evaluate the matrices C and Q. First, substituting one-year empirical defaults event correlations $\hat{\rho}^{k,\ell}_{i\to j;I\to J}$ instead of $\mathbb{C}\mathrm{orr}(\mathcal{I}^k_{i,j}, \mathcal{I}^\ell_{I,J})$ in the corresponding formulas, we obtain estimates $\widehat{v}^{k,\ell}_{i\to j;I\to J}$ of $v^{k,\ell}_{i\to j;I\to J}$. Plugging $\widehat{v}^{k,\ell}_{i\to j;I\to J}$ insted of $v^{k,\ell}_{i\to j;I\to J}$ in the expressions for $v_{i\to j;I\to J}$ and $v^k_{i\to j;I\to J}$, gives estimates, $\widehat{v}_{i\to j;I\to J}$ and $\widehat{v}^k_{i\to j;I\to J}$, of these values. Substituting in the right hand side of (7) and (8) $\widehat{v}_{i\to j;I\to J}$ and $\widehat{v}^k_{i\to j;I\to J}$, yields estimates $\widehat{c}_{i,I}$ and $\widehat{q}^k_i$ for $c_{i,\,I}$ and q^k_i.

Note that, using only default correlations, implies $j = J = M + 1$. Upgrade event correlations correspond to pairs such that $i \geq j$ and $I \geq J$. Respectively, downgrade events imply $j > i$ and $J > I$.

Since a convex combination of right-hand sides in (7) equals $c_{i,I}$, the corresponding convex combination of

$$\widehat{v}_{i\to j;I\to J} \Big/ \sqrt{\widehat{v}_{i\to s;i\to S} \widehat{v}_{I\to l;I\to L}}$$

will be an estimate for $c_{i,I}$, as well. However, as a result of averaging, the latter can be more accurate than the former. The same remark applies to convex combinations of the right hand sides in (6).

4 Estimating parameters from real data

Due to scarcity of default statistics, empirical studies, as for example Nagpal and Bahar (2001), focus on two rating classes: investment grade debtors (we will index them by "1") and non-investment grade ones (labelled below by "2"). We follow this simplification and set $M = 2$ for the rest of the paper. Like in Nagpal and Bahar (2001), investment grade firms are characterized by S&P's ratings from AAA to BBB, while non-investment grade ones occupy the ratings from BB and downward.

We base our empirical study on a data set of S&P's sample containing time series of ratings of 10413 companies from 30 countries. Applying the Standard Industry Classification (SIC) of the US Department of Labor (Occupational Safety and Health Administration) first digit industry classification, a number k from 1 to 6 was assigned to each of the following industry sectors:

k	Description	SIC code
1	Agriculture, mining and construction	0–1999
2	Manufacturing	2000–3999
3	Transportation, technology and utility	4000–4999
4	Trade	5000–5999
5	Finance	6000–6999
6	Services	7000–8999

The study covers 17 years starting with 1990. The time instants $t = 1, 2, \ldots, 17$ correspond to the years 1990, 1991, ..., 2006.

The 2×3 transition matrix P was estimated by (1) leading to

$$P = \begin{pmatrix} 0.9732 & 0.0258 & 0.0010 \\ 0.0881 & 0.8865 & 0.0254 \end{pmatrix}.$$

Hence $p_1^+ = p_{1,1} = 0.9732$ and $p_2^+ = p_{2,1} + p_{2,2} = 0.9746$. Then the required $d_{i \to j;\, I \to J}$ read:

$$\begin{aligned} d_{1\to 3;1\to 3} &= 0.0364, \\ d_{2\to 3;2\to 3} &= 1.0000, \\ d_{1\to 3;2\to 3} &= 0.1940. \end{aligned}$$

We write here P rather than $\widehat{P}$ because the corresponding values are regarded as given inputs.

As we estimate the $2 \times 6 = 12$ probabilities q_i^k and a single coefficient of correlation $c_{1,2}$ between the tendency variables for the investment grade debtors and non-investment grade ones, at least two types of moves and correlations between them have to be considered: $1 \to 3$ and $2 \to 3$. Thus, we look at default events for investment grade and non-investment grade debtors and evaluate their correlations between different industries within a credit class and between the classes.

