Extracting Systematic Factors in a Continuous-time Credit Migration Model*

Harley Thompson[†] Jonathan Harris[‡]

July 2008

Abstract

We present a methodology for estimating up and down-jump intensities in a portfolio credit ratings migration model. Emphasis is given to two particular issues that arise in practice: incomplete transitions data and time-inhomogeneity in the transition intensities. Our approach is flexible and computationally fast and should prove useful in a practical context. The model is fitted to a portfolio of S&P rated corporates.

*The views expressed in this paper are those of the authors and do not necessarily reflect those of the Commonwealth Bank of Australia.

[†]Manager, Quantitative Analytics, Commonwealth Bank of Australia, 48 Martin Place, Sydney. Email: harley.thompson@cba.com.au (corresponding author)

[‡]Associate, Quantitative Analytics, Commonwealth Bank of Australia, 48 Martin Place, Sydney. Email: jjharris@stanford.edu

1 Introduction and notation

In this paper we consider the estimation of systematic factors in a continuous–time Markov chain model of portfolio credit risk. Markov chains are a standard tool in credit risk modelling, see [JLT] among others. The portfolio in question may be a portfolio of corporate bonds rated by a ratings agency or a portfolio of internally–rated bank loans. The presence of systematic effects implies the Markov chain is time–inhomogeneous. The estimation of parameters of a time–inhomogeneous chain from discrete data is, in its full generality, a difficult problem. To make the problem tractable we therefore fix the inhomogeneity to take a piecewise constant form. This simplifies parameter estimation yet leaves enough flexibility to make the model useful in practical applications, such as portfolio stress testing and economic capital determination.

Our estimation approach makes use of the methodology of Bladt and Sorensen (see [BS1] and [BS2]). These authors develop estimation methods for multivariate, finite–state, homogeneous Markov chains observed at discrete times, where the observations consist of credit ratings of individual obligors (names). Our methodology provides simple extensions to these estimators by allowing the transition intensities to vary through time and by allowing for transition data to be incomplete. Both of these features appear to be important in practice.

The paper is organised as follows. Section 2 presents the relevant Markov chain theory and derives a portfolio credit risk model. Section 3 presents the estimation approach and discusses how our methodology can accommodate missing observations and time heterogeneity. Section 4 presents our estimation results and section 5 concludes.

Notation

$X_t^{(k)}$	Credit quality of obligor k at time t
$\mathcal{L}(t) = [\lambda_{ij}(t)]_{i,j\in S}$	Generator matrix
$C(t) = [c_{ij}(t)]_{i,j\in S}$	Systematic factors (parameter matrix)
$\widetilde{\mathcal{L}} = [\widetilde{\lambda}_{ij}]_{i,j\in S}$	Baseline transition intensities (parameter matrix)
$R_i(t_z, t_{z+1})$	Time spent in rating i on $[t_z, t_{z+1})$ (continuous data)
$N_{ij}(t_z, t_{z+1})$	Total number of transitions $i \to j$ on $[t_z, t_{z+1})$ (continuous)
$\widetilde{M}^i_{x(k,t_z),x(k,t_{z+1})}$	Expected time spent in i by k with ratings $\{x_{t_z}, x_{t_{z+1}}\}$
$\widetilde{f}^{ij}_{x(k,t_z),x(k,t_{z+1})}$	Expected no. transitions $i \to j$ by k with ratings $\{x_{t_z}, x_{t_{z+1}}\}$
$\widetilde{\mathbf{N}}(t_z, t_{z+1}) = \{\widetilde{N}_{ij}(t_z, t_{z+1})\}$	Observed transition counts on t_z, t_{z+1} (discrete)
$\widetilde{\mathbf{N}}^*(t_z, t_{z+1}) = \{\mathbb{E}(\widetilde{N}^*_{ij}(t_z, t_{z+1}))\}$	Missing transition counts on t_z, t_{z+1} (discrete)
$\widetilde{N}_i(t_z, t_D)$	No. transitions from i at t_z to non–default at t_D (discrete)
$\widetilde{N}_{ij}(t_r, t_s)$	No. observed transitions $i \to j$ between t_r and t_s, $s \geqslant r+2$ (discrete)

2 Dynamic portfolio credit quality model

We present a Markov chain model in which credit ratings are assumed to be observed state variables evolving on a discrete state space. Of primary interest is the determination of transition probabilities and their dependence on time. We provide a summary of the Markov chain theory relevant to our model and discuss how we propose to introduce time inhomogeneity in the chain.

