Bayesian Methods in Portfolio Credit Risk Management

Jonathan Wendin

7 February 2006
Introduction

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** (e.g., 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 of default dependence

Standard and Poor’s default data from 1981 to 2000 show clear evidence of cycles; we find cross-sectional as well as serial dependence in default rates.
Panel data v. cross-sectional data

Let $\mathcal{K} := \{1, \ldots, K\}$ be a set of ratings of increasing creditworthiness and set $\mathcal{K}_0 := \{0\} \cup \mathcal{K}$, where state 0 is default. Let $\{0, 1, \ldots, T\}$ be the times (e.g., days, months, or years) at which ratings are assessed.

**Objective:** statistical models for (i) correlated defaults and transitions, (ii) time series of credit ratings.

**Panel data:** Consider a portfolio consisting of $n$ obligors. For each obligor $i \in \{1, \ldots, n\}$, our data are

$$\{R_{it} : 0 \leq t \leq T\} \quad \text{and} \quad \{x_{it}, z_{it} : 1 \leq t \leq T\},$$

where $R_{it} \in \mathcal{K}_0 \cup \{\text{N.R.}\}$ is the rating at time $t$ of obligor $i$, and $x_{it}$ and $z_{it}$ are corresponding covariates.
Cross-sectional data: Let $n_t$ be the number of obligors in the portfolio in period $t$. For each period $t \in \{1, \ldots, T\}$, we have

$$\{ R_{ti} : 1 \leq i \leq n_t \} \quad \text{and} \quad \{ \kappa(t, i), x_{ti}, z_{ti} : 1 \leq i \leq n_t \}.$$

$R_{ti}$: rating of the $i$th obligor at the end of period $t$,
$\kappa(t, i)$: rating at the onset of period $t$,
$x_{ti}$, $z_{ti}$: further covariates of the $i$th obligor.

The use of cross-sectional data—in particular, for homogeneous (sub)portfolios—is mainly motivated by the widespread assumption of Markov-type rating paths.

Notice that the original rating panels cannot be recovered from a dataset of repeated cross-sectional data.
Overview of the thesis

Part I: Techniques for cross-sectional data.

1. Single-period case: generalized linear mixed models (GLMMs) for describing correlated defaults and transitions.
2. Multi-period case: state space models (in a GLMM-framework) for autocorrelated transition events.
3. First application of Gibbs sampling for statistical inference.

Part II: Techniques for rating panel data.

1. Extensions and new parameterizations of models that go beyond the Markov chain. In particular, a first account of latent systematic risk factors in models of this kind.
2. Detection and quantification of deviations from the Markov chain.
Single-period case: GLMMs

Consider a latent factor (random effect) $b_t$ (vector/scalar). Given $b_t$, we assume that $R_{t1}, \ldots, R_{tn_t}$ are conditionally independent with

$$
P \{ R_{ti} \leq \ell | b_t \} = g \left( \mu_{k(t,i),\ell} - x_{ti}' \beta - z_{ti}' b_t \right), \quad \ell \in \mathcal{K}_0
$$

where

- $g : \mathbb{R} \rightarrow (0, 1)$ is a strictly increasing response function: e.g., $\Phi(x)$ (ordered probit) or $1/(1 + e^{-x})$ (ordered logit),

- $\kappa(t,i)$ is the rating of the $i$th obligor at the outset of period $t$,

- the intercepts $(\mu_k, \ell)$ and regression coefficients $\beta$ are unknown parameters satisfying $-\infty = \mu_{k,-1} \leq \mu_{k,0} \leq \cdots \leq \mu_{k,K} = \infty$. 


Single-period case: GLMMs (cont’d)

\[
P \{ R_{ti} \leq \ell \mid b_t \} = g \left( \mu_{\kappa(t,i),\ell} - x_{ti}'\beta - z_{ti}'b_t \right)
\]

- \( x_{ti} \) and \( z_{ti} \) are additional covariates other than rating,
- \( b_t \) is latent factor with df \( F_b \) (non-degen.) and hyperparameter \( \theta \).

This defines a generalized linear mixed model (GLMM) for the ordered, categorical responses \( R_{t1}, \ldots, R_{tn_t} \). We refer to \( x_{ti}'\beta + z_{ti}'b_t \) as the systematic risk of obligor \( i \).