The one-year empirical default event correlation, $\widehat{\rho}^{k,\ell}_{i\to 3;I\to 3}$, between a debtor from sector k having credit rating i and a debtor from sector ℓ belonging to credit class I is estimated as follows. (We need these values for $\widehat{v}^{k,\ell}_{i\to 3;I\to 3} = \widehat{\rho}^{k,\ell}_{i\to 3;I\to 3}/d_{i\to 3;I\to 3}$.) For example, the one-year empirical default correlation $\widehat{\rho}^{1,3}_{1\to 3;2\to 3}$ between an investment grade issuer of the agriculture, mining and construction sector and a non-investment grade issuer of the transportation, technology and utility sector reads

$(FJD^{1,3}_{1,2} - FD^1_1 FD^3_2)/\sqrt{FD^1_1(1-FD^1_1)FD^3_2(1-FD^3_2)}$,

where

$FJD^{1,3}_{1,2} = \sum_{t=1}^{17} D^{(1,1)}_t D^{(2,3)}_t / \sum_{t=1}^{17} \mathcal{N}^{(1,1)}_t \mathcal{N}^{(2,3)}_t$,

$FD^1_1 = \sum_{t=1}^{17} D^{(1,1)}_t / \sum_{t=1}^{17} \mathcal{N}^{(1,\,1)}_t$, $FD^3_2 = \sum_{t=1}^{17} D^{(2,3)}_t / \sum_{t=1}^{17} \mathcal{N}^{(2,3)}_t$.

Here $\mathcal{N}_t^{(m,\,s)}$ denotes the number at time t of firms in sector s whose credit rating is m and $D^{(m,s)}_t$ stands for the number of defaulters at time t in sector s whose credit rating at time $t-1$ was m. That is, FD^s_m denotes the frequency of defaults (for the whole period of observation) in sector s among firms whose credit rating is m. Similarly, $FJD^{k,s}_{i,I}$ stands for the frequency of joint defaults (for the whole period of observation) of a debtor belonging to industry sector k and credit class i and a debtor from industry sector s and credit class I. Estimating correlations between debtors within a sector and a credit class, like $\widehat{\rho}^{k,k}_{m\to 3;m\to 3}$, we calculate $FJD^{k,k}_{m,m}$ using the addends $[D^{(m,k)}_t]^2$ and $[\mathcal{N}^{(m,k)}_t]^2$ rather than $D^{(m,k)}_t[D^{(m,k)}_t - 1]$ and $\mathcal{N}^{(m,k)}_t[\mathcal{N}^{(m,k)}_t - 1]$ in the corresponding sums. De Servigny and Renault (2002) suggested this modification to reduce the number negative empirical event correlations due to one or none default at some time instants. Since there are too few default events among the investment grade firms, even these estimators may imply some negative empirical default correlations.

Using the formulae (7) and (8) for our data set, we obtain that

$$c_{1,2} = 0.7843$$

and that

$$Q = \begin{pmatrix} 0.1881 & 0.1002 & 0.1830 & 0.2262 & 0.0621 & 0.2405 \\ 0.1972 & 0.1534 & 0.2775 & 0.1177 & 0.1381 & 0.1161 \end{pmatrix}.$$

With these values we may calculate the theoretical default correlations and other event correlations. They include upgrade events, given by a move $2 \to 1$, and downgrade events, when an investment grade firm becomes a non-investment grade firm.

The joint distribution of the tendency variables χ_1 and χ_2 reads

$$\begin{aligned}
\mathbb{P}\mathrm{r}\{\chi_1 &= 0, \chi_2 = 0\} = 0.0206, \\
\mathbb{P}\mathrm{r}\{\chi_1 &= 1, \chi_2 = 0\} = 0.0048, \\
\mathbb{P}\mathrm{r}\{\chi_1 &= 0, \chi_2 = 1\} = 0.0062, \\
\mathbb{P}\mathrm{r}\{\chi_1 &= 1, \chi_2 = 1\} = 0.9684.
\end{aligned}$$

5 Quick modeling of dependent credit rating transitions

After having identified the parameters, one may simulate the coupled rating processes. In this section we present a much quicker, but approximative method for finding the portfolio-wide loss distribution.