2.1 Markov chain

We consider a portfolio of K names (obligors) and denote the credit quality of the k^{th} name $\{X_t^{(k)} : t \in \mathbb{R}^+\}$, $k = \{1, ..., K\}$. Let $\mathbf{X}_t = \{\ X_t^{(1)} \quad ... \quad X_t^{(K)}\ \}_{t \geqslant 0}$ be the portfolio *credit quality process* and let $S = \{1, .., J\}$, $J \in \mathbb{Z}^+$ be the credit quality *state space*. Hence at any time t there are K obligors in the portfolio distributed across J credit quality groups, where $J << K$ in general. Set $X = 1$ to be the highest credit quality and $X = J$ to be the lowest credit quality (default). The default state is assumed to be absorbing. We assume $(X_t^{(k)})_{t \geqslant 0}$ is a Markov chain with generator matrix $\mathcal{L}(t) = [\lambda_{ij}(t)]_{i,j \in S}$ and transition probability matrix $P(t, t + \Delta) = [p_{ij}(t, t + \Delta)]_{i,j \in S}$, $\Delta \geqslant 0$. Since the transition intensities $\lambda_{ij}(t)$ depend on time the chain is *time–inhomogeneous*. Assume Δ is the observation interval of credit ratings, which from here on we will consider to be one quarter. The generator matrices along with the initial probability vector $\mathbb{P}\{X_0 = i\}$ completely characterise the chain. The transition intensities satisfy

$$\begin{aligned} \lambda_{ij}(t) &\geqslant 0, \qquad j \neq i \\ \lambda_{ii}(t) &= -\sum_{j \neq i} \lambda_{ij}(t) \end{aligned}$$

and a generator matrix satisfying these conditions is called a *valid generator matrix*. The intensities are in general unknown and represent parameters of the chain. Since the default state is absorbing, the intensity of transition from default to any non–default state is zero. Hence we also have $\lambda_{J,j}(t) = 0$ for all $j \in S$. By assumption the transition intensities of obligors with the same credit rating are equal.

In the general (time–inhomogeneous) case the transition probability matrix and the generator matrix satisfy the relation

$$(1) \qquad P(t, t + \Delta) = \mathbf{I} + \sum_{n=1}^{\infty} \int_t^{t+\Delta} \int_{s_1}^{t+\Delta} \cdots \int_{s_{n-1}}^{t+\Delta} \mathcal{L}(s_1) \dots \mathcal{L}(s_n)\ ds_n \dots ds_2 ds_1$$

on $[t, t + \Delta]$ where $\mathbf{I}$ is the identity matrix (see [BR], p. 328). In the special case $\mathcal{L}(s) = \mathcal{L}(t)\mathbf{1}\{s \in [t, t + \Delta)\}$ the chain is said to be *time–homogeneous* on $[t, t + \Delta)$ and (1) collapses to

$$P(t, t + \Delta) = \exp\left(\Delta \mathcal{L}(t)\right) = \sum\nolimits_{n=0}^{\infty} \frac{(\Delta \mathcal{L}(t))^n}{n!}.$$

This latter formula is referred to as the *matrix exponential* and is useful in calculations.

2.2 Time inhomogeneity

To model time inhomogeneity we consider a representation of the form

$$(2) \qquad \begin{aligned} \lambda_{ij}(t) &= \widetilde{\lambda}_{ij} c_{ij}(t), \quad t > 0 \\ c_{ij}(t) &= c_{ij}(t_z)\mathbf{1}\{t \in [t_z, t_{z+1})\}, \quad z = 0, ..., T - 1 \end{aligned}$$

where $t_z = \{z\Delta : z = 0, 1, ..., T-1\}$ and $\{\widetilde{\lambda}_{ij}, c_{ij}(t_z)\}$ represent parameters to be estimated. This can be seen as a special case of (1). We refer to $\widetilde{\lambda}_{ij} \geqslant 0$ as the *baseline transition intensity* and $c_{ij}(t) \geqslant 0$ as the *time inhomogeneity* (or *systematic*) *factor*. The matrix representation is $\mathcal{L}(t) = \widetilde{\mathcal{L}} \cdot C(t)$ where $\cdot$ denotes the Schur (or Hadamard) product. The usefulness of this decomposition lies in the fact that it enables an interpretation of transition intensities in terms of long-run ($\widetilde{\lambda}_{ij}$) and short–run (c_{ij}) components.[1] It also allows multi–period transition probability matrices to be expressed as the product of the matrix exponential of the generator matrix in each relevant time interval,

$$P(t_z, t_{z+r}) = \prod_{s=z}^{z+r-1} P(t_s, t_s + \Delta) = \prod_{s=z}^{z+r-1} \exp\left(\Delta \mathcal{L}(t_s)\right), \quad r \in \mathbb{Z}^+,$$

which is useful in practice.

3 Estimation

To estimate the parameters we draw upon and adapt the methodology of [BS1] and [BS2]. In general, the fact that the continuous chain is observed discretely causes complications in the estimation, in particular the maximum likelihood estimator may not exist and if it does exist it may not be unique. Furthermore even when existence and uniqueness are verified, maximisation of the likelihood requires an iterative solution due to the discrete observations and the possibility of missing ratings.[2] Thus we build our estimator in a series of steps, considering first the case where the set of ratings is complete.

