

A Simple Model for Credit Migration and Spread Curves¹

Li Chen² Damir Filipović³

May 14, 2003 (submission); May 26, 2004(this version)

¹Key words: credit risk model, affine process, equivalent change of measure. JEL Classification: G12, G13. AMS (2000) Classification: 60J25, 60J75.

²Department of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA. Email Contact: lichen@princeton.edu.

³Federal Office of Private Insurance (BPV), Switzerland. Email Contact: filipo@math.ethz.ch.

Abstract

We propose and examine a simple model for credit migration and spread curves of a single firm both under real-world and risk-neutral measures. This model is a hybrid of a structural and a reduced-form model. Default is triggered either by successive downgradings of the firm or an unpredictable jump of the state process. The default time is accordingly decomposed into predictable and totally inaccessible part.

1 Introduction

We propose and examine a simple model for credit migration and spread curves of a single firm both under real-world and risk-neutral measures. This model is based on an affine state process $Y = (Y^1, Y^2, Y^3)$ taking values in $\overline{\mathbb{R}}_+^3 = \mathbb{R}_+^3 \cup \Delta$ which is the one-point compactification of $\mathbb{R}_+^3 = \{x \in \mathbb{R}^3 \mid x_i \geq 0\}$ (Δ is the “point at infinity”).

The pair (Y_t^2, Y_t^3) represents the state of the firm, where Y^3 is a simple point process with $Y_0^3 = 0$ and jump size 1 whose intensity depends linearly on Y^2 . While Y^3 takes account of the unpredictable credit event (default), Y^2 is a credit index of the firm ranging in the non-negative real numbers with $Y_t^2 = 0$ corresponding to the best credit rating (e.g. Aaa) and $Y_t^2 = +\infty$ (that is, $Y_t = \Delta$) meaning default.

The firm has defaulted by time t if

$$Y_t \in \mathcal{D} := \{\Delta\} \cup \{y \in \mathbb{R}_+^3 \mid y_3 > 0\},$$

which is an absorbing state. Hence a default of the firm is triggered either by i) successive downgradings (explosion of Y^2) or ii) an unpredictable jump of Y^3 . The combination of i) and ii) yields a decomposition of the firm’s default time

$$T_{\mathcal{D}} := \inf\{t \mid Y_t \in \mathcal{D}\} = T_{\Delta} \wedge T_J$$

into a predictable and a totally inaccessible part, where

$$T_{\Delta} := \inf\{t \mid Y_t = \Delta\}$$

is the explosion time of Y and

$$T_J := \inf\{t \mid Y_t^3 = 1\}$$

is the first jump time of Y^3 .

The actual credit rating (e.g. Moody’s) can in principle be obtained by a monotone transformation of Y_t^2 . That is, \mathbb{R}_+ is decomposed into finitely many non-overlapping intervals I_{Aaa}, I_{Aa}, \dots with $Y_t^2 \in I_R$ meaning that the firm is R -rated, $R \in \{Aaa, Aa, \dots\}$, given that $Y_t^3 = 0$.

The component Y^1 describes the short rates r up to T_{Δ} . The process r itself follows a Cox–Ingersoll–Ross [6] (CIR) model. Our setup allows for dependence of interest rates and credit migration.

Due to the affine nature of Y and r we find explicit expressions for the real-world default probabilities and corporate bond prices. The resulting zero-recovery yield spread curve is affine in Y . The change from real-world to risk-neutral measure is specified by the market risk premium (r -dynamics) which implicitly affects the characteristics of the credit risk (Y -dynamics).

Our approach constitutes a hybrid of a structural and a reduced-form default time model. Here, with “structural” we associate any default time model which is based on the (predictable) first passage time of an underlying economic factor process, see e.g. [1, 7, 16]. The “reduced-form” on the other hand stands for any

intensity based model of a (totally inaccessible) default time, see e.g. [17, 18]. We refer the reader to [2, 12] for a recent overview of credit risk models and a comprehensive reference list. The novelty of our model lies in its explicit and tractable structure. An extension towards multi-firm models with default correlation and counterparty risk is given in [4].