While the full simulation traces the evolution of every debtor, the complexity of the approximative method depends only on the number of cells (combinations of rating classes and industry sectors). As all approximations substitute the true distribution by a mixture of Gaussians, a rounding is required for the random integer default counts.

Heuristics 1: The unconditional mean-variance approach. The first and second moments of the random counts are easily calculated as

$$\begin{aligned}
\mathbb{E}N_I^k(2) &= \sum_{i=1}^{M} N_i^k(1)p_{i,I}, \\
\mathbb{V}\mathrm{ar}(N_I^k(2)) &= \sum_{i=1}^{M} N_i^k(1)p_{i,I}(1-p_{i,I}) - \sum_{i=1}^{M} N_i^k(1)p_{i,I}(q_i^k)^2 r_{I,I}^{i,i} \\
&\quad + \sum_{i=1}^{M}[N_i^k(1)q_i^k]^2 r_{I,I}^{i,i} + 2\sum_{i=1}^{M} N_i^k(1)q_i^k \sum_{j=i+1}^{M} N_j^k(1)q_j^k c_{i,j} r_{I,I}^{i,j}, \\
\mathbb{C}\mathrm{ov}(N_I^k(2), N_J^s(2)) &= \sum_{i=1}^{M} N_i^k(1)q_i^k \sum_{j=1}^{M} N_j^s(1)q_j^s c_{i,j} r_{I,J}^{i,j}, \qquad k \neq s \\
\mathbb{C}\mathrm{ov}(N_I^k(2), N_J^k(2)) &= 6 - \sum_{i=1}^{M} N_i^k(1)p_{i,I}p_{i,J} - \sum_{i=1}^{M} N_i^k(1)(q_i^k)^2 r_{I,J}^{i,i} \\
&\quad + \sum_{i=1}^{M} N_i^k(1)q_i^k \sum_{j=1}^{M} N_j^k(1)q_j^k c_{i,j} r_{I,J}^{i,j}.
\end{aligned}$$

Here

$$r_{I,J}^{i,j} = \begin{cases}
p_{i,I}p_{j,J}\sqrt{\frac{(1-p_i^+)(1-p_j^+)}{p_i^+ p_j^+}} & \text{for } I \leq i,\ J \leq j; \\
-p_{i,I}p_{j,J}\sqrt{\frac{(1-p_i^+)p_j^+}{p_i^+(1-p_j^+)}} & \text{for } I \leq i,\ J > j; \\
-p_{i,I}p_{j,J}\sqrt{\frac{p_i^+(1-p_j^+)}{(1-p_i^+)p_j^+}} & \text{for } I > i,\ J \leq j; \\
p_{i,I}p_{j,J}\sqrt{\frac{p_i^+ p_j^+}{(1-p_i^+)(1-p_j^+)}} & \text{for } I > i,\ J > j.
\end{cases}$$

Applying the unconditional mean-variance approach, we may introduce $M \times K$ normally distributed counterparts $\widetilde{N}_I^k(2)$ of $N_I^k(2)$ having the same means, variances and covariances. Having the $M \times K$ matrix $\widetilde{\mathbf{N}}(2)$ formed by $\widetilde{N}_I^k(2)$, the number of defaults D_2 may be approximated as $\mathcal{N}_1 - \widetilde{\mathcal{N}}_2$, where $\widetilde{\mathcal{N}}_2$ stands for the sum of all entries of $\widetilde{\mathbf{N}}(2)$. Alternatively, D_2 may be estimated as $\sum_{k=1}^{K} \widetilde{N}_{M+1}^k(2)$. In this case only K normal random variables $\widetilde{N}_{M+1}^k(2)$, counterparts of $N_{M+1}^k(2)$, are required.

