Calibration of PD Term Structures: To Be Markov Or Not To Be

Christian Bluhm (Credit Suisse) and Ludger Overbeck (University of Giessen)

November 19, 2006

Abstract. Term structures of default probabilities are omnipresent in credit risk modeling: time-dynamic credit portfolio models, default times, and multi-year pricing models, they all need the time evolution of default probabilities as a basic model input. Although people tend to believe that from an economic point of view the Markov property as underlying model assumption is kind of questionable it seems to be common market practice to model PD term structures via Markov chain techniques. In this paper we illustrate that the Markov assumption carries us quite far if we allow for *nonhomogeneous* time behaviour of the Markov chain generating the PD term structures. As a 'proof of concept' we calibrate a *nonhomogeneous* time-continuous Markov chain to observed one-year rating migrations *and* multi-year default frequencies, hereby achieving convincing approximation quality.

1 Markov Chains in Credit Risk Modeling

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 masterscale a one-year PD in line with regulatory requirements; see [1]. Table 1 shows an example for default frequencies assigned to rating grades from Standard and Poor's (S&P).

	D
AAA	0.00%
AA	0.01%
A	0.04%
BBB	0.29%
BB	1.28%
B	6.24%
CCC	32.35%

Table 1: One-year default frequencies (D) assigned to S&P ratings; see [17], Table 9.

Moreover, credit risk modeling concepts like dependent default times, multi-year credit pricing, and multi-horizon economic capital require more than just one-year PDs. For multi-year credit risk modeling, banks need a whole *term structure* $(p_R^{(t)})_{t\geq 0}$ of (cumulative) PDs for every rating grade R; see, e.g., [2] for an introduction to PD term structures and [3] for their application to structured credit products.

Every bank has its own (proprietary) way to calibrate PD term structures[1] to bank-internal and external data. A look into the literature reveals that for the generation of PD term structures various Markov chain approaches, often based on time-homogeneous chains, dominate current market practice. A landmarking paper in this direction is the work by Jarrow, Lando, and Turnbull [7]. Further research has been done by various authors, see, e.g., Kadam [8], Lando [10], Sarfaraz et al. [12], Schuermann and Jafry [14, 15], Trueck and Oezturkmen [18], just to mention a few examples. A new approach via Markov mixtures has been presented recently by Frydman and Schuermann [5].

In Markov chain theory (see [11]) one distinguishes between time-discrete and time-continuous chains. For instance, a time-discrete chain can be specified by a one-year migration or transition

[1] In the literature, PD term structures are sometimes called *credit curves*.

matrix M generating multi-year transitions via powers $(M^k)_{k\geq 1}$ of M. The corresponding (yearly) time-discrete PD term structures are given by

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

where $row(R)$ denotes the row in the migration matrix M corresponding to rating R. Time-continuous chains are specified by a Q-matrix[2] Q such that $\exp(tQ)$ defines the migration matrix for the time interval $[0, t]$, where $\exp(\cdot)$ denotes the matrix exponential. Time-continuous PD term structures corresponding to a generator Q are given by

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

Time-continuous Markov chains are superior to time-discrete Markov 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 chain defined by a one-year migration matrix M we find a generator Q such that

$$M \;=\; \exp(Q), \tag{2}$$

one says that the *time-discrete chain can be embedded into a continuous-time chain.* In general, we can only expect to find approximative embeddings; see Israel, Rosenthal, and Wei [6], Jarrow, Lando, and Turnbull [7], Kreinin and Sidelnikova [9], and [2], Chapter 6. In [3], 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.* In this paper, we address this problem, *not by rejecting the Markov assumption but by dropping the homogeneity*[3] *assumption w.r.t. time.* Our results in Figure 2 show that in the context of PD term structure calibration the Markov assumption indeed is not as wrong as people sometimes claim. In fact, dropping the time-homogeneity assumption provides sufficient flexibility to calibrate a Markov process to empirical migration and default frequencies with convincing quality. Therefore, we claim that the anser to the question raised in the title of this paper is **'to be Markov'**, but **'not time-homogeneous'**.

2 Calibration of a NHCTMC for PD term structures

In the sequel, we construct a 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.

