

# Calibration of PD term structures: to be Markov or not to be

*A common discussion in credit risk modelling is the question of whether term structures of default probabilities can be satisfactorily modelled by Markov chain techniques. Christian Bluhm and Ludger Overbeck show that empirical multi-year default frequencies can be interpolated well by continuous-time Markov chains if the Markov chain is allowed to evolve with non-homogeneous behaviour in time*

approaches – most often based on homogeneous<sup>2</sup> chains – dominate current market practice. A landmark paper in this area is the work by Jarrow, Lando & Turnbull (1997). Further research has been done by various authors, such as Kadam & Lenk (2005), Lando & Skodeberg (2002), Sarfaraz, Cohen & Liberos (2004), Schuermann & Jafry (2003a and 2003b) and Trueck & Oezturkmen (2003). A new approach using Markov mixtures has been presented recently by Frydman & Schuermann (2005).

In Markov chain theory (see Noris, 1998), one distinguishes between discrete-time and continuous-time chains. For instance, a discrete-time chain can be specified by a one-year migration or transition matrix  $M$  generating multi-year transitions via powers  $(M^k)_{k \geq 1}$  of  $M$ . The corresponding (yearly) discrete-time PD term structures are given by<sup>3</sup>:

$$p_R^{(k)} = (M^k)_{\text{row}(R),8} \quad (k = 1, 2, 3, \dots)$$

where  $\text{row}(R)$  denotes the row in the migration matrix  $M$  corresponding to rating  $R$ . Continuous-time chains are specified by a Q-matrix<sup>4</sup>  $Q$  such that  $\exp(tQ)$  defines the migration matrix for the time interval  $[0, t]$ , where  $\exp(\cdot)$  denotes the matrix exponential. Continuous-time PD term structures corresponding to a generator  $Q$  are given by:

$$p_R^{(t)} = (\exp(tQ))_{\text{row}(R),8} \quad (t \geq 0) \quad (1)$$

Continuous-time Markov chains are superior to discrete-time chains because they allow for a consistent way to measure migrations and PDs for time horizons between yearly time grid points. If for a discrete-time chain defined by a one-year migration matrix  $M$  we find a generator  $Q$  with:

$$M = \exp(Q) \quad (2)$$

one says that the discrete-time chain can be embedded into a continuous-time chain. In general, we can only expect to find approximative embeddings (see Israel, Rosenthal & Wei, 2001, Jarrow, Lando & Turnbull, 1997, Kreinin & Sidelnikova, 2001, and Bluhm & Overbeck, 2003, chapter 6). In Bluhm & Overbeck (2006), section 2.3.1, we discuss an example of a generator  $Q$  almost perfectly fitted to a given one-year migration matrix from S&P (see Appendix II).

The problem is that we find that a well-fitted generator nevertheless can generate model-implied PD term structures significantly deviating from observed multi-year default frequencies.

The probability of default (PD) for a client is a fundamental risk parameter in credit risk management. It is common practice to assign to every rating grade in a bank's master scale a one-year PD in line with regulatory requirements (see Basel Committee on Banking Supervision, 2004). Table A shows an example for default frequencies assigned to rating grades from Standard & Poor's (S&P).

Moreover, credit risk modelling concepts such as dependent default times, multi-year credit pricing and multi-horizon economic capital require more than just one-year PDs. For multi-year credit risk modelling, banks need a whole term structure  $(p_R^{(t)})_{t \geq 0}$  of (cumulative) PDs for every rating grade  $R$  (see, for example, Bluhm, Overbeck & Wagner, 2003, for an introduction to PD term structures and Bluhm & Overbeck, 2006, for their application to structured credit products).

Every bank has its own (proprietary) way to calibrate PD term structures<sup>1</sup> to their internal and external data. For the generation of PD term structures, various Markov chain

A. One-year default frequencies assigned to S&P ratings	
	Default frequencies
AAA	0.00%
AA	0.01%
A	0.04%
BBB	0.29%
BB	1.28%
B	6.24%
CCC	32.35%