Now there is no need to solve the quadratic optimization problem required in KP07. In other words, this heuristics generates a portfolio path even when the original model by Kaniovski and Pflug is not feasible (because there exist no tendency variables with the given correlations).

Heuristics 2: The conditional mean-variance approach. For introducing another approximation, we recall that, conditional on a binary vector $\overrightarrow{\chi} = (\chi_1, \chi_2, \dots, \chi_M)$ of tendency variables, the random variables η_n are independent. We have that

$$\begin{aligned}
\mathbb{E}(N_I^k(2) \mid \overrightarrow{\chi}) &= \sum_{i=1}^{M} N_i^k(1)\kappa_{i,I}^k, \\
\mathbb{V}\mathrm{ar}(N_I^k(2) \mid \overrightarrow{\chi}) &= \sum_{i=1}^{M} N_i^k(1)\kappa_{i,I}^k(1-\kappa_{i,I}^k), \\
\mathbb{C}\mathrm{ov}(N_I^k(2), N_J^s(2) \mid \overrightarrow{\chi}) &= 0, \qquad k \neq s \\
\mathbb{C}\mathrm{ov}(N_I^k(2), N_J^k(2) \mid \overrightarrow{\chi}) &= -\sum_{i=1}^{M} N_i^k(1)\kappa_{i,I}^k\kappa_{i,J}^k.
\end{aligned}$$

Here

$$\kappa_I^{i,k} = [1-q_i^k]p_{i,I} + q_i^k p'_{i,I}.$$

The conditional probabilities $p'_{i,I}$ have been defined above as follows:
when $\chi_i = 1$

$$p'_{i,I} = \begin{cases} p_{i,I}/p_i^+ & \text{for } I = 1, 2, \dots, i, \\ 0 & \text{for } I = i+1, i+2, \dots, M+1; \end{cases}$$

when $\chi_i = 0$

$$p'_{i,I} = \begin{cases} p_{i,I}/(1-p_i^+) & \text{for } I = i+1, i+2, \dots, M+1, \\ 0 \text{ for } I = 1, 2, \dots, i. \end{cases}$$

Matching all moments up to the second order of the conditional on $\overrightarrow{\chi}$ joint distribution of $N_I^k(2)$, $1 \le I \le M$, $1 \le k \le K$, we may substitute it by the joint distribution of $M \times K$ normal random variables $\widehat{N}_I^k(2)$, $1 \le I \le M$, $1 \le k \le K$, having the same expected values and the same covariances matrix (conditional on $\overrightarrow{\chi}$). The number of defaults D_1 now may be approximated by $\mathcal{N}_1 - \widehat{\mathcal{N}}_2$, where $\widehat{\mathcal{N}}_2$ stands for the sum of all entries of $\widehat{\mathbf{N}}(2)$ formed by $\widehat{N}_I^k$, $1 \le I \le M$, $1 \le k \le K$. Alternatively, we may model additionally K normal random variables $\widehat{N}_{M+1}^k(2)$, $1 \le k \le K$, and estimate D_2 as $\sum_{k=1}^K \widehat{N}_{M+1}^k(2)$.

The approximation method generates random realizations of the tendency vector $\overrightarrow{\chi}$ or enumerates all possible 2^M values and represents the total number of defaulters as a mixture of normal distributions. For the realization of the tendency vector $\overrightarrow{\chi}$, the quadratic optimization problem introduced in KP07 has to be solved to find the corresponding distribution.

Heuristics 3: The Conditional Multinomial Approximation. This approach is a more refined version of the previous one.

