

AN INHOMOGENEOUS SEMI-MARKOV MODEL FOR THE TERM STRUCTURE OF CREDIT RISK SPREADS

AGLAIA VASILEIOU,* UBS Investment Bank, London

P.-C. G. VASSILIOU,** Aristotle University of Thessaloniki

Abstract

We model the evolution of the credit migration of a corporate bond as an inhomogeneous semi-Markov chain. The valuation of a defaultable bond is done with the use of the forward probability of no default up to maturity time. It is proved that, under the forward probability measure, the semi-Markov property is maintained. We find the functional relationships between the forward transition probability sequences and the real-world probability sequences. The stochastic monotonicity properties of the inhomogeneous semi-Markov model, which play a prominent role in these issues, are studied in detail. Finally, we study the term structure of credit spread, provide an algorithm for the estimation of the forward probabilities of transitions under the risk premium assumptions, and present an estimation method for the real-world probability sequences.

Keywords: Credit rating; inhomogeneous semi-Markov chain; stochastic monotonicity

2000 Mathematics Subject Classification: Primary 60J20; 60K15; 60J27

1. Introductory notes

A *default risk* is a possibility that a counterparty in a financial contract will not fulfill a contractual commitment to meet their stated obligations. If this actually happens, we say that the party defaults, or that the default event occurs. More generally, by a *credit risk* we mean the risk associated with any kind of credit-linked event, for example a change in the credit quality, a variation of credit spread, or the default event. *Corporate bonds* are debt instruments issued by corporations. We shall concentrate on a discount bond, that is, we assume that the bond pays no coupons. We shall use the term *defaultable bond* for any kind of bond with the possibility of default.

Most of the techniques presented in this paper are applicable to the valuation of general corporate liabilities, corporate loans, etc. However, merely for presentation purposes we choose to limit the discussion to corporate bonds. If a fixed fraction of a bond's face value is paid to the bondholders at the time of default (usually denoted by τ in what follows), then the recovery scheme is referred to as the *fractional recovery of par value*. There are other types of recovery schemes but, since the methodologies presented below apply similarly to any type of recovery scheme, we will not expand the discussion.

Received 20 August 2004; revision received 25 July 2005.

* Postal address: UBS Investment Bank, 100 Liverpool Street, London EC2M 2RH, UK.

Email address: aglaia.vasileiou@ubs.com

This study has been performed in a private capacity and the opinions expressed in it should not be attributed to UBS Investment Bank.

** Postal address: Statistics and Operations Research Section, Mathematics Department, Aristotle University of Thessaloniki, Thessaloniki, 54006, Greece. Email address: vasiliou@math.auth.gr

Depending on the situation, a credit spread may be expressed, for example, as the difference between respective yields to maturity or as the difference between respective instantaneous forward rates. By the *term structure of credit spreads* we will refer to the term structure of such differences. The determination of the credit spread is in fact the ultimate goal of most credit risk models. A firm's credit rating is a measure of the firm's propensity to default.

Credit derivatives are privately negotiated derivative securities that are linked to a credit-sensitive asset as the underlying asset. As estimated by J. P. Morgan, in April 2000 the total nominal value of outstanding credit derivatives exceeded US\$10¹². The common feature of all credit derivatives is that they allow for the transfer of the credit risk from one counterparty to another; they thus constitute a natural and convenient tool to control risk exposure. In contrast to standard interest-rate-sensitive derivatives, credit derivatives allow for the isolation of the firm-specific credit risk from the overall market risk. They also provide a way to synthesize assets that are otherwise not available to a particular investor.

The main objective of the quantitative models of credit risk is to provide ways to hedge financial contracts that are sensitive to credit risk. The vast majority of mathematical research devoted to credit risk is concerned with the modeling of the random time at which the default event occurs, i.e. the default time. Two competing methodologies have emerged to model the default/migration times and the recovery rates: the *structural* approach and the *reduced-form* approach. Structural models are concerned with modeling and pricing credit risk that is specific to a particular corporate obligator. Credit events are triggered by movements of the firm's value relative to some (random or nonrandom) credit-event-triggering threshold. For this reason, the structural approach is frequently referred to as the *firm value* approach. In the reduced-form approach, the value of the firm's assets and its capital structure are not modeled at all, and the credit events are specified in terms of some exogenously specified jump process. We can distinguish between the reduced-form models that are only concerned with the modeling of the default time (henceforth referred to as the *intensity-based* models), and the reduced-form models with migrations between credit rating classes (called the *credit migration* models).

Credit ratings are typically identified with elements of a finite set, referred to as the set of credit classes or credit grades. In some cases, the credit classes may correspond to credit rating systems, based on internally developed methodologies; in other cases they are attributed by a commercial rating agency, such as Moody's Investors Service, Standard and Poor's, Fitch IBCA, or Duff and Phelps.

Assume that the credit quality of a corporate bond is quantified and categorized into a finite number of disjoint credit rating classes. Each credit class is represented by an element of a finite set, denoted by $\mathbb{K} = \{1, 2, \dots, k, k + 1\}$. It is natural to distinguish a particular element $k + 1$ of the set \mathbb{K} as formally corresponding to the default event, and to assume that credit quality 1 is the highest. The main issue of the credit migrations approach is the modeling of the transition probabilities/intensities of the migration process, both under the risk-neutral and the real-world probabilities. The classical way of modeling the evolution of credit migrations is in terms of either a discrete- or continuous-time homogeneous Markov chain (or a conditionally homogeneous Markov chain), first introduced by Jarrow and Turnbull (1995) and Jarrow *et al.* (1997). Works related to the simple Markov chain approach are those by Das and Tufano (1996), Kijima (1998), Kijima and Komoribayashi (1998), Lando (1998), (2000), Arvanitis *et al.* (1999), and Duffie and Singleton (1999). An excellent review of the above papers, the mathematical foundation of their models, and the determination of their common core and interrelations are given in Bielecki and Rutkowski (2002).

In the study by Carty and Fons (1994), drawn from the Moody's Investors Service proprietary database, with a data span from 1976 to 1993, it was established that the duration of stay in a specific credit rating class followed the Weibull distribution. The estimation of the parameters of the particular Weibull distribution varied for each credit rating class. Thus, Carty and Fons (1994) in fact established, without stating it, that the appropriate model was not a simple Markov chain – in which case the duration of stay in each credit rating class would follow the exponential distribution (or the geometric distribution in the discrete case) – but rather was a semi-Markov process. The inhomogeneity in time of transition probabilities has been reported by many authors, among whom are Duffie and Singleton (2003, p. 45), Lando (2004, p. 88), Lando and Skoteberg (2001), Hamilton (2001), Fons (1991), and Jonsson and Frison (1996). As these, and other, authors have reported, it seems that the real economic activity over the business cycle is correlated with the fluctuation in transition and default probabilities. In conclusion, there was thus a need to model the evolution of the credit migration of a corporate bond as a discrete-time inhomogeneous semi-Markov chain, as a first step, and then study the stochastic, measure-theoretic, and statistical problems that would arise from such an effort. However, it is important to note that most of the present results, such as Theorem 4.1 and the stochastic monotonicity theorems, are of interest in their own right in the theory of semi-Markov processes.

In the present paper we model the evolution of the credit migration of a corporate bond as an inhomogeneous semi-Markov chain. A first attempt to define the inhomogeneous semi-Markov chain was made by Iosifescu-Manu (1972), who presented results of some simulations of various aspects of the model, in particular its asymptotic behavior. The theory of ordinary semi-Markov chains is well presented in Howard (1971, p. 577) and important further results and applications can be found in McClean (1976), (1986) and Bartholomew (1982, p. 104), (1986). A formal definition of the inhomogeneous semi-Markov chain was given in Vassiliou and Papadopoulou (1992) and the problem of asymptotic behavior studied in Papadopoulou and Vassiliou (1994). Furthermore, in Papadopoulou (1997) various other aspects of the asymptotic behavior were treated, and in Vassiliou and Symeonaki (1999) perturbations in inhomogeneous Markov systems were studied. A review of the evolution of the work on inhomogeneous semi-Markov systems can be found in Vassiliou (1997). The inputs of our model are the real-world probabilities, which are sufficient to uniquely determine an inhomogeneous semi-Markov chain.

In Section 2 we assume that there exists an arbitrage-free family of bond prices relative to the interest rate process r_t . The price at time t of a defaultable bond with maturity T is estimated, with the recovery rate of the fractional recovery of par value scheme, δ , assumed to be a constant. The valuation of the defaultable bond is given as a function of the probability of no default up to maturity time under the forward probability measure. The forward probability measure is equivalent to the probability martingale measure determined by the arbitrage-free family of nonrisky bonds. In Section 3 we provide a new definition of a \mathbb{G} -inhomogeneous semi-Markov process. In Section 4 we study the change of real-world probability to forward probability measure. In Theorem 4.1 we establish that the semi-Markov property is maintained, and we find, in closed analytic form, the sequences of transition probabilities under the forward probability measure as functions of the sequences of real-world transition probabilities. In Section 5 we study stochastic monotonicities with the use of the general inhomogeneous semi-Markov model. The case of stochastic monotonicities for the Markov model was studied by Kijima (1998). In Theorem 5.1 we provide necessary conditions for the time of default of a bond in credit class i to be stochastically greater than the time of default of a bond in credit class $i + 1$. In Theorem 5.4 we provide necessary conditions for the time of default of a bond

in credit class i to be greater in the hazard sense than the time of default of a bond in credit class $i + 1$. In Theorem 5.5 we provide necessary conditions for the time of default of a bond in credit class 1 to be *new worse than used*, and of a bond of credit class k to be *new better than used*. In Section 6 we study the term structure of credit spread and provide an algorithm for the estimation of the forward probabilities of transitions under the risk premium assumptions. Finally we provide a method of estimating the real-world transition probability sequences for the semi-Markov processes, and statistics for testing their homogeneity over time.

2. The price of a defaultable bond

Let the real-world filtered probability space be $(\Omega, \mathbb{G}, \mathbb{Q})$, where \mathbb{Q} is interpreted as the real-world probability measure, and let the filtration $\mathcal{G}_t \subseteq \mathbb{G}$, $t = 0, 1, 2, \dots$, model the flow of all observations available to traders. Let $T^* > 0$ be a fixed horizon date. We will present a model that accounts for the migration of a defaultable bond between rating grades.

Let us assume that B_t is the savings account, opened with one unit of money. In accordance with simple compounding convention in the discrete-time case, we have

$$B_t = \prod_{u=0}^t (1 + r_u),$$

where r_t denotes the interest rate process for $t = 0, 1, 2, \dots$

By a zero-coupon bond (a discount bond) of maturity T , we mean a financial security paying its holder one unit of money at a specified date in the future. This means that, by convention, the bond’s principal (known also as face value or nominal value) is one unit. The price of a zero-coupon bond of maturity T will, at any instant $t \leq T$, be denoted by $B(t, T)$; it is thus obvious that $B(T, T) = 1$. In our framework, we assume throughout that, for any fixed maturity T , the price process $B(t, T)$, $t \in [0, T]$, follows a strictly positive and adapted process on the filtered probability space $(\Omega, \mathbb{G}, \mathbb{Q})$.

We are interested in the absence of arbitrage opportunities between all bonds with different maturities and a savings account.

Definition 2.1. A family $B(t, T)$, $t \leq T \leq T^*$, of adapted processes is called an arbitrage-free family of bond prices relative to r_t if the following conditions are satisfied:

- (a) $B(T, T) = 1$ for every $T \in [0, T^*]$,
- (b) there exists a probability measure \mathbb{Q}^* on $(\Omega, \mathcal{G}_{T^*})$, equivalent to \mathbb{Q} , such that, for any maturity $T \in [0, T^*]$, the relative bond price

$$\frac{B(t, T)}{B_t}, \quad t \in [0, T],$$

follows a martingale under \mathbb{Q}^* .

Any probability measure \mathbb{Q}^* of the type in Definition 2.1 is called a *martingale measure for the family $B(t, T)$ relative to r_t* or, briefly, a *martingale measure for the family $B(t, T)$* if no confusion may arise. The existence of such a martingale measure for the family $B(t, T)$ is often called the fundamental theorem of asset pricing (Shreve (2004b, p. 231)). It is implicitly assumed that we have already constructed an arbitrage-free model of a market in which all bonds of different maturities, as well as a certain number of other assets, are primary traded

securities. In addition, we assume that the market is complete, which makes the martingale measure Q^* unique and moreover allows us to postulate the existence of a replicating strategy for any contingent claim in the default-free market.

From the above definition we find (see Musiela and Rutkowski (1997b, p. 283)) that

$$B(t, T) = B_t E_{Q^*}(B_T^{-1} \mid \mathcal{G}_t), \quad t \in [0, T]. \tag{2.1}$$

Let $D^\delta(t, T)$ be the price at time t of a defaultable bond with a face value of one unit of money. Then, as in (2.1), we obtain

$$D^\delta(t, T) = B_t E_{Q^*}(B_T^{-1}(\delta 1_{\{T \geq \tau\}} + 1_{\{T < \tau\}}) \mid \mathcal{G}_t), \tag{2.2}$$

where $1_{\{T \geq \tau\}}$ equals 1 if $T \geq \tau$ and equals 0 otherwise.