Starting point for our construction is the generator $Q = (q_{ij})_{1\leq i,j\leq 8}$ from Table 4. But now we do no longer assume that the transition rates q_{ij} are constant over time, leading to a HTCMC. Instead, we replace the time-homogeneous generator Q leading to migration matrices $\exp(tQ)$ for the time interval $[0, t]$ by the time-dependent generator

$$Q_t \;=\; \Phi(t) * Q \tag{3}$$

where '$*$' 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

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

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 $\varphi_{\alpha,\beta}$ w.r.t. parameters α and β are defined as follows. Set

$$\tilde{\varphi}_{\alpha,\beta} : [0, \infty) \to [0, \infty), \; t \mapsto \tilde{\varphi}_{\alpha,\beta}(t) = (1 - e^{-\alpha t})t^{\beta-1}$$

[2] A square matrix Q is a Q-matrix/generator if $\sum_{j=1}^{N} q_{ij} = 0 \;\forall\; i$, $0 \leq -q_{ii} < \infty \;\forall\; i$, and $q_{ij} \geq 0 \;\forall\; i \neq j$.

[3] A Markov chain is *time-homogeneous* if transition probabilities (the generator) do not depend on time.

for nonnegative constants α and β. We want to normalize the functions in a way such that at time $t = 1$ the functions take on the value 1. Therefore, we define $\varphi_{\alpha,\beta}$ as

$$\varphi_{\alpha,\beta} : [0, \infty) \to [0, \infty),\ t \mapsto \varphi_{\alpha,\beta}(t) = \frac{\tilde{\varphi}_{\alpha,\beta}(t)}{\tilde{\varphi}_{\alpha,\beta}(1)}.$$

Figure 1 illustrates the functions $t \mapsto t\varphi_{\alpha,\beta}(t)$. They have the following properties:

1. $\varphi_{\alpha,\beta}(1) = 1$ (normalized at time $t = 1$; holds by construction) and
2. $t\varphi_{\alpha,\beta}(t)$ is increasing in the time parameter $t \geq 0$.
3. The first part of $\tilde{\varphi}_{\alpha,\beta}$, namely $(1 - e^{-\alpha t})$, is the distribution function of an *exponentially distributed* random variable with intensity α; the second part of $\tilde{\varphi}_{\alpha,\beta}$, namely t^{β}, can be considered[4] as a *convexity or concavity adjustment* term, respectively.

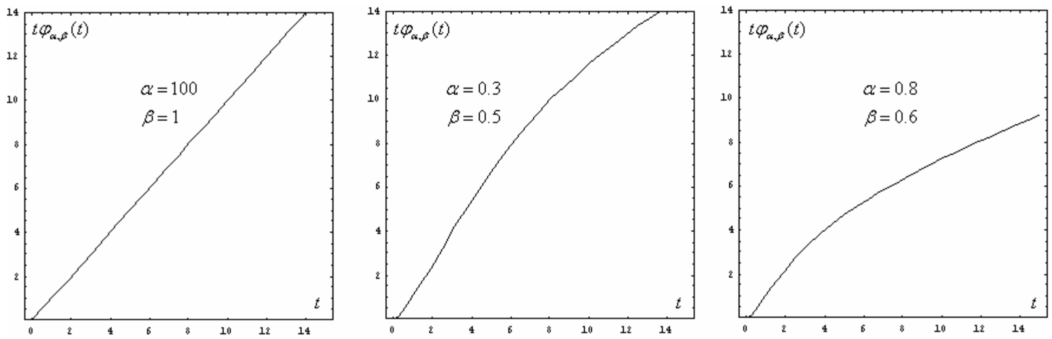

Figure 1: Illustration of the functions $\varphi_{\alpha,\beta}$ for different α and β

Property 1 is necessary to guarantee consistency at time $t = 1$ between the given one-year migration matrix $M = \exp(Q)$ and its nonhomogeneous modification $\exp(Q_1)$. Property 2 is necessary for keeping the direction of time (moving into the future and not into the past). Property 3 is meant as a remark to make the point that the special form of the functions $\varphi_{\alpha,\beta}$, while it has the flavour of an 'ad hoc' parameterization, is not completely arbitrary but 'close' to well-known functions used in probability.