The remainder of the paper is as follows. In Section 2 we introduce the basic affine state process Y and discuss some of its properties, citing results from [10]. In Section 3 an explicit expression for the real-world default probabilities is derived. In Section 4 we obtain expressions for Treasury and corporate bond prices, with zero-recovery and fractional recovery at maturity. The zero-recovery yield spread curve is given as an explicit affine function of Y (Lemma 4.2). Section 5 provides an equivalent change of measure which links the real-world and the risk-neutral model from Sections 3 and 4, respectively. In Section 6 we explore the empirical performance of the proposed model using real data. The appendix contains the proof of Lemma 4.1 which allows to embed Y^1 in a global CIR short rate model.

2 The Basic State Process

We will frequently make use of the notation and the general results for affine processes, which can be found in [10]. Let $\alpha_1, \alpha_2, b_1, b_2, \beta_{21}, c, \gamma_1, \gamma_2, \ell, \lambda_1, \lambda_2 \in \mathbb{R}_+, \beta_1, \beta_{22} \in \mathbb{R}, \{e_1, e_2, e_3\}$ be the standard basis in \mathbb{R}^3 and

$$\mu_\theta(d\xi) := \frac{\theta}{\Gamma(1-\theta)} \frac{1}{\xi^{1+\theta}} d\xi \quad (1)$$

for some $\theta \in (0, 1)$. Our basic state process is the unique $\overline{\mathbb{R}_+^3}$ -valued regular affine Markov process Y with generator

$$\begin{aligned} \mathcal{A}f(y) &= \alpha_1 y_1 \partial_{y_1}^2 f(y) + (b_1 + \beta_1 y_1) \partial_{y_1} f(y) \\ &\quad + \alpha_2 y_2 \partial_{y_2}^2 f(y) + (b_2 + \beta_{21} y_1 + \beta_{22} y_2) \partial_{y_2} f(y) \\ &\quad + \int_{\mathbb{R}_{++}} (f(y + \xi e_2) - f(y)) (\ell + \lambda_1 y_1 + \lambda_2 y_2) \mu_\theta(d\xi) \\ &\quad + (f(y + e_3) - f(y)) (c + \gamma_1 y_1 + \gamma_2 y_2). \end{aligned} \quad (2)$$

In what follows, we let Y be realized as a càdlàg process on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$, which is rich enough to carry a Brownian motion W (this holds, for instance, for the product of the Wiener space with the canonical space of càdlàg paths in $\overline{\mathbb{R}_+^3}$). The measure \mathbb{P} stands for either the real-world or the risk-neutral measure.

It is shown in [5] that for every stopping time $\tau < T_\Delta$ the stopped process Y^τ is a semimartingale with characteristics determined by the property that

$$M_t^f := f(Y_{t \wedge \tau}) - f(Y_0) - \int_0^{t \wedge \tau} \mathcal{A}f(Y_s) ds \quad (3)$$

is a local martingale for all $f \in C_b^2(\mathbb{R}_+^3)$ (bounded C^2 -functions). We refer the reader to [15] for the notion of the characteristics of a semimartingale, in particular Theorem II.2.42. Hence $Y_{t \wedge \tau}^1$ is a continuous semimartingale with drift $(b_1 + \beta_1 Y_s^1)1_{\{s \leq \tau\}}$ and diffusion $\alpha_1 Y_s^1 1_{\{s \leq \tau\}}$.

If not otherwise stated, we shall henceforth assume that

$$b_1 > 0, \quad \lambda_1 > 0, \quad \lambda_2 > 0 \quad \text{and} \quad Y_0^3 = 0 \quad (\text{hence } Y_0 \neq \Delta). \quad (4)$$

Remark 2.1. Every measurable function f on \mathbb{R}_+^3 is extended to $\overline{\mathbb{R}_+^3}$ by the convention $f(\Delta) = 0$. This is standard in the theory of Feller semigroups (see e.g. [13]). In particular, we write

$$e^{\langle 0, y \rangle} = 1_{\{y \neq \Delta\}}. \quad (5)$$

With this convention, the basic affine property of Y reads

$$\mathbb{E} \left[e^{\langle v, Y_T \rangle} \mid \mathcal{F}_t \right] = e^{\phi(T-t, v) + \langle \psi(T-t, v), Y_t \rangle} \quad (6)$$