We assume throughout that the price $B(t, T)$ of a zero-coupon bond of maturity T follows an Itô process under the martingale measure Q^* (see Musiela and Rutkowski (1997a)):

$$dB(t, T) = B(t, T)(r_t dt + b(t, T) d\mathbb{W}_t^*).$$

Here \mathbb{W}^* denotes a d -dimensional standard Brownian motion under the martingale measure Q^* defined on the filtered probability space (Ω, \mathbb{G}, Q) . The σ -algebra \mathbb{G} encompasses the filtration generated by \mathbb{W}^* . In our case, of discrete time, the above equation changes accordingly.

A *forward contract* is an agreement, established at a time $t < T$, to pay or receive at the settlement time T a preassigned payoff, say X , at an agreed forward price. We now provide the definition of the forward probability measure introduced by Geman (1989) and further developed by Geman *et al.* (1995), who observed that the forward price of any financial asset follows a local martingale under the forward neutral probability associated with the settlement date of a forward contract.

Definition 2.2. A probability measure Q_T on (Ω, \mathcal{G}_T) , equivalent to Q^* , with Radon–Nikodým derivative given by the formula

$$\frac{dQ_T}{dQ^*} = \frac{B_T^{-1}}{E_{Q^*}(B_T^{-1})} = \frac{1}{B_T B(0, T)} \quad Q^*\text{-almost surely (} Q^*\text{-a.s.)} \tag{2.3}$$

is called the forward martingale measure (or, briefly, the forward measure) for the settlement time T .

Notice that the above Radon–Nikodým derivative, when restricted to the σ -algebra \mathcal{G}_t , satisfies

$$\eta_t := \left. \frac{dQ_T}{dQ^*} \right|_{\mathcal{G}_t} = E_{Q^*} \left(\left. \frac{1}{B_T B(0, T)} \right|_{\mathcal{G}_t} \right) = \frac{B(t, T)}{B_t B(0, T)} \tag{2.4}$$

for every $t \in [0, T]$.

The next result, which is known as the abstract version of the Bayes formula, applies in the situation in which two mutually equivalent probability measures Q^* and Q_T , say, are defined on a common measurable space (Ω, \mathbb{G}) . Suppose that the Radon–Nikodým derivative of Q_T with respect to Q^* is

$$\frac{dQ_T}{dQ^*} = \eta \quad Q^*\text{-a.s.} \tag{2.5}$$

Note that the random variable η is strictly positive Q^* -a.s. and that η is Q^* -integrable (that is, $E_{Q^*}(|\eta|) < \infty$) with $E_{Q^*}(\eta) = 1$.

Equation (2.5) implies (Shreve (2004a, p. 62)) the apparently stronger condition that, for any Q_T -integrable random variable ψ ,

$$E_{Q_T}(\psi) = E_{Q^*}(\psi\eta).$$

Lemma 2.1. *Let \mathcal{G}_t be a σ -subalgebra of the σ -algebra \mathbb{G} and let ψ be a random variable integrable with respect to Q_T . Then*

$$E_{Q_T}(\psi \mid \mathcal{G}_t) = \frac{E_{Q^*}(\psi\eta \mid \mathcal{G}_t)}{E_{Q^*}(\eta \mid \mathcal{G}_t)}. \tag{2.6}$$

The proof of Lemma 2.1 can be found in Musiela and Rutkowski (1997b, p. 458).

Now let

$$\psi = (\delta 1_{\{T \geq \tau\}} + 1_{\{T < \tau\}}).$$

From (2.6) we then obtain

$$\begin{aligned} E_{Q_T}((\delta 1_{\{T \geq \tau\}} + 1_{\{T < \tau\}}) \mid \mathcal{G}_t) &= \frac{E_{Q^*}([1/(B_T B(0, T))](\delta 1_{\{T \geq \tau\}} + 1_{\{T < \tau\}}) \mid \mathcal{G}_t)}{E_{Q^*}([1/B_T B(0, T)] \mid \mathcal{G}_t)} \\ &= \frac{E_{Q^*}(B_T^{-1}(\delta 1_{\{T \geq \tau\}} + 1_{\{T < \tau\}}) \mid \mathcal{G}_t)}{B_t^{-1} B(t, T)}, \end{aligned} \tag{2.7}$$

using (2.4) and (2.1), and from (2.2) and (2.7) we find that

$$\begin{aligned} D^\delta(t, T) &= B(t, T) E_{Q_T}((\delta 1_{\{T \geq \tau\}} + 1_{\{T < \tau\}}) \mid \mathcal{G}_t) \\ &= B(t, T)(\delta + (1 - \delta) Q_T\{\tau > T \mid \mathcal{G}_t\}). \end{aligned} \tag{2.8}$$

3. Discrete-time inhomogeneous semi-Markov process

In what follows, we model the evolution of the credit migrations of a defaultable bond as an inhomogeneous semi-Markov process. We provide a new and more general definition of an inhomogeneous semi-Markov process than that which first appeared in the literature in Vassiliou and Papadopoulou (1992). Note that, apart from the generalization, the delicate essential differences in the cores of the two definitions are those that make possible the change of probability measure which follows later. Let the real-world filtered probability space (Ω, \mathbb{G}, Q) be as described in Section 2. Let $\{X_t\}_{t=0}^\infty$ be a stochastic process on (Ω, \mathbb{G}, Q) taking values in $\mathbb{K} = \{1, 2, \dots, k, k + 1\}$ and describing the state the defaultable bond enters at time t . Also, let \mathbb{F}^X be the natural filtration generated by the process $\{X_t\}_{t=0}^\infty$, i.e. $\mathbb{F}_t^X = \sigma(X_n, n = 0, 1, \dots, t)$. We assume that \mathbb{F}^X is a subfiltration of \mathbb{G} , i.e. $\mathbb{F}^X \subseteq \mathbb{G}$. Now let S_{t+} be the random variable valued in \mathbb{K} that describes the selection at time t^+ of the next transition of the bond, given that its last entrance to a class was at time t . We assume that $t^+ = t + \varepsilon \in (t, t + 1)$, where for simplicity ε is a small positive constant independent of t . Let \mathbb{F}^S be the natural filtration generated by the process $\{S_{t+}\}_{t=0}^\infty$ and assume that $\mathbb{F}^S \vee \mathbb{F}^X \subseteq \mathbb{G}$. We assume that the evolution in \mathbb{K} of a defaultable bond follows a \mathbb{G} -inhomogeneous semi-Markov process the definition of which we now provide.

Definition 3.1. The pair $\{X_t, S_{t+}\}_{t=0}^\infty$ is a discrete-time \mathbb{G} -inhomogeneous semi-Markov process if the following conditions hold.

(a) The stochastic process $\{S_t\}_{t=0}^\infty$ is a discrete inhomogeneous Markov process adapted to the stochastic process $\{X_t\}_{t=0}^\infty$ under Q with respect to \mathbb{G} , i.e. for any function $f: \mathbb{K} \rightarrow \mathbb{R}$ we have

$$E_Q(f(S_{(t+v)^+}) | \mathcal{G}_t) = E_Q(f(S_{(t+v)^+}) | \sigma(X_t)), \quad t, v = 0, 1, \dots \tag{3.1}$$

(b) Let $X_t = i$ and $S_{t+} = j$. Then the times of transition, say $\varpi_{ij}(t)$, for any $i, j \in \mathbb{K}$, are positive, integer-valued random variables each governed by a probability density function $h_{ij}(t, m)$, where obviously

$$h_{ij}(t, m) = Q\{\varpi_{ij}(t) = m | S_{t+} = j, X_t = i\}, \quad t, m = 0, 1, 2, \dots$$

The $h_{ij}(t, m)$ are ‘estimated’ (known) one ‘click’ of time after t^+ , say at t^{++} with $t < t^+ < t^{++} < t + 1$.

(c) For any function $f: \mathbb{K} \times \mathbb{K} \times \mathbb{N}^* \rightarrow \mathbb{R}$, we have

$$\begin{aligned} & Q\{f(X_t, S_{t+}, \varpi_{X_t, S_{t+}}(t)) | \mathcal{G}_{t^+}\} \\ &= Q\{f(X_t, S_{t+}, \varpi_{X_t, S_{t+}}(t)) | \sigma(X_t) \vee \sigma(S_{t^+})\}, \end{aligned} \tag{3.2}$$

where $\mathbb{N}^* = \{0, 1, 2, \dots\}$.

We will call $\{S_{t^+}\}_{t=0}^\infty$ the *inherent \mathbb{G} -inhomogeneous Markov process*.

If a stochastic process $\{S_{t^+}\}_{t=0}^\infty$ is a \mathbb{G} -inhomogeneous Markov chain under Q it is also an \mathbb{F}^S -inhomogeneous Markov chain under Q , but the converse implication is not true, in general. Since the set \mathbb{K} is finite, condition (a) above is equivalent (Shreve (2004a, p. 44)) to the following ones.

- (i) $Q\{S_{t^+} = j | \mathcal{G}_t\} = Q\{S_{t^+} = j | \sigma(X_t)\}$.
- (ii) Let \mathbb{K}_v be the Cartesian product of v copies of \mathbb{K} . For any function $f: \mathbb{K}_v \rightarrow \mathbb{R}$ for which $E_Q(|f(S_{(t+1)^+}, \dots, S_{(t+v)^+})|) < \infty$, we have

$$E_Q(f(S_{(t+1)^+}, \dots, S_{(t+v)^+}) | \mathcal{G}_t) = E_Q(f(S_{(t+1)^+}, \dots, S_{(t+v)^+}) | \sigma(X_t)).$$

- (iii) For every $A_{(t+1)^+}, A_{(t+2)^+}, \dots, A_{(t+v)^+} \subset \mathbb{R}$, we have

$$\begin{aligned} & Q\{S_{(t+1)^+} \in A_{(t+1)^+}, S_{(t+2)^+} \in A_{(t+2)^+}, \dots, S_{(t+v)^+} \in A_{(t+v)^+} | \mathcal{G}_t\} \\ &= Q\{S_{(t+1)^+} \in A_{(t+1)^+}, S_{(t+2)^+} \in A_{(t+2)^+}, \dots, S_{(t+v)^+} \in A_{(t+v)^+} | \sigma(X_t)\}. \end{aligned}$$

Now let $p_{ij}(t) = Q\{S_{t^+} = j | X_t = i\}$ and denote by $\tilde{P}(t) = \{p_{ij}(t)\}_{i,j \in \mathbb{K}}$ the matrix of transition probabilities, which will be of the form

$$\tilde{P}(t) = \begin{pmatrix} p_{11}(t) & p_{12}(t) & \cdots & p_{1k}(t) & p_{1,k+1}(t) \\ p_{21}(t) & p_{22}(t) & \cdots & p_{2k}(t) & p_{2,k+1}(t) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ p_{k1}(t) & p_{k2}(t) & \cdots & p_{kk}(t) & p_{k,k+1}(t) \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix} = \begin{pmatrix} P(t) & P_{k+1}^\top(t) \\ \mathbf{0} & 1 \end{pmatrix}. \tag{3.3}$$

Note that $p_{k+1,k+1}(t) = 1$; thus, the default state is an absorption state for the process $\{X_t\}_{t=0}^\infty$. According to Definition 3.1, whenever a defaultable bond enters a state i it selects the next credit class j to which it will move with probability $p_{ij}(t) = Q\{S_{t^+} = j | X_t = i\}$.

However, after j has been selected, but before making the transition from class i to class j , the process ‘holds’ for a time $\varpi_{ij}(t)$ in state i . Recall that the holding times $\varpi_{ij}(t)$ are positive, integer-valued random variables each governed by a probability density function $h_{ij}(t, m)$. The matrix $\tilde{\mathbf{H}}(t, m) = \{h_{ij}(t, m)\}_{i,j \in \mathbb{K}}$ is of the form

$$\begin{aligned} \tilde{\mathbf{H}}(t, m) &= \begin{pmatrix} h_{11}(t, m) & h_{12}(t, m) & \cdots & h_{1k}(t, m) & h_{1,k+1}(t, m) \\ h_{21}(t, m) & h_{22}(t, m) & \cdots & h_{2k}(t, m) & h_{2,k+1}(t, m) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ h_{k1}(t, m) & h_{k2}(t, m) & \cdots & h_{kk}(t, m) & h_{k,k+1}(t, m) \\ 0 & 0 & \cdots & 0 & 1_{\{m=t+1\}} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{H}(t, m) & \mathbf{h}_{k+1}^\top(t, m) \\ \mathbf{0} & 1_{\{m=t+1\}} \end{pmatrix}. \end{aligned} \tag{3.4}$$

We assume that the means of all holding time distributions are finite and that all holding times are nonzero, i.e. $h_{ij}(t, 0) = 0$ for every t .