Recall that $N_i^k(1)$ is the number of debtors in rating class i and in sector k. We draw first a random version of the tendency vector χ_i using the correlation matrix C (or use complete enumeration). Then we sample for every category (i,k) a Binomial random variable $\nu_i^k \sim Binomial(N_i^k, q_i^k)$. Given the tendency χ_i for this rating class i, we sample $V_i^k = (V_i^k(1), \ldots, V_i^k(M{+}1)) \sim Multinomial(1, p')$, where $p'_{i,j}$ is the conditional migration probability given the tendency. Notice that this multinomial distribution has $L = M + 1$ components. We also sample $Y_i^k = (Y_i^k(1), \ldots, Y_i^k(M + 1)) \sim Multinomial(N_i^k - \nu_i^k, p)$. The new number of debtors in category (j,k) is then

$$N_j^k(2) = \sum_{i=1}^{M} [\nu_i^k \cdot V_i^k(j) + Y_i^k(j)] \tag{9}$$

and the newly defaulted are

$$N_{M+1}^k(2) = \sum_{i=1}^{M} [\nu_i^k \cdot V_i^k(M+1) + Y_i^k(M+1)]. \tag{10}$$

The multinomial vectors may be approximated by normal random variables as indicated in the Appendix.

If only the number of defaulters matters, the variable $N_{M+1}^k(2)$ can be represented as an independent sum of compounds of Binomial variables. The variables ν_i^k, $V_i^k(M{+}1)$ and $Y_i^k(M+1)$ are conditionally independent given the tendency and distributed as follows:

$$\begin{aligned} \nu_i^k &\sim Binomial(N_i^k, q_i^k) \\ V_i^k(M+1) &\sim Binomial(1, p'_{i,M+1}) \\ Y_i^k(M+1) &\sim Binomial(N_i^k - \nu_i^k, p_{i,M+1}) \end{aligned}$$

By the Lemma of the Appendix, the variable $\nu_i^k \cdot V_i^k(M+1) + Y_i^k(M+1)$ is distributed according to a compound Binomial distribution, which may be approximated by a compound normal distribution. This approximation allows a rather accurate yet efficient approximation.

6 Simulations of a portfolio evolution governed by empirically measured parameters

We have run 2000 paths of a portfolio containing initially 1200 debtors: 100 for each combination of a non-default credit class and an industry sector. The transition matrix and the coefficient of correlation equal to the above empirically measured values. In the tables below the average number of defaults as well as the 95 percentile of the number of defaults after 3, 5 and 7 years are given. The figures give the sample distributions of the number of defaults after 7 years. The simulations are done using the simulation of the full coupled MC model (called exact method here) and the above three heuristics.

Algorithm	Mean	95 percentile
Heuristics 1	44	86
Heuristics 2	44	125
Heuristics 3	44	126
Full simulation	44	127

Table 1. Time horizon 3 years.

Algorithm	Mean	95 percentile
Heuristics 1	68	119
Heuristics 2	67	148
Heuristics 3	68	149
Full simulation	68	151

Table 2. Time horizon 5 years.

Algorithm	Mean	95 percentile
Heuristics 1	89	148
Heuristics 2	89	169
Heuristics 3	88	169
Full simulation	89	170

Table 3. Time horizon 7 years.

As a conclusion, one sees that the simple matching of moments is not quite accurate, the Heuristics 2 gives very good and Heuristics 3 even better approximations and allow to assess the loss distributions to large and even huge portfolios in a way, which is much more efficient than the full simulation of all individual coupled rating trajectories.

7 Appendix: Approximations

Approximating a Binomial distribution. It is well known that a $Binomial(N, p)$ distribution has the same first two moments as a $Normal(Np, Np(1-p))$ distribution.