Conditional transition probabilities:

\[
P \{ R_{ti} = \ell \mid b_t \} = P \{ R_{ti} \leq \ell \mid b_t \} - P \{ R_{ti} \leq \ell - 1 \mid b_t \}.
\]
Rating transitions as multinomial trials

Define the rating indicator $Y_{ti} := \left( \mathbb{I}\{R_{ti}=0\}, \ldots, \mathbb{I}\{R_{ti}=K\} \right)'$.

Given $b_t; Y_{t1}, \ldots, Y_{tn_t}$ are conditionally independent with

$$Y_{ti} | b_t \sim \text{Multinomial} \left\{ 1, p_{k(t,i)}(x'_{ti} \beta + z'_{ti} b_t) \right\},$$

where $p_k(z) = (p_{k,0}(z), \ldots, p_{k,K}(z))'$ with $p_{k,\ell}(z)$ being the probability of a transition $k \rightarrow \ell$, if the systematic risk is $z$.

Unconditionally, the rating indicators are dependent:

$$P \{ Y_{t1} = y_1, \ldots, Y_{tn_t} = y_{n_t} \} = \int \prod_{i=1}^{n_t} P \{ Y_{ti} = y_i | b_t \} \; dF_b(b_t).$$

Important special case: $Y_{ti} := \mathbb{I}\{R_{ti}=0\}$ corresponds to GLMMs for binary responses, i.e., models for default risk only.
Interpretation as an asset value model

Let $\varepsilon_{t1}, \ldots, \varepsilon_{tn_t}$ be iid rvs with df $g$ (independent of $b_t$). Given $b_t$, we define $V_{ti} := x'_{ti}\beta + z'_{ti}b_t + \varepsilon_{ti}$, $i = 1, \ldots, n_t$, and notice that

$$R_{ti} = \ell \iff V_{ti} \in \left(\mu_{\kappa(t,i),\ell-1}, \mu_{\kappa(t,i),\ell}\right].$$

**Interpretation:** $V_{ti}$ is the asset value and $(\mu_{\kappa(t,i),\ell})_\ell$ are critical liability levels. We refer to $\text{corr}(V_{ti}, V_{tj})$ as the implied asset correlation between obligors $i$ and $j$. 
Multi-period case: Dynamic credit risk mgmt

Serial dependence in transition probabilities can be achieved with autocorrelated latent factors $b_1, \ldots, b_T$.

Let $Y_t := \{Y_{ti} : 1 \leq i \leq n_t\}$ summarize the transitions taking place during period $t$. We assume that:

1. Given $(b_t)_{t=1}^T$, the indicators $\{(Y_{ti})_{i=1}^{n_t} : 1 \leq t \leq T\}$ are conditionally independent.
   Moreover, $Y_t$ depends on no latent factor other than $b_t$.

2. The latent factors $(b_t)_{t=1}^T$ form a Markov chain.

Remark: The above assumptions define a state space model (hidden Markov model) for the time series $\{Y_t : 1 \leq t \leq T\}$. 
The Markov property of rating panel data

If \((R_{it})_{t \in T}\) is a Markov chain, then the transitions in different time periods are independent.

**Important consequence:** If migrations in different time periods are correlated, then \((R_{it})_{t \in T}\) is not a Markov chain.

Hence, autocorrelation in the latent factors \((b_t)\) is a source of non-Markovian effects in rating panel data.
Model calibration: Likelihood inference

Let $\mathbf{b}_1, \ldots, \mathbf{b}_T$ have joint df $F_b(\mathbf{b}_1, \ldots, \mathbf{b}_T; \theta)$. The likelihood function given the observed data, $\{(\mathbf{y}_{ti})_{t=1}^{n_t} : 1 \leq t \leq T\}$, takes the form

$$L(\mu, \beta, \theta \mid \text{observed data}) = \int \cdots \int_{\mathbb{R}^{T \times p}} \prod_{t=1}^{T} \prod_{i=1}^{n_t} \mathbb{P}_{\mu, \beta} \{Y_{ti} = y_{ti} \mid \mathbf{b}_t\} \ dF_b(\mathbf{b}_1, \ldots, \mathbf{b}_T; \theta),$$

where $p = \text{dim}(\mathbf{b}_t)$.

To evaluate this expression we have an integral over $\mathbb{R}^{T \times p}$, which makes standard maximum likelihood difficult.
Model calibration: Bayesian inference

We distinguish between observed model quantities \( D := (Y_t^T)_{t=1}^T \) and unobserved ones \( \vartheta := (\mu, \beta, \theta, b_1, \ldots, b_T) \). The latter are treated as rvs.