4. Change of real-world probability to forward martingale measure

For the fixed time horizon $T < \infty$, we have defined the real-world probability measure \mathbb{Q} and the martingale measure \mathbb{Q}^* to be equivalent on (Ω, \mathcal{G}_T) . Let the Radon–Nikodým derivative be given by

$$\left. \frac{d\mathbb{Q}^*}{d\mathbb{Q}} \right|_{\mathcal{G}_T} = \psi_T \quad \mathbb{Q}\text{-a.s.}, \tag{4.1}$$

where the \mathcal{G}_T -measurable random variable ψ_T is strictly positive \mathbb{Q} -a.s. and $E_{\mathbb{Q}}(\psi_T) = 1$. Then the density process $\psi_t = E_{\mathbb{Q}}(\psi_T \mid \mathcal{G}_t)$, $t = 0, 1, \dots, T$, follows a strictly positive martingale under \mathbb{Q} .

From (2.3) and (4.1) we find that the real-world probability measure \mathbb{Q} and the forward martingale measure \mathbb{Q}_T are equivalent on (Ω, \mathcal{G}_T) , and that their Radon–Nikodým derivative will be given by

$$\left. \frac{d\mathbb{Q}_T}{d\mathbb{Q}} \right|_{\mathcal{G}_T} = \frac{\psi_T}{B_T B(0, T)} = \theta_T \quad \mathbb{Q}\text{-a.s.}, \tag{4.2}$$

where the \mathcal{G}_T -measurable random variable θ_T is strictly positive \mathbb{Q} -a.s. and $E_{\mathbb{Q}}(\theta_T) = 1$. Note that the density process $\theta_t = E_{\mathbb{Q}}(\theta_T \mid \mathcal{G}_t)$, $t = 0, 1, \dots, T$, follows a strictly positive martingale under \mathbb{Q} , i.e.

$$E_{\mathbb{Q}}(\theta_{t+1} \mid \mathcal{G}_t) = E_{\mathbb{Q}}(E_{\mathbb{Q}}(\theta_T \mid \mathcal{G}_{t+1}) \mid \mathcal{G}_t) = E_{\mathbb{Q}}(\theta_T \mid \mathcal{G}_t) = \theta_t.$$

Our next goal is to examine whether a \mathbb{G} -inhomogeneous semi-Markov process under \mathbb{Q} remains a \mathbb{G} -inhomogeneous semi-Markov process under \mathbb{Q}_T . If the \mathbb{G} -inhomogeneous semi-Markov property is preserved under \mathbb{Q}_T , we would also like to relate the real-world transition probabilities $p_{ij}(t) = \mathbb{Q}\{S_{t+} = j \mid X_t = i\}$ and density functions

$$h_{ij}(t, m) = \mathbb{Q}\{\varpi_{ij}(t) = m \mid S_{t+} = j, X_t = i\}$$

to the respective probabilities under \mathbb{Q}_T . We provide the result in the following theorem.

Theorem 4.1. *Let the random variable $\theta_t^{-1}\theta_{t+}$ be $(\sigma(X_t) \vee \sigma(S_{t+}))$ -measurable for any $t = 0, 1, \dots, T - 1$. Consequently, for every such t we have*

$$\theta_t^{-1}\theta_{t+} = g_t(X_t, S_{t+})$$

for some function $g_t: \mathbb{K} \times \mathbb{K} \rightarrow \mathbb{R}$. In addition, assume that $\theta_{t+}^{-1}\theta_{t++}$ is $(\sigma(X_t) \vee \sigma(S_{t+}) \vee \sigma(\varpi_{X_t, S_{t+}}(t)))$ -measurable for any $t = 0, 1, \dots, T - 1$. In other words,

$$\theta_{t+}^{-1}\theta_{t++} = f_{t+}(X_t, S_{t+}, \varpi_{X_t, S_{t+}}(t))$$

for some function $f_{t+}: \mathbb{K} \times \mathbb{K} \times \mathbb{N}^* \rightarrow \mathbb{R}$. If the pair $\{X_t, S_{t+}\}_{t=0}^\infty$ follows a discrete-time \mathbb{G} -inhomogeneous semi-Markov process under \mathbb{Q} then it also follows a discrete-time \mathbb{G} -inhomogeneous semi-Markov process under \mathbb{Q}_T and, in addition,

- (a) ${}_F p_{ij}(t) := \mathbb{Q}_T\{S_{t+} = j \mid X_t = i\} = g_t(i, j)p_{ij}(t)$ and
- (b) ${}_F h_{ij}(t, m) := \mathbb{Q}_T\{\varpi_{ij}(t) = m \mid S_{t+} = j, X_t = i\} = f_{t+}(i, j, m)h_{ij}(t, m)$.

Proof. Using Lemma 2.1, we fix a $t \in \mathbb{N}^*$ and, for any state $j \in \mathbb{K}$, obtain

$$\begin{aligned} \mathbb{Q}_T\{S_{t+} = j \mid \mathcal{G}_t\} &= \mathbb{E}_{\mathbb{Q}_T}(1_{\{S_{t+}=j\}} \mid \mathcal{G}_t) \\ &= \frac{\mathbb{E}_{\mathbb{Q}}(\theta_T 1_{\{S_{t+}=j\}} \mid \mathcal{G}_t)}{\mathbb{E}_{\mathbb{Q}}(\theta_T \mid \mathcal{G}_t)} \quad (\text{from Lemma 2.1 and (4.2)}) \\ &= \mathbb{E}_{\mathbb{Q}}(\theta_t^{-1}\theta_T 1_{\{S_{t+}=j\}} \mid \mathcal{G}_t) \\ &= \mathbb{E}_{\mathbb{Q}}(\mathbb{E}_{\mathbb{Q}}(\theta_t^{-1}\theta_T 1_{\{S_{t+}=j\}} \mid \mathcal{G}_{t+}) \mid \mathcal{G}_t) \\ &= \mathbb{E}_{\mathbb{Q}}(\theta_t^{-1}\mathbb{E}_{\mathbb{Q}}(\theta_T \mid \mathcal{G}_{t+}) 1_{\{S_{t+}=j\}} \mid \mathcal{G}_t) \\ &= \mathbb{E}_{\mathbb{Q}}(\theta_t^{-1}\theta_{t+} 1_{\{S_{t+}=j\}} \mid \mathcal{G}_t) \\ &= \mathbb{E}_{\mathbb{Q}}(g_t(X_t, S_{t+}) 1_{\{S_{t+}=j\}} \mid \mathcal{G}_t) \\ &= \mathbb{E}_{\mathbb{Q}}(g_t(X_t, S_{t+}) 1_{\{S_{t+}=j\}} \mid \sigma(X_t)) \\ &\quad (\text{from (3.1) with } f(S_{t+}) = g_t(X_t, S_{t+}) 1_{\{S_{t+}=j\}}). \end{aligned} \tag{4.3}$$

The third equality follows from the facts that $\mathbb{E}_{\mathbb{Q}}(\theta_T \mid \mathcal{G}_t) = \theta_t$ is a martingale under \mathbb{Q} and θ_t is \mathcal{G}_t -measurable. Equation (4.3) shows that the conditional probability is a $\sigma(X_t)$ -measurable random variable. Since $\sigma(X_t) \subseteq \mathcal{G}_t$, we conclude that

$$\mathbb{Q}_T\{S_{t+} = j \mid \mathcal{G}_t\} = \mathbb{Q}_T\{S_{t+} = j \mid \sigma(X_t)\},$$

which in fact proves that $\{S_{t+}\}_{t+}$ is a \mathbb{G} -inhomogeneous Markov process under \mathbb{Q}_T .

We will now prove part (a) of the theorem. It follows simply from

$$\begin{aligned} {}_F p_{ij}(t) &= \mathbb{Q}_T\{S_{t+} = j \mid X_t = i\} \\ &= \mathbb{E}_{\mathbb{Q}}(\theta_t^{-1}\theta_{t+} 1_{\{S_{t+}=j\}} \mid X_t = i) \\ &= \mathbb{E}_{\mathbb{Q}}(g_t(X_t, S_{t+}) 1_{\{S_{t+}=j\}} \mid X_t = i) \\ &= g_t(i, j)p_{ij}(t). \end{aligned}$$

We will now prove that the pair $\{X_t, S_{t+}\}_{t=0}^\infty$ is a discrete-time \mathbb{G} -inhomogeneous semi-Markov process under Q_T . In this respect, we have

$$\begin{aligned} Q_T\{\varpi_{ij}(t) = m \mid \mathcal{G}_{t+}\} &= E_{Q_T}(1_{\{\varpi_{ij}(t)=m\}} \mid \mathcal{G}_{t+}) \\ &= \frac{E_Q(\theta_T 1_{\{\varpi_{ij}(t)=m\}} \mid \mathcal{G}_{t+})}{E_Q(\theta_T \mid \mathcal{G}_{t+})} \quad (\text{from Lemma 2.1}) \\ &= E_Q(\theta_{t+}^{-1} \theta_T 1_{\{\varpi_{ij}(t)=m\}} \mid \mathcal{G}_{t+}) \\ &= E_Q(\theta_{t+}^{-1} E_Q(\theta_T \mid \mathcal{G}_{t++}) 1_{\{\varpi_{ij}(t)=m\}} \mid \mathcal{G}_{t+}) \\ &= E_Q(\theta_{t+}^{-1} \theta_{t++} 1_{\{\varpi_{ij}(t)=m\}} \mid \mathcal{G}_{t+}) \\ &= E_Q(f_{t+}(X_t, S_{t+}, \varpi_{X_t S_{t+}}(t)) 1_{\{\varpi_{ij}(t)=m\}} \mid \mathcal{G}_{t+}) \\ &= E_Q(f_{t+}(X_t, S_{t+}, \varpi_{X_t S_{t+}}(t)) 1_{\{\varpi_{ij}(t)=m\}} \mid \sigma(X_t) \vee \sigma(S_{t+})) \quad (\text{from (3.2)}). \end{aligned} \tag{4.4}$$

Equation (4.4) shows that the conditional probability $Q_T\{\varpi_{ij}(t) = m \mid \mathcal{G}_{t+}\}$ is a $(\sigma(X_t) \vee \sigma(S_{t+}))$ -measurable random variable. Since $\sigma(X_t) \vee \sigma(S_{t+}) \subseteq \mathcal{G}_{t+}$, we conclude that X_t is a \mathbb{G} -inhomogeneous semi-Markov process under Q_T , and that

$$Q_T\{\varpi_{ij}(t) = m \mid \mathcal{G}_{t+}\} = Q_T\{\varpi_{ij}(t) = m \mid \sigma(X_t) \vee \sigma(S_{t+})\}.$$

Now, from (4.4) it also follows that

$$\begin{aligned} {}_F h_{ij}(t, m) &= Q_T(\varpi_{ij}(t) = m \mid S_{t+} = j, X_t = i) \\ &= E_Q(\theta_{t+}^{-1} \theta_{t++} 1_{\{\varpi_{ij}(t)=m\}} \mid S_{t+} = j, X_t = i) \\ &= f_{t+}(i, j, m) h_{ij}(t, m), \end{aligned}$$

which concludes the proof of the theorem.

Remark 4.1. It is well known (Howard (1971, p. 579)) that the homogeneous semi-Markov process is a generalization of the homogeneous Markov chain model used by Jarrow *et al.* (1997) for the evolution of a defaultable bond. It is easy to see (Howard (1971, p. 586)) that the discrete-time homogeneous Markov chain is a discrete-time homogeneous semi-Markov process for which all holding times are exactly one unit in length. A simple extension of the above argument is that a discrete-time \mathbb{G} -homogeneous Markov chain is a discrete-time \mathbb{G} -inhomogeneous semi-Markov process for which

$$\begin{aligned} p_{ij}(t) &= p_{ij} && \text{for every } i, j \in \mathbb{K} \text{ and } t = 0, 1, 2, \dots, \\ h_{ij}(t, 1) &= h_{ij}(1) = 1 && \text{for every } t = 0, 1, 2, \dots, \\ h_{ij}(t, m) &= 0 && \text{for } m > 1 \text{ and for every } t = 0, 1, 2, \dots. \end{aligned}$$

For the model of Jarrow *et al.* (1997), Theorem 4.1, where now $\{X_t\}_{t=0}^\infty$ is a discrete-time \mathbb{G} -homogeneous Markov chain under Q , can be stated as follows.

Let the random variable $\theta_t^{-1} \theta_{t+1}$ be $\sigma(X_t, X_{t+1})$ -measurable for any $t = 0, 1, \dots, T - 1$, meaning that for every such t we have $\theta_t^{-1} \theta_{t+1} = g_t(X_t, X_{t+1})$ for some function $g_t: \mathbb{K}^2 \rightarrow \mathbb{R}$. Then $\{X_t\}_{t=0}^\infty$ follows a discrete-time \mathbb{G} -inhomogeneous Markov chain under Q_T and, in addition, ${}_F p_{ij}(t) = p_{ij} g_t(i, j)$.