Note: see Standard & Poor's (2005), table 9

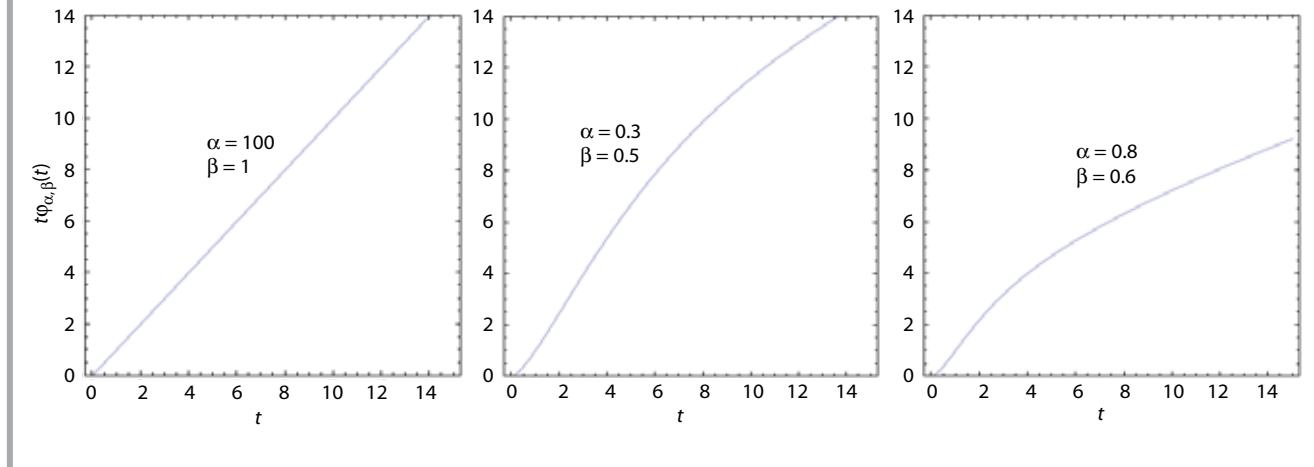
<sup>1</sup> In the literature, PD term structures are sometimes called credit curves

<sup>2</sup> A Markov chain is called homogeneous if transition probabilities do not depend on time

<sup>3</sup> The second index '8' in the notation below refers to the eighth (default) column in S&P's  $8 \times 8$ -migration matrices

<sup>4</sup> A square matrix  $Q$  is a Q-matrix/generator if  $\sum_{j=1}^N q_{ij} = 0 \forall i$ ,  $q_{ii} \leq 0 \forall i$  and  $q_{ij} \geq 0 \forall i \neq j$

1 Illustration of the functions  $\varphi_{\alpha,\beta}$  for different  $\alpha$  and  $\beta$



In this article, we address this problem, not by rejecting the Markov assumption but by dropping the homogeneity assumption and working with non-homogeneous continuous-time Markov chains (NHCTMCs). Our results in figure 2 show that in the context of PD term structure calibration the Markov assumption is not as questionable as people sometimes claim. In fact, dropping the homogeneity assumption provides sufficient flexibility to calibrate a Markov process to empirical migration and default frequencies with convincing quality. Therefore, we answer the question raised in the title of this article by ‘to be Markov’, but ‘not homogeneous’.

#### Calibration of an NHCTMC for PD term structures

In the following, we construct an NHCTMC, which we use for the generation of PD term structures. In Appendix I, we provide some comments on the stochastic rationale of the approach. Appendix III provides information on the data underlying figure 2 used for the calibration of the model.

The starting point is the generator  $Q = (q_{ij})_{1 \leq i, j \leq 8}$  from table D explained in the example in Appendix II. In contrast to the time-homogeneous case, we no longer assume that the transition rates  $q_{ij}$  are constant over time, as is the case for homogeneous continuous-time Markov chains (HCTMCs). Instead, we replace the homogeneous generator  $Q$  leading to migration matrices  $\exp(tQ)$  for the time interval  $[0, t]$  by the time-dependent generator:

$$Q_t = \Phi(t) \times Q \quad (3)$$

B. Optimal choices for $\alpha$ and $\beta$ vectors		
	$\alpha$	$\beta$
AAA	0.34	0.89
AA	0.11	0.26
A	0.81	0.65
BBB	0.23	0.30
BB	0.32	0.56
B	0.23	0.40
CCC	2.15	0.46

where  $\times$  denotes matrix multiplication and  $\Phi(t) = (\varphi_{ij}(t))_{1 \leq i, j \leq 8}$  is the diagonal matrix in  $\mathbb{R}^{8 \times 8}$  with:

$$\varphi_{ij}(t) = \begin{cases} 0 & \text{if } i \neq j \\ \varphi_{\alpha_i, \beta_i}(t) & \text{if } i = j \end{cases} \quad (4)$$