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

Inference in our model is based on the posterior distribution:

\[
p(\vartheta | D) = \frac{p(D | \vartheta)p(\vartheta)}{p(D)} = \frac{p(D | \vartheta)p(\vartheta)}{\int p(D | \vartheta')p(\vartheta') \, d\vartheta'}.
\]

**Problem:** finding \( p(\vartheta | D) \! \).

**Possible solution:** Simulation from \( p(\vartheta | D) \) with Markov chain Monte Carlo (i.e., construction of a Markov chain that has \( p(\vartheta | D) \) as its stationary distribution).
Model calibration: Bayesian inference (cont’d)

**Notation:** Let \( \underline{x} := (x_1, \ldots, x_T)' \).

Joint distribution of data \( \underline{Y} \), latent factors \( \underline{b} \), hyperparameters \( \theta \), thresholds \( \mu \), and regression coefficients \( \beta \) reads

\[
p(\underline{Y}, \underline{b}, \theta, \mu, \beta) = \prod_{t=1}^{T} \prod_{i=1}^{n_t} p(Y_{ti} | b_t, \mu, \beta)p(b_t | \theta)p(\theta, \mu, \beta),
\]

where

\[
p(b | \theta) = p(b_1 | \theta) \prod_{t=2}^{T} p(b_t | b_{t-1}, \theta).
\]

The joint distribution factors into a product of conditional distributions.
Advantages of MCMC

• calibration of complex models with multivariate, serially correlated latent factors; implementation straightforward; simulation fast;

• point estimates, standard errors and (joint) confidence sets of primary parameters and derived model parameters (e.g., implied asset correlations) are easily extracted from the output;

• posterior path of latent factors \((b_t)\) can be compared with other macro-economic variables;

• prior information about parameters governing portfolio dependence can be entered in the analysis.
Part I: Empirical conclusions

- Residual, cyclical, latent component in the systematic risk; even after accounting for observed business cycle covariates.

- Implied asset correlation of obligors sharing industry sector is 10.9 %, whereas the across-sector counterpart is only 6.5 %. \(^1\) (Single-factor v. multi-factor framework, concentration risk.)

- Implied asset correlations do not appear to fall monotonely with increasing probability of default. (The relationship between “signal-to-noise ratio” and probability of default is not unambiguous).

---

\(^1\) A model without sector-specific latent factors yields an overall implied asset correlation of 7.5 % for the same dataset.
Part I: Empirical conclusions (cont’d)

- Accounting for serial correlation is important for prediction of failure rates.
Posterior path of the latent factors

• posterior mean of \((b_t)\) with 95% credible set.

• posterior mean of \((b_t)\) and CFNAI \((x_t/\beta)\).
Part II: Models for the rating path

The Markov chain is the industry standard:

\[
P \{ R_{t+1} = \ell \mid R_t = k, (R_{s-1} = r_{s-1})_{s=1}^t \} = \mathbf{P} \{ R_{t+1} = \ell \mid R_t = k \} =: u_{k,\ell}.
\]

Violations of the Markov property have been documented in the literature. Examples include:

1. “downgrade momentum”: recently downgraded obligors have a higher probability of further downgrades, or default, than other obligors in the same rating class.

2. The rare occurrence of “rating reversals”.


Extensions of the Markov chain

Consider a Markov chain \( \{S_t = (R_t, I_t) : t \geq 0\} \) evolving on \( \mathcal{K}_0 \times \{0, 1\} \), where \( I_t \) serves as a (possibly) unobserved risk flag. Set

\[
u_{k,\ell} := \mathbb{P} \{ R_{t+1} = \ell | R_t = k, I_t = 0 \}
\]

\[
u_{k,\ell} := \mathbb{P} \{ R_{t+1} = \ell | R_t = k, I_t = 1 \}.
\]

Clearly,

\[
\mathbb{P} \{ R_{t+1} = \ell | R_t = k, (R_s)_{s=1}^t \} = u_{k,\ell} \mathbb{P} \{ I_t = 0 | (R_s)_{s=0}^t \} + v_{k,\ell} \mathbb{P} \{ I_t = 1 | (R_s)_{s=0}^t \},
\]

which is a convex combination of \( u_{k,\ell} \) and \( v_{k,\ell} \). Thus, \( (R_t)_{t \geq 0} \) is not a Markov chain; unless, for instance, \( u_{k,\ell} = v_{k,\ell} \) for all \( k, \ell \in \mathcal{K}_0 \).