This is in full agreement with assumption (6) of Jarrow *et al.* (1997), showing that their assumption was based on sound intuition.

5. Stochastic monotonicity

In order to study various forms of stochastic monotonicity, we first need to define the hazard function of default for a defaultable bond that at time 0 is in class j . Thus, let us denote by τ_j the absorption time of X_t when $X_0 = j$. The hazard function of τ_j is then defined to be

$$d_j(t) = Q\{\tau_j = t \mid \tau_j \geq t\}, \quad t = 1, 2, \dots$$

Let $\mathbf{d}^\top(t) = [d_1(t), d_2(t), \dots, d_k(t)]^\top$ and, as in Vassiliou and Papadopoulou (1992), define the *interval transition probability*, i.e. the probability of a bond that entered state i at time s being in state j at time t , $s \leq t$:

$$\tilde{q}_{ij}(s, t) = Q\{\text{the bond is in state } j \text{ at time } t \mid X_s = i\},$$

$$i, j = 1, 2, \dots, k, k + 1, s, t = 0, 1, \dots$$

Then define the matrix $\tilde{\mathbf{Q}}(s, t) = \{\tilde{q}_{ij}(s, t)\}_{i,j \in \mathbb{K}}$, which will be of the form

$$\tilde{\mathbf{Q}}(s, t) = \begin{pmatrix} \mathbf{Q}(s, t) & \mathbf{q}_{k+1}^\top(s, t) \\ \mathbf{0} & 1 \end{pmatrix},$$

where our notation is similar to that in (3.3). Also define

$$\mathbf{W}(s, t) = \text{diag}\{W_1(s, t), W_2(s, t), \dots, W_k(s, t)\},$$

where $W_i(s, t)$ is the survival function of a bond's credit class i .

We now turn to the matter of developing an expression for the interval transition probabilities. How can a process that started in state i at time s be in state j at time t ? One way this can happen is for i and j to be the same state and for the process never to leave state i throughout the period (s, t) . This requires that the process makes its first transition after time t . Every other way to get from state i to state j in the interval (s, t) requires that the process makes at least one transition within this interval. By taking into account all the possible states and times that constitute mutually exclusive paths, we obtain

$$\mathbf{Q}(s, t) = \mathbf{W}(s, t) + \sum_{m=0}^{t-s} [\mathbf{P}(s) \Delta \mathbf{H}(s, m)] \mathbf{Q}(s + m, t), \tag{5.1}$$

where $\mathbf{P}(s) \Delta \mathbf{H}(s, m)$ is the Hadamard product of the two matrices (see Horn and Johnson (1991, Chapter 5)). We similarly obtain

$$\mathbf{q}_{k+1}^\top(s, t) = \sum_{m=0}^{t-s} \{[\mathbf{P}(s) \Delta \mathbf{H}(s, m)] \mathbf{q}_{k+1}^\top(s + m, t) + \mathbf{p}_{k+1}^\top(s) \Delta \mathbf{h}_{k+1}^\top(s, m)\}. \tag{5.2}$$

We now provide the following proposition.

Proposition 5.1. *The vector of hazard functions $\mathbf{d}^\top(t)$ can be given as a function of the interval transition probabilities, as follows:*

$$\mathbf{d}^\top(t) = \mathbf{1}^\top - (\mathbf{I} - \text{diag}\{q_{i,k+1}(0, t - 1)\}_{i=1}^k)^{-1}$$

$$\times \left(\mathbf{1}^\top - \sum_{m=0}^t \{[\mathbf{P}(0) \Delta \mathbf{H}(0, m)] \mathbf{q}_{k+1}^\top(m, t) + \mathbf{p}_{k+1}^\top(0) \Delta \mathbf{h}_{k+1}^\top(0, m)\} \right).$$

Here, $\mathbf{1}^\top$ is the column vector consisting of 1s and \mathbf{I} is the identity matrix.

Proof. Because state $k + 1$ is absorbing, we have

$$q_{j,k+1}(0, t) = Q\{\tau_j \leq t\}, \quad j = 1, 2, \dots, k, t = 1, 2, \dots$$

In addition it is inherently assumed that

$$Q\{\tau_j \leq 0\} = 0 \quad \text{for } j = 1, 2, \dots, k.$$

We have $Q\{\tau_j \geq t\} = 1 - Q\{\tau_j \leq t - 1\}$ and, thus,

$$Q\{\tau_j = t\} = Q\{\tau_j \leq t\} - Q\{\tau_j \leq t - 1\}.$$

The hazard function for default is

$$\begin{aligned} d_j(t) &= Q\{\tau_j = t \mid \tau_j \geq t\} \\ &= \frac{Q\{\tau_j = t\}}{Q\{\tau_j \geq t\}} \\ &= \frac{Q\{\tau_j \leq t\} - Q\{\tau_j \leq t - 1\}}{Q\{\tau_j \geq t\}} \\ &= 1 - \frac{1 - q_{j,k+1}(0, t)}{1 - q_{j,k+1}(0, t - 1)}, \end{aligned} \tag{5.3}$$

where $q_{j,k+1}(0, t - 1) < 1$ for $i = 1, 2, \dots, k$ and $t = 1, 2, \dots$. From (5.2) and (5.3) we obtain the desired result.

Let

$$e_{ij}(s, t) = Q\{\text{the bond enters credit class } j \text{ at time } t \mid \text{it entered credit class } i \text{ at time } s\} \tag{5.4}$$

for $i, j \in \mathbb{K}$, and by $\mathbf{E}(s, t) = \{e_{ij}(s, t)\}$, $i, j = 1, 2, \dots, k$, and $\tilde{\mathbf{E}}(s, t)$ define the matrices of entrance probabilities such that

$$\tilde{\mathbf{E}}(s, t) = \begin{pmatrix} \mathbf{E}(s, t) & \mathbf{e}_{k+1}^\top(s, t) \\ \mathbf{0} & 1_{\{t-s=1\}} \end{pmatrix}, \tag{5.5}$$

with our usual notation. Just as we proved (5.1), we can prove that

$$\mathbf{E}(s, t) = \sum_{m=0}^{t-s} [\mathbf{P}(s) \Delta \mathbf{H}(s, m)] \mathbf{E}(s + m, t) \quad \text{for } s < t, \tag{5.6}$$

with $\mathbf{E}(s, s) = \mathbf{1}$ for $s = 0, 1, \dots$. Furthermore, we can also prove that

$$\mathbf{e}_{k+1}^\top(s, t) = \sum_{m=0}^{t-s} [\mathbf{P}(s) \Delta \mathbf{H}(s, m)] \mathbf{e}_{k+1}^\top(s + m, t) + \mathbf{p}_{k+1}^\top(s) \Delta \mathbf{h}_{k+1}^\top(s, t - s) \tag{5.7}$$

and that

$$\mathbf{Q}(s, t) = \sum_{m=0}^{t-s} \mathbf{E}(s, s + m) \mathbf{W}(s + m, t).$$

We now continue, with the formal definition of *first-order stochastic dominance* (Kijima (1997, p. 129)).

Definition 5.1. For random variables X and Y , X is said to be stochastically greater than Y (written $X \geq_{st} Y$) if

$$E(f(X)) \geq E(f(Y))$$

for all increasing functions $f(x)$ for which the expectation exists.

In the study of stochastic monotonicity, the following matrices have been proved useful (see Kijima (1997, p. 103)):

$$V = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 0 & 1 & 1 & \dots & 1 \\ 0 & 0 & 1 & \dots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}, \quad V^{-1} = \begin{pmatrix} 1 & -1 & 0 & \dots & 0 \\ 0 & 1 & -1 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}. \tag{5.8}$$

We will now prove that bonds of lower credit classes are riskier than their respective upper classes in terms of their times to default.

Theorem 5.1. If, for $i = 1, 2, \dots, k - 1, v = 1, 2, \dots, k$, and $s = 0, 1, \dots$,

$$\sum_{j=1}^v p_{ij}(s)h_{ij}(s, s + 1) \geq \sum_{j=1}^v p_{i+1,j}(s)h_{i+1,j}(s, s + 1), \tag{5.9}$$

then $\tau_j \geq_{st} \tau_{j+1}$ for all j .

Proof. Consider two inhomogeneous semi-Markov chains $\{X_t\}_{t=0}^\infty$ and $\{Y_t\}_{t=0}^\infty$ whose evolution is generated by the same transition probability matrices $\{P(t)\}_{t=0}^\infty$ and $\{H(t, m)\}_{t=0}^\infty$, with $X_0 = j$ and $Y_0 = j + 1$. Let

$$\pi_X(t) = [\pi_{1X}(t), \pi_{2X}(t), \dots, \pi_{k+1,X}(t)],$$

where

$$\pi_{iX}(t) = Q\{X_t \text{ enters state } i \text{ at time } t \mid X_0 = j\},$$

and define

$$\pi_Y(t) = [\pi_{1Y}(t), \pi_{2Y}(t), \dots, \pi_{k+1,Y}(t)]$$

analogously.

We can easily prove that

$$\pi_X(t) = \pi_X(0)\tilde{E}(0, t) = \pi_X(t - 1)\tilde{E}(t - 1, t),$$

and consequently obtain

$$\pi_X(t + 1)V - \pi_Y(t + 1)V = (\pi_X(t)V - \pi_Y(t)V)V^{-1}\tilde{E}(t, t + 1)V. \tag{5.10}$$

For $t = 0$ this equation becomes

$$\pi_X(1)V - \pi_Y(1)V = (\pi_X(0)V - \pi_Y(0)V)V^{-1}\tilde{E}(0, 1)V.$$

Since $\pi_{iX}(0) = 1_{\{i=j\}}$ and $\pi_{iY}(0) = 1_{\{i=j+1\}}$, we easily find that $\pi_X(0)V \geq \pi_Y(0)V$. Let

$$\mathbf{1}_-^\top = [0, 0, \dots, -1]^\top.$$

From (5.8) we then obtain

$$\begin{aligned} \mathbf{V}^{-1} \tilde{\mathbf{E}}(0, 1) \mathbf{V} &= \begin{pmatrix} \mathbf{V}^{-1} & \mathbf{1}^\top \\ \mathbf{0} & 1 \end{pmatrix} \begin{pmatrix} \mathbf{E}(0, 1) & \mathbf{e}_{k+1}^\top(0, 1) \\ \mathbf{0} & 1 \end{pmatrix} \begin{pmatrix} \mathbf{V} & \mathbf{1}^\top \\ \mathbf{0} & 1 \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{V}^{-1} \mathbf{E}(0, 1) \mathbf{V} & \mathbf{0}^\top \\ \mathbf{0} & 1 \end{pmatrix}. \end{aligned} \tag{5.11}$$

Thus, in order to prove that $\pi_X(1)\mathbf{V} \geq \pi_Y(1)\mathbf{V}$, it is sufficient to show that

$$\mathbf{V}^{-1} \mathbf{E}(0, 1) \mathbf{V} \geq \mathbf{0}.$$

From (5.6) we obtain

$$\begin{aligned} \mathbf{V}^{-1} \mathbf{E}(0, 1) \mathbf{V} &= \mathbf{V}^{-1} \left\{ \sum_{m=0}^1 [\mathbf{P}(0) \Delta \mathbf{H}(0, m)] \mathbf{E}(m, 1) \right\} \mathbf{V} \\ &= \mathbf{V}^{-1} [\mathbf{P}(0) \Delta \mathbf{H}(0, 1)] \mathbf{V}, \end{aligned} \tag{5.12}$$

since $\mathbf{H}(0, 0) = \mathbf{0}$, $\mathbf{E}(0, 0) = \mathbf{I}$, and $\mathbf{E}(1, 1) = \mathbf{I}$.

The (i, r) th element of the matrix $\mathbf{V}^{-1} [\mathbf{P}(0) \Delta \mathbf{H}(0, 1)] \mathbf{V}$, for $i = 1, 2, \dots, k - 1$ and $r = 1, 2, \dots, k$, is

$$\sum_{j=1}^r p_{ij}(0) h_{ij}(0, 1) - \sum_{j=1}^r p_{i+1,j}(0) h_{i+1,j}(0, 1),$$

which, according to (5.9), is nonnegative. The remaining elements of the matrix (in the last row) are also nonnegative and, consequently, the matrix is nonnegative.