Because  $\Phi(t)$  is a diagonal matrix,  $Q_t$  is a Q-matrix (scaling rows of a Q-matrix gives a Q-matrix). The functions  $\varphi_{\alpha, \beta}$  with respect to parameters  $\alpha$  and  $\beta$  are defined as follows. Set:

$$\varphi_{\alpha, \beta} : [0, \infty) \rightarrow [0, \infty), t \mapsto \varphi_{\alpha, \beta}(t) = \frac{(1 - e^{-\alpha t}) t^{\beta-1}}{1 - e^{-\alpha}}$$

for non-negative constants  $\alpha$  and  $\beta$ . Figure 1 illustrates the functions  $t \mapsto t\varphi_{\alpha, \beta}(t)$ . They have the following properties:

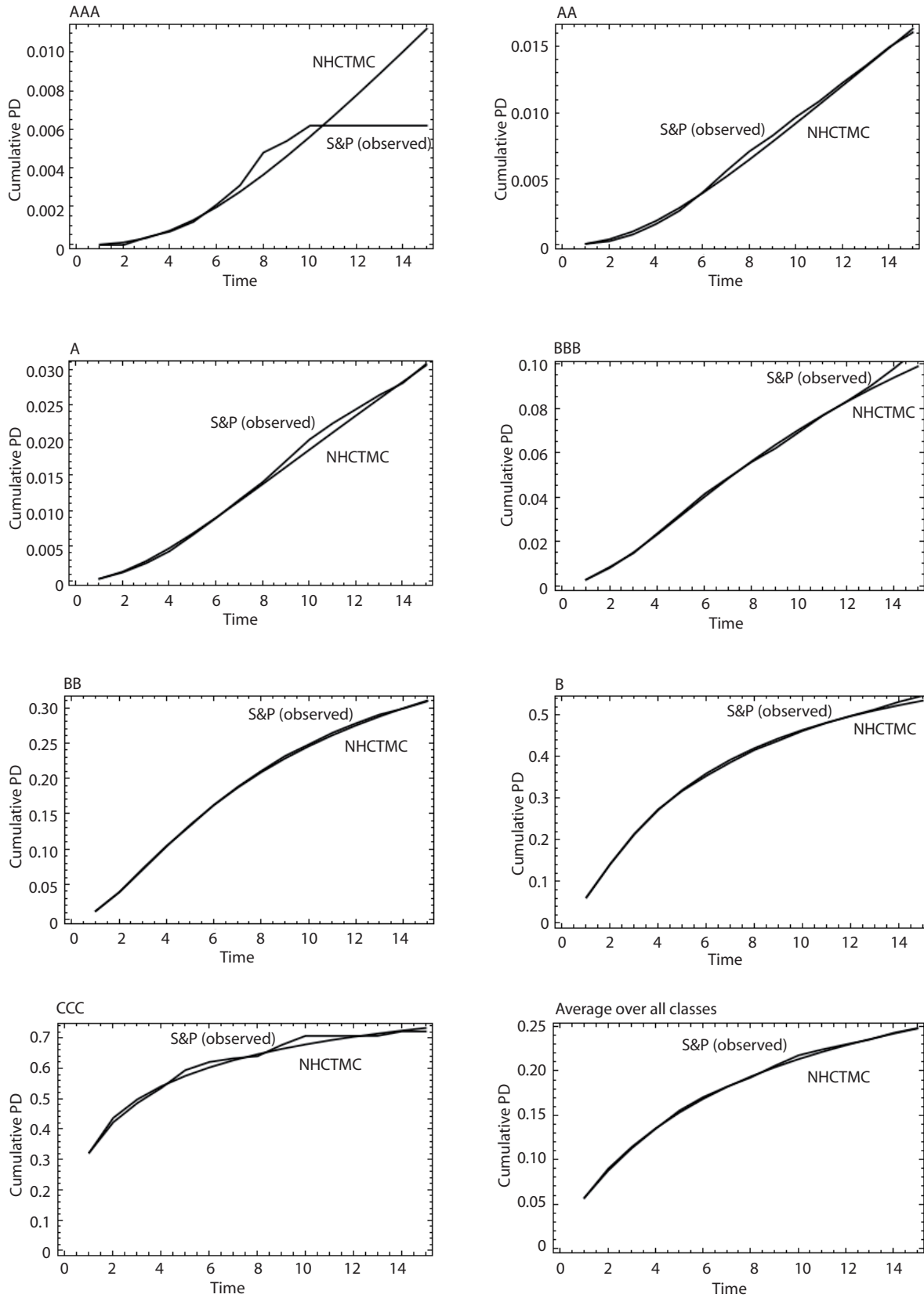
- 1.  $\varphi_{\alpha, \beta}(1) = 1$  (normalised at time  $t = 1$ ; holds by construction) and
- 2.  $t\varphi_{\alpha, \beta}(t)$  is increasing in the time parameter  $t \geq 0$ .
- 3. In the numerator of  $t\varphi_{\alpha, \beta}$ , the first factor  $(1 - e^{-\alpha t})$  is the distribution function of an exponentially distributed random variable with intensity  $\alpha$ ; the second factor, namely  $t^\beta$ , can be considered<sup>5</sup> as a convexity or concavity adjustment term, respectively.

