Statistical Inference in Models of Dependent Defaults and Credit Migrations

Alexander J. McNeil
Department of Mathematics
Federal Institute of Technology
ETH Zentrum
CH-8092 Zurich

www.math.ethz.ch/~mcneil   www.math.ethz.ch/finance/

25th February 2005, Cambridge

Forthcoming Book

Quantitative Risk Management: Concepts Techniques and Tools
by Alexander McNeil, Rüdiger Frey and Paul Embrechts
Princeton University Press, to appear in 2005

For more information visit the book website at
www.math.ethz.ch/∼mcneil/book.html .

Contents

A. Motivation

1. On Statistics

2. On Dependence

3. Literature

A1. On Statistics

This talk concerns statistical calibration of portfolio credit risk models under the real-world measure. We tacitly assume the model is to be used for credit risk measurement and management purposes, such as

- Calculation of credit VaR or expected shortfall,
- Calculation of shortfall contributions for allocation of risk capital.

Industry approaches to calibration of portfolio models have generally not been formally statistical. There are good reasons for this, mainly the lack of relevant, historical data, particularly for higher-rated companies.

Industry Calibration Approaches

Industry models generally separate the problems of estimating (i) default probabilities and (ii) model parameters describing the dependence of defaults.

1. Default probabilities are usually estimated by a historical default rate for "similar" companies, where the similarity metric may be based on ratings (CreditMetrics) or a proprietary measure like distance-to-default (KMV).

2. Dependence is usually described by a macro-economic or fundamental factor model. Parameters of factor models are often simply "assigned" by economic argument or derived from factor analyses of proxy variables (e.g. equity returns for asset value returns in the Merton-style models).

Ad Hoc Calibration and Model Risk

The ad hoc nature of some of the attempts to model dependence raises the issue of model risk.

For example, in KMV/CreditMetrics, how confident are we that we can correctly determine the size of the systematic component of risk (loadings on the common factors)?

Changes to this part of model can have drastic effect on the tail of portfolio loss distribution.

Our philosophy. As historical default and migration data improve in quantity and quality over time, the use of formal statistical inference will become more viable and should complement existing approaches.

A2. On Dependence

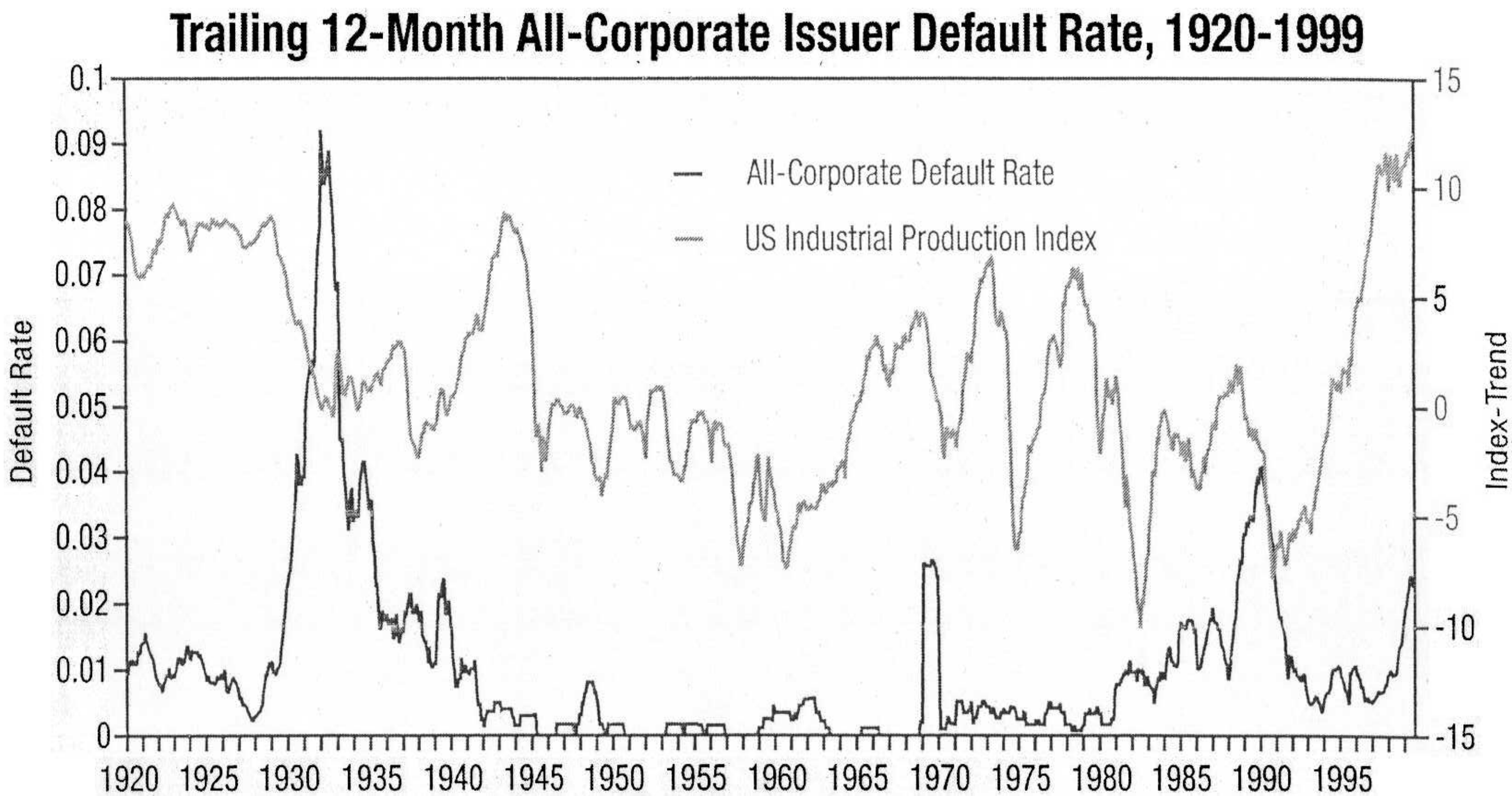
Dependence between defaults is key issue in credit risk management.