Now, for $t = 1$ (5.10) becomes

$$\pi_X(2)\mathbf{V} - \pi_Y(2)\mathbf{V} = (\pi_X(1)\mathbf{V} - \pi_Y(1)\mathbf{V}) \mathbf{V}^{-1} \tilde{\mathbf{E}}(1, 2) \mathbf{V}.$$

We have already proved that $\pi_X(1)\mathbf{V} \geq \pi_Y(1)\mathbf{V}$; thus, in order to obtain $\pi_X(2)\mathbf{V} \geq \pi_Y(2)\mathbf{V}$ it remains to show that $\mathbf{V}^{-1} \tilde{\mathbf{E}}(1, 2) \mathbf{V} \geq \mathbf{0}$. This can be proved similarly to (5.12). By arguing recursively, we consequently arrive at the conclusion that, according to (5.9),

$$\pi_X(t)\mathbf{V} \geq \pi_Y(t)\mathbf{V} \quad \text{for } t = 0, 1, \dots,$$

from the k th component of which we obtain

$$\sum_{j=1}^k \pi_{jX}(t) \geq \sum_{j=1}^k \pi_{jY}(t)$$

or, equivalently,

$$\begin{aligned} 1 - \mathbb{Q}\{X_t = k + 1 \text{ and enters at time } t \mid X_0 = j \text{ and enters at time } 0\} \\ \geq 1 - \mathbb{Q}\{Y_t = k + 1 \text{ and enters at time } t \mid Y_0 = j + 1 \text{ and enters at time } 0\}. \end{aligned}$$

Thus,

$$\mathbb{Q}\{\tau_j = t\} \leq \mathbb{Q}\{\tau_{j+1} = t\} \quad \text{for } t = 1, 2, \dots,$$

from which the theorem follows.

Our next proposition reflects an interesting property of the matrix $V^{-1}[\tilde{P}(t)\Delta\tilde{H}(t, m)]V$, which, as we have seen, plays an important role.

Proposition 5.2. *We have*

$$\sum_{m=0}^{\infty} V^{-1}[\tilde{P}(t)\Delta\tilde{H}(t, m)]V = \begin{pmatrix} V^{-1}P(t)V & \mathbf{0}^\top \\ \mathbf{0} & 1 \end{pmatrix}.$$

Proof. From (3.4) and (3.3) we have

$$\begin{aligned} V^{-1}[\tilde{P}(t)\Delta\tilde{H}(t, m)]V &= \begin{pmatrix} V^{-1} & \mathbf{1}^\top \\ \mathbf{0} & 1 \end{pmatrix} \\ &\quad \times \begin{pmatrix} P(t)\Delta H(t, m) & p_{k+1}^\top(t)\Delta h_{k+1}^\top(t) \\ \mathbf{0} & 1_{\{m=t+1\}} \end{pmatrix} \begin{pmatrix} V & \mathbf{1}^\top \\ \mathbf{0} & 1 \end{pmatrix} \\ &= \begin{pmatrix} V^{-1}[P(t)\Delta H(t, m)]V & V^{-1}w^\top(t, m) + \mathbf{1}^\top 1_{\{m=t+1\}} \\ \mathbf{0} & 1_{\{m=t+1\}} \end{pmatrix}, \end{aligned}$$

where

$$w^\top(t, m) = [P(t)\Delta H(t, m)]\mathbf{1}^\top + p_{k+1}^\top(t)\Delta h_{k+1}^\top(t)$$

is the column vector of the probabilities of a defaultable bond moving out of a specific credit class. From

$$\begin{aligned} \sum_{m=0}^{\infty} V^{-1}[P(t)\Delta H(t, m)]V &= V^{-1}P(t)V, \\ \sum_{m=0}^{\infty} V^{-1}w^\top(t, m) + \sum_{m=0}^{\infty} \mathbf{1}^\top 1_{\{m=t+1\}} &= \mathbf{0}, \end{aligned}$$

and

$$\sum_{m=0}^{\infty} 1_{\{m=t+1\}} = 1,$$

the result then follows.

Now define by $\tau_j(\nu)$ the time to default of a bond which at time 0 entered state j and made ν transitions up to the time at which it entered the default state $k + 1$. The question that immediately arises, and which we will answer in what follows, is, under what condition does $\tau_j(\nu) \geq_{st} \tau_{j+1}(\nu)$ hold for every $j = 1, 2, \dots, k$ and $\nu = 0, 1, \dots$? To answer this, we first let

$$e_{ij}(0, \nu, t) = Q\{X_t = j, \text{tr}(t) = \nu \mid X_0 = i\},$$

where $\text{tr}(t) = \nu$ means that the process makes ν transitions in t time steps. Let $\tilde{E}(s, \nu, t) = \{e_{ij}(s, \nu, t)\}_{i, j \in \mathbb{K}}$ and $E(s, \nu, t) = \{e_{ij}(s, \nu, t)\}$, $i, j = 1, 2, \dots, k$, be the matrices of counting transitions entrance probabilities. The matrix $\tilde{E}(s, \nu, t)$ can be written in the form

$$\tilde{E}(s, \nu, t) = \begin{pmatrix} E(s, \nu, t) & e_{k+1}^\top(s, \nu, t) \\ \mathbf{0} & 1_{\{t-s=1\}} 1_{\{\nu=1\}} \end{pmatrix},$$

with our usual notation. Just as we proved (5.6), we can prove that

$$\begin{aligned}
 E(s, v, t) &= \sum_{m=0}^{t-s-v+1} [P(s)\Delta H(s, m)]E(s+m, v-1, t), \\
 e_{k+1}^\top(s, v, t) &= \sum_{m=0}^{t-s-v+1} [P(s)\Delta H(s, m)]e_{k+1}^\top(s+m, v-1, t) \\
 &\quad + 1_{\{v=1\}}[p_{k+1}^\top(s)\Delta h_{k+1}^\top(s, t-s)].
 \end{aligned}$$

Define

$$\pi_X(0, v, t) = [\pi_{1X}(0, v, t), \pi_{2X}(0, v, t), \dots, \pi_{kX}(0, v, t)],$$

where

$$\pi_{iX}(0, v, t) = Q\{X_t = i, \text{tr}(t) = v \mid X_0 = j\}.$$

It can be proved that

$$\pi_X(0, v, t) = \pi_X(0, 0, 0)E(0, v, t).$$

These definitions give us the tools to prove the following theorem along the same lines as Theorem 5.1.

Theorem 5.2. *If, for $i = 1, 2, \dots, k - 1, \mu = 1, 2, \dots, k$, and $s = 0, 1, \dots$,*

$$\sum_{j=1}^{\mu} p_{ij}(s)h_{ij}(s, s+1) \geq \sum_{j=1}^{\mu} p_{i+1,j}(s)h_{i+1,j}(s, s+1),$$

then $\tau_j(v) \geq_{st} \tau_{j+1}(v)$ for all j .

We will now present in a compact way some other known results on various stochastic monotonicities (for details, see Kijima (1997, p. 121), (1998)), which will prove to be useful later.

Definition 5.2. Suppose that the random variables X and Y have respective distribution functions $F(x)$ and $G(y)$, probability density functions $f(x)$ and $g(y)$, and survival functions $\mathfrak{F}(x)$ and $\mathfrak{G}(y)$.

- (i) X is said to be greater than Y in the sense of likelihood ratio ordering, written $X \geq_{lr} Y$, if $f(x)g(y) \geq f(y)g(x)$ for all $x > y$.
- (ii) X is said to be greater than Y in the sense of hazard rate, written $X \geq_{hr} Y$, if $\mathfrak{F}(x)\mathfrak{G}(y) \leq \mathfrak{F}(y)\mathfrak{G}(x)$ for all $x < y$.
- (iii) X is said to be greater than Y in the sense of reversed hazard rate, written $X \geq_{rh} Y$, if $F(x)G(y) \leq F(y)G(x)$ for all $x < y$.

Definition 5.3. A nonnegative matrix A is called totally positive of order 2, written $A \in TP_2$, if its 2×2 minors are all nonnegative.

Proposition 5.3. *Let X and Y be discrete random variables defined on \mathbb{N} . Then*

$$X \geq_{hr} Y \quad \text{if and only if} \quad d_{Xi} \leq d_{Yi}.$$

Proof. The hazard function is

$$d_{Xi} = \frac{Q\{X = i\}}{Q\{X \geq i\}}.$$

From the definition above, for $x = i$ and $y = i + 1$ we have $X \geq_{hr} Y$ if and only if

$$\frac{\mathfrak{F}(i + 1)}{\mathfrak{F}(i)} \geq \frac{\mathfrak{G}(i + 1)}{\mathfrak{G}(i)} \Leftrightarrow d_{Xi} \leq d_{Yi}.$$

Lemma 5.1. (Kijima (1997, p. 122).) *Suppose that X and Y are discrete random variables defined on \mathbb{N} with respective probability vectors $\mathbf{p} = (p_i)$ and $\mathbf{q} = (q_j)$. Then*

- (i) $X \geq_{lr} Y$ if and only if $\begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix} \in TP_2$,
- (ii) $X \geq_{hr} Y$ if and only if $\begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix} \mathbf{V}^\top \in TP_2$,
- (iii) $X \geq_{rh} Y$ if and only if $\begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix} \mathbf{V} \in TP_2$.

Corollary 5.1. (Kijima (1997, p. 111).) *Let \mathbf{A} and \mathbf{B} be finite, nonnegative matrices.*

- (i) *Suppose that $\mathbf{A}, \mathbf{B} \in TP_2$. Then $\mathbf{AB} \in TP_2$.*
- (ii) *Suppose that \mathbf{AV} and \mathbf{BV} have no null rows or columns. If $\mathbf{AV}, \mathbf{BV} \in TP_2$ and $\mathbf{V}^{-1}\mathbf{BV} \geq \mathbf{0}$, then $\mathbf{ABV} \in TP_2$.*
- (iii) *Suppose that \mathbf{AV}^\top and \mathbf{BV}^\top have no null rows or columns. If $\mathbf{AV}^\top, \mathbf{BV}^\top \in TP_2$ and $(\mathbf{V}^\top)^{-1}\mathbf{BV} \geq \mathbf{0}$, then $\mathbf{ABV}^\top \in TP_2$.*

Theorem 5.3. (Kijima (1997, p. 123).) *For two discrete random variables X and Y , $X \geq_{lr} Y$ implies that both $X \geq_{hr} Y$ and $X \geq_{rh} Y$. Both $X \geq_{hr} Y$ and $X \geq_{rh} Y$ imply that $X \geq_{st} Y$.*

Let $\hat{\mathbb{K}} = \{1, 2, \dots, k\}$ be the set of nonabsorbing states. Define by \hat{X}_t the stochastic process with state space $\hat{\mathbb{K}}$ given that absorption has not occurred up to time t . Let $\pi_{\hat{X}}(0)$ be the initial probability distribution of \hat{X}_0 and suppose that it is concentrated on $\hat{\mathbb{K}}$, i.e. $\pi_{\hat{X}}(0) \neq \mathbf{0}$. It is easy to see that

$$Q\{\tau \geq t \mid \pi_{\hat{X}}(0)\} = \pi_{\hat{X}}(0)\mathbf{E}(0, t - 1)\mathbf{1}^\top.$$

Now let

$$\pi_{\hat{X}}(t) = [Q\{X_t = 1 \mid \tau \geq t, \pi_{\hat{X}}(0)\}, \dots, Q\{X_t = k \mid \tau \geq t, \pi_{\hat{X}}(0)\}].$$

It is then easy to prove that

$$\pi_{\hat{X}}(t) = \frac{\pi_{\hat{X}}(0)\mathbf{E}(0, t)}{\pi_{\hat{X}}(0)\mathbf{E}(0, t)\mathbf{1}^\top} = \pi_{\hat{X}}(0)\hat{\mathbf{E}}(0, t).$$

We will refer to \hat{X}_t as the *conditional inherent inhomogeneous semi-Markov chain* of X_t . Since state $k + 1$ is absorbing, we have

$$Q\{\tau = t \mid X_0 = j\} = \sum_{i \neq k+1} Q\{X_{t-1} = i, X_t = k + 1 \mid X_0 = j\}$$

and it easy to prove that the hazard function of τ_j is given by

$$d_j(t) = E(f(\hat{X}_{t-1}) \mid X_0 = j),$$

where $f(\hat{X}_{t-1} = i) = e_{i,k+1}(t - 1, t)$. We can now present the following theorem.

Theorem 5.4. *For the conditional inherent inhomogeneous semi-Markov chain $\{\hat{X}_t\}$, assume that $V^{-1}\hat{E}(t, t + 1)V \geq \mathbf{0}$ for every t . If, in addition, $\hat{E}(t, t + 1)V \in TP_2$, then $\tau_j \geq_{hr} \tau_{j+1}$ for all j .*

Proof. Let \hat{X}_t be the conditional inherent inhomogeneous semi-Markov process of X_t starting at j , and let \hat{Y}_t be the conditional inherent inhomogeneous semi-Markov process of X_t starting at $j + 1$. Then obviously

$$\hat{\pi}_X(t) = \hat{\pi}_X(0)\hat{E}(0, t) \quad \text{and} \quad \hat{\pi}_Y(t) = \hat{\pi}_Y(0)\hat{E}(0, t).$$

From Definition 5.3 and (5.8), we easily obtain

$$\begin{pmatrix} \hat{\pi}_X(0) \\ \hat{\pi}_Y(0) \end{pmatrix} V \in TP_2.$$

As an induction hypothesis, assume that

$$\begin{pmatrix} \hat{\pi}_X(t) \\ \hat{\pi}_Y(t) \end{pmatrix} V \in TP_2. \tag{5.13}$$

From the assumptions of the theorem and of Corollary 5.1, we find that

$$\begin{pmatrix} \hat{\pi}_X(t) \\ \hat{\pi}_Y(t) \end{pmatrix} \hat{E}(t, t + 1)V \in TP_2$$

and, for every t , from (5.13) we immediately obtain

$$\begin{pmatrix} \hat{\pi}_X(t + 1) \\ \hat{\pi}_Y(t + 1) \end{pmatrix} V \in TP_2.$$

From Lemma 5.1 we then have $\hat{Y}_t \geq_{rh} \hat{X}_t$, and, since $\hat{Y}_t \geq_{rh} \hat{X}_t$ implies $\hat{Y}_t \geq_{st} \hat{X}_t$, it follows that

$$d_j(t + 1) = E(f(\hat{X}_t) \mid X_0 = j) \leq E(f(\hat{Y}_t) \mid Y_0 = j + 1) = d_{j+1}(t + 1),$$

from which the result follows.

