A Note on Fitting Markov Operator Credit Risk Models*

Harley Thompson† Jonathan Harris‡

June 2008

Abstract

We estimate a Markov operator credit migration model in which credit conditions vary through time in response to underlying macroeconomic factors. Emphasis is given to practical issues arising when fitting the model to a portfolio of risk rated credits, including the treatment of incomplete data, accounting for portfolio regeneration and aggregation issues.

*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

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

1 Introduction

In a series of recent papers, Albanese *et al* developed Markov analytic methods for the treatment of lattice–based models of portfolio credit risk (see [A], [AC1] and [ACDV] and references therein). In these models credit qualities evolve as continuous–time Markov chains on an N–element state space. Overlaid on this is a set of M "credit barriers" which bucket the chains according to their credit ratings, where $M << N$ in general. In this note we consider the estimation of such models under the statistical (real world) probability measure. Specifically, we consider estimation of transition matrices for a portfolio of risk–rated corporate loans when there is time variation in underlying causal factors. Particular emphasis is given to:

- Aggregation issues
- The effects of portfolio regeneration
- Treatment of irregular observation periods and ratings assignment errors.

We assume a framework in which credit qualities are assigned to M ratings, the first $M-1$ of which are non–default. Creditworthiness is assumed to be distinguished solely by rating. Of primary importance is the definition of default. From the Basel definition, default is defined as an obligor being:

(a) at least 90 days past due on a material credit obligation *or*

(b) considered by the debt holder to be unlikely to be able to meet credit obligations in full without recourse by the holder to actions such as realising security.

Under (a) the default time is considered to be the time when the payment first becomes 90 days in arrears. Under (b) however default status need not bear a necessary relation to the time when a payment is missed – indeed an obligor may be considered to be in default even if they are not in arrears on a contractual payment. Conversely a missed payment need not imply the obligor is in default – for instance an obligor may be behind on a payment at the end of a given quarter but make good these payments during the following quarter while remaining solvent during the whole period.

In practice credits tend to be re–rated at irregular intervals ranging between one or two months to a year or more. Thus if a model is specified on (say) a quarterly basis a distinction needs to be made between a "current" rating (a rating that is the result of a re–assessment of credit quality made in the corresponding quarter) and a "carried forward" rating (a rating that retains its previous value because new financial information has not become available). Only "current" ratings should be used to calibrate the model, with carried forward ratings being essentially missing data. Under this convention observed ratings transitions are indexed by four parameters: previous rating, current rating, calendar time and time between ratings.

A generic modeling approach involves treating the credit quality of the k^{th} name as a stochastic process $\{X_t^{(k)} : t \geqslant 0\}$, $k = 1, ..., K$ evolving on state space $S \subseteq \mathbb{R}$ with default occurring when the process first hits a lower absorbing boundary. Such a definition of default allows transition densities to be defined. Fix $S = [0, 1]$ and let the non–default part of the state space be the set $S^* = S \backslash \{0\}$. Let $a_{t+\Delta}^{(k)}$ be the minimum value of the k^{th} credit quality process on $(t, t+\Delta]$, $\Delta > 0$

$$a_{t+\Delta}^{(k)} = \inf_{s \in (t, t+\Delta]} \left\{ X_s^{(k)} \,|\mathcal{F}_t \right\},$$

where $\mathbb{F} = (\mathcal{F}_t)_{t \geqslant 0}$ is some reference filtration (containing say information on the obligor's credit ratings history and economic conditions). Since default is an absorbing state the transition density f is

$$f(X^{(k)}_{t+\Delta}\,\Big|X^{(k)}_t = x_t > 0\,) = \begin{cases} \mathbb{P}(a^{(k)}_{t+\Delta} > 0, X^{(k)}_{t+\Delta}\,|\mathcal{F}_t\,) & : X^{(k)}_{t+\Delta} \in S^* \\ \mathbb{P}(a^{(k)}_{t+\Delta} \leqslant 0\,|\mathcal{F}_t\,) & : X^{(k)}_{t+\Delta} = 0. \end{cases} \tag{1}$$

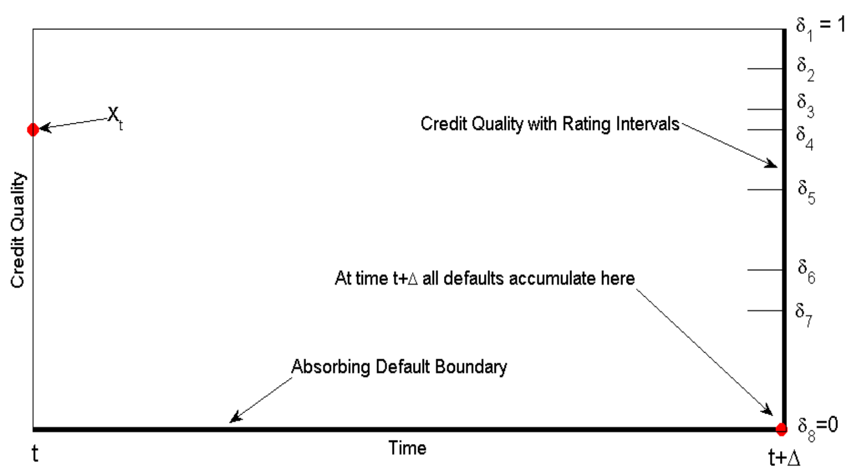
The segment of the transition density on S^* represents the probability of a transition from non–default credit quality $x > 0$ at t to non–default $X^{(k)}_{t+\Delta} > 0$ at time $t + \Delta$ without hitting the absorbing barrier at any time between t and $t + \Delta$. Conversely, the transition density at zero represents the probability that credit quality hits the lower default barrier *at some time* between t and $t+\Delta$ (i.e., the probability that the first hitting time is less than or equal to $t+\Delta$). It is worth noting that empirical studies tend to favour the specification of heavy–tailed transition densities (see for instance [AC1] and [GH], among others).

While the transition density completely specifies the model at a continuous level, in practice credit qualities are measured as ratings. Thus a link needs to be defined between the underlying credit quality process $X^{(k)}_t$ and the set of credit ratings. We do this via a *credit barrier mapping function* of the form:

$$\begin{aligned} \text{Obligor } k \text{ has rating } m \quad &\in \quad \{1, ..., M-1\} \text{ at time } t \iff X^{(k)}_t \in [\delta_m, \delta_{m+1}) \\ \text{Obligor } k \text{ has rating } m \quad &= \quad M \text{ at time } t \iff X^{(k)}_t = 0 \end{aligned}$$

where the constants $1 = \delta_1 > \delta_2 > ... > \delta_M = 0$ denote the boundaries of M ratings categories (see Figure 1 for the case $M = 8$).

FIGURE 1: Credit quality space



The remainder of the paper is organised as follows. Section 2 describes a Markov operator model for portfolio credit risk. Sections 3 and 4 present our estimation approach. Section 5 presents empirical results and Section 6 concludes.

2 Model of Credit Quality

We now turn to the specification of a model of credit quality. As mentioned earlier, our approach uses the lattice–based Markov operator method. The basic tool in our construction is the method of *stochastic time change.* Stochastic time change methods enable a rich but tractable class of heavy–tailed processes to be specified. Using the setup described above, we assume credit quality $\{X_t : t \geqslant 0\}$ follows a continuous–time Markov process on S^* with zero being an absorbing state. A discretisation is then applied to obtain a lattice representation of the model.