3.1 Basic estimator

For the basic estimator we have from (2)

$$\lambda_{ij}(t) = \widetilde{\lambda}_{ij} c_{ij}(t_z) \mathbf{1}\{t \in [t_z, t_{z+1})\}, \quad z = 0, ..., T-1. \tag{3}$$

Furthermore we assume that state inhomogeneity depends on j according only to whether j is less than or greater than i,

$$c_{ij}(t_z) = \begin{cases} c^u(t_z) & : j < i \\ c^d(t_z) & : j > i \end{cases} \tag{4}$$

where c^u represents the up jump intensity and c^d the down jump intensity.[3] Let $\theta = \{\widetilde{\lambda}_{ij}, c^u(t_z), c^d(t_z)\}$ be the set of model parameters. For the continuous time–

[1] Note that in the estimator discussed in the next section it is not strictly necessary to model inhomogeneity in this way; the model could be estimated equivalently on a sequence of independent one-period intervals with restrictions placed directly on the admissible class of generator matrices. However we adopt the conventions discussed here for ease of interpretation.

[2] Incomplete data is not likely to arise in commercially available databases such as S&P CreditPro however it is likely to occur in datasets which use an internal ratings–based approach, such as those maintained by retail banks.

[3] A more general estimator is of course possible. Note that an identification condition, such as $c_{ij}(t_0) = 1$, is also required.

inhomogeneous Markov chain defined by (3) the *continuous likelihood*[4] is

$$\begin{aligned} L(\theta) &= \prod_{z=0}^{T-1}\prod_{i=1}^{J-1}\prod_{j\neq i}\lambda_{ij}(t_z)^{N_{ij}(t_z,t_{z+1})}\exp(-\lambda_{ij}(t_z)R_i(t_z,t_{z+1})) \\ &= \prod_{z=0}^{T-1}\prod_{i=1}^{J-1}\prod_{j\neq i}\left[\widetilde{\lambda}_{ij}c_{ij}(t_z)\right]^{N_{ij}(t_z,t_{z+1})}\exp(-\widetilde{\lambda}_{ij}c_{ij}(t_z)R_i(t_z,t_{z+1})) \end{aligned}$$

where $N_{ij}(t_z,t_{z+1})$ is the total number of transitions $i\to j$ by all obligors on $[t_z,t_{z+1})$ and

$$R_i(t_z,t_{z+1}) = \sum_{k=1}^{K}\int_{t_z}^{t_{z+1}}\mathbf{1}\{X_u^{(k)}=i\}du$$

is the total time spent in rating i by all obligors on the same period. Taking logs and substituting (4) gives

$$\ln L(\theta) = \sum_{z=0}^{T-1}\sum_{i=1}^{J-1}\Bigg[\sum_{j<i}\left[N_{ij}(t_z,t_{z+1})\ln\left(c^u(t_z)\right) - c^u(t_z)R_i(t_z,t_{z+1})\widetilde{\lambda}_{ij} + N_{ij}(t_z,t_{z+1})\ln(\widetilde{\lambda}_{ij})\right]$$
$$+\sum_{j>i}\left[N_{ij}(t_z,t_{z+1})\ln(c^d(t_z)) - c^d(t_z)R_i(t_z,t_{z+1})\widetilde{\lambda}_{ij} + N_{ij}(t_z,t_{z+1})\ln(\widetilde{\lambda}_{ij})\right]\Bigg]$$

and maximising obtains the three conditional MLEs

$$\widetilde{\lambda}_{ij}\,\Big|\left\{c^u(t_z),c^d(t_z)\right\}_{z=0}^{T-1} = \frac{\sum_{z=0}^{T-1}N_{ij}(t_z,t_{z+1})}{\sum_{z=0}^{T-1}c_{ij}(t_z)R_i(t_z,t_{z+1})},\quad i\neq J,\ j\neq i \tag{5}$$

$$c^u(t_z)\,\Big|\widetilde{\lambda}_{ij} = \frac{\sum_{i=1}^{J-1}\sum_{j<i}N_{ij}(t_z,t_{z+1})}{\sum_{i=1}^{J-1}\sum_{j<i}R_i(t_z,t_{z+1})\widetilde{\lambda}_{ij}},\quad z=0,...,T-1 \tag{6}$$

$$c^d(t_z)\,\Big|\widetilde{\lambda}_{ij} = \frac{\sum_{i=1}^{J-1}\sum_{j>i}N_{ij}(t_z,t_{z+1})}{\sum_{i=1}^{J-1}\sum_{j>i}R_i(t_z,t_{z+1})\widetilde{\lambda}_{ij}},\quad z=0,...,T-1. \tag{7}$$

While these estimators are unproblematic a difficulty arises in that they depend on $N_{ij}(t_z,t_{z+1})$ and $R_i(t_z,t_{z+1})$, which are unobservable since ratings are measured only at the discrete times t_z. To address this [BS1] suggest the use of an EM algorithm, in which the continuous data are replaced by their conditional expectations and the likelihood is maximised iteratively. Below we provide a slight generalisation of the algorithm in [BS2] to the case where the transition intensity is given by (3)–(4).