Definition 5.4. A random variable X with survival function \mathfrak{F} is called *new better than used*, written $X \in NBU$, if

$$\mathfrak{F}(x + y) \leq \mathfrak{F}(x)\mathfrak{F}(y). \tag{5.14}$$

If the opposite inequality is valid then the random variable X is called *new worse than used*, written $X \in NWU$.

From the above equation and its physical meaning it is apparent why the name NBU is used.

Theorem 5.5. *Assume that*

$$U^{-1}[P(s)\Delta H(s, m)]U \geq \mathbf{0} \quad \text{for every } s, m = 0, 1, \dots, \tag{5.15}$$

where $U = V^T$. Let τ be the time of default. If the initial state is 1 then $\tau \in NWU$, and if the initial state is k then $\tau \in NBU$.

Proof. Let $\hat{\pi}_X(0)$ have a 1 in state 1. Then

$$Q\{\tau \geq t \mid \hat{\pi}_X(0)\} = \hat{\pi}_X(0)E(0, t - 1)\mathbf{1}^\top.$$

From (5.14), we have

$$U^{-1}E(0, 1)U = U^{-1}[P(0)\Delta H(0, 1)]U \geq \mathbf{0}.$$

Now consider the matrix of entrance probabilities $E(s, t)$, for $t - s = 2$. For easier exposition we set $s = 0$. Then, using (5.6), we obtain

$$\begin{aligned} U^{-1}E(0, 2)U &= U^{-1}[P(0)\Delta H(0, 1)]UU^{-1}E(1, 2)U \\ &\quad + U^{-1}[P(0)\Delta H(0, 2)]U \\ &\geq \mathbf{0}. \end{aligned} \tag{5.16}$$

Working by induction, assume that

$$U^{-1}E(s, t)U \geq \mathbf{0} \quad \text{for every } t \text{ and } s \text{ such that } t - s = 1, 2, \dots, k.$$

From (5.6) we then obtain

$$\begin{aligned} U^{-1}E(s, t + 1)U &= \sum_{m=0}^{t-s} U^{-1}[P(s)\Delta H(s, m)]UU^{-1}E(s + m, t + 1)U \\ &\quad + U^{-1}[P(s)\Delta H(s, t - s + 1)]U \\ &\geq \mathbf{0}. \end{aligned}$$

Thus, from (5.6) and (5.16),

$$U^{-1}E(s, t)U \geq \mathbf{0} \quad \text{for every } s, t = 0, 1, \dots$$

We also have

$$U^{-1}E(0, t)\mathbf{1}^\top = U^{-1}E(0, t)UU^{-1}\mathbf{1}^\top = U^{-1}E(0, t)U\hat{\pi}_X^\top(0) \geq \mathbf{0}^\top,$$

from which we conclude that the sums of the rows of $E(0, t)$ are increasing as the number of the row i ($i = 1, 2, \dots, k$) increases, and that result holds for every $t = 1, 2, \dots$. From the last conclusion, (5.15), and since $\hat{\pi}_X(0) = [1, 0, \dots, 0]$, we obtain

$$E(0, t - 1)\mathbf{1}^\top \geq Q\{\tau \geq t \mid \hat{\pi}_X(0)\}\mathbf{1}^\top \quad \text{for every } t = 1, 2, \dots$$

Finally, we have

$$\mathfrak{F}_0(x + y) := Q\{\tau \geq x + y \mid \hat{\pi}_X(0)\} = \hat{\pi}_X(0)E(0, x + y - 1)\mathbf{1}^\top$$

and, as can be proved,

$$E(0, x + y - 1) \geq E(0, x - 1)E(x, x + y - 1).$$

From this it follows that

$$\begin{aligned} \mathfrak{F}_0(x + y) &\geq \hat{\pi}_X(0)E(0, x - 1)E(x, x + y - 1)\mathbf{1}^\top \\ &\geq \hat{\pi}_X(0)E(0, x - 1)Q\{\tau \geq y \mid \hat{\pi}_X(x)\}\mathbf{1}^\top \\ &\geq Q\{\tau \geq y \mid \hat{\pi}_X(x)\}\hat{\pi}_X(0)E(0, x - 1)\mathbf{1}^\top \\ &\geq Q\{\tau \geq x \mid \hat{\pi}_X(0)\}Q\{\tau \geq y \mid \hat{\pi}_X(x)\} \\ &= \mathfrak{F}_0(x)\mathfrak{F}_0(y) \end{aligned}$$

for every $x, y = 1, 2, \dots$, whence $\tau \in \text{NWU}$. The other case follows similarly.

6. Term structure of credit spread and model calibration

We have proved that the valuation process of a defaultable bond is given by (2.8). By definition, the one-step forward rate of the defaultable bond at the future time T , as seen at time $t \leq T$, is given by

$$f_D(t, T) = -\log\left(\frac{D^\delta(t, T + 1)}{D^\delta(t, T)}\right).$$

Likewise, for the default-free bond, we define the one-step forward rate by

$$f(t, T) = -\log\left(\frac{B(t, T + 1)}{B(t, T)}\right).$$

It follows that the credit spread process will be given by

$$s(t, T) = f_D(t, T) - f(t, T) = \log\left(\frac{\delta + (1 - \delta) Q_T\{\tau > T \mid \mathcal{G}_t\}}{\delta + (1 - \delta) Q_{T+1}\{\tau > T + 1 \mid \mathcal{G}_t\}}\right).$$

The above formula is apparently valid for any credit state i that the defaultable bond is in at time t . For the value at time t of the defaultable bond with maturity T that is in credit state i at time t , we then have

$$D_i^\delta(t, T) := B(t, T)(\delta + (1 - \delta) Q_T\{\tau > T \mid X_t = i\}), \quad i = 1, 2, \dots, k. \quad (6.1)$$

Note that different values are brought into the formula each time by the calculation of the probability $Q_T\{\tau > T \mid X_t = i\}$.

We will now discuss the issue of calibration of the inhomogeneous semi-Markov model. The calibration of the simple Markov chain model by Jarrow *et al.* (1997) was discussed by Bielecki and Rutkowski (2002, p. 360) and was extended, taking into consideration some empirical problems motivated by actual values of the real-world probabilities, by Kijima and Komoribayashi (1998). In the present case, modifications arise from two sources: the change of the model into a semi-Markov model and the inhomogeneity in time imposed in the proposed general model, which also alters the data needed to calibrate the model.

We assume that we have data for a time window $[0, T^*]$, where T^* is the present. The inputs in our problem are assumed to be the following.

(i) Data estimated from the available historical data on credit migrations for real-world probability matrix sequences $\{P(t)\}_{t=0}^{T^*-1}$ and $\{H(t, m)\}_{t,m=0}^{T^*-1}$ during our time window $[0, T^*]$. Methods of estimation with good properties for the estimators for inhomogeneous semi-Markov models are provided in Subsection 6.1.

(ii) The term structure of default-free bonds, that is, the market values of $B(t, T)$, for $t = 0, 1, \dots, T^*$ and $T = t + 1, \dots, T^*$.

(iii) The observed term structures of defaultable bonds from various credit classes $D_i^\delta(t, T)$, for $i = 1, 2, \dots, k, t = 1, 2, \dots, T^*$, and $T = t + 1, \dots, T^*$.

(iv) Data estimated from the available historical data on recovery rates, δ , of corporate bonds. Note that the methods presented allow for δ to be a function of the initial credit state, say i . All that is needed in such cases is the replacement of δ by δ_i in our formulae.

Our aim is to identify the probability matrix sequences, $\{ {}_F P(t) \}_{t=0}^{T^*-1}$ and $\{ {}_F H(t, m) \}_{t,m=0}^{T^*-1}$, of the probabilities ${}_F p_{ij}(t)$ and ${}_F h_{ij}(t, m)$ defined in Theorem 4.1, that cause the observed market prices $D_i^\delta(t, T)$ within our time window $[0, T^*]$ to coincide with the theoretical values given by the inhomogeneous semi-Markov model, through (6.1).

At this stage of the calibration process, we make assumptions on the inhomogeneous semi-Markov model analogous to those made by Jarrow *et al.* (1997) on the simple Markov model. Furthermore, we will also take into consideration the risk premium adjustments proposed by Kijima and Komoribayashi (1998). The following assumptions are compatible with Theorem 4.1.

Assumption 6.1. (The risk premium assumptions.) (i) For every $t \in [0, T^* - 1]$ and $i, j = 1, 2, \dots, k$, assume that

$${}_F p_{ij}(t) = \pi_i(t)p_{ij}(t) \quad \text{and} \quad {}_F p_{i,k+1}(t) = 1 - \pi_i(t)(1 - p_{i,k+1}(t)). \tag{6.2}$$

(ii) For every $t \in [0, T^* - 1]$, $i = 1, 2, \dots, k$, and $j = 1, 2, \dots, k + 1$, assume that

$${}_F h_{ij}(t, m) = \eta_i(t, m)h_{ij}(t, m). \tag{6.3}$$

We call the functions $\pi_i(t)$ and $\eta_i(t, m)$ the risk premium adjustments for the probabilities ${}_F p_{ij}(t)$ and ${}_F h_{ij}(t, m)$, respectively. The conditions that the risk premium adjustments should satisfy are

$$0 < \pi_i(t) \leq \frac{1}{1 - p_{i,k+1}(t)} \quad \text{for all } t \in [0, T^* - 1] \text{ and } i = 1, 2, \dots, k, \tag{6.4}$$

which is a consequence of the facts that

$$\sum_{j=1}^{k+1} {}_F p_{ij}(t) = 1 \quad \text{for all } t \in [0, T^* - 1] \text{ and } i = 1, 2, \dots, k$$

and

$$\sum_{m=0}^{\infty} \eta_i(t, m)h_{ij}(t, m) = 1 \quad \text{for all } t \in [0, T^* - 1], i = 1, 2, \dots, k, \text{ and } j = 1, 2, \dots, k + 1.$$

Note here that the $p_{i,k+1}(t)$ values are typically very small numbers (Carty and Fons (1994), Kijima and Komoribayashi (1998)) and, thus, (6.4) is easily satisfied in practice.

Now, from (5.4) and (5.5) we have

$$\begin{aligned} Q_T\{\tau > T \mid X_t = i\} &= 1 - Q_T\{\tau \leq T \mid X_t = i\} \\ &=: 1 - \sum_{s=t}^T {}_F e_{i,k+1}(s, T), \quad i = 1, 2, \dots, k. \end{aligned}$$

Using (6.1) and the above formula we easily find that

$${}_F e_{i,k+1}(t, t + 1) = \frac{B(t, t + 1) - D_i^\delta(t, t + 1)}{(1 - \delta)B(t, t + 1)} \quad \text{for every } t \in [0, T^* - 1].$$

Since by Theorem 4.1 the change of measure retains the semi-Markov process properties, we can use (5.7) and the risk premium assumptions (6.2) and (6.3) to show that, for every $t \in [0, T^* - 1]$,

$${}^F e_{i,k+1}(t, t + 1) = {}^F p_{i,k+1}(t) {}^F h_{i,k+1}(t, 1) = \pi_i(t) \eta_i(t, 1) p_{i,k+1}(t) h_{i,k+1}(t, 1)$$

and

$$\pi_i(t) \eta_i(t, 1) = \frac{B(t, t + 1) - D_i^\delta(t, t + 1)}{(1 - \delta)B(t, t + 1) p_{i,k+1}(t) h_{i,k+1}(t, 1)}.$$

Having thus found the products of the risk premiums $\pi_i(t) \eta_i(t, 1)$, $t \in [0, T^* - 1]$, and knowing the corresponding real-world probabilities $\mathbf{P}(t)$ and $\mathbf{H}(t, 1)$, we can immediately find ${}^F \mathbf{P}(t) \Delta_F \mathbf{H}(t, 1)$ and, from (5.6), obtain

$${}^F \mathbf{E}(t, t + 1) = {}^F \mathbf{P}(t) \Delta_F \mathbf{H}(t, 1), \quad t \in [0, T^* - 1].$$

These considerations suggest a method of evaluating the risk premiums, which we summarize in the following algorithm.