2.1 Stochastic time change and CGMY process

Let V_t be a continuous Lévy process[1] with characteristic function

$$\mathbb{E}\left(e^{iuV_t}\right) = \exp(\Psi(u)t)$$

and let H_t be a non–decreasing Lévy process with moment generating function

$$\mathbb{E}\left(e^{uH_t}\right) = \exp(h(u)t)$$

for some functions Ψ and h (respectively the *characteristic exponent* of V and *Laplace exponent* of H). We state the following without proof (see [CT], p.108).

Proposition 2.1. *The time–changed process X formed by sampling V at H_t−time,*

$$X_t = V_{H_t},$$

is also a Lévy process and has characteristic function

$$\mathbb{E}\left(e^{iuX_t}\right) = \exp(h\left(\Psi\left(u\right)\right)t).$$

Thus the characteristic exponent of X is obtained by composition of the characteristic exponent $\Psi(u)$ of the underlying process V with the Laplace exponent (or Bernstein function) $h(\lambda)$ of the time change H. Process H is known as the *subordinator* and X is said to be *subordinate* to V.

An appealing class of processes for our purposes is the (symmetric) CGMY process of [CGMY]. The symmetric CGMY process is a Lévy process with measure[2]

$$c(x) = \begin{cases} \widetilde{C}\exp(-\widetilde{G}x)x^{-(1+\widetilde{Y})} & : x > 0 \\ \widetilde{C}\exp(-\widetilde{G}\,|x|)\,|x|^{-(1+\widetilde{Y})} & : x < 0 \end{cases}$$

and characteristic function

$$\mathbb{E}\left(e^{iuX_t^{CGMY}}\right) = \exp\left(\widetilde{C}\Gamma(-\widetilde{Y})\left[\left(\widetilde{G}+iu\right)^{\widetilde{Y}} + \left(\widetilde{G}-iu\right)^{\widetilde{Y}} - 2\widetilde{G}^{\widetilde{Y}}\right]t\right), \tag{2}$$

where $\widetilde{C} > 0, \widetilde{G} \geqslant 0$ and $-1 < \widetilde{Y} < 2$ are parameters. It is a special case of the tempered stable class of Lévy processes (see [CT], p.116-123). A (slightly informal) interpretation of the parameters is as follows:

---

[1] A Lévy process is a stationary, independent increment process satisfying the property of stochastic continuity (for all $\varepsilon > 0$, $\lim_{a\to 0}\mathbb{P}\left(|X_{t+a} - X_t| \geqslant \varepsilon\right) = 0$). Note that the stochastic continuity property does not imply that the sample paths are continuous, rather it rules out jumps at fixed times.

[2] The Lévy measure $c(x)$ counts the expected number of jumps of size x per unit time. It is not in general a probability measure. A Lévy measure is said to be completely monotone if all the derivatives of $c(x)$ exist and satisfy $(-1)^k \frac{d^k c(u)}{du^k} > 0$ for all $k \geq 1$.

• $\widetilde{C}$ controls the "height" of the Lévy density and hence the overall intensity of jumps
• $\widetilde{G}$ controls the rate of decay of the tails of the measure and hence the size of large jumps
• $\widetilde{Y}$ controls the arrival frequency of small jumps.

The CGMY process is undefined for $\widetilde{Y} = 0, 1, 2$. In the limit as $\widetilde{Y} \to 0$ the Variance–Gamma process is obtained. Other special cases include:
• $\widetilde{Y} = -1$ and $\widetilde{C} = \widetilde{G}/2$: symmetric double–exponential jump density
• $\widetilde{Y} \to 2$ with $\widetilde{C} = 1 - \widetilde{Y}/2$: Brownian motion.

A range of further properties implied by different values of $\widetilde{Y}$ are summarised in Table 2.

TABLE 1: Properties of the time–changed process by $\widetilde{Y}$

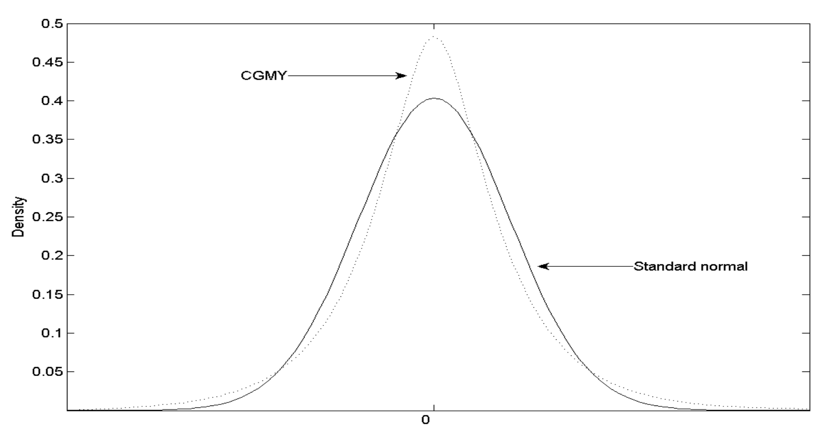
	Variation	Quad.Variation	Activity[5]	Lévy measure
$\widetilde{Y} < -1$	finite	finite	finite	not completely monotone
$\widetilde{Y} \in (-1, 0)$	finite	finite	finite	completely monotone
$\widetilde{Y} \in (0, 1)$	finite	finite	infinite	completely monotone
$\widetilde{Y} \in (1, 2)$	infinite	finite	infinite	completely monotone

The CGMY process may be obtained as time–changed Brownian Motion. Letting the underlying process be $V_t = W_t$, the Bernstein function implied by the CGMY characteristic function (2) is

$$h(\lambda) = 2\widetilde{C}\Gamma(-\widetilde{Y}) \left[\widetilde{G}^{\widetilde{Y}} - \left(2\lambda + \widetilde{G}^2\right)^{\widetilde{Y}/2} \cos\left(\widetilde{Y} \arctan\left(\frac{\sqrt{2\lambda}}{\widetilde{G}}\right)\right)\right] \tag{3}$$

(see Appendix 1). The CGMY process differs from the underlying Brownian process by having a higher peak in the centre of the distribution and fatter tails (see Figure 2).

FIGURE 2: CGMY density time–changed from the standard normal



2.2 Transition matrix construction

To implement our model of credit quality we fit a CGMY process to a space and time–discretised version of the state space. Let $N > 0$ be a positive integer and let S now comprise a discretisation of $[0, 1]$ comprising equally spaced points $\{\frac{N-i}{N-1} : i = 1, ..., N\}$. Assume that the discrete times $T = \{r\Delta : r = 0, 1, 2, ..\}$ correspond to the observation times of the ratings and common risk factors and let $P_X(t, t+r\Delta)$ be the transition matrix on S describing the dynamics of X over $[t, t+r\Delta]$, where t is constrained to the set T.