for all $v \in \mathbb{C}_-^3 = \{v \in \mathbb{C}^3 \mid \operatorname{Re} v \in \mathbb{R}_-^3\}$ and $0 \leq t \leq T$, where the \mathbb{C}_- -valued jointly continuous functions $\phi = \phi(t, v)$ and $\psi_i = \psi_i(t, v)$ solve the generalized Riccati equations (GREs)

$$\begin{aligned} \partial_t \phi &= b_1 \psi_1 + b_2 \psi_2 - \ell(-\psi_2)^\theta + c(e^{\psi_3} - 1) \\ \phi(0, v) &= 0 \\ \partial_t \psi_1 &= \alpha_1 \psi_1^2 + \beta_1 \psi_1 + \beta_{21} \psi_2 - \lambda_1(-\psi_2)^\theta + \gamma_1(e^{\psi_3} - 1) \\ \psi_1(0, v) &= v_1 \\ \partial_t \psi_2 &= \alpha_2 \psi_2^2 + \beta_{22} \psi_2 - \lambda_2(-\psi_2)^\theta + \gamma_2(e^{\psi_3} - 1) \\ \psi_2(0, v) &= v_2 \\ \partial_t \psi_3 &= 0 \\ \psi_3(0, v) &= v_3. \end{aligned} \quad (7)$$

In particular, we have $\psi_3(t, v) = v_3$ and $\phi(t, v)$ is an ordinary integral. This explicit form of the GREs follows since

$$\int_{\mathbb{R}_{++}} (e^{v\xi} - 1) \mu_\theta(d\xi) = -(-v)^\theta, \quad v \in \mathbb{C}_-$$

(see also Example 9.3 in [10]).

An explicit expression for ψ_2 is available for particular parameter choices.

Lemma 2.2. *If $\alpha_2 = \gamma_2 = 0$ then*

$$\psi_2(t, v) = - \left(e^{(1-\theta)\beta_{22}t} (-v_2)^{1-\theta} + \frac{\lambda_2}{\beta_{22}} \left(e^{(1-\theta)\beta_{22}t} - 1 \right) \right)^{\frac{1}{1-\theta}}. \quad (8)$$

If $\lambda_2 = 0$ then

$$\psi_2(t, v) = - \frac{2\gamma_2(1 - e^{v_3})(e^{\rho t} - 1) - (\rho(e^{\rho t} + 1) + \beta_{22}(e^{\rho t} - 1))v_2}{\rho(e^{\rho t} + 1) - \beta_{22}(e^{\rho t} - 1) - 2\alpha_2(e^{\rho t} - 1)v_2} \quad (9)$$

where $\rho := \sqrt{\beta_{22}^2 + 4\alpha_2\gamma_2(1 - e^{v_3})}$.

Proof. If $\alpha_2 = \gamma_2 = 0$ then ψ_2 solves a Bernoulli equation (see [3, Exercise 14.2]). If $\lambda_2 = 0$ then ψ_2 solves a classical Riccati equation. \square

The solution of the GREs (7) is unique for $\operatorname{Re} v_2 < 0$. However, the right hand side of (7) is not Lipschitz continuous at $\operatorname{Re} \psi_2 = 0$ because of the term $\lambda(-\psi_2)^\theta$. Indeed, $\phi(t, 0)$ and $\psi_i(t, 0)$ solve (7) for $v = 0$, but so does the zero function. We uniquely obtain $\phi(t, 0)$ and $\psi_i(t, 0)$ by continuity

$$\begin{aligned}\phi(t, 0) &= \lim_{s \downarrow 0} \phi(t, -se_2) \\ \psi_i(t, 0) &= \lim_{s \downarrow 0} \psi_i(t, -se_2), \quad i = 1, 2, 3.\end{aligned}\tag{10}$$

In view of (4) thus $\phi(t, 0), \psi_1(t, 0), \psi_2(t, 0) < 0$ for all $t > 0$, see (8) for a special case. Consequently,

$$\mathbb{E}[e^{\langle 0, Y_t \rangle}] = e^{\phi(t, 0) + \langle \psi(t, 0), Y_0 \rangle} < 1 \quad \forall t > 0.$$

Hence Y is non-conservative and $T_\Delta < \infty$ almost surely. By the Feller property of Y , we have that $Y_t = \Delta$ for all $t \geq T_\Delta$ a.s. (see [21, Proposition III.2.9]), so that Δ , and hence \mathcal{D} , is an absorbing state as required.