Algorithm 6.1. (*Risk premium algorithm.*)

Step 1. For $t = 0, 1, \dots, T^* - 1$, for $i = 1, 2, \dots, k$, find

$${}^F e_{i,k+1}(t, t + 1) = \frac{B(t, t + 1) - D_i^\delta(t, t + 1)}{(1 - \delta)B(t, t + 1)},$$

from which can be obtained

$$\pi_i(t) \eta_i(t, 1) = \frac{B(t, t + 1) - D_i^\delta(t, t + 1)}{(1 - \delta)B(t, t + 1) p_{i,k+1}(t) h_{i,k+1}(t, 1)}.$$

Since the change to forward measure preserves the semi-Markov process, it is straightforward to find the products ${}^F \mathbf{P}(t) \Delta_F \mathbf{H}(t, 1)$ and, consequently, the probabilities:

$${}^F \mathbf{E}(t, t + 1) = {}^F \mathbf{P}(t) \Delta_F \mathbf{H}(t, 1).$$

Step 2. For $t = 0, 1, \dots, T^* - 2$, for $i = 1, 2, \dots, k$, find

$${}^F e_{i,k+1}(t, t + 2) = \frac{B(t, t + 2) - D_i^\delta(t, t + 2)}{(1 - \delta)B(t, t + 2)} - {}^F e_{i,k+1}(t, t + 1). \tag{6.5}$$

From (5.7), since the change to forward measure preserves the semi-Markov process, we find that

$$\begin{aligned} {}^F \mathbf{e}_{k+1}^\top(t, t + 2) &= [{}^F \mathbf{P}(t) \Delta_F \mathbf{H}(t, 1)] {}^F \mathbf{e}_{k+1}^\top(t + 1, t + 2) + {}^F \mathbf{p}_{k+1}^\top(t) \Delta_F \mathbf{h}_{k+1}^\top(t, 2) \\ &= [{}^F \mathbf{P}(t) \Delta_F \mathbf{H}(t, 1)] {}^F \mathbf{e}_{k+1}^\top(t + 1, t + 2) \\ &\quad + \pi_i(t) \eta_i(t, 2) [{}^F \mathbf{p}_{k+1}^\top(t) \Delta_F \mathbf{h}_{k+1}^\top(t, 2)]. \end{aligned} \tag{6.6}$$

Solving the simple system of equations (6.5) and (6.6) yields $\pi_i(t) \eta_i(t, 2)$.

From (5.6), since the change to forward measure preserves the semi-Markov process, we have

$${}^F \mathbf{E}(t, t + 2) = [{}^F \mathbf{P}(t) \Delta_F \mathbf{H}(t, 1)] [{}^F \mathbf{P}(t + 1) \Delta_F \mathbf{H}(t + 1, 1)] + {}^F \mathbf{P}(t) \Delta_F \mathbf{H}(t, 2),$$

from which we calculate the probabilities ${}^F \mathbf{E}(t, t + 2)$.

We continue in the way step 2 indicates to find $\pi_i(t)\eta_i(t, 3)$, $\pi_i(t)\eta_i(t, 4)$, and so on, until all possible products of the form $\pi_i(t)\eta_i(t, m)$, $t \in [0, T^* - 1]$, have been calculated. Note that, in our time window $[0, T^*]$, for a specific value of t the maximum value of m for which we can find the risk premium products $\pi_i(t)\eta_i(t, m)$ is $m = T^* - t$. As we have seen from the algorithm, the risk premium products $\pi_i(t)\eta_i(t, m)$ provide sufficient information to find the probabilities ${}_F P(t) \Delta_F H(t, m)$ and ${}_F \mathbf{p}_{k+1}^\top(t) \Delta_F \mathbf{h}_{k+1}^\top(t, m)$, which in turn provide sufficient information to find the probabilities ${}_F E(t, t + m)$ and ${}_F \mathbf{e}_{k+1}^\top(t, t + m)$.

In addition, there is a way to find the separate values of $\pi_i(t)$ and $\eta_i(t, m)$ for which sufficient information has been calculated with the proposed algorithm from our time window $[0, T^*]$. Assume that, for a specific t , $M_{ij}(t)$ is the maximum value of m for which $h_{ij}(t, m)$ is positive, and then set

$$h_{ij}(t, M_{ij}(t)) = 1 - \sum_{m=1}^{M_{ij}(t)-1} h_{ij}(t, m).$$

Since, by Theorem 4.1, the semi-Markov property is preserved, $\sum_{m=0}^\infty {}_F h_{ij}(t, m) = 1$ for every $t \in [0, T^* - 1]$ and $i, j = 1, 2, \dots, k$, and, consequently, $\sum_{m=0}^{M_{ij}(t)} \eta_i(t, m)h_{ij}(t, m) = 1$. From this we obtain

$$\pi_i(t) = \sum_{m=0}^{M_{ij}(t)} \pi_i(t)\eta_i(t, m)h_{ij}(t, m),$$

which provides the values $\pi_i(t)$ and, consequently, the corresponding values

$$\eta_i(t, m), \eta_i(t, m + 1), \dots, \eta_i(t, M_{ij}(t)).$$

6.1. Estimation

One of the questions which naturally arises for the practitioner is that of how to estimate, from the available data, the real-world probabilities for the inhomogeneous semi-Markov process. In the studies by Carty and Fons (1994), Lando (2004, p. 95), and Christensen *et al.* (2004), the rating transition histories were taken from the complete ‘Moody’s Corporate Bond Default Database’, that is, the edition containing complete issuer histories since 1970. In all studies seven rating categories were distinguished: Aaa, Aa, A, Baa, Ba, B, Caa, and Default.

The rating category Aaa is the most senior and we denote it by 1; Aa is the next most senior category and is denoted by 2, and so on along the list. In statistical analysis, the estimation method is closely connected to the sampling scheme, i.e. the type of data available. In the Moody’s Investors Service database the data are collected as shown in Table 1, which contains two typical examples from Lando (2004, p. 94) and Carty and Fons (1994).

The entry with rating WR reports that the issuers had their rating withdrawn at the date referred to. A senior rating might be withdrawn for any number of reasons, from retirement of all rated debt to completion of an exchange offer for all rated debt. We will treat the ‘withdrawn’ flow as Type III censoring (Lee (1992, p. 26)), as is very common with biomedical data, or as right censoring, which is common elsewhere. We will also assume that our censoring scheme is independent, which in our case means that the reason for the rating WR is not that they understand that their possibilities of default have increased.

In the semi-Markov chain model we have provided, one important variable to be estimated is the conditional density function of transition of a credit rating from state i to any other credit rating or default, given either that a specific credit rating or default has been selected for transition, or, equivalently, that transition to a specific credit rating or default will eventually

TABLE 1: Database examples.

Example 1		Example 2	
Date	Rating	Date	Rating
29 May 1998	Ba	11 January 1984	A
27 December 2000	B	20 June 1991	Baa
1 October 2001	Default	7 February 1993	WR

occur. This problem is essentially the same as that, in manpower planning, of people moving among the grades of an organization, or leaving it, as discussed in Bartholomew *et al.* (1991, p. 184). It is also essentially the same as the competing risk model in the medical literature, as discussed by Kalbfleisch and Prentice (2002, Chapter 8), Lee (1992, p. 26), and Cox and Oakes (1984, Chapter 4), where the event of interest is frequently time elapsed prior to either the appearance of various symptoms or death. Similar problems also appear in reliability theory, where the event of interest is time elapsed until a component breaks down, and actuarial studies, as discussed in Elandt-Johnson and Johnson (1980, Chapter 7). For all the above kinds of data, much work has been done on both parametric and nonparametric estimation. For our present purposes, we will start by providing a nonparametric estimation method for the probabilities $p_{ij}(t)$ and $h_{ij}(t, m)$, $i, j = 1, 2, \dots, k + 1, t, m = 1, 2, \dots$, that is an extension of the competing risks estimation methods for semi-Markov processes discussed by Bartholomew *et al.* (1991, p. 184).

It is well known that our semi-Markov process can be completely specified by using the transition probabilities $p_{ij}(t)$ and either the probability distributions $h_{ij}(t, m)$, $i, j = 1, 2, \dots, k + 1, t, m = 1, 2, \dots$, or the hazard rates $r_{ij}(t, m)$, which correspond to the probability distributions $h_{ij}(t, m)$. The two ways are equivalent, so which one is chosen is purely a matter of convenience and is often determined by ease of estimation.

We assume the data to be grouped into cohorts of bonds entered into a specific rating category at about the same time, from 1970 onwards, at intervals of three months, six months, or a year (see Hamilton (2001, Exhibit 43)). In this way, we assume that each cohort is a homogeneous group of entrants to credit rating i at time t . For each (i, t) cohort, for $i = 1, 2, \dots, 7$ and $t = 1, 2, \dots, T^* - 1$, collect the data into intervals $[m, m + 1)$, where the length of the interval is equal to the time interval for $t = 1, 2, \dots, T^* - 1$ and $m = 1, 2, \dots, M$. Here M is the maximum length of stay in any credit rating before revaluation. Let $R_{ij}^{(m)}(t)$ be the number of bonds that leave credit rating i to go to credit rating j in the m th interval, given that they entered credit rating i at time t . Also, let $W_i^{(m)}(t)$ be the number of bonds withdrawn in the time interval m , given that they entered credit rating i at time t . Then the bond movements for each (i, t) cohort can be tabulated in the typical form seen in the following array:

	1	2	3	...	$M - 1$	M
$R_{i1}^{(1)}(t)$	$R_{i1}^{(2)}(t)$	$R_{i1}^{(3)}(t)$...	$R_{i1}^{(M-1)}(t)$	$R_{i1}^{(M)}(t)$	
$R_{i2}^{(1)}(t)$	$R_{i2}^{(2)}(t)$	$R_{i2}^{(3)}(t)$...	$R_{i2}^{(M-1)}(t)$	$R_{i2}^{(M)}(t)$	
\vdots	\vdots	\vdots	\ddots	\vdots	\vdots	
$R_{i,k+1}^{(1)}(t)$	$R_{i,k+1}^{(2)}(t)$	$R_{i,k+1}^{(3)}(t)$...	$R_{i,k+1}^{(M-1)}(t)$	$R_{i,k+1}^{(M)}(t)$	
$W_i^{(1)}(t)$	$W_i^{(2)}(t)$	$W_i^{(3)}(t)$...	$W_i^{(M-1)}(t)$	$W_i^{(M)}(t)$	

Define

$$W_i(t) = \sum_{m=1}^M W_i^{(m)}(t), \quad i = 1, 2, \dots, k, \quad t = 1, 2, \dots, T.$$

This is the number of bonds withdrawn (right censored) over the entire time of observation of cohort (i, t) or, equivalently, the number of bonds eventually censored from cohort (i, t) . Then define

$$N_{ij}(t) = \sum_{m=1}^M R_{ij}^{(m)}(t), \quad i = 1, 2, \dots, k, \quad j = 1, 2, \dots, k + 1, \quad t = 1, 2, \dots, T.$$

This is the total number of bonds that move from credit rating i to credit rating j in cohort (i, t) or, equivalently, the number of bonds that eventually move from credit rating i to credit rating j . The total size of cohort (i, t) is thus

$$N_i(t) = \sum_{j=1}^{k+1} \sum_{m=1}^M R_{ij}^{(m)}(t) + \sum_{m=1}^M W_i^{(m)}(t).$$

The probability of a bond that enters credit rating i at time t selecting to move to credit rating j or, equivalently, of a bond that enters credit rating i at time t to eventually move to credit rating j is then estimated (using multinomial flow) as

$$\hat{p}_{ij}(t) = \frac{N_{ij}(t)}{N_i(t) - W_i(t)}, \quad i = 1, 2, \dots, k, \quad j = 1, 2, \dots, k + 1, \quad t = 1, 2, \dots, T.$$

We now estimate the probability distribution $h_{ij}(t, m)$ for specific values of i and j , and for $m = 1, 2, \dots, M$. As estimator, we take $\hat{h}_{ij}(t, m)$, which is the probability of a bond moving from credit rating i to credit rating j during the interval $[m - 1, m)$, given that the bond entered credit rating i at time t and selected to move to credit rating j in its next transition. Equivalently, this is the probability of a bond moving from credit rating i to credit rating j during the interval $[m - 1, m)$, given that the bond entered credit rating i at time t and eventually moves to credit rating j . Thus,

$$\begin{aligned} \hat{h}_{ij}(t, m) &= \frac{R_{ij}^{(m)}(t)/N_i(t)}{N_{ij}(t)/(N_i(t) - W_i(t))} \\ &= \frac{R_{ij}^{(m)}(t) N_i(t) - W_i(t)}{N_{ij}(t) N_i(t)} \\ &= \frac{R_{ij}^{(m)}(t)}{N_{ij}(t)} P_t(i), \end{aligned}$$

where $P_t(i)$ is the probability of a bond entering credit rating i at time t to be uncensored or, equivalently, not to be withdrawn.