Key to our construction is being able to obtain an expression for the infinitesimal generator[6] $\mathcal{L}_X$ of X on S since this forms the basis of the construction of the transition matrix P_X. For a process X_t subordinate to V_t if the generator $\mathcal{L}_V$ of V is known then $\mathcal{L}_X$ may be obtained via

$$\mathcal{L}_X = -h(-\mathcal{L}_V) \tag{4}$$

where, in the CGMY case, h is the Bernstein function (3) (see [ACDV]). Since for CGMY V is standard Brownian Motion, the finite difference approximation of its generator on S is (at any fixed time)[7]

$$\left(\mathcal{L}_V f\right)(v_i) = (N-1)^2 \left[\frac{1}{2} f(v_{i+1}) - f(v_i) + \frac{1}{2} f(v_{i-1})\right], \quad i = 2, ..., N-1, \quad v_i \in S.$$

This implies $\mathcal{L}_V$ is a tri–diagonal matrix with elements

$$\begin{aligned} \mathcal{L}_V(i, i+1) &= \mathcal{L}_V(i, i-1) = \frac{1}{2}(N-1)^2 \\ \mathcal{L}_V(i, i) &= -(N-1)^2. \end{aligned}$$

We refer to this discrete approximation as the *diffusion generator matrix*. The generator matrix satisfies conservation of probability and non–negativity conditions

$$\begin{aligned} \mathcal{L}_V(i, j) &\geqslant 0; \qquad j \neq i \\ \mathcal{L}_V(i, i) &= -\sum_{j \neq i} \mathcal{L}_V(i, j). \end{aligned} \tag{5}$$

and is also required to satisfy boundary conditions

$$\begin{aligned} \mathcal{L}_V(N, j) &= 0 \text{ for all } j \\ \mathcal{L}_V(1, j) &= \begin{cases} -(N-1)^2 & : j = 1 \\ -\mathcal{L}_V(1,1) & : j = 2 \\ 0 & : j = 3, ..., N \end{cases} \end{aligned} \tag{6}$$

which represent, respectively, absorption at the default boundary ($v_N = 0$) and reflection at the upper boundary ($v_1 = 1$). The upper boundary corresponds to "perfect" credit quality.

The generator matrix $\mathcal{L}_V$ admits the diagonalisation $\mathcal{L}_V = U\Lambda U^{-1}$ where $U = \{\mathbf{u}_1, .., \mathbf{u}_N\}$ are the eigenvectors of $\mathcal{L}_V$ and $\Lambda = \text{diag}(\lambda_0, ..., \lambda_N)$ are the eigenvalues. This property in conjunction with (4) implies we may obtain the symmetric CGMY generator matrix $\mathcal{L}_X$ as

$$\begin{aligned} \mathcal{L}_X &= -h\left(-\mathcal{L}_V\right) \\ &= -h\left(-U\Lambda U^{-1}\right) \\ &= -U h\left(-\Lambda\right) U^{-1}. \end{aligned}$$

[6] The operator $\mathcal{L}$ satisfying $\lim_{\Delta \to 0} \frac{1}{\Delta} \mathbb{E}\left(f(V_{t+\Delta}) - f(v_t)\right) = \left(\mathcal{L} f\right)(v_t)$.

[7] The generator on continuous space is $(\mathcal{L} f)(v) = \frac{1}{2} \frac{\partial^2 f(v)}{\partial v^2}$, $v \in \mathbb{R}$.

Note that the mean and variance implied by the generator are, by state

$$\begin{aligned}\mu(x_i) &= \sum\nolimits_{j=1}^{N}(x_j - x_i)\mathcal{L}_X(i,j)\\ var(x_i) &= \sum\nolimits_{j=1}^{N}(x_j - x_i)^2\mathcal{L}_X(i,j)\end{aligned}$$

for all $x_i = \frac{N-i}{N-1}$.

The final step remains to extend the jump generator matrix from infinitesimal time to finite time $\Delta > 0$. This may be achieved via matrix exponentiation. When X is a time–homogeneous Markov chain the transition and generator matrices satisfy the relation

$$\begin{aligned}P_X(t, t+\Delta) &= \exp(\Delta\mathcal{L}_X(t))\\ &= U\exp(-\Delta h\,(-\Lambda))U^{-1}.\end{aligned} \tag{7}$$

Note that the transition matrix captures the probability of making a transition from non–default lattice point x_i at t to non–default lattice point x_j at $t+\Delta$ without hitting zero any time in between. Furthermore, the transition probability from non–default point x_i to the default point $x_j = 0$ is the equal to the probability that the process hits zero at some time between t and $t+\Delta$. Thus the transition matrix obtained in (7) corresponds to the lattice version of (1) for jump process X, as required.

2.3 State and time dependence

At this stage we have a transition matrix on the lattice S describing the dynamics of a symmetric CGMY jump process X over $[t, t+\Delta]$. In practice however the generator matrix comprises jumps whose frequency and magnitude vary across time and state. To introduce such asymmetry we calculate generators for two separate symmetric CGMY processes $\mathcal{L}_{up}$ and $\mathcal{L}_{dn}$ then obtain an asymmetric jump process by weighting the lower triangle of $\mathcal{L}_{up}$ and the upper triangle of $\mathcal{L}_{dn}$.[8] The two symmetric processes are generated using CGMY Bernstein functions with time and state dependent values of $\widetilde{C}$, $\widetilde{G}$ and $\widetilde{Y}$. We assume the following state and time dependence properties:

- the magnitude (distribution) of jumps depends on the state but not on time
- the intensity of jumps (of any magnitude) depends on time but not on state.

These assumptions are made so as to yield a model that is flexible yet remains parametrically identified. Since the $\widetilde{G}$ and $\widetilde{Y}$ parameters determine the shape of the Lévy measure, and hence the distribution of jumps, state dependence is introduced into the up and down generators $\mathcal{L}_{up}$ and $\mathcal{L}_{dn}$ by letting these parameters depend on the state

$$\begin{aligned}\widetilde{G}_{up}(i) &= f_{up}(x_i;\boldsymbol{\varphi}_{up}), \qquad \widetilde{G}_{dn}(i) = f_{dn}(x_i;\boldsymbol{\varphi}_{dn})\\ \widetilde{Y}_{up}(i) &= g_{up}(x_i;\boldsymbol{\psi}_{up}), \qquad \widetilde{Y}_{dn}(i) = g_{dn}(x_i;\boldsymbol{\psi}_{dn}),\end{aligned}$$

where the f and g are given functions and $\boldsymbol{\varphi}$ and $\boldsymbol{\psi}$ are parameters to be estimated. This implies that each row of the generator matrix is generated by a different CGMY process. Letting $h_i^{up}(-\Lambda)$ and $h_i^{dn}(-\Lambda)$ be the Bernstein functions associated with parameters $\left\{\widetilde{G}_{up}(i), \widetilde{Y}_{up}(i)\right\}$ and $\left\{\widetilde{G}_{dn}(i), \widetilde{Y}_{dn}(i)\right\}$ and letting $\mathbf{u}(i)$ be the i^{th} row of U, the i^{th} rows of

[8] This approach makes use of the property that if $\mathcal{L}_1$ and $\mathcal{L}_2$ are generator matrices then so is $\mathcal{L} = \mu_1\mathcal{L}_1 + \mu_2\mathcal{L}_2$ for any $\mu_1, \mu_2 \geqslant 0$.