- In large balanced loan portfolios main risk is occurrence of many joint defaults – this might be termed extreme credit risk.
- Dependence between default critically affects performance of many basket credit derivatives

Sources for dependence between defaults

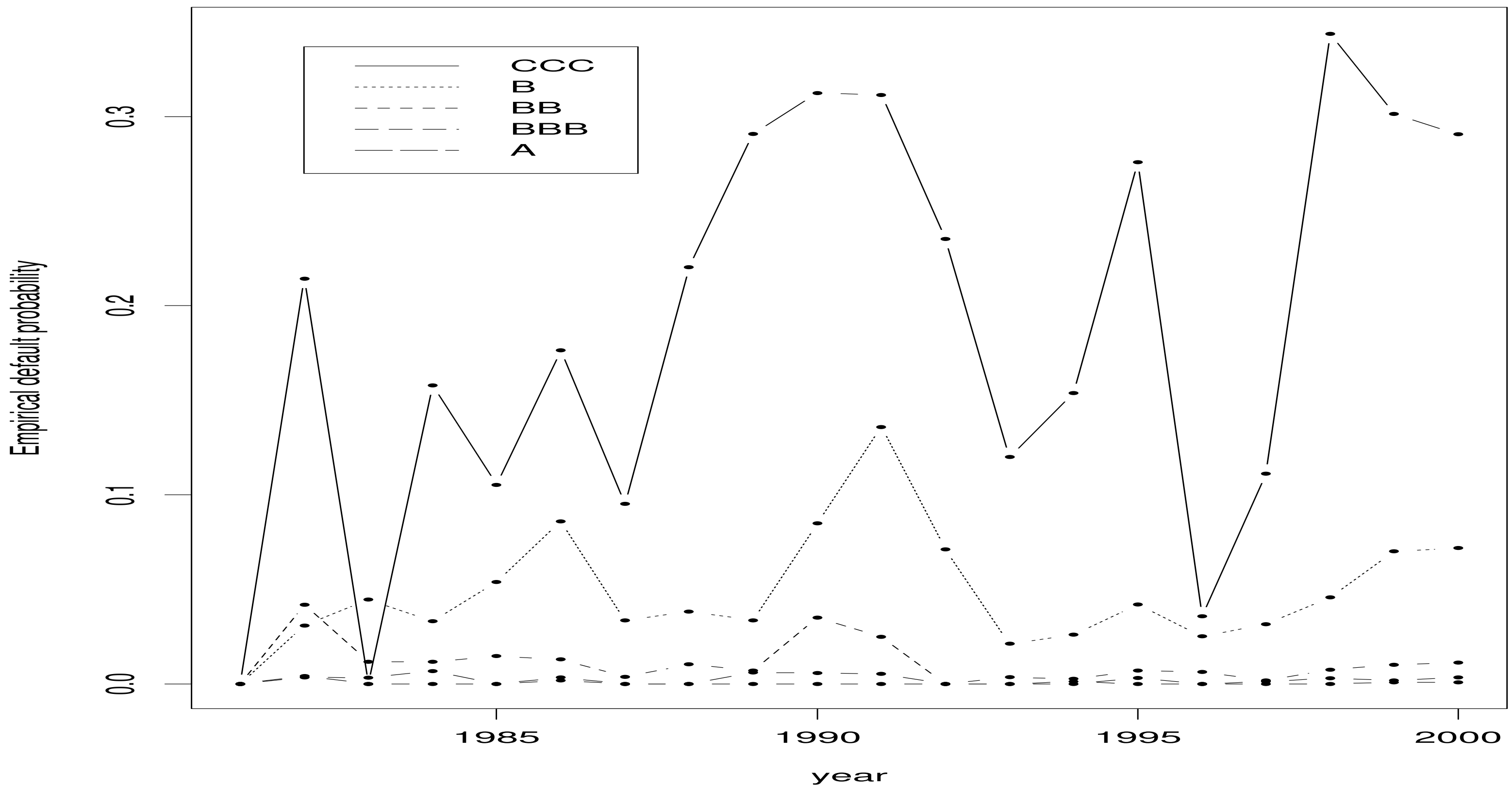
- Dependence caused by common factors (eg. interest rates and changes in economic growth) affecting all obligors
- Default of company A may have direct impact on default probability of company B and vice versa because of direct business relations, a phenomenon known as counterparty risk or contagion.

Empirical Evidence I



Moodys' annual default rates (defaulted companies/overall number of rated companies) and changes in economic growth from 1920 – 1999; changes in economic growth clearly affect default rates.

Empirical Evidence II



Standard and Poor's default data from 1980 to 2000 show clear evidence of cycles; we expect within-year and between-year dependence.

A3. Literature

There are various strands in the literature on statistical analysis of credit models.

- GLM analysis of migration count data using ordered probit model; this extends logistic regression analysis of defaults. See, for example, [Nickell et al., 2000] and [Hu et al., 2002].

- Markov chain methods for estimating rating transition matrices; see, for example, [Lando, 2004], [Lando and Skodeberg, 2002] and [Schuerman and Jafry, 2003]. Evidence for ratings momentum generally contradicts the Markov assumption.

Literature II

- Models with latent structure to capture the dynamics of systematic risk. An example is [Crowder et al., 2003] who use a two-state hidden Markov structure to capture periods of high and low default risk. See also [Gagliardini and Gouriéroux, 2004].

We essentially work in the GLM ordered probit/logit framework but add random effects to capture default and migration dependence. We incorporate the thinking of state space models by trying to make our random effects dynamic.

The data we consider are repeated cross-sectional data, which are readily available from the rating agencies.

B. Modelling Dependent Defaults

1. Dependence Through Mixing

2. From Linear Models to GLMMs

3. Fitting GLMMs

4. Example: Simplified KMV/CreditMetrics

5. Example: Model with Economic Cycle Effect

B1. Dependence Through Mixing

Consider a set of m obligors. Let $Y_1, \ldots, Y_m$ be their default indicators for the next time period, i.e. for all $i \in \{1, \ldots, m\}$

$$Y_i = \begin{cases} 1 & \text{if obligor } i \text{ defaults in the next time period} \\ 0 & \text{otherwise.} \end{cases}$$

For the time being, we assume that our set of obligors is homogenous so that the probability law of each Y_i is identical. We assume

$$Y_i \,|\, Q \overset{\text{iid}}{\sim} \mathrm{Be}(Q), \quad i = 1, \ldots, m,$$

where Q is a mixing variable with distribution on $[0, 1]$.

Distribution of Defaults

The conditional distribution of $\mathbf{Y} = (Y_1, \ldots, Y_m)'$ is given by the conditional independence property:

$$P(\mathbf{Y} = \mathbf{y} \,|\, Q = q) = \prod_{i=1}^{m} q^{y_i}(1-q)^{1-y_i}, \quad \mathbf{y} \in \{0,1\}^m.$$

The unconditional distribution of $\mathbf{Y}$ is obtained by integrating out Q:

$$P(\mathbf{Y} = \mathbf{y}) = \int P(\mathbf{Y} = \mathbf{y} \,|\, Q = q)\, \mathrm{d}G(q).$$

This two-stage stochastic model creates dependence among the responses $Y_1, \ldots, Y_m$.

Statistical evidence for dependence

Consider the two cases $Q = q_0 = \text{const}$ vs. $Q \sim \beta(a, b)$. Since $Q = \text{const}$ is a degenerate special case of $Q \sim \beta(a, b)$, we may use a likelihood-ratio test to test the hypothesis

$$H_0 : \text{the model } Q = q_0 \text{ is adequate.}$$

Under H_0, we have

$$2 \log \left(\frac{L(\hat{a}, \hat{b})}{L(\hat{q}_0)} \right) \sim \chi_1^2.$$

For e.g. rating class on earlier slide, we have `P-value 7.0 e-12`. The null hypothesis H_0 is clearly rejected.