Let $x(k,t_z)$ denote the rating of name k at time t_z. Then to obtain the conditional expectations of the continuous data on $[t_z,t_{z+1})$:

(i) Set $b\geqslant\max_{i\in S}\left(-\lambda_{ii}(t_z)\right)$ and $\mathbf{B}=\mathbf{I}+b^{-1}\mathcal{L}(t_z)$ and calculate for all $i\in S\backslash\{J\}$,

[4] The likelihood assuming the chain is observed continuously in time.

$j \in S\backslash\{i\}$ and k

$$\begin{aligned}
\mathbf{M}^{i}(t_z) &= \exp\left(-b\Delta\right) b^{-1} \sum_{u=0}^{\infty} \frac{(\Delta b)^{u+1}}{(u+1)!} \sum_{s=0}^{u} \mathbf{B}^{s}(\mathbf{e}_i \mathbf{e}_i') \mathbf{B}^{u-s} \\
\widetilde{M}^{i}_{x(k,t_z),x(k,t_{z+1})}(t_z) &= \frac{M^{i}_{x(k,t_z),x(k,t_{z+1})}(t_z)}{\mathbf{e}'_{x(k,t_z)} \exp(\Delta\widetilde{\mathcal{L}} \cdot C(t_z))\mathbf{e}_{x(k,t_{z+1})}} \\
\mathbf{f}^{ij}(t_z) &= \lambda_{ij}(t_z) \exp\left(-b\Delta\right) b^{-1} \sum_{u=0}^{\infty} \frac{(\Delta b)^{u+1}}{(u+1)!} \sum_{s=0}^{u} \mathbf{B}^{s}(\mathbf{e}_i \mathbf{e}_j') \mathbf{B}^{u-s} \\
\widetilde{f}^{ij}_{x(k,t_z),x(k,t_{z+1})}(t_z) &= \frac{f^{ij}_{x(k,t_z),x(k,t_{z+1})}(t_z)}{\mathbf{e}'_{x(k,t_z)} \exp(\Delta\widetilde{\mathcal{L}} \cdot C(t_z))\mathbf{e}_{x(k,t_{z+1})}}
\end{aligned}$$

where $M^{i}_{x(k,t_z),x(k,t_{z+1})}$ is the $(x(k,t_z), x(k,t_{z+1}))^{th}$ element of $\mathbf{M}^i$ and $f^{ij}_{x(k,t_z),x(k,t_{z+1})}$ is the equivalent element of $\mathbf{f}^{ij}$.

(ii) Calculate for all relevant i, j the conditional expectations

$$\mathbb{E}\left(N_{ij}(t_z, t_{z+1}) \middle| \widetilde{\mathbf{N}}(t_z, t_{z+1})\right) = \sum_{k=1}^{K} \widetilde{f}^{ij}_{x(k,t_z),x(k,t_{z+1})}(t_z) \tag{8}$$

$$\mathbb{E}\left(R_{i}(t_z, t_{z+1}) \middle| \widetilde{\mathbf{N}}(t_z, t_{z+1})\right) = \sum_{k=1}^{K} \widetilde{M}^{i}_{x(k,t_z),x(k,t_{z+1})}(t_z) \tag{9}$$

where $\widetilde{\mathbf{N}}(t_z, t_{z+1}) = \{\widetilde{N}_{ij}(t_z, t_{z+1}) : i, j \in S\}$ is the matrix of observed transition counts for all (i, j) pairs between t_z and t_{z+1}.

Note the distinction between $\widetilde{N}_{ij}(t_z, t_{z+1})$ and $N_{ij}(t_z, t_{z+1})$: the former is the number of transitions *actually observed* between discrete times t_z and t_{z+1} while the latter is the number of unobserved continuous–time transitions on the interval $[t_z, t_{z+1})$. Clearly $\mathbb{E}\left(N_{ij}(t_z, t_{z+1})\right) \geqslant \widetilde{N}_{ij}(t_z, t_{z+1})$ since if a given number of obligors are observed to transition $i \to j$ then the actual number of transitions on the same interval, which also includes transitions not observed between t_z and t_{z+1}, must be at least this great.