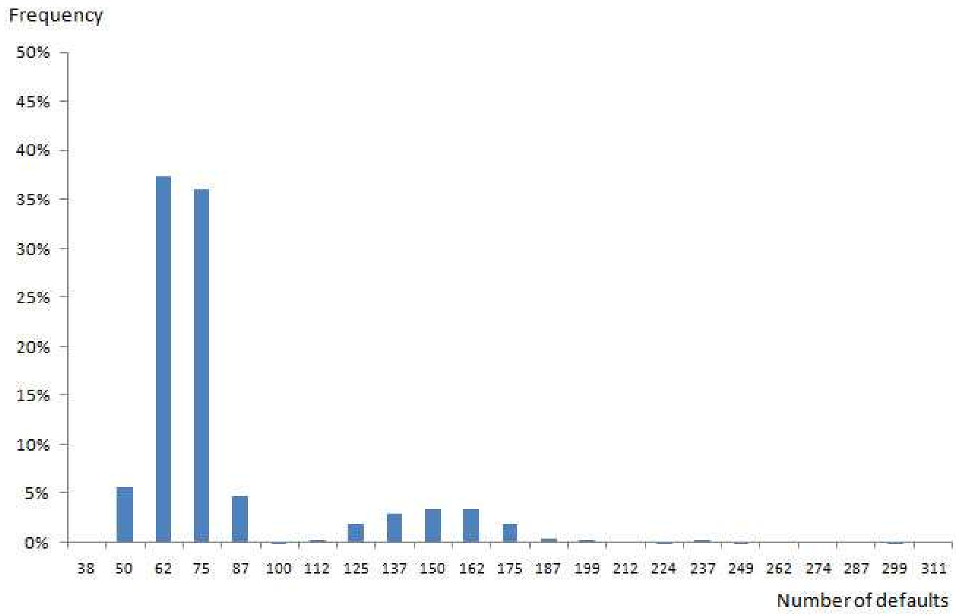

Figure 1: Full simulation, distribution of the number of defaults after 7 years.

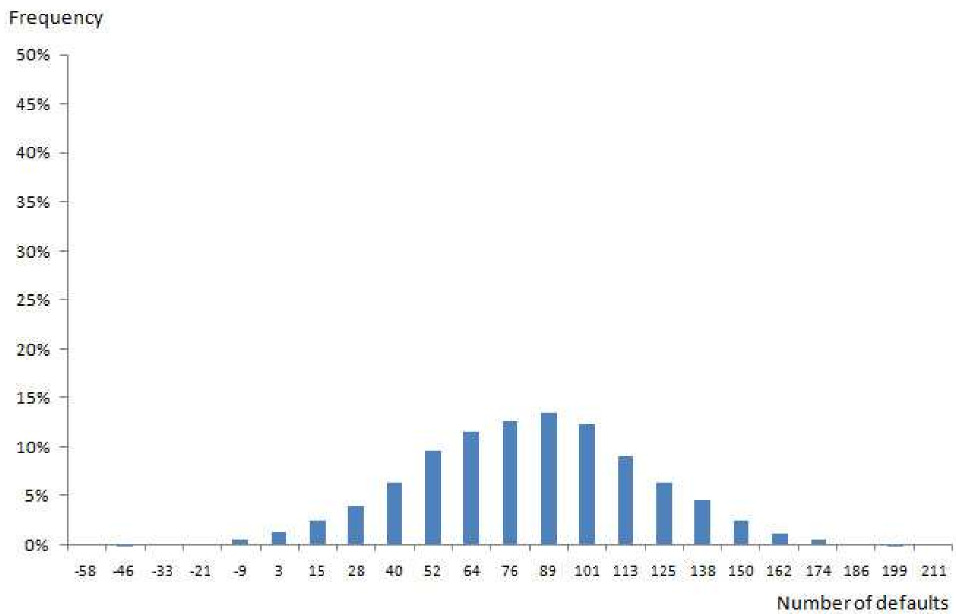

Figure 2: Heuristics 1, distribution of the number of defaults after 7 years.