A more general mixing structure

Conditional on random effects $\mathbf{b}$ we assume

$$Y_i \,|\, \mathbf{b} \sim \mathrm{Be}(p_i(\mathbf{b})), \quad i = 1, \ldots, m, \quad \text{where} \tag{1}$$
$$p_i(\mathbf{b}) = g\left(\mathbf{x}_i' \boldsymbol{\beta} + \mathbf{z}_i' \mathbf{b}\right).$$

- $g(\cdot)$ is a monotone function, typically a mapping from $\mathbb{R}$ to $(0,1)$ like a distribution function (e.g. $g = \Phi$).

- $\mathbf{x}_i$ and $\mathbf{z}_i$ are explanatory variables (covariates) for ith obligor, such as indicators for rating category or sector, or firm-specific information from balance sheet.

- $\boldsymbol{\beta}$ are unknown parameters (including generally an intercept).

B2. From Linear Models to GLMMs

Linear Model

$$\mathbf{Y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

- $\mathbf{Y} = (Y_1, \ldots, Y_m)'$ is (multivariate) Gaussian,
- $E(Y_i) = \mu_i = \mathbf{x}_i'\boldsymbol{\beta}$,
- $Var(\mathbf{Y}) = \sigma^2 I_{m \times m}$.

This model is not suitable for binary and count data.

Generalized Linear Model (GLM)

- $(Y_1, \ldots, Y_m)$ are independent following the same exponential family distribution (e.g. Bernoulli, Poisson, Normal),

- $E(Y_i) = \mu_i$ with $g(\mu_i) = \mathbf{x}_i' \boldsymbol{\beta}$,

- $Var(Y_i) = k_i v(\mu_i)$.

$g(\cdot)$ is the link function, $v(\cdot)$ is the variance function and k_i is a constant.

GLMs offer interesting possibilities for e.g. count data, but the responses $Y_1, \ldots, Y_m$ are independent.

Generalized Linear Mixed Model (GLMM)

- Given a realisation of $\mathbf{b}$, $(Y_1, \ldots, Y_m)$ are conditionally independent following the same exponential family distribution. $\mathbf{b}$ is a random effect following a distribution of our choice. We denote by $\boldsymbol{\theta}$ any hyperparameters of $\mathbf{b}$,

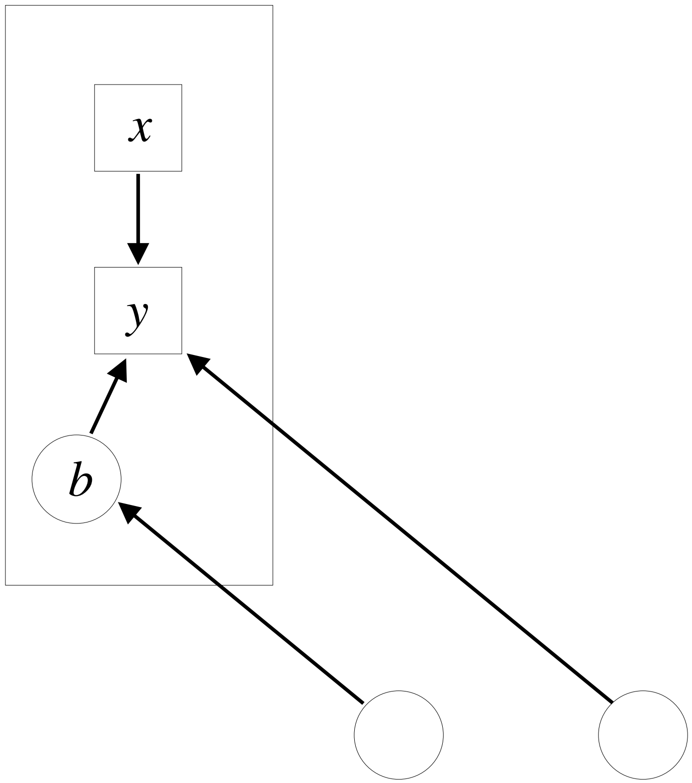
- $E(Y_i \,|\, \mathbf{b}) = \mu_i$ with $g(\mu_i) = \mathbf{x}_i'\boldsymbol{\beta} + \mathbf{z}_i'\mathbf{b}$,

- $Var(Y_i \,|\, \mathbf{b}) = k_i v(\mu_i)$.

By integrating out the effect of $\mathbf{b}$, the responses $Y_1, \ldots, Y_m$ are no longer independent.

$$\boxed{g(\mu_i) = \text{fixed effects} + \text{random effects.}}$$

GLMM as DAG (Directed Acyclic Graph): one unit



Multiple Units

To estimate the random effects and the parameters of their distribution (so-called hyperparameters) we need data corresponding to multiple realisations of the random effect.