Property 1 is necessary to guarantee consistency at time  $t = 1$  between the given one-year migration matrix  $M = \exp(Q)$  and its non-homogeneous modification  $\exp(Q_t)$ . Property 2 is necessary for keeping the direction of time (moving into the future and not into the past). Property 3 points out that the special form of the functions  $\varphi_{\alpha, \beta}$ , while it has the flavour of an *ad hoc* parameterisation, is not an arbitrary choice but is related to well-known functions used in probability theory. Below in the text we summarise our findings and at this point emphasise that our model is an interpolating approach: it relies on a suitable parametric framework to interpolate empirically given cumulative default rates. The attribute ‘suitable’ does not mean ‘unique’ or ‘naturally given’. It just means that we found functions  $\varphi_{\alpha, \beta}$  sufficiently reasonable to be applied in the definition of  $Q_t$ , for which we get very good interpolation results (see figure 2).

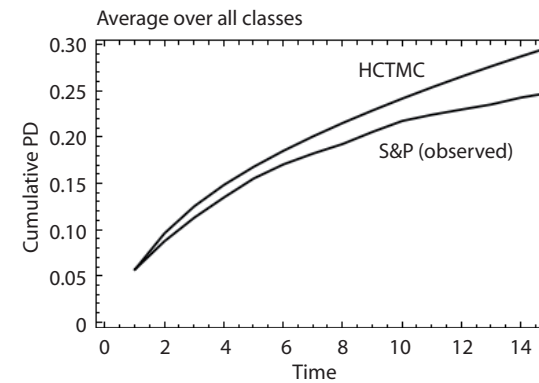