$\mathcal{L}_{up}$ and $\mathcal{L}_{dn}$ are obtained by

$$\mathcal{L}_{up}(i) = -\mathbf{u}(i)h_i^{up}(-\Lambda)\,U^{-1} \tag{8}$$
$$\mathcal{L}_{dn}(i) = -\mathbf{u}(i)h_i^{dn}(-\Lambda)\,U^{-1}. \tag{9}$$

Time dependence is introduced by conditioning the overall jump intensity $\widetilde{C}$ on a vector of observed common risk factors $\mathbf{Z}_t$

$$\widetilde{C}_{up}(t) = \exp(\boldsymbol{\alpha}'_{up}\mathbf{Z}_t)$$
$$\widetilde{C}_{dn}(t) = \exp(\boldsymbol{\alpha}'_{dn}\mathbf{Z}_t)$$

where $\boldsymbol{\alpha}$ are parameters. The final state and time–dependent generator matrix is then written

$$\mathcal{L}_X(t) = \widetilde{C}_{up}(t)\widehat{\mathcal{L}}_{up} + \widetilde{C}_{dn}(t)\widehat{\mathcal{L}}_{dn} \tag{10}$$

where $\widehat{\mathcal{L}}_{up}$ and $\widehat{\mathcal{L}}_{dn}$ are the lower and upper triangles of $\mathcal{L}_{up}$ and $\mathcal{L}_{dn}$.[9]

As well as allowing rich dynamic properties, the generator formed in this way has several practical advantages:

- correlation structures in very high dimensional portfolios may be represented by a relatively small number of parameters
- model outcomes are intuitive and easily interpretable
- multi–factor stress tests are straightforward to perform.

To illustrate the application to stress testing, suppose that the systematic factor comprises contemporaneous values of real activity (A) and the real long bond rate (L). The up and down jump intensities may then be written as (say)

$$\widetilde{C}_{up}(t) = \exp\left(\alpha_{1,up}A_t + \alpha_{2,up}L_t\right), \tag{11}$$
$$\widetilde{C}_{dn}(t) = \exp\left(\alpha_{1,dn}A_t + \alpha_{2,dn}L_t\right) \tag{12}$$

where the parameters $(\alpha_{1,up}, \alpha_{2,up}, \alpha_{1,dn}, \alpha_{2,dn})$ require estimation from empirical observations. This can be done using any simple method such as least squares. Depending on the values of these parameters a variety of model properties may be obtained (see Table 2).

TABLE 2: Impact on the credit quality process of an increase in A_t, by sgn(α_1) parameter

sgn($\alpha_{1,u}$)	sgn($\alpha_{1,d}$)	Drift	Variance
+	+	Increase in absolute value	Increase
+	–	Increase	Either increase or decrease depending on parameters
–	+	Decrease	Either increase or decrease depending on parameters
–	–	Decrease in absolute value	Decrease

It is clear that even with this small number of risk factors the model is capable of generating a variety of dynamic properties; adding more risk factors naturally increases the range of possible model outcomes.

[9] Note that because $\widetilde{C}$ enters linearly in (3) it can be taken out of the Bernstein function when calculating (8)–(9). That is we can first compute (8)–(9) using a standardised Bernstein function (say by setting $\widetilde{C}_{(\cdot)} = 1$) then in a second step multiply the resultant $\mathcal{L}_{up}$ and $\mathcal{L}_{dn}$ by $\widetilde{C}_{up}(t_k)$ and $\widetilde{C}_{dn}(t_k)$ as in (10).

3 Aggregation methodology

To estimate the model one must calculate transition probabilities at the rating category level (and the corresponding numbers of rating to rating transitions). However since the underlying model is specified in terms of one–period lattice–level transition matrices $P_X(t, t+\Delta)$ an aggregation scheme is required to convert the transition probabilities from the lattice level to the rating level, and from single to multi–period transitions.

3.1 Converting from lattice to rating level

The sequence of steps to construct the rating–level transition matrix is as follows:

1. Convert the one–period lattice–level transition matrix $P_X(t, t+\Delta)$ into a multi–period lattice–level transition matrix $P_X(t, t+r\Delta)$, $r > 1$

2. Convert $P_X(t, t+r\Delta)$ into a multi–period lattice–level transition matrix $P_Y(t, t+r\Delta)$ with measurement errors taken into account

3. Aggregate $P_Y(t, t+r\Delta)$ into a multi–period rating–level transition matrix $\widehat{P}_Y(t, t+r\Delta)$.

Step 1: Converting from one–period to multi–period lattice–level measured transition probabilities[10] In practice several quarters may elapse between ratings re–assessments. Due to the presence of the absorbing barrier at zero the r–period transition probability, $r > 1$, cannot be obtained simply by multiplying r one–period transition matrices. Two considerations are relevant in this regard:

• Default is an absorbing state so observed multi–period transitions between non–default ratings must traverse via non–default ratings on all periods in between.

• Defaults are assumed to be recorded as soon as they occur, thus if a contiguous rating pair, separated by several periods in which no ratings is recorded, comprises a non–default and a default rating, the transition to default is assumed to have occurred in the final period.

Let $\widetilde{\mathbf{x}} = \left[\begin{array}{ccc} x_1 & \cdots & x_{N-1} \end{array}\right]$ be the non–default segment of the state space and $\mathbf{x} = \left[\begin{array}{c} \widetilde{\mathbf{x}} \vdots x_N \end{array}\right]$ the full state space. Then given the one–period matrix $P_X(t, t+\Delta)$ the $r-$period matrix $P_X(t, t+r\Delta)$ satisfies the recursion

$$\begin{aligned} P_X(\widetilde{\mathbf{x}}, t; \mathbf{x}, t+r\Delta) &= P_X(\widetilde{\mathbf{x}}, t; \widetilde{\mathbf{x}}, t+(r-1)\Delta) P_X(\widetilde{\mathbf{x}}, t+(r-1)\Delta; \mathbf{x}, t+r\Delta) \\ P_X(x_N, t; \mathbf{x}, t+r\Delta) &= \left[\begin{array}{c} \mathbf{0} \vdots 1 \end{array}\right]. \end{aligned}$$

This expression says that the $r-$period transition probability may be calculated as the probability of not defaulting over the first $r-1$ periods multiplied by the one–period transition probability in the r^{th} period. Note this implies that as the length of time between ratings increases the transition probability between pairs of non–default ratings will tend to zero and multi–period transition probabilities will not sum to one across rows.