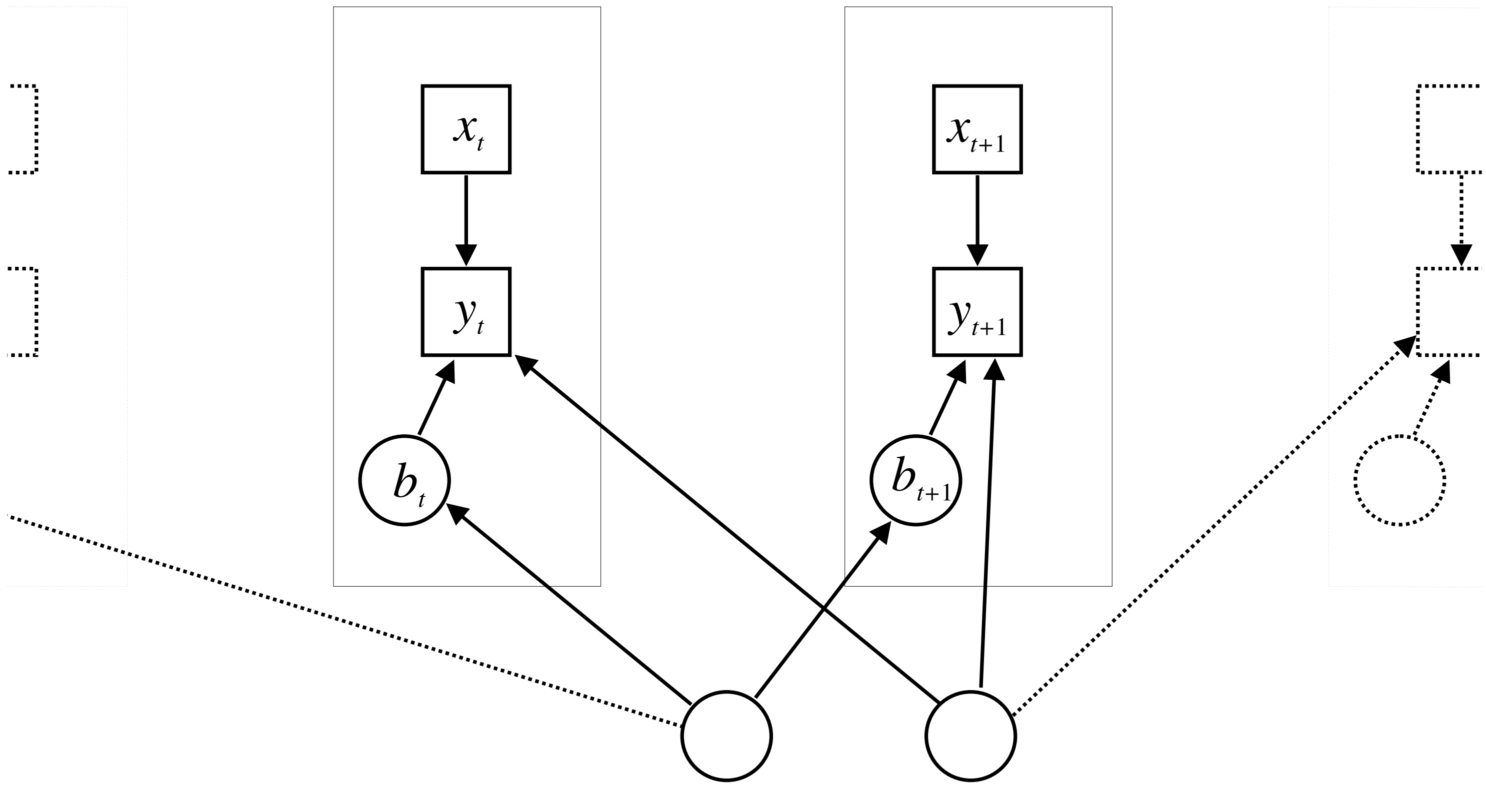
We introduce the idea of units. In our application units will correspond to years. Classical examples: patients in hospital where hospital is unit; children in schools where school is unit.

In each unit we have a random effect $\mathbf{b}_t$ generating the dependence for all observations on that unit. In our applications $\mathbf{b}_t$ can be thought of as representing stochastic state of economy in year t.

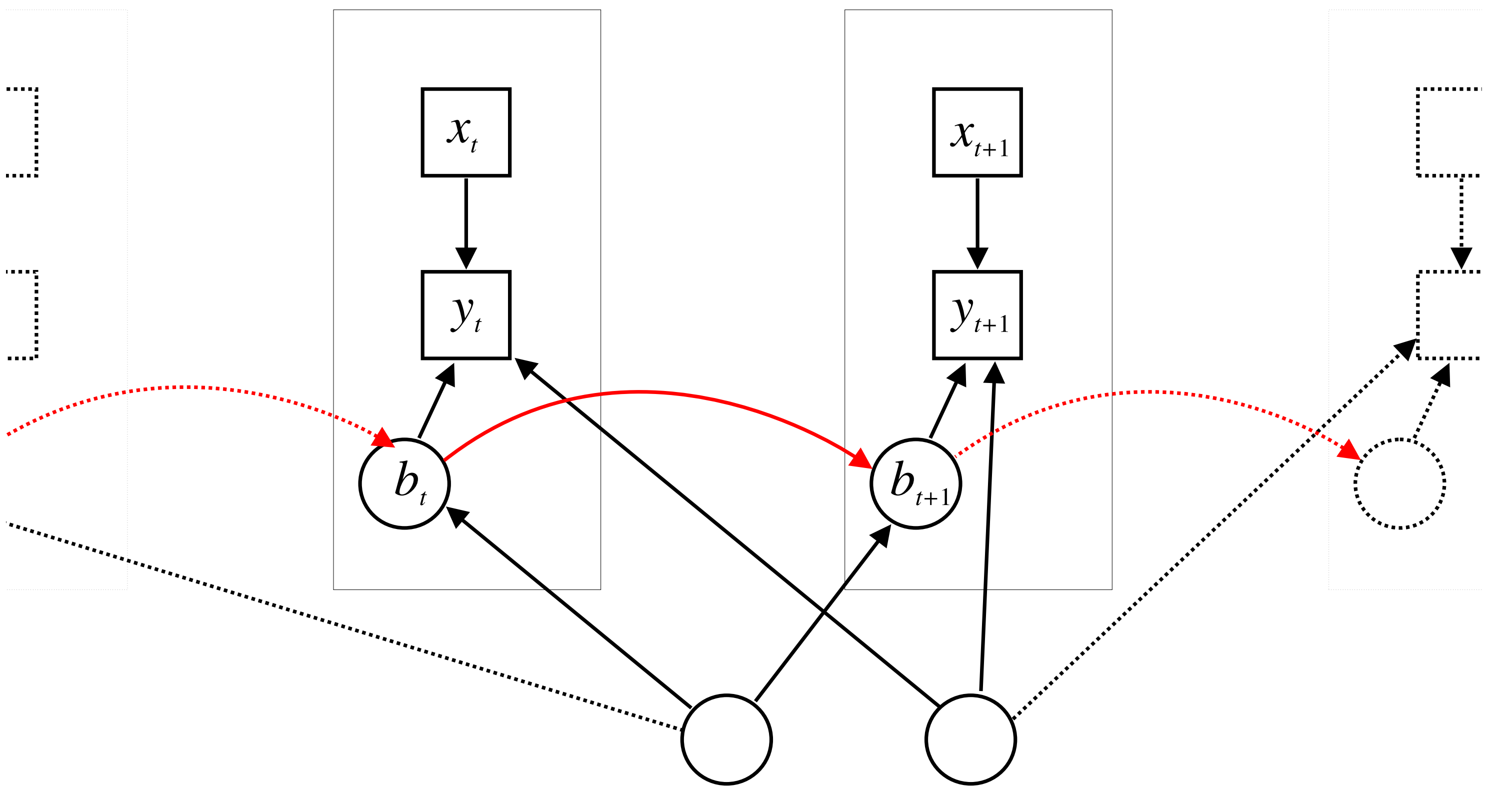
$$Y_{ti} \,|\, \mathbf{b}_t \sim \mathrm{Be}(p_{ti}(\mathbf{b}_t)), \quad i = 1, \ldots, m_t, \quad t = 1, \ldots, n, \quad \text{where}$$

$$p_{ti}(\mathbf{b}_t) = g\left(\mathbf{x}_{ti}'\boldsymbol{\beta} + \mathbf{z}_{ti}'\mathbf{b}_t\right).$$