Since the functional form of the time-dependent generators $(Q_t)_{t\geq 0}$ is fixed by Equation (4), the generators Q_t are solely determined by two vectors $(\alpha_1, ..., \alpha_8)$ and $(\beta_1, ..., \beta_8)$ in $[0, \infty)^8$. 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

$$M_t \;=\; \exp(tQ_t) \qquad (t \geq 0). \tag{5}$$

The last step we have to make is to optimize[5] $(\alpha_1, ..., \alpha_8)$ and $(\beta_1, ..., \beta_8)$ for the best fit of the term structure generated by the default column of the migration matrices (5) to S&P's [17] empirical term structures of default frequencies. As distance measure for our optimization we use the mean-squared distance. Table 2 and Figure 2 show the outcome of best-fitting α- and β-vectors as well as the resulting (NHCTMC-implied) credit curves in comparison to the empirically observed multi-year default frequencies from S&P.

Summarizing, *we parameterized a Markov chain approach for calibrating model-implied PD term structures in continuous time, which fit empirical observed default frequencies very well. Crucial in our approach was the acceptance of a nonhomogeneous time evolution of the chain.* The choice of parameters involved a one-year migration matrix as well as observed default frequencies. It is an *interpolating* not an *extrapolating* approach because the fit can only be exercised withing the time window of observations[6].

[4] Note that $\varphi_{\alpha,\beta}$ exhibits some simularity to the *gamma distribution*, frequently applied in the context of queuing theory and reliability analysis.

[5] Note that α_8 and β_8 have no meaning and can be fixed at some arbitrary value.

[6] In contrast to homogeneous Markov chains where extrapolation can be done quite naturally.