With the above expectations in place an EM algorithm may be employed to maximise the likelihood:

Set $iter = 0$ and initialise the model parameters $\theta^{(0)}$ at an arbitrary interior value, then:

(i) [**E step**]: Set $iter = iter + 1$ and for all $z \in \{0, ..., T-1\}$, $i \in S\backslash\{J\}$ and $j \in S\backslash\{i\}$ compute

$$\begin{aligned}
&\mathbb{E}\left(N_{ij}(t_z, t_{z+1}) \middle| \theta^{(iter-1)}, \widetilde{\mathbf{N}}\left(t_z, t_{z+1}\right)\right) \\
&\mathbb{E}\left(R_{ij}(t_z, t_{z+1}) \middle| \theta^{(iter-1)}, \widetilde{\mathbf{N}}\left(t_z, t_{z+1}\right)\right)
\end{aligned}$$

using (8)–(9).

(ii) [**M step**]: For all $z \in \{0, ..., T-1\}$, $i \in S\backslash\{J\}$ and $j \in S\backslash\{i\}$ compute (5)–(7) with $N_{ij}(t_z, t_{z+1})$ and $R_i(t_z, t_{z+1})$ replaced by their expectations calculated in (i); update $\theta^{(iter)}$.

(iii) If $\left\|\theta^{(iter)} - \theta^{(iter-1)}\right\|$ is less than tolerance set $\widehat{\theta} = \theta^{(iter)}$ and exit, otherwise return to (i).

3.2 Missing ratings estimator

We now consider an extension of the above methodology to enable missing ratings to be accommodated. Let $\widetilde{\mathbf{N}}^*(t_z, t_{z+1})$ be the counterpart to $\widetilde{\mathbf{N}}(t_z, t_{z+1})$ comprising the transitions that were *not* recorded in the period t_z, t_{z+1} (the missing ratings). $\widetilde{\mathbf{N}}^*(t_z, t_{z+1})$ may exist for two reasons:

(a) There are obligors in the dataset whose date of departure t_D occurs after the date of their final rating.[5] It is known in such cases only that the obligor does not default at the date of departure and that they were still active clients at the departure date.

(b) There may be obligors that are rated at some time before t_z but not rated again until after t_{z+1}.[6] Thus a transition may have occurred at times t_z, t_{z+1} but is not recorded because the obligor was not reviewed during the period.

Clearly $\widetilde{\mathbf{N}}^*(t_z, t_{z+1})$ cannot be observed directly but its conditional expectation may be computed. To compute (a), let $\widetilde{N}_i(t_z, t_D)$, $t_D > t_z$ be the number of obligors that were assigned the rating i at time t_z but were not subsequently re–rated and left the books at the end of quarter t_D.[7] The model can only accept transitions with both an initial and terminal rating (i.e., $\widetilde{N}_{ij}(t_z, t_D)$) as input. Thus to incorporate this information one must apportion $\widetilde{N}_i(t_z, t_D)$ across $\{\widetilde{N}_{ij}(t_z, t_D) : j \neq J\}$ on the basis of the expected number of transitions to each non–default j at t_D. That is

$$\mathbb{E}\left(\widetilde{N}_{ij}(t_z, t_D)\,\middle|\,\widetilde{N}_i(t_z, t_D)\right) = p_{ij}(t_z, t_D\,|j \neq J\,)\widetilde{N}_i(t_z, t_D), \quad j \neq J \tag{10}$$

where

$$p_{ij}(t_z, t_D\,|j \neq J\,) \triangleq \mathbb{P}\{X_{t_D} = j\,|X_{t_z} = i, X_{t_D} \neq J\,\} = \frac{p_{ij}(t_z, t_D)}{1 - p_{i,J}(t_z, t_D)}.$$

To compute (b) we take the sum over all initial and terminal states and periods:

$$\begin{aligned} \mathbb{E}\left(\widetilde{N}^*_{ij}(t_z, t_{z+1})\right) &= \sum_{r=0}^{z} \sum_{s=(r+2)\vee(z+1)}^{T} \sum_{i_1=1}^{J-1} \sum_{j_1 \in S} \mathbb{E}\left(\widetilde{N}^*_{ij}(t_z, t_{z+1})\,\middle|\,\widetilde{N}_{i_1,,j_1}(t_r, t_s)\right) \\ &= \sum_{r=0}^{z} \sum_{s=(r+2)\vee(z+1)}^{T} \sum_{i_1=1}^{J-1} \sum_{j_1 \in S} p_{ij}(t_z, t_{z+1}\,|i_1,, j_1\,)\widetilde{N}_{i_1,,j_1}(t_r, t_s) \end{aligned} \tag{11}$$

for all $i \neq J$ and $j \in \{1, ..., J-1\}$ where $\widetilde{N}_{i_1,,j_1}(t_r, t_s)$ is observed and

$$p_{ij}(t_z, t_{z+1}\,|i_1,, j_1\,) = \frac{p_{i_1,i}(t_r, t_z) p_{ij}(t_z, t_{z+1}) p_{j,j_1}(t_{z+1}, t_s)}{p_{i_1,,j_1}(t_r, t_s)} \tag{12}$$

is the probability of observing the rating pair (i, j) on t_z, t_{z+1} conditional on observing a transition from i_1 at t_r to j_1 at t_s. Note that the transition probabilities calculated in this step are conditional on survival in each period. The proof of (12) is as follows: if θ is available then for all relevant t_z the $\Delta-$ period transition probabilities of the chain may

[5] These observations can be expected to be especially numerous in the final few quarters of a dataset because obligors rated in very recent quarters will not have had time to be re–rated.