GLMM as DAG: several independent units



GLMM as DAG: several serially dependent units



B3. Fitting GLMMs

Independent random effects

The unconditional density or mass function of $\mathbf{Y}_t = (Y_{t1}, \ldots, Y_{tm_t})'$:

$$f(\mathbf{y}_t \,|\, \boldsymbol{\beta}, \boldsymbol{\theta}) = \int_{\mathbb{R}^p} \left(\prod_{i=1}^{m_t} P(Y_{ti} = y_{ti} \,|\, \mathbf{b}_t, \boldsymbol{\beta}) \right) f_{b_t}(\mathbf{b}_t \,|\, \boldsymbol{\theta}) \, \mathrm{d}\mathbf{b}_t,$$

where $p = \dim(\mathbf{b}_t)$ and $f_{b_t}(\mathbf{b}_t|\boldsymbol{\theta})$ is the density of $\mathbf{b}_t$. The likelihood function with $\mathbf{b}_1, \ldots, \mathbf{b}_n$ independent is

$$L(\boldsymbol{\beta}, \boldsymbol{\theta} \,|\, \text{observed data}) = \prod_{t=1}^{n} f(\mathbf{y}_t \,|\, \boldsymbol{\beta}, \boldsymbol{\theta}). \tag{2}$$

There is no between-year dependence.

Fitting GLMMs

Dependent random effects

Let $\mathbf{b}_1, \ldots, \mathbf{b}_n$ have joint density $f_b(\mathbf{b}_1, \ldots, \mathbf{b}_n \,|\, \boldsymbol{\theta})$. The likelihood function $L(\boldsymbol{\beta}, \boldsymbol{\theta} \,|\, \text{observed data})$ now takes the form

$$\int \cdots \int \prod_{t=1}^{n} \prod_{i=1}^{m_t} P(Y_{ti} = y_{ti} \,|\, \mathbf{b}_t, \boldsymbol{\beta}) f_b(\mathbf{b}_1, \ldots, \mathbf{b}_n \,|\, \boldsymbol{\theta}) \, \mathrm{d}\mathbf{b}_1 \cdots \mathrm{d}\mathbf{b}_n,$$

and in particular numerical integration over $\mathbb{R}^{n \times p}$, where $p = \dim(\mathbf{b}_t)$.

The high-dimensional integrals make standard maximum likelihood difficult.

Fitting GLMMs

Bayesian Statistics

We distinguish between observed quantities $D := (\mathbf{x}_t, \mathbf{z}_t, \mathbf{y}_t)_{t=1}^n$ and unobserved quantities $\vartheta := (\boldsymbol{\theta}, \boldsymbol{\beta}, \mathbf{b}_1, \ldots, \mathbf{b}_n)$.

The prior distribution $p(\vartheta)$ expresses a state of knowledge (or ignorance) about the unobserved elements ϑ before the data D are obtained.

Inference in our model is based on the posterior distribution $p(\vartheta \,|\, D)$

$$p(\vartheta \,|\, D) = \frac{p(D \,|\, \vartheta)p(\vartheta)}{p(D)} = \frac{p(D \,|\, \vartheta)p(\vartheta)}{\int p(D \,|\, \vartheta)p(\vartheta)\, \mathrm{d}\vartheta}.$$

Problem: finding $p(\vartheta \,|\, D)$!

Fitting GLMMs

Markov Chain Monte Carlo (MCMC) Methods

Assume we want to simulate from a (multivariate) distribution $p(\mathbf{x})$.

Idea: Construct an ergodic Markov chain with p as its stationary distribution. Regard a sample of the Markov chain (possibly after a certain burn-in) as a sample from p. Constructing such a Markov chain turns out to be surprisingly simple:

- Metropolis-Hastings algorithm
- Gibbs sampler (special case)

MCMC can be used to simulate $p(\vartheta \mid D)$ even in complex cases. [Robert and Casella, 1999, Clayton, 1996]

B4. CreditMetrics-style model

We fit a model to S&P default data (ratings classes A, BBB, BB, B and C) with a scalar random effect b_t in each year and no serial dependence.

$$Y_{ti} \,|\, b_t \sim \mathrm{Be}(p_{ti}(b_t)), \quad i = 1, \ldots, m_t, \quad t = 1, \ldots, n, \quad \text{where}$$
$$p_{ti}(b_t) = \Phi\left(\mu_{r(t,i)} + b_t\right), \quad b_t \sim N(0, \sigma^2),$$

where $r(t,i)$ gives rating of firm i in year t. Fixed effects $\boldsymbol{\beta} = (\mu_1, \ldots, \mu_k)'$, hyperparameter $\theta = \sigma$.

This model can be fitted by Gibbs sampling or by brute-force maximum likelihood (integrating out random effects numerically) [Frey and McNeil, 2003, McNeil and Wendin, 2003].

Results

From fitted model we infer estimates of default probabilities as well as within-group and between-group default correlations.

Parameter	A	BBB	BB	B	CCC
μ_r (mean)	-7.84	-6.13	-4.64	-2.94	-1.53
μ_r (median)	-7.79	-6.11	-4.62	-2.93	-1.52
s.e.(μ_r)	0.429	0.477	0.443	0.433	0.436
σ	(mean)	0.158		(median)	0.0633
s.e.(σ)	0.380				
π_r (mean)	0.0004	0.0022	0.0099	0.0542	0.2220
π_r (median)	0.0004	0.0022	0.0099	0.0536	0.1986
π_r (ML)	0.0004	0.0022	0.0098	0.0503	0.2066

π_r stands for implied estimate of default probability in rating group r based on fitting of a 5-group model to S&P data

Results II

Within and between group correlations

$\rho_Y^{(r,s)}$	A	BBB	BB	B	C
A	0.00022	0.00047	0.00103	0.00166	0.00256
BBB	0.00047	0.00103	0.00223	0.00361	0.00564
BB	0.00103	0.00223	0.00484	0.00791	0.01226
B	0.00166	0.00361	0.00791	0.01303	0.02048
C	0.00256	0.00564	0.01226	0.02048	0.03270

Implied estimates of within-group and between-group default correlations based on fitting of a 5-group model to S&P data