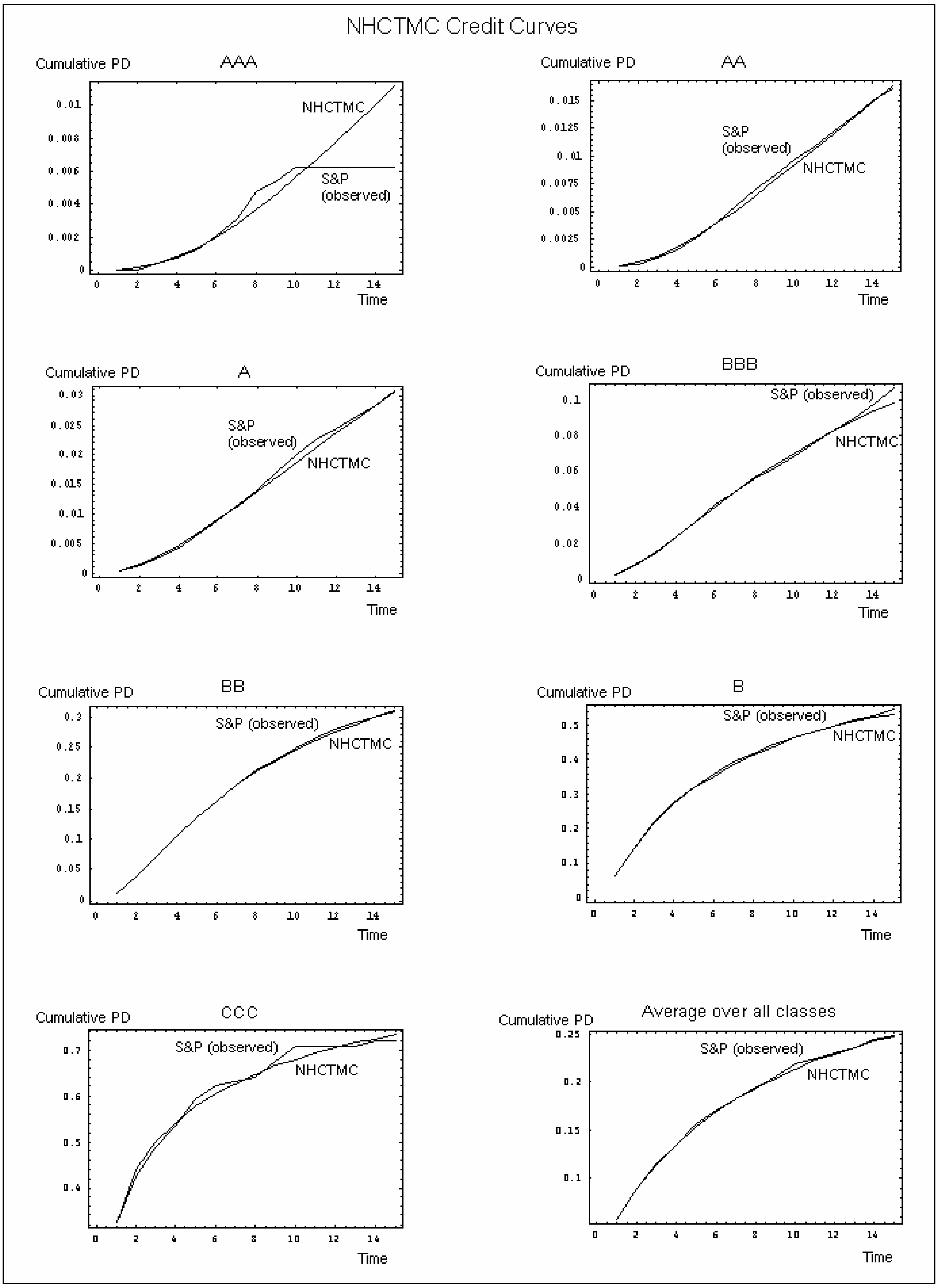

Figure 2: PD term structures based on a NHCTMC approach

	α	β
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

Table 2: Optimal choices for α- and β-vectors

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 a more convenient notation, let us denote by $\Psi(t)$ the diagonal matrix with diagonal elements

$$\psi_{ii}(t) \ = \ t\varphi_{\alpha_i,\beta_i}(t) \qquad (i = 1, ..., 8;\ t \geq 0).$$

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

$$M_t \ = \ \exp(\Psi(t) * Q) \qquad (t \geq 0). \tag{6}$$

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

$$\begin{aligned} \frac{\partial}{\partial t} P_{0,t} &= \sum_{k=0}^{\infty} \frac{\partial}{\partial t} \frac{(\Psi(t) * Q)^k}{k!} \\ &= \sum_{k=1}^{\infty} \Big(\frac{\partial}{\partial t} \Psi(t) * Q \Big) * \frac{(\Psi(t) * Q)^{(k-1)}}{(k-1)!} \\ &= \Big(\frac{\partial}{\partial t} \Psi(t) * Q \Big) * P_{0,t}. \end{aligned} \tag{7}$$

Because $\Psi(t)$ is a diagonal matrix, $(\partial/\partial t)\Psi(t)$ is the diagonal matrix with entries $\psi'_{ii}(t)$. Therefore, the matrix $(\partial/\partial t)\Psi(t) * Q$ is a Q-matrix, arguing in the same way as above where we said that $\Psi(t) * Q$ is a Q-matrix and taking into account that $\psi'_{ii}(t) \geq 0$ at all times[7] t. As a consequence of general Markov theory (see Ethier and Kurtz [4], Theorem 7.3 in Chapter 4, Lando and Skodeberg [10], and Schoenbucher [13]), Equation (7) is part of the *forward equation* of a time-inhomogenous Markov chain $(X_t)_{t\geq 0}$ with state space $\{1, 2, ..., 8\}$ corresponding to a *semigroup* $\{P_{s,t} \mid 0 \leq s \leq t\}$ satisfying the *Kolmogorov backward* and *forward equations* associated with the family $\{(\partial/\partial t)\Psi(t) * Q \mid t \geq 0\}$ defining the *infinitesimal generator* of the Markov process. Equation (7) shows that the NHCTMC $(X_t)_{t\geq 0}$ induces the PD term structures illustrated in Figure 2 via the default column of kernel-based transition matrices $P_{0,t} = M_t = \exp(\Psi(t) * Q)$.

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

The following example is taken from [3], Section 2.3.1. We start with the adjusted[8] average one-year migration matrix $M = (m_{ij})_{i,j=1,...,8}$ shown in Table 3, based on Table 9 in [17].