Since there is no potential term in (2), the transition of Y to Δ occurs by explosion (see [5]). An explosion of Y is due to the jump characteristics $(\ell + \langle \lambda, Y_t \rangle) \mu_\theta(d\xi)$ of Y^2 , which induces jumps of large size and with an intensity depending linearly on Y^2 . This feedback effect leads Y to exploding in finite time (we will analyze the behavior of Y^1 at T_Δ in more detail in Lemma A.1 below). The explosion time T_Δ accordingly is predictable with announcing sequence $T_n < T_\Delta$ and $\lim_n T_n = T_\Delta$ a.s. given by

$$T_n := \inf\{t \mid \|Y_{t-}\| \geq n \text{ or } \|Y_t\| \geq n\}, \quad n \in \mathbb{N}.\tag{11}$$

T_Δ is the appropriate model for the default time of a low-rated firm. Indeed, the larger Y_t^2 the more likely are consecutive downgradings of the firm with eventual default (explosion of Y).

In contrast, T_J is totally inaccessible and hence the appropriate model for an unpredictable, sudden default of a highly rated firm.

3 Credit Migration

In this section we let \mathbb{P} denote the real-world measure. In view of Remark 2.1 we have

$$\mathbf{1}_{\{T < T_D\}} = \mathbf{1}_{\{Y_T \neq \Delta\}} \mathbf{1}_{\{Y_T^3 = 0\}} = e^{\langle 0, Y_T \rangle} \lim_{k \rightarrow \infty} e^{-kY_T^3} = \lim_{k \rightarrow \infty} e^{-kY_T^3}.\tag{12}$$

Hence

$$\begin{aligned}\mathbb{P}[T < T_D \mid \mathcal{F}_t] &= \mathbb{E} \left[\lim_{k \rightarrow \infty} e^{-kY_T^3} \mid \mathcal{F}_t \right] = \lim_{k \rightarrow \infty} e^{\phi(T-t, -ke_3) + \langle \psi(T-t, -ke_3), Y_t \rangle} \\ &= e^{\tilde{\phi}(T-t, 0) + \tilde{\psi}_1(T-t, 0)Y_t^1 + \tilde{\psi}_2(T-t, 0)Y_t^2} \mathbf{1}_{\{t < T_D\}}\end{aligned}\tag{13}$$

where $\tilde{\phi} = \tilde{\phi}(t, v)$ and $\tilde{\psi}_i = \tilde{\psi}_i(t, v)$ solve the GREs

$$\begin{aligned}
\partial_t \tilde{\phi} &= b_1 \tilde{\psi}_1 + b_2 \tilde{\psi}_2 - \ell(-\tilde{\psi}_2)^\theta - c \\
\tilde{\phi}(0, v) &= 0 \\
\partial_t \tilde{\psi}_1 &= \alpha_1 \tilde{\psi}_1^2 + \beta_1 \tilde{\psi}_1 + \beta_{21} \tilde{\psi}_2 - \lambda_1(-\tilde{\psi}_2)^\theta - \gamma_1 \\
\tilde{\psi}_1(0, v) &= v_1 \\
\partial_t \tilde{\psi}_2 &= \alpha_2 \tilde{\psi}_2^2 + \beta_{22} \tilde{\psi}_2 - \lambda_2(-\tilde{\psi}_2)^\theta - \gamma_2 \\
\tilde{\psi}_2(0, v) &= v_2.
\end{aligned} \tag{14}$$

This follows since the right-hand side of the GREs (7) converges uniformly on compacts to the right-hand side of (14) as $k \rightarrow \infty$. Notice that $\tilde{\phi}(t, 0)$ and $\tilde{\psi}_i(t, 0)$ are given according to (10). Moreover, if $c = \gamma_i = 0$ then $\tilde{\phi} = \phi$ and $\tilde{\psi}_i = \psi_i$, $i = 1, 2$.