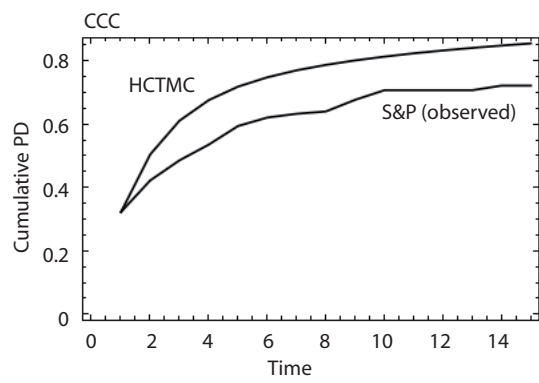
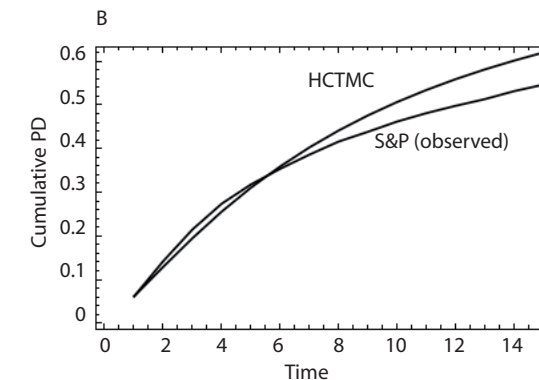
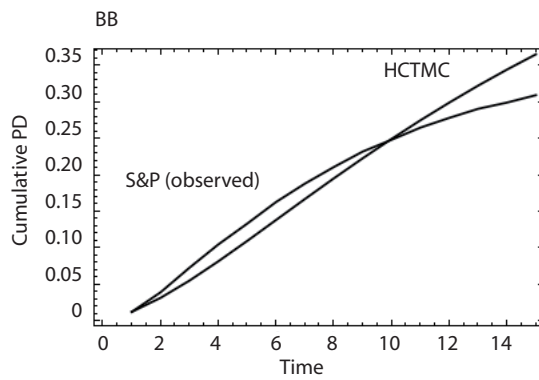
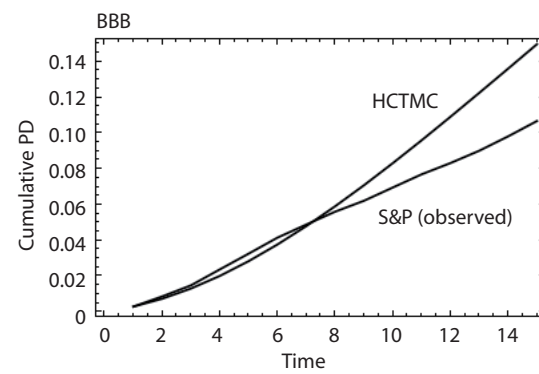
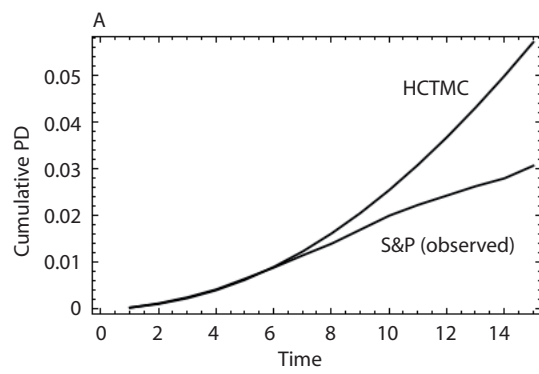
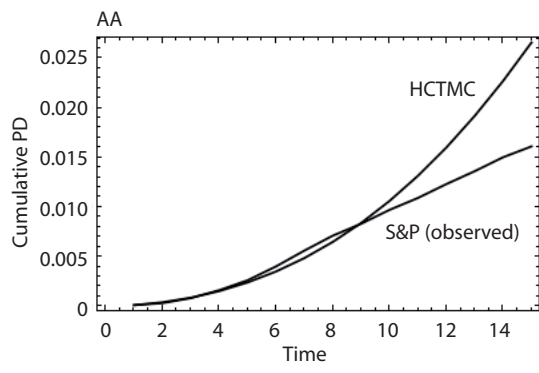
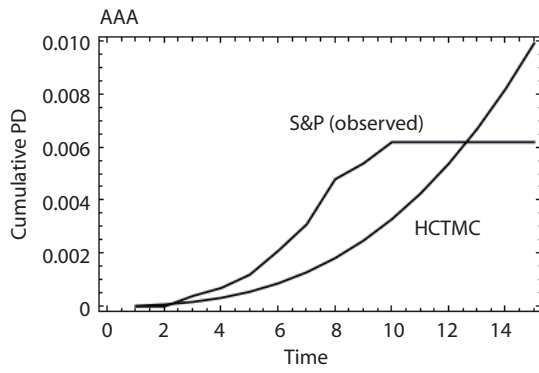
Since the functional form of the time-dependent generators

<sup>5</sup> Note that  $\varphi_{\alpha, \beta}$  exhibits some similarity to the gamma distribution, which is frequently applied in the context of queuing theory and reliability analysis

2 PD term structures based on a non-homogeneous continuous-time Markov chain (NHCTMC) approach



3 PD term structures based on a homogeneous continuous-time Markov chain (HCTMC) approach



Appendix I: stochastic rationale of the NHCTMC approach

In this appendix, we briefly comment on the stochastic rationale of our approach. For the sake of convenient notation, let us denote by  $\Psi(t)$  the diagonal matrix with diagonal elements:

$$\Psi_{ii}(t) = t\phi_{\alpha_i, \beta_i}(t) \quad (i = 1, \dots, 8; t \geq 0)$$