Step 2: Converting from the 'true' to 'measured' lattice–level transition probabilities To introduce measurement errors we make a distinction between measured and true credit quality. Let true credit quality X_t be as given by the preceding model and let $Y_t \in S$ represent measured credit quality. Suppose that ratings assessments comprise

[10] For some portfolios, such as the S&P portfolio considered later in the paper, $r = 1$ in all periods and this step becomes redundant. However in more general cases, such as for internally rated commercial bank portfolios, step 1 is likely to be relevant.

the sum of true credit quality and a measurement error ε with time–homogeneous density f_ε (not necessarily zero mean), $Y_t = X_t + \varepsilon_t$. Let $x_t, y_t \in S$ be the true and measured lattice point credit qualities respectively. Then in the presence of measurement errors the lattice transition probabilities for the measured credit qualities are equal to the transition probabilities of the true credit qualities convoluted with the measurement errors,

$$p(y_t; y_{t+r\Delta}) = \sum\nolimits_{i=1}^{N} \sum\nolimits_{j=1}^{N} f_\varepsilon(y_t - x_{i,t}) p(x_{i,t}; x_{j,t+r\Delta}) f_\varepsilon(y_{t+r\Delta} - x_{j,t+r\Delta}) \tag{13}$$

for $r > 0$. Letting E be the matrix of measurement error densities

$$E = \begin{bmatrix} f_\varepsilon(y_1 - x_1) & \cdots & f_\varepsilon(y_1 - x_N) \\ \vdots & \ddots & \vdots \\ f_\varepsilon(y_N - x_1) & \cdots & f_\varepsilon(y_N - x_N) \end{bmatrix}$$

we may write (13) in matrix form

$$P_Y(t, t + r\Delta) = E_1 P_X(t, t + r\Delta) E_2', \quad r > 0 \tag{14}$$

where E_1 and E_2 are scaled versions of E in which the rows and columns respectively sum to one.[11] The measurement error probabilities are assumed to be double exponentially distributed with weights that depend on the state

$$f_\varepsilon(y_i - x_l) \propto \begin{cases} (1 - x_l)\beta \exp(-\beta\,|y_i - x_l|) & : y_i > x_l \\ x_l \beta \exp(-\beta\,|y_i - x_l|) & : y_i \leqslant x_l. \end{cases}$$

This density allows measurement errors to be parameterised in terms of a single parameter β.[12]

Step 3: Aggregating from lattice to rating category level The final step involves aggregating the multi–period lattice–level transition matrix to the rating level. Aggregation involves summing the time–$t + r\Delta$ probabilities and probability–weighting the time–t probabilities. In matrix terms the rating–to–rating transition matrix $\widehat{P}_Y(t, t + r\Delta)$ can be written

$$\widehat{P}_Y(t, t + r\Delta) = W(t) P_Y(t, t + r\Delta) A'$$

where A is an $M \times N$ aggregation matrix

$$A = \begin{bmatrix} [1 \ \cdots \ 1] & & & & \\ & [1 \ \cdots \ 1] & & & \\ & & \ddots & & \\ & & & [1 \ \cdots \ 1] & \\ & & & & 1 \end{bmatrix} \begin{array}{l} \leftarrow x_i(t + r\Delta) \in [1, \delta_2) \\ \leftarrow x_i(t + r\Delta) \in [\delta_2, \delta_3) \\ \\ \leftarrow x_i(t + r\Delta) \in [\delta_{M-1}, \delta_M) \\ \leftarrow x_i(t + r\Delta) = 0 \ \ \text{(default)} \end{array}$$

and $W(t) \propto A \times \mathrm{diag}(\mathbf{p}_Y^*(t))$ is an $M \times N$ weighting matrix[13] with $\mathbf{p}_Y^*(t)$ the unconditional density of credit quality at time t. Note that the weighting and aggregation matrices share

---

[11] That is they are scaled to form probability matrices.

[12] Since the double exponential distribution is defined on $\mathbb{R}$ it will not in general integrate to 1 over the region $[0, 1]$. In practice however we find that β is sufficiently large (estimated measurement errors are sufficently small) that any difference from 1 is negligible.

[13] It is re-scaled every period so that its rows sum to one. $\mathrm{diag}(\mathbf{p}_Y\,(t_{k+\Delta}))$ is the matrix formed by placing the vector $\mathbf{p}_Y\,(t_{k+\Delta})$ on the diagonal of an $N \times N$ identity matrix

a common structure (i.e. non–zeros occur in the same parts of the matrix). The number of non–zero columns in the m^{th} row of A and $W(t)$ is determined by the number of lattice points that fall in the interval $[\delta_m, \delta_{m+1})$.[14] Therefore by changing the positions of the ratings interval boundaries, the number of elements in each row changes and hence also the number of obligors being weighted/aggregated into the corresponding rating group. Substitution from (14) allows steps 2 and 3 to be combined in a single equation

$$\widehat{P}_Y(t, t + r\Delta) = W_1(t) P_X(t, t + r\Delta) A_2',$$

where $W_1(t) = W(t)E_1$ and $A_2 = AE_2$.

The difficulty with this step lies in the calculation of the weighting matrix, which relies on knowledge of $\mathbf{p}_Y^*$. The weight in the m^{th} row and i^{th} column of $W(t)$ is the unconditional probability of being at lattice point $x_i \in [\delta_m, \delta_{m+1})$ at time t. This probability can be obtained by transitioning the previous period's unconditional density. However in each period an adjustment is required to account for the effect of commencing and departing obligors.[15] To derive this adjustment we first define the following unconditional densities:

- $\mathbf{p}_Y^*(t)$: lattice level density not adjusted for commencements/departures
- $\mathbf{p}_Y(t)$: lattice level density adjusted for commencements/departures
- $\widehat{\mathbf{p}}_Y^*(t)$: rating level density not adjusted for commencements/departures
- $\widehat{\mathbf{p}}_Y(t)$: rating level density adjusted for commencements/departures
- $\widehat{\mathbf{c}}(t)$: empirical density of commencements at ratings level

Assuming an initial condition for $\mathbf{p}_Y^*(0)$ the lattice–level unconditional density $\{\mathbf{p}_Y(t) : t = \Delta, 2\Delta, \ldots, R\Delta\}$ with commencements and departures taken into account may be computed recursively as follows:

$$\begin{aligned} \widehat{\mathbf{p}}_Y^*(t) &= A\mathbf{p}_Y^*(t) \\ \widehat{\mathbf{p}}_Y(t) &= (1 - \gamma(t))\widehat{\mathbf{p}}_Y^*(t) + \gamma(t)\widehat{\mathbf{c}}(t) \\ \mathbf{p}_Y(t) &= W'(t)\widehat{\mathbf{p}}_Y(t) \\ \mathbf{p}_Y^*(t + \Delta) &= P_Y'(t, t + \Delta)\mathbf{p}_Y(t) \end{aligned} \tag{15}$$

where $\gamma(t)$ is the proportion of the portfolio that commenced in t,[16]

$$\gamma(t) = \frac{\#\{\text{Obligors commencing in } t\}}{\#\{\text{Obligors in portfolio in } t\}}.$$

The intuition behind the recursion is as follows. First the unadjusted density is aggregated to the rating level and weighted with the empirical density of the newly

[14] In practice the matrix is slightly modified so that the end point values in each row overlap in adjacent rows and sum to one down the columns. These overlapping values are weighted according to their distance from the nearest δ. The reason for this is twofold. First it allows the aggregation matrix to vary continuously in the boundaries δ thus making the likelihood more numerically stable. Second it makes the discrete model a better approximation to the underlying continuous model.

[15] This effect becomes important when there is a high degree of portfolio turnover. If no distinction is made between continuing and newly arriving/departing obligors the fitted transition matrix will reflect the combined effect of the true credit quality transition and the densities of new and departing obligors. Since in general the empirical densities of new and departing obligors will differ from the density of continuing obligors (and may reflect, for instance, lending policy rather than economic factors) if no adjustment is made the estimated correlation with economic factors will be biased.