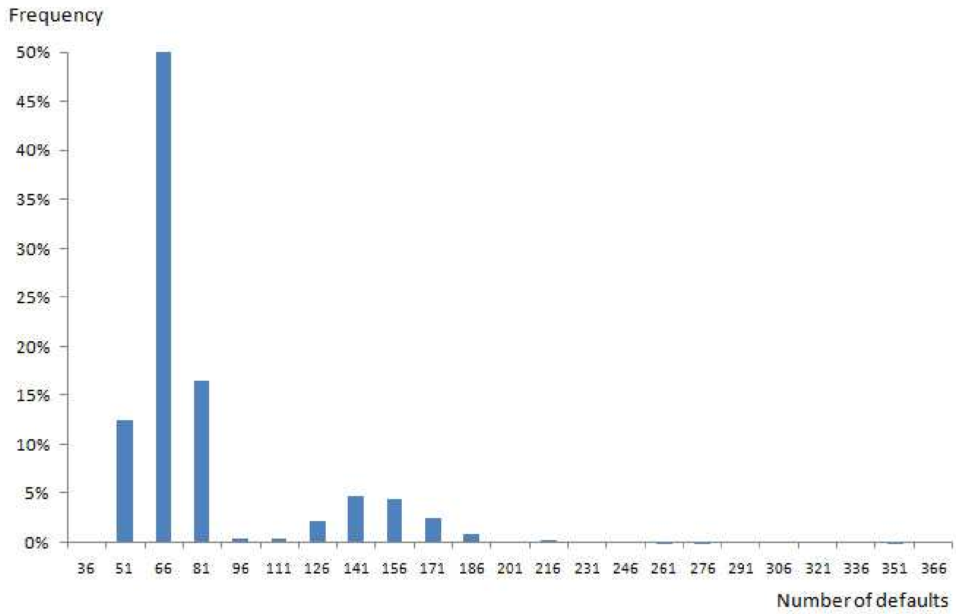

Figure 3: Heuristics 2, distribution of the number of defaults after 7 years.

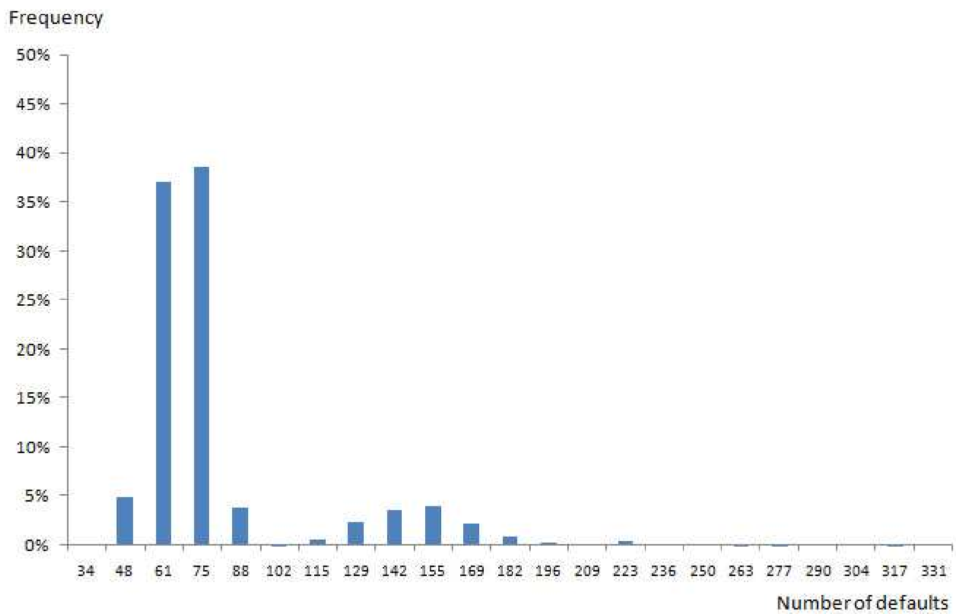

Figure 4: Heuristics 3, distribution of the number of defaults after 7 years.

Approximating a Multinomial distribution. Let $V = (V_1, \ldots, V_L)^\top$ be a $Multinomial(N, p_1, \ldots, p_L)$ distribution. Recall that this distribution has the following moments:

$$\mathbb{E}[V] = \begin{pmatrix} Np_1 \\ \vdots \\ Np_L \end{pmatrix}$$

$$\mathbb{C}\text{ov}[V] = \begin{pmatrix} Np_1(1-p_1) & -Np_1p_2 & \cdots & -Np_1p_L \\ -Np_2p_1 & Np_2(1-p_2) & \cdots & -Np_2p_L \\ \vdots & \vdots & \ddots & \vdots \\ -Np_Lp_1 & -Np_Lp_2 & \cdots & Np_L(1-p_L) \end{pmatrix}$$

We may approximate the random vector V by the normal random vector W. Here is how one may construct W: Let $Z_i \sim Normal(0, Np_i), i = 1, \ldots, L$ and let

$$W_i = Z_i - p_i \sum_{m=1}^{L} Z_i + Np_i.$$

Then V and W coincide in the first two moments.