By looking at $R_{ij}^{(m)}(t)$ as a multinomial sample from the population $N_{ij}(t)$ with probability $h_{ij}(t, m)$ of moving to state j within the interval $[m - 1, m)$, we obtain the standard error of $\hat{h}_{ij}(t, m)$ to be

$$se(\hat{h}_{ij}(t, m)) = \left[\frac{\hat{h}_{ij}(t, m)(1 - \hat{h}_{ij}(t, m))}{N_{ij}(t)} \right]^{1/2} P_t(i).$$

Now consider the null hypothesis that the transition probabilities $p_{ij}(t)$ remain constant with respect to the time t of entering credit rating i : $p_{ij}(t) = p_{ij}$ for a specific j , given i .

An application of the results of Anderson and Goodman (1957) and Vassiliou (1976), considering the flow of bonds which eventually move from credit rate i to credit rate j as a multinomial random variable, then leads to the conclusion that the statistic

$$\chi^2(i, j) = \sum_{t=0}^{T^*-1} N_i(t) \frac{(\hat{p}_{ij}(t) - \hat{p}_{ij})^2}{\hat{p}_{ij}},$$

where

$$\hat{p}_{ij} = \frac{\sum_{t=0}^{T^*-1} N_{ij}(t)}{\sum_{t=0}^{T^*-1} [N_i(t) - W_i(t)]}$$

is the maximum likelihood estimate under the null hypothesis, is chi-square distributed with $T^* - 1$ degrees of freedom.

In an obvious analogous way we can test the constancy over time of the probabilities $h_{ij}(t, m)$ for specific values of i, j , and m . Alternatively, if we estimate the survivor function by

$$\hat{\mathcal{H}}_{ij}(t, m) = 1 - \sum_{x \leq m} \hat{h}_{ij}(t, x), \quad t = 1, 2, \dots, T,$$

then we can use the statistics designed for biomedical data testing survivor distributions under $T^* - 1$ different treatments; see, for example, Lee (1992, p. 26) or Elandt-Johnson and Johnson (1980, Chapter 7). However, in such a case we lose information from the sample since, for example, for the 2002 cohort we will not have an estimate of the survivor function for all the possible values of m .

The way the ‘Moody’s Corporate Bond Default Database’ is organized permits an alternative estimation of the survivor distributions $\mathcal{H}_{ij}(t, m)$ through the use of the Kaplan–Meyer estimators for semi-Markov processes (see Kalbfleisch and Prentice (2002, Chapter 8), McClean (1976), (1986), McClean and Gribbin (1987), (1991), McClean *et al.* (1997), and McClean and Montgomery (1999)). To do so we must extend the existing results for the inhomogeneous case in an analogous way as previously. However, there will still be a need to find the discrete estimates of the survivor functions at each time $t = 1, 2, \dots, T^* - 1$. This makes the Kaplan–Meyer estimators more appropriate to the continuous-time version of the present theory.

The inhomogeneity in time of transition probabilities has been reported by many authors, such as Duffie and Singleton (2003, p. 45), Lando (2004, p. 88), Lando and Skoteberg (2001), Hamilton (2001), Fons (1991), and Jonsson and Frison (1996). As these, and other, authors reported, it seems that the real economic activity over the business cycle is correlated with the fluctuation in transition and default probabilities. Duffie and Singleton (2003, p. 47) stated that the correlation between the four-quarter moving averages of default rates and national gross domestic product growth rates for the sample period 1983–1997 was -0.78 . Fons (1991), Jonsson and Frison (1996), and Hamilton (2001) included various additional time covariates, such as industrial productivity, as predictors.

It is possible that the practitioner will need to predict the real-world transition probabilities in the semi-Markov process model based on the available data. Two possible solutions are suggested in such a case. First, let $x_{1ij}(t), x_{2ij}(t), \dots, x_{nij}(t)$ be possible time covariates correlated with the transition probabilities $\hat{p}_{ij}(t)$ and $\hat{h}_{ij}(t, m)$, for $t = 1, 2, \dots, T^* - 1$. Then

logistic stepwise regression is an appropriate model for these probabilities. For example, for the $p_{ij}(t)$, we assume that

$$\hat{p}_{ij}(t) = \frac{\exp[a_0 + a_1x_{1ij}(t) + a_2x_{2ij}(t) + \cdots + a_nx_{nij}(t)]}{1 + \exp[a_0 + a_1x_{1ij}(t) + a_2x_{2ij}(t) + \cdots + a_nx_{nij}(t)]}.$$

If we define

$$\text{logit}(x) = \log\left(\frac{x}{1-x}\right),$$

then we may write

$$\text{logit}(\hat{p}_{ij}(t)) = a_0 + a_1x_{1ij}(t) + a_2x_{2ij}(t) + \cdots + a_nx_{nij}(t)$$

and stepwise regression techniques can be applied to estimate the a_i . In a similar way, a predictive stochastic model can be created for the probabilities $\hat{h}_{ij}(t, m)$. Second, existing time series models can be used for the purposes of predicting the probabilities $p_{ij}(t)$, $h_{ij}(t, m)$, $\pi_i(t)\eta_i(t, m)$, $\pi_i(t)$, and $\eta_i(t, m)$.

Acknowledgement

The authors would like to thank the referee for his constructive comments and remarks.

References

- ANDERSON, T. W. AND GOODMAN, L. A. (1957). Statistical inference about Markov chains. *Ann. Math. Statist.* **28**, 89–110.
- ARVANITIS, A., GREGORY, J. AND LAURENT, J.-P. (1999). Building models for credit spreads. *J. Derivatives* **6**, 27–43.
- BARTHOLOMEW, D. J. (1982). *Stochastic Models for Social Processes*, 3rd edn. John Wiley, Chichester.
- BARTHOLOMEW, D. J. (1986). Social applications of semi-Markov processes. In *Semi-Markov Models* (Brussels, 1984), ed. J. Janssen, Plenum, New York, pp. 463–474.
- BARTHOLOMEW, D. J., FORBES, A. F. AND MCCLEAN, S. (1991). *Statistical Techniques for Manpower Planning*. John Wiley, Chichester.
- BIELECKI, T. R. AND RUTKOWSKI, M. (2002). *Credit Risk: Modelling, Valuation and Hedging*. Springer, Berlin.
- CARTY, L. V. AND FONS, J. S. (1994). Measuring changes in corporate credit quality. *J. Fixed Income* **4**, 27–41.
- COX, D. R. AND OAKES, D. (1984). *Analysis of Survival Data*. Chapman and Hall, London.
- CHRISTENSEN, J. H. E., HANSEN, E. AND LANDO, D. (2004). Confidence sets for continuous-time rating transition probabilities. Working paper, Copenhagen Business School and University of Copenhagen.
- DAS, S. R. AND TUFANO, P. (1996). Pricing credit-sensitive debt when interest rates, credit ratings, and credit spreads are stochastic. *J. Financial Eng.* **5**, 161–198.
- DUFFIE, D. AND SINGLETON, K. (1999). Modeling term structures of defaultable bonds. *Rev. Financial Stud.* **12**, 687–720.
- DUFFIE, D. AND SINGLETON, K. (2003). *Credit Risk*. Princeton University Press.
- ELANDT-JOHNSON, R. C. AND JOHNSON, N. L. (1980). *Survival Models and Data Analysis*. John Wiley, New York.
- FONS, J. S. (1991). An approach to forecasting default rates. Working paper, Moody's Investors Service.
- GEMAN, H. (1989). The importance of the forward neutral probability in a stochastic approach of interest rates. Working paper, ESSEC.
- GEMAN, H., EL KAROUI, N. AND ROCHET, J.-C. (1995). Changes of numéraire, changes of probability measure and option pricing. *J. Appl. Prob.* **32**, 443–458.
- HAMILTON, D. T. (2001). Default and recovery rates of corporate bond issuers: 2000. Special comment, Moody's Investors Service.
- HORN, R. A. AND JOHNSON, C. R. (1991). *Topics in Matrix Analysis*. Cambridge University Press.
- HOWARD, R. A. (1971). *Dynamic Probabilistic Systems*, Vols I and II. John Wiley, Chichester.
- IOSIFESCU-MANU, A. (1972). Processe semimarkovien neomgene. *Stud. Lere. Mat.* **24**, 229–233.
- JARROW, R. AND TURNBULL, S. (1995). Pricing derivatives on financial securities subject to credit risk. *J. Finance* **50**, 53–86.

- JARROW, R., LANDO, D. AND TURNBULL, S. (1997). A Markov model for the term structure of credit risk spreads. *Rev. Financial Stud.* **10**, 481–523.
- JONSSON, J. G. AND FRISON, M. S. (1996). Forecasting default rates on high-yield bonds. *J. Fixed Income* **6**, 69–77.
- KALBFLEISCH, J. D. AND PRENTICE, R. L. (2002). *The Statistical Analysis of Failure Time Data*, 2nd edn. John Wiley, New York.
- KIJIMA, M. (1997). *Markov Processes for Stochastic Modeling*. Chapman and Hall, London.
- KIJIMA, M. (1998). Monotonicity in a Markov chain model for valuing coupon bond subject to credit risk. *Math. Finance* **8**, 229–247.
- KIJIMA, M. AND KOMORIBAYASHI, K. (1998). A Markov chain model for valuing credit risk derivatives. *J. Derivatives* **6**, 97–108.
- LANDO, D. (1998). On Cox processes and credit-risky securities. *Rev. Derivatives Res.* **2**, 99–120.
- LANDO, D. (2000). Some elements of rating-based credit risk modeling. In *Advanced Fixed-Income Valuation Tools*, eds N. Jegadeesh and B. Tuckman, John Wiley, Chichester, pp. 193–215.
- LANDO, D. (2004). *Credit Risk Modeling*. Princeton University Press.
- LANDO, D. AND SKOTEBERG, T. (2001). Analysing rating transitions and rating drift with continuous observations. *J. Banking Finance* **26**, 423–444.
- LEE, E. T. (1992). *Statistical Methods for Survival Data Analysis*, 2nd edn. John Wiley, New York.
- MCCLEAN, S. I. (1976). Semi-Markov models for human resource modelling. *J. Math. Appl. Business Industry* **4**, 307–315.
- MCCLEAN, S. I. (1986). Semi-Markov models for manpower systems. In *Semi-Markov Models: Theory and Applications*, ed. J. Janssen, Plenum, New York, pp. 283–300.
- MCCLEAN, S. I. AND GRIBBIN, J. O. (1987). Estimation for incomplete manpower data. *Appl. Stoch. Models Data Anal.* **3**, 13–25.
- MCCLEAN, S. I. AND GRIBBIN, J. O. (1991). A non-parametric competing risks model for manpower planning. *Appl. Stoch. Models Data Anal.* **7**, 327–341.
- MCCLEAN, S. I. AND MONTGOMERY, E. (1999). Estimation for semi-Markov manpower models in a stochastic environment. In *Semi-Markov Models and Applications*, eds J. Janssen and N. Limnios, Kluwer, Dordrecht, pp. 219–227.
- MCCLEAN, S. I., MONTGOMERY, E. AND UGWUOWO, F. (1997). Inhomogeneous continuous-time Markov and semi-Markov manpower models. *Appl. Stoch. Models Data Anal.* **13**, 191–198.
- MUSIELA, M. AND RUTKOWSKI, M. (1997a). Continuous-time term structure models: forward measure approach. *Finance Stoch.* **1**, 261–291.
- MUSIELA, M. AND RUTKOWSKI, M. (1997b). *Martingale Methods in Financial Modelling*. Springer, Berlin.
- PAPADOPOULOU, A. A. (1997). Counting transitions—entrance probabilities in inhomogeneous semi-Markov systems. *Appl. Stoch. Models Data Anal.* **13**, 199–206.
- PAPADOPOULOU, A. A. AND VASSILIOU, P.-C. G. (1994). Asymptotic behavior of nonhomogeneous semi-Markov systems. *Linear Algebra Appl.* **210**, 153–198.
- SHREVE, S. E. (2004a). *Stochastic Calculus for Finance*, Vol. I, *The Binomial Asset Pricing Model*. Springer, New York.
- SHREVE, S. E. (2004b). *Stochastic Calculus for Finance*, Vol. II, *Continuous-Time Models*. Springer, New York.
- VASSILIOU, P.-C. G. (1976). A Markov chain model for wastage in manpower systems. *Operat. Res. Quart.* **27**, 57–70.
- VASSILIOU, P.-C. G. (1997). The evolution of the theory of non-homogeneous Markov systems. *Appl. Stoch. Models Data Anal.* **13**, 159–177.
- VASSILIOU, P.-C. G. AND PAPADOPOULOU, A. A. (1992). Nonhomogeneous semi-Markov systems and maintainability of the state sizes. *J. Appl. Prob.* **29**, 519–534.
- VASSILIOU, P.-C. G. AND SYMEONAKI, M. A. (1999). The perturbed non-homogeneous semi-Markov system. *Linear Algebra Appl.* **289**, 319–332.