Table 4 shows the calibration of a generator (Q-matrix) Q based on the log-expansion of M and a so-called *diagonal adjustment*; see [9]. The approximation of the original matrix M by $\exp(Q)$ is

[7] We have $(1 - \exp(-\alpha))\psi'_{ii}(t) = \alpha \exp(-\alpha t)t^{\beta} + (1 - \exp(-\alpha t))\beta t^{\beta-1} \geq 0$ for all $t \geq 0$.

[8] Rows are normalized in order to get a stochastic matrix and the PD for AAA is set equal to 0.2 bps, based on a linear regression of PDs on a logarithmic scale.

	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%

Table 3: Modified average one-year migration matrix M based on S&P data [17]

	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%

Table 4: Approximative generator Q for M

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 continuous time-homogeneous 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 [17]. 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 time-homogeneous Markov chain approach. However, Figure 2 shows that the picture can completely change to the positive if we drop the time-homogeneity assumption.

References

[1] Basel Committee on Banking Supervision; *International Convergence of Capital Measurement and Capital Standards*; Bank for International Settlements, June (2004)

[2] Bluhm, C., Overbeck, L., Wagner, C.; *An Introduction to Credit Risk Modeling*; Chapman & Hall/CRC Financial Mathematics Series; 2nd Reprint; CRC Press (2003)

[3] Bluhm, C., Overbeck, L.; *Structured Credit Portfolio Analysis, Baskets & CDOs*; Chapman & Hall/CRC Financial Mathematics Series; CRC Press (2006)

[4] Ethier, S. N., Kurtz, T. G.; *Markov Processes Characterization and Convergence* John Wiley and Sons (2005)

[5] Frydman, H., Schuermann, T.; *Credit Rating Dynamics and Markov Mixture Models*; Working Paper, June (2005)

[6] Israel, R., Rosenthal, J., Wei, J.; *Finding Generators for Markov Chains via Empirical Transition Matrices with Application to Credit Ratings*; Mathematical Finance **11** (2), 245-265 (2001)