B5. Model with Economic Cycles

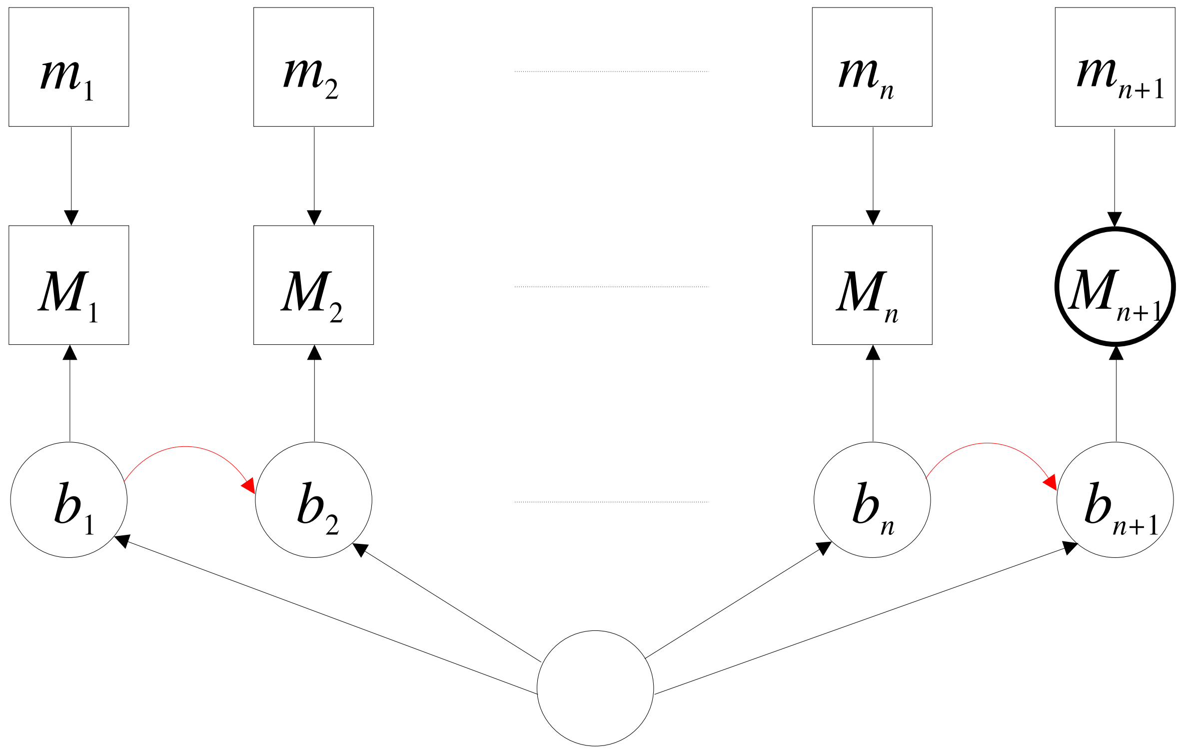
We give the random effects (b_t) an autoregressive structure

$$b_t = \alpha b_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma^2),$$

which introduces an additional hyperparameter α.

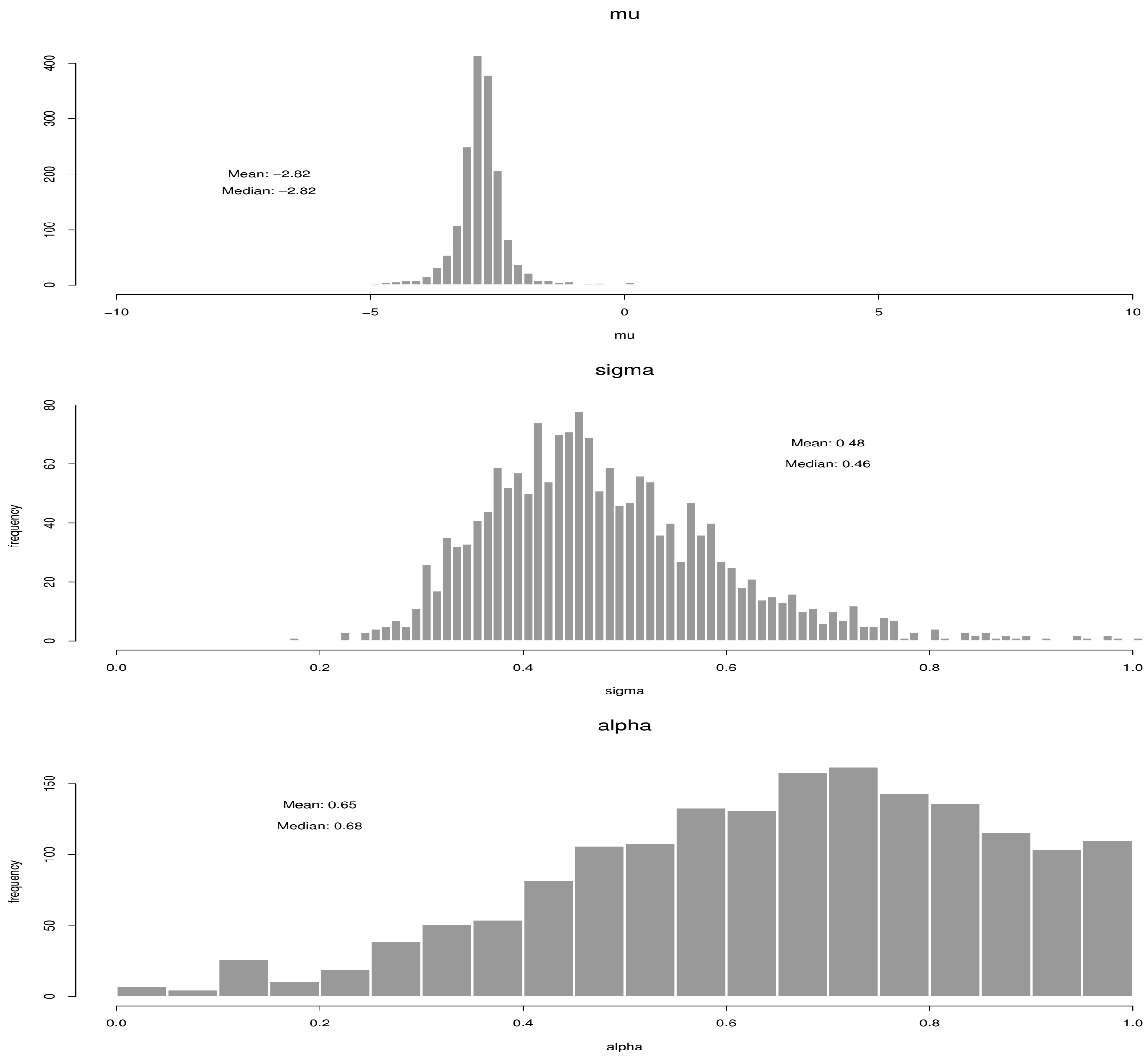
For simplicity consider a single homogeneous group and write $M_t = \sum_{i=1}^{m_t} Y_{ti}$ for the number of defaults in year t. We would like to use the model predictively to say something about M_{n+1}, the number of defaults in the next year period. [McNeil and Wendin, 2003]