The transition matrix  $M_t$  in (5) for the time period  $[0, t]$  can then be written as:

$$M_t = \exp(\Psi(t) \times Q) \quad (t \geq 0) \quad (6)$$

Writing the exponential matrix as a power series and using the typical Markov kernel notation  $P_{0,t} = M_t$ , term-by-term differentiation yields:

$$\begin{aligned} \frac{\partial}{\partial t} P_{0,t} &= \sum_{k=0}^{\infty} \frac{\partial (\Psi(t) \times Q)^k}{\partial t} \frac{1}{k!} \\ &= \sum_{k=1}^{\infty} \left( \frac{\partial}{\partial t} \Psi(t) \times Q \right) \times \frac{(\Psi(t) \times Q)^{k-1}}{(k-1)!} \\ &= \left( \frac{\partial}{\partial t} \Psi(t) \times Q \right) \times P_{0,t} \end{aligned} \quad (7)$$

Because  $\Psi(t)$  is a diagonal matrix:

$$\frac{\partial \Psi(t)}{\partial t}$$

is the diagonal matrix with entries  $\Psi'_{ii}(t)$ . Therefore, the matrix:

$$\frac{\partial \Psi(t)}{\partial t} \times Q$$

is a Q-matrix, arguing in the same way as above where we said that  $\Psi(t) \times Q$  is a Q-matrix and taking into account that  $\Psi'_{ii}(t) \geq 0$  at all times<sup>1</sup>. As a consequence of general Markov theory (see Ethier & Kurtz, 2005, theorem 7.3 in chapter 4, Lando & Skodeberg, 2002, and Schoenbucher, 2005), equation (7) is part of the forward equation of a non-homogeneous Markov chain  $(X_t)_{t \geq 0}$  with state space  $\{1, 2, \dots, 8\}$  corresponding to a semigroup  $\{P_{s,t} \mid 0 \leq s \leq t\}$  satisfying the Kolmogorov backward and forward equations associated with the family:

$$\left\{ \frac{\partial \Psi(t)}{\partial t} \times Q \mid t \geq 0 \right\}$$

defining the infinitesimal generator of the Markov process. Equation (7) shows that the non-homogeneous continuous-time Markov chain  $(X_t)_{t \geq 0}$  induces the PD term structures illustrated in figure 2 via the default column of kernel-based transition matrices  $P_{0,t} = M_t = \exp(\Psi(t) \times Q)$ .

<sup>1</sup> We have  $(1 - \exp(-\alpha))\Psi'_{ii}(t) = \alpha \exp(-\alpha)t^{\beta} + (1 - \exp(-\alpha))\beta t^{\beta-1} \geq 0$  for all  $t \geq 0$

Appendix II: example of a generator well fitted to migrations but poorly fitting observed default frequencies

The following example is taken from Bluhm & Oberbeck (2006), section 2.3.1. We start with the adjusted<sup>1</sup> average one-year migration matrix  $M = (m_{ij})_{i,j=1,\dots,8}$  shown in table C, based on table 9 in Standard & Poor's (2005), which reports on average historic annual migration rates observed by S&P.

Table D shows the calibration of a generator (Q-matrix)  $Q$  based on the log-expansion of  $M$  and a so-called diagonal adjustment. The method we used is a standard procedure for calibrating generator matrices (see, for example, Kreinin & Sidelnikova, 2001). The approximation of the original matrix  $M$  by  $\exp(Q)$  is very much acceptable, based on the following small approximation error:

$$\|M - \exp(Q)\|_2 = \sqrt{\sum_{i,j=1}^8 (m_{ij} - (\exp(Q))_{ij})^2} \approx 0.00023$$

We can generate PD term structures based on the homogeneous continuous-time Markov chain generated by  $Q$  via:

$$p_R^{(t)} = (\exp(tQ))_{row(R),8} \quad (t \geq 0)$$

as in equation (1) in the introduction. Figure 3 compares the result of this calculation with empirically observed default frequencies, also taken from the S&P report (2005). The picture we get is quite disappointing: despite the good fit of the Q-matrix exponential to  $M$ , empirical default frequencies are not reflected by the model-implied PD term structures derived from the chosen non-homogeneous Markov chain approach. However, figure 2 shows that the picture can completely change to the positive if we drop the homogeneity assumption.