[16] This recursion is only strictly correct when the distribution of departing obligors is the same as the iterated unconditional distribution $\widehat{\mathbf{p}}_Y^*$; we maintain this assumption here because the distribution of departing obligors is in general unknown.

arriving obligors (which are measured only at the rating level). Then the combined density is disaggregated back to the lattice level using the weighting matrix. Finally the lattice–level unconditional density is transitioned using the transition matrix and the recursion repeated. Since aggregation performs a lattice–to–rating translation and weighting performs a rating–to–lattice translation the two matrices can be thought of as quasi inverses of one another. Note that if $\gamma(t) = 0$ (15) collapses to the standard recursion $\mathbf{p}_Y(t+\Delta) = P_Y'(t, t+\Delta)\mathbf{p}_Y(t)$. Note also that W is computed from the unadjusted density $\mathbf{p}_Y^*$ rather than the adjusted density $\mathbf{p}_Y$ because only the unadjusted density is available at the time W needs to be calculated.

4 Estimation

Estimation involves identifying the set of parameters that define the underlying generator from ratings transitions recorded over discrete time intervals. In general the question of estimation of continuous–time Markov chain from discrete data is not trivial. In particular, the MLE may not exist and if it does exist it may not be unique. We provide a brief treatment of existence and uniqueness issues, based on [BS1], in Appendix 2. The log likelihood function for a continuous, time–inhomogeneous Markov chain[17] is, assuming continuous observation of the chain,

$$L = \sum_{t=0}^{T-1} \sum_{m=1}^{M-1} \sum_{n \neq m} \mathbf{N}_{m,n}(t, t+\Delta) \ln \lambda_{m,n}(t) - \lambda_{m,n}(t)\mathbf{R}_m(t, t+\Delta)$$

where $\mathbf{N}_{m,n}(t, t+\Delta)$ is the total number of transitions $m \to n$ observed on the interval $(t, t+\Delta]$, $\mathbf{R}_m(t, t+\Delta)$ is the total time spent in rating m (by all obligors) over the same period and $\lambda_{m,n}(t)$ is the $(m,n)^{th}$ element of $\mathcal{L}_X(t)$. The difficulty here is that the likelihood depends on the continuous data, which are unobservable since ratings are measured only quarterly. Additionally the measured quarterly transitions are themselves incomplete because for each obligor the interval between ratings dates is often greater than one quarter. Two possible approaches to addressing this problem are:

(1) Obtain $\lambda_{m,n}(t)$ directly via an iterative scheme in which the continuous data are replaced by their conditional expectations (see [BS1], [BS2]).

(2) Write the likelihood directly in terms of $r\Delta-$interval observable quantities, in this case the observed ratings transitions.

Given the analysis from the preceding sections we adopt the latter approach. We assume the times between ratings (quarters) are equidistant and set $\Delta = 1$ to represent a single quarter. Then employing the conditional independence and Markov properties of the model the discrete likelihood is

$$\begin{aligned} L &= \sum_{t=0}^{T-1} \sum_{r=1}^{T-t} \sum_{m=1}^{M-1} \sum_{n=1}^{M} \mathbf{N}_{m,n}(t, t+r) \ln q_{m,n}(t, t+r; \boldsymbol{\theta}) \\ &\quad + \sum_{t=0}^{T-1} \sum_{r=1}^{T-t} \sum_{m=1}^{M-1} \mathbf{N}_{m,/M}(t, t+r) \ln(1 - q_{m,n}(t, t+r; \boldsymbol{\theta})) \end{aligned}$$

where $\mathbf{N}_{m,n}(t, t+r)$ is as described previously and

[17] The inhomogeneity here is assumed to be of a special "piecewise" form whereby the process is homogeneous on any given interval $[t, t+\Delta]$ but inhomogeneous across t.

• $\mathbf{N}_{m,/M}(t,t+r)$ is the number of obligors whose last recorded rating is at time t and who departed without default at time $t+r$, or were still in the portfolio and not in default at the end of the sample

• $q_{m,n}(t,t+r)$ is the $(m,n)^{th}$ element of $\widehat{P}_Y(t,t+r)$.

Thus it is sufficient to group the observations in contiguous pairs weighted by the number of observations in each group, with the indices in the likelihood being the ratings pairs (m,n) and the calendar times of the ratings $(t,t+r)$. The parameters $\boldsymbol{\theta}$ comprise up and down jump parameters, credit quality boundaries, CGMY parameters and the measurement error parameter (see Table 3).

TABLE 3: Model parameters $\boldsymbol{\theta}$ by purpose

Parameter	Description
$\boldsymbol{\varphi}_{up}$	$\widetilde{G}$ parameters (up jumps)
$\boldsymbol{\varphi}_{dn}$	$\widetilde{G}$ parameters (down jumps)
$\boldsymbol{\psi}_{up}$	$\widetilde{Y}$ parameters (up jumps)
$\boldsymbol{\psi}_{dn}$	$\widetilde{Y}$ parameters (down jumps)
$\boldsymbol{\alpha}_{up}$	$\widetilde{C}$ parameters on common risk factors (up jumps)
$\boldsymbol{\alpha}_{dn}$	$\widetilde{C}$ parameters on common risk factors (down jumps)
$\boldsymbol{\delta}$	Credit quality boundaries
$\mathbf{s}$	Seasonal factors
β	Measurement error

Rather than estimate the risk factor parameters directly we estimate them in two steps. First $\widetilde{C}_{up}(t)$ and $\widetilde{C}_{dn}(t)$ are estimated as free parameters (with no risk factors present) subject to the restriction that

$$\begin{aligned}\sum\nolimits_{k=0}^{K} \ln \widetilde{C}_{up}(t) &= 0 \\ \sum\nolimits_{k=0}^{K} \ln \widetilde{C}_{dn}(t) &= 0.\end{aligned}$$

This approach ensures that the variability of $\widetilde{C}_{up}(t)$ and $\widetilde{C}_{dn}(t)$ over time is not restricted by the choice of risk factors thus ensuring that the obtained time series is flexible and accurate. Then these times series are fitted to observed macroeconomic/credit factors using equations of the form (11)–(12).

5 Estimation results

For the set of observation times $\{r\Delta : r = 0,1,\ldots,R\}$ we set $\Delta = 1$ to represent one quarter (3 months). The model is fitted to transitions from the S&P CreditPro database (US companies only) from 1981:1 to 2007:4.[18] Data are binned into quarterly rating–to–rating transition counts with companies that transition to the "not rated" category ignored. Parameters are estimated iteratively in six blocks: $\boldsymbol{\delta}$,$\{\boldsymbol{\psi}_{up},\boldsymbol{\psi}_{dn}\}$, $\{\boldsymbol{\varphi}_{up},\boldsymbol{\varphi}_{dn}\}$, $\mathbf{s}$, $\boldsymbol{\alpha}_{up}$, $\boldsymbol{\alpha}_{dn}$ with the estimation converging fully in around 20 iterations (see Figure 3).

[18] Note that for this database Step 1 is essentially redundant since $r = 1$ in all cases. However in more general cases, such as internally rated commercial bank portfolios, Step 1 will remain relevant.