DAG Representation of Model

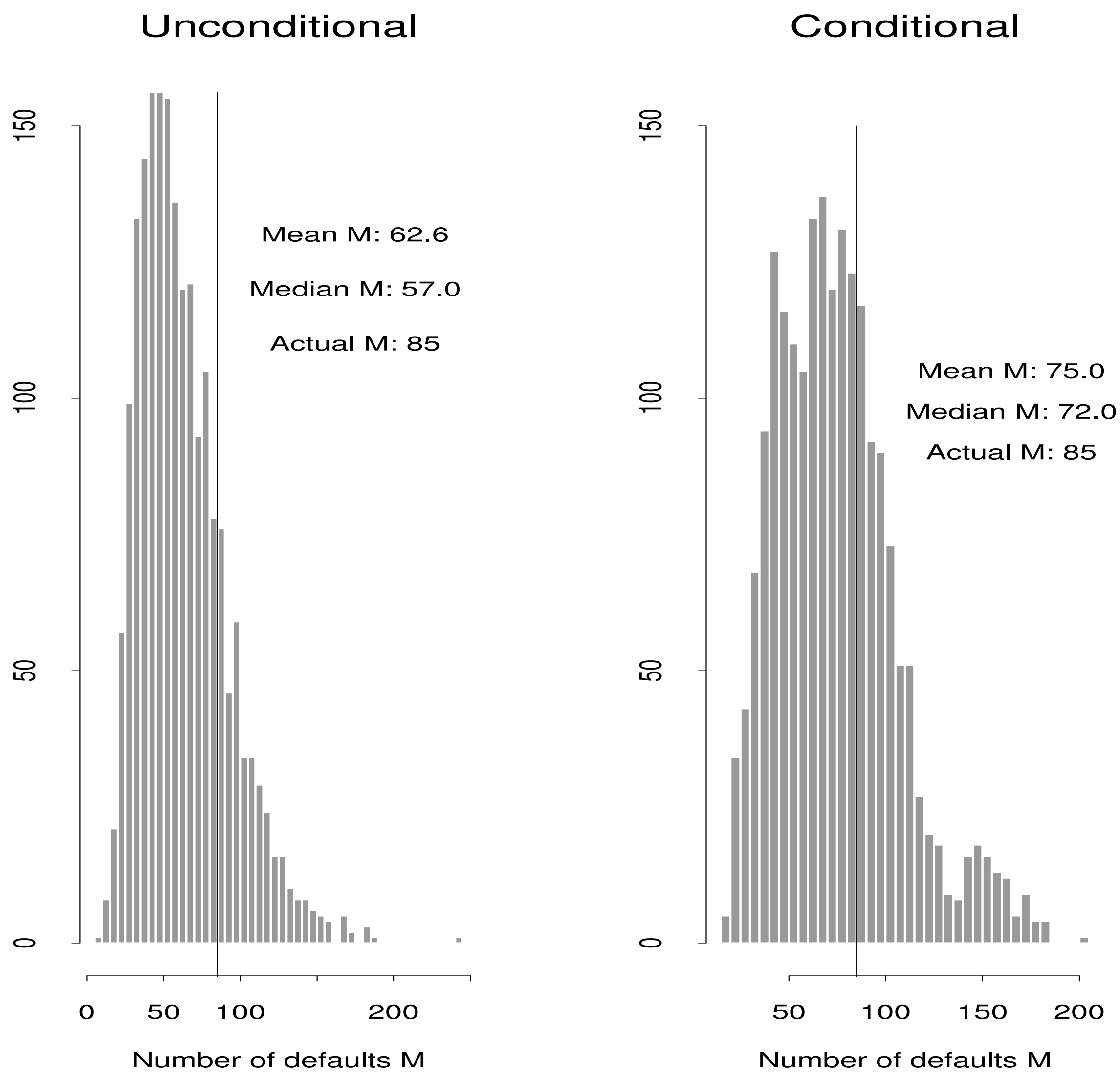


Here, the unknowns are $\vartheta = (\boldsymbol{\theta}, b_1, \ldots, b_n, b_{n+1}, M_{n+1})'$.
MCMC techniques simulate the posterior distribution of all these quantities, although only $\boldsymbol{\theta}$ and M_{n+1} are of prime interest.

Posterior distributions of μ, σ and α



Unconditional vs. Conditional distribution for M_{n+1}



C. Modelling Dependent Migrations

1. Migration Models in GLMM Framework

2. Example

C1. The GLMM Framework + Gibbs Sampling

The statistical framework we have chosen allows relatively complicated models, where for example:

- Random effects capture common economic effects on rating migrations for several firms;
- Individual covariates, including current and possibly previous ratings, modify the migration risk of each firm;
- Serially dependent random effects can capture economic cycles.

Migration Models in GLMM Framework

Let $S_1, \ldots, S_m$ be random variables represent the rating classes of m obligors during the "next time period". These take values on a scale of increasing credit quality $\{0, 1, \ldots, k\}$ where 0 represents default.

We can generalize the idea in (1). Conditional on $\mathbf{b}$ we assume that $S_1, \ldots, S_m$ are independent and multinomially distributed so that for $r \in \{0, 1, \ldots, k\}$

$$P\left(S_i = r \mid \mathbf{b}\right) = p_{ir}(\mathbf{b})$$

for some functions $p_{ir}(\mathbf{b})$ such that

$$p_{i0}(\mathbf{b}) + p_{i1}(\mathbf{b}) + \cdots + p_{ik}(\mathbf{b}) = 1.$$

Linking to covariates

The conditional migration probabilities are related to covariates and random effects by assuming that for $r \in \{0, 1, \ldots, k\}$

$$P(S_i \leq r \,|\, \mathbf{b}) = \sum_{j=0}^{r} p_{ij}(\mathbf{b}) = g\left(\mu_{rl(i)} + \mathbf{x}_i'\boldsymbol{\beta} + \mathbf{z}_i'\mathbf{b}\right),$$

where the μ_{rl} are threshold parameters for a firm with current rating l and satisfy $\mu_{0l} \leq \cdots \leq \mu_{kl}$ for all l. The $\mathbf{x}_i$ are additional covariates other than rating.

The probability that obligor i migrates to state r conditional on $\mathbf{b}$ is

$$p_{ir}(\mathbf{b}) = g\left(\mu_{rl(i)} + \mathbf{x}_i'\boldsymbol{\beta} + \mathbf{z}_i'\mathbf{b}\right) - g\left(\mu_{(r-1)l(i)} + \mathbf{x}_i'\boldsymbol{\beta} + \mathbf{z}_i'\mathbf{b}\right),$$

where $\mu_{-1} = -\infty$.

Interpretation as Asset Value Model

This formulation links very naturally to the structural model approach to credit risk. Suppose $g = \Phi$ and let V_i represent be a standard normally distributed variate representing standardized asset value of firm i. Then conditional on $\mathbf{b}$

$$\{S_i = r\} = \left\{V_i \in \left(\mu_{(r-1)l(i)} + \mathbf{x}_i'\boldsymbol{\beta} + \mathbf{z}_i'\mathbf{b}, \mu_{rl(i)} + \mathbf{x}_i'\boldsymbol{\beta} + \mathbf{z}_i'\mathbf{b}\right]\right\}.$$

The required asset range for a rating r for a company with current rating $l(i)$ is thus modified by additional covariates $\mathbf{x}_i$ and the state of the economy represented by $\mathbf{b}$.