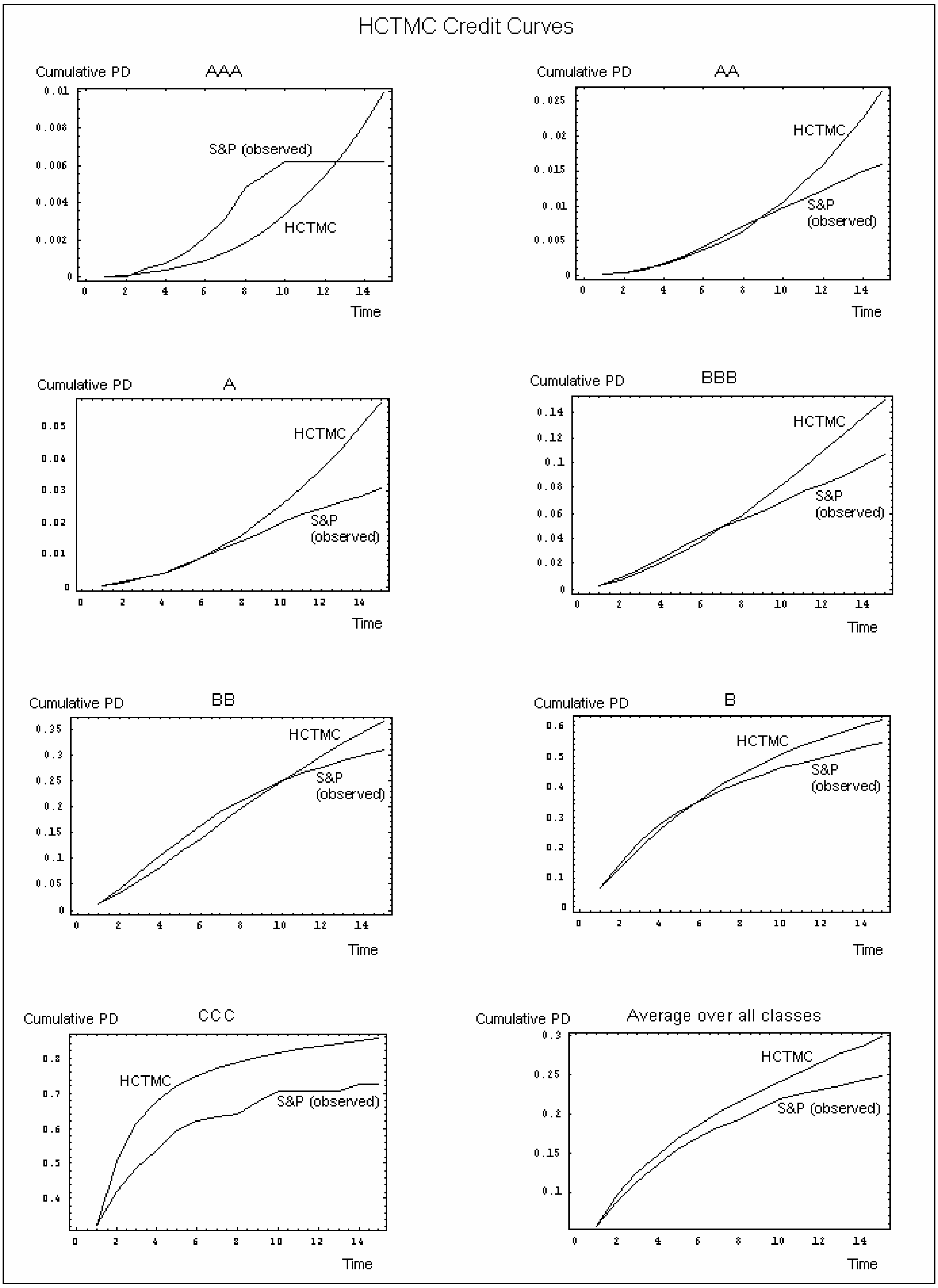

Figure 3: PD term structures based on a homogeneous continuous-time Markov chain (HCTMC)

[7] Jarrow, R. A., Lando, D., Turnbull, S. M.; *A Markov Model for the Term Structure of Credit Risk Spreads*; Review of Financial Studies **10**, 481-523 (1997)

[8] Kadam, A., Lenk, P.; *Heterogeneity in Ratings Migration*; Working Paper, October (2005)

[9] Kreinin, A., Sidelnikova, M.; *Regularization Algorithms for Transition Matrices*; Algo Research Quarterly **4** (1/2), 25-40 (2001)

[10] Lando, D., Skodeberg, T. M.; *Analyzing Rating Transitions and Rating Drift with Continuous Observations*; Journal of Banking and Finance **26**, No. 2-3, 423-444 (2002)

[11] Noris, J.; *Markov Chains*; Cambridge Series in Statistical and Probabilistic Mathematics; Cambridge University Press (1998)

[12] Sarfaraz, A., Cohen, M., Libreros, S.; *Use of Transition Matrices in Risk Management and Valuation*; Fair Isaac White Paper, September (2004)

[13] Schoenbucher, P.; *Portfolio Losses and the Term Structure of Loss Transition Rates: A New Methodology for the Pricing of Portfolio Credit Derivates*; NCCR FinRisk Working Paper No. **264**, September (2005)

[14] Schuermann, T., Jafry, Y.; *Measurement and Estimation of Credit Migration Matrices*; Wharton School Center for Financial Institutions, University of Pennsylvania, Working Paper 03-08 (2003)

[15] Schuermann, T. Jafry, Y.; *Metrics for Comparing Credit Migration Matrices*; Wharton School Center for Financial Institutions, University of Pennsylvania, Working Paper 03-09 (2003)

[16] Sklar, A.; *Fonction de Repartition à n Dimension et Leur Marges*; Publications de l'Insitute Statistique de l'Université de Paris **8**, 229-231 (1959)

[17] Standard & Poor's; *Annual Global Corporate Default Study: Corporate Defaults Poised to Rise in 2005*; S&P Global Fixed Income Research, January (2005)

[18] Trueck, S., Oezturkmen, E.; *Adjustment and Application of Transition Matrices in Credit Risk Models*; Working Paper, University of Karlsruhe, September (2003)