C. Modified average one-year migration matrix  $M$  (%)

	AAA	AA	A	BBB	BB	B	CCC	D
AAA	91.68	7.69	0.48	0.09	0.06	0.00	0.00	0.00
AA	0.62	90.49	8.10	0.60	0.05	0.11	0.02	0.01
A	0.05	2.16	91.34	5.77	0.44	0.17	0.03	0.04
BBB	0.02	0.22	4.07	89.72	4.68	0.80	0.20	0.29
BB	0.04	0.08	0.36	5.78	83.38	8.05	1.03	1.28
B	0.00	0.07	0.22	0.32	5.84	82.53	4.78	6.24
CCC	0.09	0.00	0.36	0.45	1.52	11.17	54.06	32.35
D	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00

Note: based on S&P data (Standard & Poor's, 2005)

D. Approximative generator  $Q$  for  $M$  (%)

	AAA	AA	A	BBB	BB	B	CCC	D
AAA	-8.73	8.44	0.15	0.07	0.06	0.00	0.00	0.00
AA	0.68	-10.13	8.91	0.38	0.02	0.12	0.02	0.00
A	0.05	2.37	-9.31	6.37	0.33	0.15	0.03	0.02
BBB	0.02	0.19	4.49	-11.17	5.39	0.65	0.22	0.21
BB	0.04	0.08	0.24	6.68	-18.71	9.63	1.15	0.88
B	0.00	0.08	0.22	0.12	7.01	-20.06	7.09	5.55
CCC	0.13	0.00	0.47	0.54	1.61	16.59	-62.22	42.88
D	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

<sup>1</sup> Rows are normalised to get a stochastic matrix and the PD for AAA is set equal to 0.2 basis points, based on a linear regression of PDs on a logarithmic scale

$(Q_t)_{t \geq 0}$  is fixed by equation (4), the generators  $Q_t$  are solely determined by two vectors  $(\alpha_1, \dots, \alpha_8)$  and  $(\beta_1, \dots, \beta_8)$  in  $[0, 1]^8$ . For any chosen pair of parameter vectors, we can now generate a term structure of cumulative PDs by calculating migration matrices  $M_t$  for the time period  $[0, t]$  via:

$$M_t = \exp(tQ_t) \quad (t \geq 0) \quad (5)$$

The last step we have to make is to optimise<sup>6</sup>  $(\alpha_1, \dots, \alpha_8)$  and  $(\beta_1, \dots, \beta_8)$  for the best fit of the term structure generated by the default column of the migration matrices (5) to S&P's (Standard & Poor's, 2005) empirical term structures of default frequencies. As a distance measure for our optimisation, we use the mean-squared distance. Table B and figure 2 show the outcome of best-fitting  $\alpha$ - and  $\beta$ -vectors as well as the resulting (NHCTMC-implied) credit curves in comparison with the empirically observed multi-year default frequencies from S&P.

Summarising, we parameterised a Markov chain approach for calibrating model-implied PD term structures in continuous time, which fit empirical observed default frequencies very well. Crucial to our approach was the acceptance of a non-homogeneous time evolution of the chain. The choice of parameters involved a one-year migration matrix as well as observed default frequencies. As mentioned before, it is an interpolating and not an extrapolating approach because the fit can only be exercised within the time window of observations.<sup>7</sup> ■

Christian Bluhm heads the credit portfolio management unit at Credit Suisse in Zurich. Ludger Overbeck is a professor of mathematics at the University of Giessen. Email: christian.bluhm@credit-suisse.com, ludger.overbeck@web.de

<sup>6</sup> Note that  $\alpha_8$  and  $\beta_8$  have no meaning and can be fixed at some arbitrary value  
<sup>7</sup> In contrast to homogeneous Markov chains, where extrapolation can be done quite naturally

### Appendix III: data and calibration underlying figure 2

Figure 2 shows the nice model-based fit of empirical observed cumulative default frequencies we get from the non-homogeneous continuous-time Markov chain (NHCTMC) approach described in the main article. In this appendix, we briefly comment on underlying data and the model calibration leading to figure 2.

As explained, the parametric model we use has two major components: a time-homogeneous generator  $Q$  (see Appendix II for comments and references) and the functions  $\phi_{\alpha, \beta}$  for which we need to calibrate  $\alpha$ - and  $\beta$ -vectors. This is done as follows.