Multi-Period Data

The data come in the form of m rating panels $(S_{it})_{t \in T_i}$ where T_i is a set of times for which ratings are available for firm i. (We may well have the problem that these differ from firm to firm and that default is an absorbing state.)

A multi-period model would be of the form

$$P(S_{it} \le r \,|\, \mathbf{b}_t) = g\left(\mu_{rl(i)} + \mathbf{x}'_{ti}\boldsymbol{\beta} + \mathbf{z}'_{ti}\mathbf{b}_t\right),$$

where the random effects $b_1, \ldots, b_n$ could be either iid or serially correlated as before.

Fitting in either case can be achieved by MCMC methods.

C2. Practical Example

We consider a simple model where

$$P(S_{it} \leq r \,|\, \mathbf{b}_t) = g\left(\mu_{rl(i)} + b_t\right),$$

for iid random effects satisfying $b_t \sim N(0, \sigma^2)$. Thus in this model the only covariate determining the transition probabilities is current rating.

We use the logit link function, i.e. the df of the logistic distribution $g(x) = (1 + e^{-x})^{-1}$. We have a matrix of thresholds to estimate as well as σ^2. As before we use Standard and Poor's yearly default and migration data. [Wendin and McNeil, 2004]

Parameter Estimates

μ_{rl} (l, r)	D	CCC	B	BB	BBB	A	AAA
AAA	-41.89 (7.71)	-29.98 (9.17)	-17.97 (7.69)	-8.76 (1.27)	-6.90 (0.53)	-5.29 (0.24)	-2.71 (0.09)
AA	-9.87 (1.31)	-8.21 (0.63)	-6.79 (0.32)	-6.33 (0.26)	-5.01 (0.15)	-2.50 (0.07)	5.06 (0.15)
A	-8.01 (0.42)	-7.73 (0.36)	-6.04 (0.17)	-4.92 (0.11)	-2.77 (0.07)	3.78 (0.08)	7.34 (0.33)
BBB	-6.11 (0.22)	-5.58 (0.17)	-4.43 (0.11)	-2.83 (0.07)	2.95 (0.07)	5.89 (0.20)	8.24 (0.62)
BB	-4.56 (0.13)	-3.86 (0.10)	-2.26 (0.07)	2.58 (0.07)	5.27 (0.18)	7.13 (0.43)	8.53 (0.81)
B	-2.80 (0.07)	-2.26 (0.07)	2.66 (0.07)	4.74 (0.14)	5.49 (0.19)	6.99 (0.39)	14.82 (3.08)
CCC	-1.13 (0.11)	1.85 (0.13)	3.43 (0.23)	4.42 (0.36)	5.70 (0.65)	7.38 (1.32)	8.38 (1.70)
σ	0.235 (0.045)						

Table containing parameter estimates and standard errors obtained by Gibbs sampling.

The Implied Migration Probabilities

$l \backslash r$	AAA	AA	A	BBB	BB	B	CCC	D
AAA	9.36 e−01	5.89 e−02	4.11 e−03	8.73 e−04	1.61 e−04	5.75 e−08	7.41 e−11	4.97 e−16
AA	6.45 e−03	9.16 e−01	7.07 e−02	4.97 e−03	6.64 e−04	8.79 e−04	2.25 e−04	5.32 e−05
A	6.66 e−04	2.22 e−02	9.17 e−01	5.29 e−02	5.02 e−03	1.97 e−03	1.23 e−04	3.42 e−04
BBB	2.72 e−04	2.57 e−03	4.78 e−02	8.92 e−01	4.51 e−02	8.23 e−03	1.60 e−03	2.27 e−03
BB	2.04 e−04	6.18 e−04	4.42 e−03	6.66 e−02	8.32 e−01	7.49 e−02	1.06 e−02	1.07 e−02
B	3.77 e−07	9.41 e−04	3.25 e−03	4.68 e−03	5.80 e−02	8.37 e−01	3.81 e−02	5.85 e−02
CCC	2.37 e−04	4.01 e−04	2.77 e−03	8.84 e−03	1.98 e−02	1.06 e−01	6.15 e−01	2.47 e−01
D	0	0	0	0	0	0	0	1

- Since the migrations of two companies in the same time period are not independent we can also calculate migration correlations, or upgrade and downgrade correlations between companies; this generalizes the concept of default correlation.

- The model may be used to estimate the distribution of ratings for a particular cohort of firms in the next time period, or the financial loss distribution associated with the change in rating composition.

References

[Clayton, 1996] Clayton, D. (1996). Generalized linear mixed models. In Gilks, W., Richardson, S., and Spiegelhalter, D., editors, *Markov Chain Monte Carlo in Practice*. Chapman & Hall, London.

[Crowder et al., 2003] Crowder, M., Davis, M., and Giampieri, G. (2003). A hidden Markov model of default interaction. Preprint, Imperial College, London.

[Frey and McNeil, 2003] Frey, R. and McNeil, A. (2003). Dependent defaults in models of portfolio credit risk. *J. Risk*, 6(1):59–92.

[Gagliardini and Gouriéroux, 2004] Gagliardini, P. and Gouriéroux, C. (2004). Stochastic migration models with application to corporate risk. Working paper, HEC Montreal.

[Hu et al., 2002] Hu, Y.-T., Kiesel, R., and Perraudin, W. (2002). The estimation of transition matrices for sovereign credit ratings. *J. Banking Finance*, 26:1383–1406.

[Lando, 2004] Lando, D. (2004). *Credit Risk Modeling: Theory and Applications*. Princeton University Press, Princeton, New Jersey.

[Lando and Skodeberg, 2002] Lando, D. and Skodeberg, T. (2002). Analyzing rating transitions and rating drift with continuous observations. *Journal of Banking and Finance*, 26:423–444.

[McNeil and Wendin, 2003] McNeil, A. and Wendin, J. (2003). Generalised linear mixed models in portfolio credit risk modelling. Working paper.

[Nickell et al., 2000] Nickell, P., Perraudin, W., and Varotto, S.

(2000). Stability of ratings transitions. *J. Banking Finance*, 24:203–227.

[Robert and Casella, 1999] Robert, C. and Casella, G. (1999). *Monte Carlo Statistical Methods*. Springer, New York.

[Schuerman and Jafry, 2003] Schuerman, T. and Jafry, Y. (2003). Measurement and estimation of credit migration matrices. Working paper, Federal Reserve Bank of New York.

[Wendin and McNeil, 2004] Wendin, J. and McNeil, A. (2004). Dependent credit migrations. Technical Report 182, NCCR FINRISK Working Paper Series.