FIGURE 3: Convergence of parameters (split between large and small values)

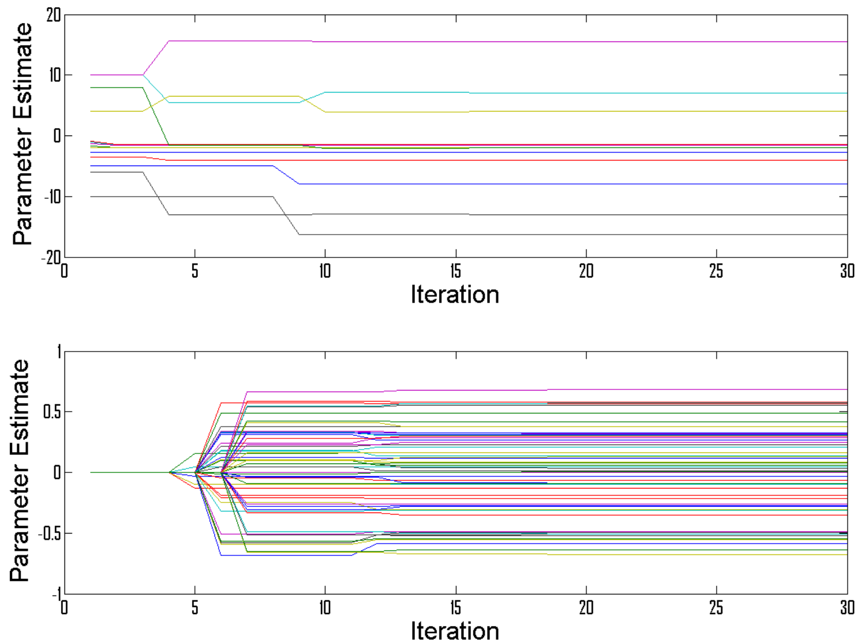


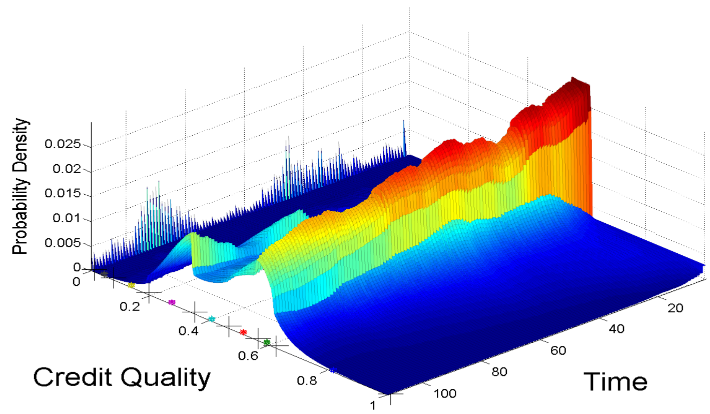
Table 4 shows the average fitted (annualised) transition matrices for the universe of companies over the full estimation period 1980 to 2007.

TABLE 4: Baseline annual transition matrix

		TO:							
	%	AA/AAA	A	BBB	BB	B	CCC	C–DDD	Def
	AA/AAA	93.8	4.7	1.3	0.2	0	0	0	0
	A	3.5	88.7	7.0	0.8	0.1	0	0	0
F	BBB	0.4	2.7	90.9	5.5	0.5	0	0	0
R	BB	0	0.2	3.4	89.1	6.7	0.4	0	0.2
O	B	0	0	0.4	5.2	84.8	5.1	0.5	3.9
M:	CCC	0.1	0.1	0.3	1.6	10.7	52.1	1.7	33.4
	C–DDD	0.1	0.1	0.3	0.8	2.9	5.4	16.4	74.1
	Default	0	0	0	0	0	0	0	100

Figure 4 shows the fitted unconditional distribution of S&P credit quality with the lattice comprising $N = 70$ points. At a basic level this diagram gives a quick visual representation of the distribution of obligors across rating grades and time. Note the accumulation of defaults at the lower boundary, the number of which fluctuates through time with the economic cycle. Most of the observed variation in the distribution reflects the impact of portfolio regeneration, that is obligors entering and departing the portfolio at a different distribution from the current unconditional distribution. Notwithstanding this, the unconditional distribution remains reasonably steady through time.

FIGURE 4: Unconditional distribution of credit quality through time (quarters since 1980:1)



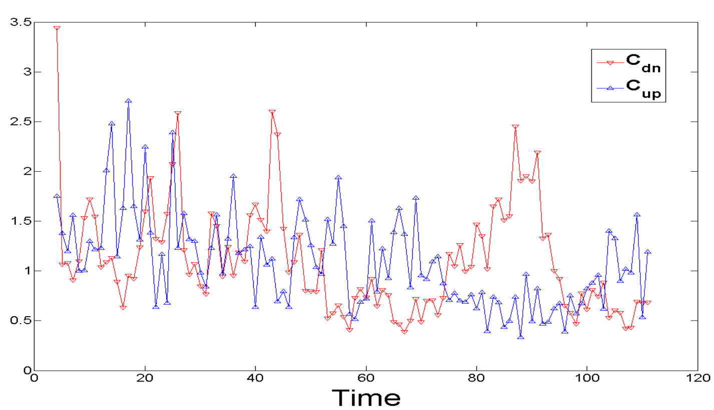
The credit quality regions associated with the distribution in Figure 4 are shown in Table 5.

TABLE 5: Boundaries of Credit Quality Regions

	Credit quality interval
AA/AAA	$[1, 0.63)$
A	$[0.63, 0.57)$
BBB	$[0.57, 0.47)$
BB	$[0.47, 0.35)$
B	$[0.35, 0.19)$
CCC	$[0.19, 0.07)$
C–DDD	$[0.07, 0)$
Default	0

Up and down systematic factors $\widetilde{C}_{up}(t)$ and $\widetilde{C}_{dn}(t)$ are shown in Figure 5. It is notable that the net systematic factor corresponds well with US business cycle fluctuations over the same period.

FIGURE 5: Systematic factor intensity through time (quarters since 1980:1)



6 Conclusion

We have developed a highly flexible yet tractable lattice–based credit migration model and shown how estimation may be performed in the presence of various data irregularities, including incomplete data, measurement errors and partial observability of the credit quality process. Our results indicate the feasibility of using transition densities with jump dynamics and complicated state and time dependence properties and should prove useful for model based stress testing and economic capital determination. Properly incorporating second–order (but important) feedback effects such as default contagion is a potential area for further research. Extending the model to include loss–given–default will also form the basis of future work.