[6] That is, there is a gap between the obligor's rating dates of two or more quarters and the interval straddles the period $[t_z, t_{z+1}]$.

[7] Note that $\widetilde{N}_i(t_z, t_D)$ can be distinguished from $\widetilde{N}_{ij}(t_z, t_D)$ by the lack of the second subscript j.

be computed by exponentiation of $\mathcal{L}(t_z)$. Hence on homogeneous intervals the probability of any given path can be determined,

$$\begin{aligned}
p_{ij}(t_z, t_{z+1} \,|i_{1,}, j_1) &\triangleq \mathbb{P}\left\{X_{t_z} = i, X_{t_{z+1}} = j \,|X_{t_r} = i_1, X_{t_s} = j_1\right\} \\
&= \frac{\mathbb{P}\left\{X_{t_s} = j_1, X_{t_{z+1}} = j, X_{t_z} = i \,|X_{t_r} = i_1\right\}}{\mathbb{P}\left\{X_{t_s} = j_1 \,|X_{t_r} = i_1\right\}} \\
&= \frac{\mathbb{P}\left\{X_{t_z} = i \,|X_{t_r} = i_1\right\} \mathbb{P}\left\{X_{t_{z+1}} = j \,|X_{t_z} = i\right\} \mathbb{P}\left\{X_{t_s} = j_1 \left|X_{t_{z+1}} = j\right.\right\}}{\mathbb{P}\left\{X_{t_s} = j_1 \,|X_{t_r} = i_1\right\}} \\
&= \frac{p_{i_1,i}(t_r, t_z) p_{ij}(t_z, t_{z+1}) p_{j,j_1}(t_{z+1}, t_s)}{p_{i_{1,},j_1}(t_r, t_s)},
\end{aligned}$$

where each term on the RHS can be computed explicitly.

It is important to note that (a) must be computed before (b). The reason is that in (b) it is assumed that both an initial and a terminal rating exist but in (a) no terminal rating is available, therefore it must first be calculated. The counterparts to (8)–(9) are

$$\mathbb{E}\left(N_{ij}(t_z, t_{z+1}) \left|\widetilde{\mathbf{N}}(t_z, t_{z+1}), \widetilde{\mathbf{N}}^*(t_z, t_{z+1})\right.\right) = \sum\nolimits_{k=1}^{K} \widetilde{f}^{ij}_{x(k,t_z),x(k,t_{z+1})}(t_z) \tag{13}$$

$$\mathbb{E}\left(R_i(t_z, t_{z+1}) \left|\widetilde{\mathbf{N}}(t_z, t_{z+1}), \widetilde{\mathbf{N}}^*(t_z, t_{z+1})\right.\right) = \sum\nolimits_{k=1}^{K} \widetilde{M}^{i}_{x(k,t_z),x(k,t_{z+1})}(t_z) \tag{14}$$

where $\widetilde{\mathbf{N}}^*(t_z, t_{z+1}) = \{\mathbb{E}(\widetilde{N}^*_{ij}(t_z, t_{z+1})) : i, j \in S\}$ and $\widetilde{f}^{ij}_{x(k,t_z),x(k,t_{z+1})}$ and $\widetilde{M}^{i}_{x(k,t_z),x(k,t_{z+1})}$ are as previous but with both observed and expected numbers of missing transitions taken into account.

With the above expectations in place an EM algorithm can again be employed to maximise the likelihood. The procedure is similar to the basic estimator except there is now an extra expectation step:

Set $iter = 0$ and initialise the model parameters $\theta^{(0)}$ at arbitrary interior values, then:

(i) **[E step]**: Set $iter = iter + 1$ and for all $z \in \{0, ..., T-1\}$ compute

$$\widetilde{\mathbf{N}}^*(t_z, t_{z+1}) = \left\{\mathbb{E}\left(\widetilde{N}^*_{ij}(t_z, t_{z+1}) \left|\theta^{(iter)}, \left\{\widetilde{\mathbf{N}}\left(t_r, t_s\right)\right\}^{z,T}_{r=0,s=z+1}\right.\right) : i \in S\backslash\{J\}, \quad j \in S\backslash\{i\}\right\}$$

from (10)–(12).