In the same Standard & Poor's report (2005) where we found migration data for the calibration of  $Q$ , we also find empirical observed cumulative average default rates (15 yearly cumulative values based on observations of default rates from the years 1981–2004); see table 11 in the S&P report. In our test calculations, we also experimented with other data sets. In all considered cases we

were able to calibrate  $\alpha$ - and  $\beta$ -vectors for our NHCTMC model, leading to comparably satisfying interpolation results as in figure 2. The model contains a sufficient degree of flexibility in its parameterisation to make the interpolation well fitting. The quantity to be minimised for the determination of  $\alpha$ - and  $\beta$ -vectors is:

$$distance \left[ \left( \hat{p}_R^{(t)} \right)_{t,R}, \left( (M_t)_{row(R),s} \right)_{t,R} \right] = small \quad (8)$$

where  $M_t$  is defined in equation (5) and  $\hat{p}_R^{(t)}$  denotes the empirical S&P cumulative average default rate in year  $t$  for rating grade  $R$ . As already mentioned, as a distance measure for the optimisation problem specified by equation (8) we used the mean-squared distance. The optimisation can easily be done with standard mathematical software such as Mathematica or Matlab.

### References

- |   |  |  |  |
|---|--|--|--|
| <p><b>Basel Committee on Banking Supervision, 2004</b><br/> <i>International convergence of capital measurement and capital standards</i><br/>         Bank for International Settlements, June</p> <p><b>Bluhm C and L Overbeck, 2006</b><br/> <i>Structured credit portfolio analysis, baskets &amp; CDOs</i><br/>         Chapman &amp; Hall/CRC Financial Mathematics Series, CRC Press</p> <p><b>Bluhm C, L Overbeck and C Wagner, 2003</b><br/> <i>An introduction to credit risk modeling</i><br/>         Chapman &amp; Hall/CRC Financial Mathematics Series, second reprint, CRC Press</p> <p><b>Ethier S and T Kurtz, 2005</b><br/> <i>Markov processes characterization and convergence</i><br/>         John Wiley and Sons</p> <p><b>Frydman H and T Schuermann, 2005</b><br/> <i>Credit rating dynamics and Markov mixture models</i><br/>         Working paper, June</p> | <p><b>Israel R, J Rosenthal and J Wei, 2001</b><br/> <i>Finding generators for Markov chains via empirical transition matrices with application to credit ratings</i><br/>         Mathematical Finance 11(2), pages 245–265</p> <p><b>Jarrow R, D Lando and S Turnbull, 1997</b><br/> <i>A Markov model for the term structure of credit risk spreads</i><br/>         Review of Financial Studies 10, pages 481–523</p> <p><b>Kadam A and P Lenk, 2005</b><br/> <i>Heterogeneity in ratings migration</i><br/>         Working paper, October</p> <p><b>Kreinin A and M Sidelnikova, 2001</b><br/> <i>Regularization algorithms for transition matrices</i><br/>         Algo Research Quarterly 4(1/2), pages 25–40</p> | <p><b>Lando D and T Skodeberg, 2002</b><br/> <i>Analyzing rating transitions and rating drift with continuous observations</i><br/>         Journal of Banking and Finance 26(2–3), pages 423–444</p> <p><b>Noris J, 1998</b><br/> <i>Markov chains</i><br/>         Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press</p> <p><b>Sarfaraz A, M Cohen and S Libreros, 2004</b><br/> <i>Use of transition matrices in risk management and valuation</i><br/>         Fair Isaac white paper, September</p> <p><b>Schoenbucher P, 2005</b><br/> <i>Portfolio losses and the term structure of loss transition rates: a new methodology for the pricing of portfolio credit derivatives</i><br/>         NCCR FinRisk working paper 264, September</p> | <p><b>Schuermann T and Y Jafry, 2003a</b><br/> <i>Measurement and estimation of credit migration matrices</i><br/>         Working paper 03-08, Wharton School Center for Financial Institutions, University of Pennsylvania</p> <p><b>Schuermann T and Y Jafry, 2003b</b><br/> <i>Metrics for comparing credit migration matrices</i><br/>         Working paper 03-09, Wharton School Center for Financial Institutions, University of Pennsylvania</p> <p><b>Standard &amp; Poor's, 2005</b><br/> <i>Annual global corporate default study: corporate defaults poised to rise in 2005</i><br/>         S&amp;P Global Fixed Income Research, January</p> <p><b>Trueck S and E Oezturkmen, 2003</b><br/> <i>Adjustment and application of transition matrices in credit risk models</i><br/>         Working paper, University of Karlsruhe, September</p> |
|---|--|--|--|