Appendix 1 (CGMY Bernstein function)
The characteristic function of the symmetric CGMY process X is

$$\mathbb{E}\left(e^{iuX_t}\right) = \exp\left(\widetilde{C}\Gamma(-\widetilde{Y})\left[\left(\widetilde{G}+iu\right)^{\widetilde{Y}} + \left(\widetilde{G}-iu\right)^{\widetilde{Y}} - 2\widetilde{G}^{\widetilde{Y}}\right]t\right). \tag{16}$$

Thus if the underlying process is $V_t = W_t$ and H_t is a subordinator, the aim is to find the Bernstein function $h(u)$ of H such that $X_t = V_{H_t}$ is a CGMY process. We may condition on the time change to write

$$\begin{aligned}
\mathbb{E}\left(e^{iuX_t}\right) &= \mathbb{E}\left(e^{iuV_{H_t}}\right) \\
&= \mathbb{E}\left(\mathbb{E}\left(e^{iuW_{H_t}}\right)|H_t\right) \\
&= \mathbb{E}\left(e^{-\frac{1}{2}u^2 H_t}\right) \\
&= \mathbb{E}\left(e^{-\lambda H_t}\right) \\
&= e^{-h(\lambda)t} \qquad (17)
\end{aligned}$$

where $\lambda = \frac{1}{2}u^2 = \Psi(u)$. The solutions to this latter equation are $u(\lambda) = \pm\sqrt{2\lambda}$ so matching terms in (16) and (17) we obtain

$$h(\lambda) = \widetilde{C}\Gamma(-\widetilde{Y})\left[2\widetilde{G}^{\widetilde{Y}} - \left(\widetilde{G}+iu(\lambda)\right)^{\widetilde{Y}} + \left(\widetilde{G}-iu(\lambda)\right)^{\widetilde{Y}}\right].$$

The RHS can be solved explicitly: taking either root

$$\widetilde{G} \pm iu(\lambda) = \widetilde{G} \pm i\sqrt{2\lambda},$$

we have the Bernstein function

$$\begin{aligned}
h(\lambda) &= \widetilde{C}\Gamma(-\widetilde{Y})\left[2\widetilde{G}^{\widetilde{Y}} - \left(\widetilde{G}+i\sqrt{2\lambda}\right)^{\widetilde{Y}} + \left(\widetilde{G}-i\sqrt{2\lambda}\right)^{\widetilde{Y}}\right] \\
&= 2\widetilde{C}\Gamma(-\widetilde{Y})\left[\widetilde{G}^{\widetilde{Y}} - \left(2\lambda+\widetilde{G}^2\right)^{\widetilde{Y}/2}\cos\left(\widetilde{Y}\arctan\left(\frac{\sqrt{2\lambda}}{\widetilde{G}}\right)\right)\right].
\end{aligned}$$

Appendix 2 (Existence and uniqueness of MLE)

Let G be the set of matrices satisfying the conditions of (5)–(6) and let $\wp = \{\exp(\Delta\mathcal{L}) : \mathcal{L} \in G\}$ be the set of transition matrices corresponding to Δ time observation of a continuous–time Markov chain. Let P_X^Δ be the empirical transition matrix calculated from the data. If $P_X^\Delta \in \wp$, it is known that there exists an $\widetilde{\mathcal{L}} \in G$ such that $\exp(\widetilde{\mathcal{L}}) = P_X^\Delta$ and the MLE attains its maximum at $\widetilde{\mathcal{L}}$. Hence the estimation problem appears straightforward. However there are two complications:

- It is sometimes possible for the MLE to exist even when $P_X^\Delta \notin \wp$, so determination of the existence of the MLE requires characterisation of the set $\wp$
- The matrix exponential function is not an injection on all parts of its domain so the MLE need not be unique.

Existence Results on the existence of the MLE may be summarised as (see [BS1]):

- If $P_X^\Delta \in \wp$ there exists an $\widetilde{\mathcal{L}} \in G$ such that $\exp(\widetilde{\mathcal{L}}) = P_X^\Delta$ and the MLE attains its maximum at $\widetilde{\mathcal{L}}$
- If $P_X^\Delta \notin \wp$ then either:

  (a) the MLE does not exist and the likelihood function has no maximum in G; or

  (b) the MLE exists and satisfies $\exp(\widetilde{\mathcal{L}}) \in \delta\wp$, where $\delta\wp$ denotes the boundary of $\wp$ relative to $\wp_2 = \left\{P_X^\Delta : \det(P_X^\Delta) > 0\right\}$[19]
- If the Markov chain is ergodic with generator satisfying $\exp(\mathcal{L}) \in \operatorname{int}(\wp)$ then the MLE exists with probability one as the sample size tends to infinity.

The problem of characterising the set $\wp$ is known as the *embedding problem* for finite Markov chains. For the case $M = 2$ it is known that $\wp = \wp_2$, while for $M \geqslant 3$ it is known that $\wp$ is a (relatively) closed subset of $\wp_2$ with non–convex geometric shape and dimension $M(M-1)$. An explicit description for $M = 3$ is available, however for the dimension of our problem ($M = 8$) an explicit description does not appear possible. [BS1] suggest using the convergence of their EM algorithm to determine existence of the MLE . They construct a sequence of estimates of $\mathcal{L}$ such that convergence to a stationary point implies existence of the MLE. If on the other hand $\det(\exp(\mathcal{L})) \longrightarrow 0$ it suggests the MLE does not exist. Using their (adapted) algorithm on our data yields convergence to a stationary point, providing some validation for our ML approach.

Uniqueness Uniqueness is concerned with determining whether there are two or more generators $\mathcal{L} \in G$ for which the discrete sample $\{X_t : t \in \omega\}$ has the same distribution. This amounts to asking whether the real logarithm of the transition matrix is unique or, in the context of ML estimation, whether the parameter set is identifiable. A condition for the uniqueness of the generator for P_X^Δ is the *Cuthbert criterion* (see [BS1], p.398):

$$\inf_i \left\{P_X^\Delta(i,t;i,t+\Delta)\right\} \det\left(P_X^\Delta(t;t+\Delta)\right) - \exp(-\pi)\prod_{i=1}^{M} P_X^\Delta(i,t;i,t+\Delta) > 0, \quad \vee\, t, \Delta$$

which may be tested readily.

[19] This boundary is defined as $\delta\wp = (\cup_{i\neq j} E_{ij}) \cup \varepsilon$ for E_{ij} a non–empty subset of the set of exponentials of intensity matrices with $q_{ij} = 0$ and ε a non–empty subset of the $M \times M$ transition matrices with fewer than M distinct eigenvalues.

References

[AC1] Albanese, C. and O. Chen (2006), ‘Implied Migration Rates From Credit Barrier Models’, *Journal of Banking and Finance*, 30(2), p.607-626.

[AC2] Albanese, C. and O. Chen (2005), ‘Discrete Credit Barrier Models’, *Quantitative Finance*, 5, p.247-256.

[AK] Albanese, C. and A. Kuznetsov (2003), ‘Discretization Schemes for Subordinated Processes’, *preprint.*

[ACDV] Albanese, C., O. Chen, A. Dalessandro, and A. Vidler (2006), ‘Dynamic Credit Correlation Modeling’, *preprint.*

[A] Albanese, C. (2007), ‘Operator Methods, Abelian Processes and Dynamic Conditioning’, *preprint.*

[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.*

[CGMY] Carr, P., H. Geman, D. Madan and M. Yor (2002), ‘The Fine Structure of Asset Returns: An Empirical Investigation’, *Journal of Business*, 75, p.305-332.

[CT] Cont, R., and P., Tankov (2004), *Financial Modelling with Jump Processes*. Chapman and Hall/CRC Financial Mathematics series.

[GH] Gordy, M. and E. Heitfield (2001), ‘Of Moody’s and Merton: A Structural Model of Bond Rating Transitions’, Board of Governors of the Federal Reserve System.

[MY] Madan, D. and M. Yor (2005), ‘CGMY and Meixner Subordinators are Absolutely Continuous with respect to One Sided Stable Subordinators’, *preprint.*