A result on compound Binomial distributions.

Let $\mathbb{C}(Distribution1, Distribution2, p)$ be the compound distribution (which is distributed according to $Distribution1$ with probability p and according to $Distribution2$ with probability $1-p$). Our credit rating model uses compound Bernoulli variables. We show a result on these compounds.

Let ν, V, Y be independent random variables with the following Binomial distributions:

Lemma.

$$\begin{aligned} \nu &\sim Binomial(N, q) \\ V &\sim Binomial(1, p') \\ Y &\sim Binomial(N-\nu, p) \end{aligned}$$

Then the r.v. $A = \nu \cdot V + Y$ is distributed according to a compound Binomial distribution.

$$A \sim \mathcal{C}(Binomial(N, (1-q)p+q), Binomial(N, (1-q)p), p') \tag{11}$$

Proof. Conditional on ν, the moment generating function of A is

$$\begin{aligned} \mathbb{E}(u^A|\nu) &= (1-p')[(1-p)+pu]^N \left(\frac{1}{(1-p)+pu}\right)^\nu \\ &+ p'[(1-p)+pu]^N \left(\frac{u}{(1-p)+pu}\right)^\nu \end{aligned}$$

Taking the expectation w.r.t. ν we arrive after some calculation at

$$(1-p')[(1-q)(1-p)+q+(1-q)pu]^N + p'[(1-q)(1-p)+[(1-q)p+q]u]^N$$

which is the generating function of the compound distribution given in (11).

Notice that (11) has expectation $\mathbb{E}(A) = N[p'((1-q)p+q)) + (1-p')(1-q)p) = Nqp' + N(1-q)p$.

To approximate A for large N, we may use the compound normal distribution

$$\begin{aligned} A \sim\ & \mathcal{C}(Normal(N[(1-q)p_2+q], N[(1-q)p_2+q](1-q)p_2), \\ & Normal(N(1-q)p_2, N(1-q)p_2[(1-q)(1-p_2)+q]), p_1, 1-p_1). \end{aligned}$$

References

[1] Credit Metrics, 1997. The Benchmark for Understanding Risk. Technical Document. J.P. Morgan Inc., New York, NY.

[2] De Servigny, A. and Renault, O., 2002. Default Correlation: Empirical Evidence. S&P Working Paper.

[3] Devis, M., and Lo, V. 2001. Infectious defaults. Quantitative Finance, Vol. 1, No.4, pp. 382-387.

[4] Hull, J., and White, A. 2004. Valuation of a CDO and an n-th to default CDS without Monte-Carlo simulation. Journal of Derivatives, Vol. 12, No. 2, pp. 18-23.

[5] Hurd, T. R., and Kuznetsov, A. 2007. Affine Markov chain model of multifirm credit migration. Journal of Risk, Vol. 3, No.1, pp. 3-29.

[6] Jarrow, R.A., Lando, D., and Turnbull, S.M., 1997. A Markov Model for the Term Structure of Credit Risk Spreads. Review of Financial Studies, Vol. 10, No. 2, pp. 481-523.

[7] Jarrow, R.A. and Turnbull, S.M., 1995. Pricing derivatives on financial securities subject to credit risk. Journal of Finance, Vol. 50, No. 1, pp. 53-85.

[8] Kaniovski, Y.M, and Pflug, G. Ch., 2007. Risk assessment for credit portfolios: A coupled Markov chain model. Journal of Banking and Finance, Vol. 31, No. 8, pp. 2303-2323.

[9] Li, D. 2000. On default correlation: a copula function approach. Journal of Fixed Income, Vol. 9, No. 4, pp. 43-54.

[10] Merton, R. 1974. On the pricing of corporate debt: the risk structure of interest rates. Journal of Finance, Vol. 29, No. 2, pp. 449-470.

[11] Nagpal, K., and Bahar, R., 2001. Measuring Default Correlation. Risk, Vol. 14, No. 3, pp. 129-132.