(ii) **[E step]**: For all $z \in \{0, ..., T-1\}$ and $i \in S\backslash\{J\}$ and $j \in S\backslash\{i\}$ compute

$$\mathbb{E}\left(N_{ij}(t_z, t_{z+1}) \left|\theta^{(iter-1)}, \widetilde{\mathbf{N}}\left(t_z, t_{z+1}\right), \widetilde{\mathbf{N}}^*(t_z, t_{z+1})\right.\right)$$
$$\mathbb{E}\left(R_{ij}(t_z, t_{z+1}) \left|\theta^{(iter-1)}, \widetilde{\mathbf{N}}\left(t_z, t_{z+1}\right), \widetilde{\mathbf{N}}^*(t_z, t_{z+1})\right.\right)$$

using (13)–(14).

(iii) **[M step]**: For all $z \in \{0, ..., T-1\}$ compute (5)–(7) with $N_{ij}(t_z, t_{z+1})$ and $R_i(t_z, t_{z+1})$ replaced with their expectations calculated in (ii); update $\theta^{(iter)}$.

(iv) If $\left\|\theta^{(iter)} - \theta^{(iter-1)}\right\|$ is less than tolerance set $\widehat{\theta} = \theta^{(iter)}$ and exit, otherwise return to (i).

4 Estimation results

The model is fitted to quarterly transitions from Standard & Poor's CreditPro database (US companies only) from 1981:1 to 2007:4. Data are binned into quarterly rating–to–rating transition counts with companies that transition to the "not rated" category ignored[8]. The S&P dataset is "complete" in the sense that ratings are available for every obligor in every quarter.

4.1 Basic estimator

Tables 1 and 2 show the estimated baseline generator matrix and the annual transition matrix associated with the generator.[9] Estimated annual migration probabilities match well the average empirical transition rates over the same period. A single pass of the EM algorithm can be implemented in MATLAB in only a few seconds on a 3.8GHz computer and highly accurate results can be obtained in under one minute. From a practical standpoint this is an important point in favor of the method.

TABLE 1: Baseline transition intensities for credit quality process ($\widetilde{\lambda}_{ij}$) ($\times 10^{-2}$)

		TO:							
		AA/AAA	A	BBB	BB	B	CCC	C–DDD	Def
	AA/AAA	-1.75	1.61	0.09	0.02	0.02	0	0	0
	A	0.71	-2.53	1.66	0.10	0.04	0	0.01	0.01
F	BBB	0.07	1.46	-3.23	1.51	0.15	0.02	0	0.02
R	BB	0.04	0.11	2.11	-5.38	2.85	0.14	0.03	0.1
O	B	0.02	0.07	0.11	2.15	-5.18	1.96	0.16	0.72
M:	CCC	0	0.13	0.24	0.24	5.68	-19.07	2.33	10.45
	C–DDD	0	1.45	0	1.49	5.74	9.72	-55.42	37.03
	Default	0	0	0	0	0	0	0	0

TABLE 2: Baseline annual transition matrix

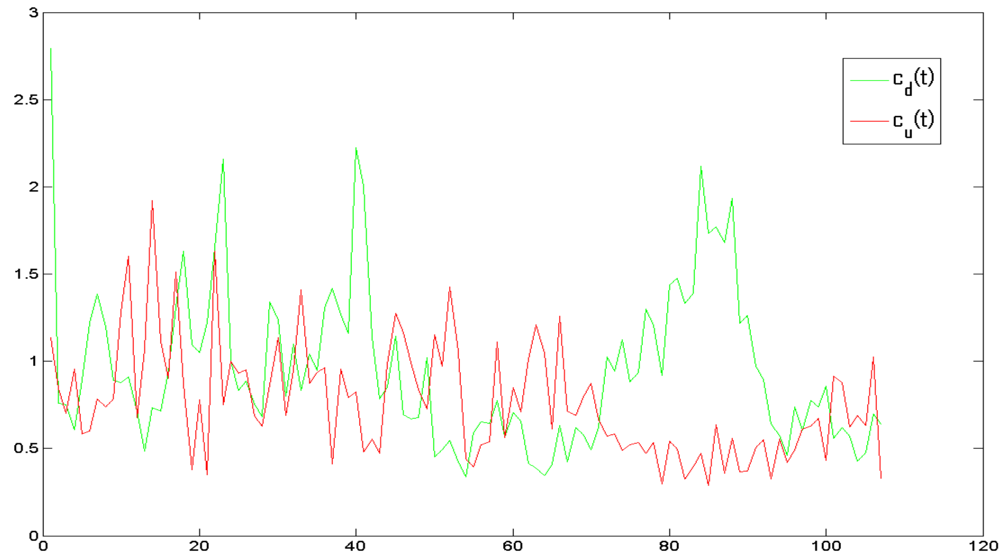
		TO:							
	%	AA/AAA	A	BBB	BB	B	CCC	C–DDD	Def
	AA/AAA	93	6	1	0	0	0	0	0
	A	3	91	5	1	0	0	0	0
F	BBB	1	5	88	5	1	0	0	0
R	BB	0	1	7	81	9	1	0	1
O	B	0	0	1	7	82	5	1	4
M:	CCC	0	0	1	1	15	48	2	33
	C–DDD	0	2	1	3	10	10	11	63
	Default	0	0	0	0	0	0	0	100