**Role of \( I_t \):** account for heterogeneity in exposure to migration risk.
Example I: Markov mixture models

Let us assume that

\[ P \{ S_{t+1} = (\ell, 1 - y) \,|\, S_t = (k, y) \} = 0 \quad \text{for} \quad y \in \{0, 1\}, \]

\[ P \{ I_0 = 1 \} = s \in [0, 1], \]

i.e., the process \((I_t)_{t \geq 0}\) is constant and unobserved.

The model is thus fully specified by \(U = (u_{k,\ell}), V = (v_{k,\ell}),\) and \(s.\)

This is the two-point Markov mixture model for \((R_t)_{t \geq 0}\).
Example II: HMM with partly observed risk flags

We assume that for $y \in \{0, 1\}$

\[
P \{ S_{t+1} = (\ell, 1) | S_t = (k, y) \} = 0 \quad \text{if } \ell > k \quad \text{(upgrade)},
\]

\[
P \{ S_{t+1} = (\ell, 0) | S_t = (k, y) \} = 0 \quad \text{if } \ell < k \quad \text{(downgrade)},
\]

\[
P \{ S_{t+1} = (k, 1) | S_t = (k, 0) \} = 0,
\]

\[
P \{ I_0 = 1 \} = s \in [0, 1].
\]

Observe that $I_t$ equals 1, if the obligor is downgraded during period $t$.

With this model, we are able to capture the effect of previous migrations on the transition probabilities.
Sources of non-Markovian effects in rating panels

Possible sources of non-Markovian effects in rating panel data include

1. autocorrelated systematic risk factors,

2. unobserved heterogeneity in the exposure to migration risk,

3. past migrations having a genuine impact on the probability of future migrations.

From the perspective of risk management, the measures to be taken are different in all three cases.

It is therefore necessary to control for 1. and 2. upon investigating 3.
Accounting for correlation effects in rating panels

Let $T_i \subseteq T$ be the times at which the rating of obligor $i$ is assessed.

The issue of correlation and autocorrelation can be addressed by introducing serially correlated, latent systematic risk factors $(b_t)$.

**Assumptions:**

**Part I:** given $(b_t)_{t=1}^T$, the rating paths $\{(R_{it})_{t \in T_i} : 1 \leq i \leq n\}$ are conditionally independent Markov chains.

**Part II:** given $(b_t)_{t=1}^T$, the state processes $\{(S_{it})_{t \in T_i} : 1 \leq i \leq n\}$ are conditionally independent Markov chains.

In both cases, the transitions taking place in period $t$ depend on no other latent factor than $b_t$, and the time series $(b_t)$ is given the dynamics of a Markov chain.
Part II: Empirical conclusions

- The hypothesis of Markovian rating paths is rejected—even after accounting for serially correlated systematic risk through \((b_t)\).

- There is substantial evidence of downgrade momentum—even after accounting for serially correlated systematic risk and unobserved heterogeneity among obligors (through obligor-specific random intercept shifts).

- The momentum effects are most pronounced for obligors that are downgraded into rating class BBB and below. The adverse persistence of a downgrade, assuming no intermediate change of rating, is then about one year.
Illustration of downgrade momentum

Consider two obligors 1 and 2, and assume that

- obligor 1 reaches state \( k \in \mathcal{K} \) at time 0 through a \textit{downgrade},
- obligor 2 reaches state \( k \in \mathcal{K} \) at time 0 through an \textit{upgrade}.

The additional risk of the downgraded obligor, and its evolution as a function of time-in-state, can be visualized with

\[
\nu(t, k) := \frac{\mathbb{P} \{ R_{1,t} = 0 | R_{1,t-1} = \cdots = R_{1,0} = k \}}{\mathbb{P} \{ R_{2,t} = 0 | R_{2,t-1} = \cdots = R_{2,0} = k \}}.
\]

Notice that a Markov chain \( (R_{it}) \) has \( \nu(t, k) \equiv 1 \).
Illustration of downgrade momentum (cont’d)

Plot of \( \{(t, \nu(t, k)) : 1 \leq t \leq 50\} \) with one-month time period for \( k \in \{\text{CCC}, \ldots, \text{BBB}\} \). Obtained with the HMM in Example II. (The triangle-edged line shows the one-sided 95\% credible set.)
Conclusions

Central to many models for portfolio credit risk is the assumption of exchangeability, which is liable to be violated unless care is taken.

Sources of heterogeneity include

- industry sector, country, or geographical region,
- time-varying economic conditions,
- recent migrations experienced.

This thesis provides a unified statistical framework for joint consideration of these effects. Moreover, its main source of information for inference is historical rating transition data.