Estimated up and down systematic factors (4) are shown in Figure 1. Note that the factors are negatively correlated with one another in general. This is as expected, since changes in credit conditions that increase the up jump intensity should concurrently decrease the down jump intensity and vice–versa. The net systematic factor $c^u(t_z) - c^d(t_z)$ also corresponds well to United States business cycle fluctuations (and particularly credit cycle fluctuations) over the same period.

[8] Ratings with a "+" or "-" suffix are counted in the group they would be in if they were without the suffix.

[9] Given by $P = \exp(4\Delta\widetilde{\mathcal{L}})$.

FIGURE 1: Quarterly systematic factors in S&P corporate bond portfolio, 1981:1-2007:4



4.2 Missing ratings estimator

To test the missing ratings estimator we simulate an incomplete S&P dataset by randomly selecting obligors and removing transitions at random locations in time. We then estimate the model using the missing–ratings EM algorithm of section 3.2. Our estimation results, which are available on request, confirm the validity of the estimator.

5 Concluding remarks

This paper has developed a highly flexible yet tractable method for extracting systematic factors in a portfolio credit risk model. Several useful extensions to the methodology could form the basis for further work.

• While the model has been fitted here to a portfolio of corporate credits, it is straightforward to apply the methodology to other portfolios that have different structural properties. For example in mortgage portfolios the intensity of prepayment is also of interest. This possibility can be easily incorporated in our model by adding an upper boundary as a second absorbing state. It is also straightforward to decompose the model by cohort. This is potentially an important factor when the credit underwriting standards have changed through time.

• Systematic effects in the model are estimated as free parameters. This means sharp changes in transition intensities and correlations from one period to the next, such as may occur in a contagion environment, can be readily captured. However contagion feedback effects could also be incorporated directly into the model by conditioning the transition intensities on a default count variable, or some function of it; there are several ways of doing this, some examples include [FB], [JY] and [GA]. Since the default count itself follows a continuous time Markov chain with transition intensities that depend on the default transition intensity, such a model would in this context comprise a system of transition intensities. It would be interesting to consider extensions of the estimators to this multi–intensity case.

• The model addresses only the case when the systematic factors are estimated on the same frequency as the ratings. In practice it may be useful to extract systematic factors at a higher frequency. Extending the model to allow for intra–period variation in the factors is in principal straightforward. For example if $t_{z,r} = \{t_z + \frac{\Delta r}{m} : r = 0, ..., m-1\}$, $m \geqslant 1$ is the set of equally spaced points between t_z and t_{z+1} one could seek to estimate $c_{ij}(t_{z,r})$ using the same approach as in section 3. However some subtleties arise. For example survivorship of obligors on intra–period intervals should no longer be assumed, c.f. (12). Also restrictions are required on the $c_{ij}(t_{z,r})$ to ensure the model is identified. There are several possible ways of ensuring identification, including representing the model in a non–linear filtering framework. A more detailed examination of these issues will form the basis of a future paper.

References

[GA] Azizpour, S. and K. Giesecke (2008), 'Self–Exciting Corporate Defaults: Contagion vs. Frailty', *preprint*, Stanford University.

[BR] Bielecki, T. and M. Rutkowski (2004), *Credit Risk: Modeling, Valuation and Hedging*, 2nd edition, Springer–Verlag.

[BS1] Bladt, M. and M. Sørensen (2005), 'Statistical Inference for Discretely Observed Markov Jump Processes', *Journal of the Royal Statistical Society B*, 67, p.395-410.

[BS2] Bladt, M. and M. Sørensen (2007), 'Efficient Estimation of Transition Rates Between Credit Ratings from Observations at Discrete Time Points', *preprint*, University of Leipzig.

[FB] Frey, R. and J. Backhaus (2008), 'Pricing and Hedging of Portfolio Credit Derivatives with Interacting Default Intensites', *preprint*, University of Leipzig.

[JLT] Jarrow, R., D. Lando and S. Turnbull (1997), 'A Markov Model for the Term Structure of Credit Risk Spreads', *The Review of Financial Studies*, 10(2), p.481-523.

[JY] Jarrow, R. and F. Yu (2001), 'Counterparty Risk and the Pricing of Defaultable Securities', *Journal of Finance*, 56, p.1765-1799.