Equation (13) yields an explicit expression for the \mathcal{F}_t -conditional default probability by T of the firm as a function of its current credit state (Y_t^2, Y_t^3) and the short rate Y_t^1

$$\mathbb{P}[T_{\mathcal{D}} \leq T \mid \mathcal{F}_t] = 1 - e^{\tilde{\phi}(T-t, 0) + \tilde{\psi}_1(T-t, 0)Y_t^1 + \tilde{\psi}_2(T-t, 0)Y_t^2} \mathbf{1}_{\{t < T_{\mathcal{D}}\}}. \tag{15}$$

The \mathcal{F}_t -conditional transition probability from current credit state (Y_t^2, Y_t^3) into the interval $I_R \times \{0\}$ at time $T > t$ can be derived by numerical Fourier inversion of (6). Of course, there is an infinite degree of freedom to calibrate the model to a given transition matrix (e.g. Moody's) since one has to specify the correspondence between rating classes Aaa, Aa, \dots and intervals I_{Aaa}, I_{Aa}, \dots of \mathbb{R}_+ . But the default state, $Y_t \in \mathcal{D}$, is unique and the explicit expression (15) allows to calibrate the model parameters α_1, \dots, θ to the actual (e.g. Moody's) default probabilities.

4 Credit Spread Curves

In this section we calculate the corporate bond prices of a firm with given credit rating. In what follows we interpret \mathbb{P} as risk-neutral measure.

The process Y^1 describes the short rates only up to T_Δ since $Y_t = \Delta$ for $t \geq T_\Delta$ a.s. Before we can value a Treasury bond we first have to embed Y^1 in a global CIR model with generator

$$\mathcal{A}_1 g(r) := \alpha_1 r g''(r) + (b_1 + \beta_1 r) g'(r). \tag{16}$$

For the notion of a martingale problem we refer to [13].

Lemma 4.1. *There exists a continuous adapted process r which is a solution of the martingale problem for \mathcal{A}_1 (and hence is a CIR short rate process) and satisfies*

$$r_t = Y_t^1 \quad \forall t < T_\Delta.$$

The proof of this lemma can be found in the appendix.

4.1 Treasury Bond Pricing

Since r is the CIR short rate process with generator \mathcal{A}_1 we obtain for the time t -price of a zero-coupon Treasury bond with maturity $T \geq t$

$$P^{\text{tr}}(t, T) = \mathbb{E} \left[e^{-\int_t^T r_s ds} \mid \mathcal{F}_t \right] = e^{\phi^{\text{tr}}(T-t) + \psi^{\text{tr}}(T-t)r_t} \quad (17)$$

where

$$\begin{aligned} \phi^{\text{tr}}(t) &= \frac{b_1}{\alpha_1} \log \left(\frac{2\rho e^{\frac{1}{2}(\rho-\beta_1)t}}{(\rho-\beta_1)(e^{\rho t} - 1) + 2\rho} \right), \\ \psi^{\text{tr}}(t) &= -\frac{2(e^{\rho t} - 1)}{(\rho-\beta_1)(e^{\rho t} - 1) + 2\rho}, \end{aligned}$$

with $\rho = \sqrt{\beta_1^2 + 4\alpha_1}$, see (9), and the corresponding yield is given by

$$y^{\text{tr}}(t, T) = -\frac{1}{T-t} \log P^{\text{tr}}(t, T) = -\frac{1}{T-t} (\phi^{\text{tr}}(T-t) + \psi^{\text{tr}}(T-t)r_t).$$

4.2 Defaultable Bond Pricing

We consider a zero-coupon corporate bond with zero recovery and with partial recovery at maturity.

4.2.1 Zero Recovery

The payoff at maturity T is

$$1_{\{T < T_D\}} = \lim_{k \rightarrow \infty} e^{-kY_T^3},$$

see (12). Define the measurable function $\Pi : \overline{\mathbb{R}_+^3} \rightarrow \mathbb{R}_+$ by $\Pi(y) = y_1 1_{\{y \neq \Delta\}}$, which is consistent with Remark 2.1. It follows literally as in [10, Section 11.1] that the following Feynman–Kac formula holds:

$$\mathbb{E} \left[e^{-\int_t^T \Pi(Y_u) du} e^{\langle v, Y_T \rangle} \mid \mathcal{F}_t \right] = e^{\phi^{\text{co}}(T-t, v) + \langle \psi^{\text{co}}(T-t, v), Y_t \rangle} \quad (18)$$

for all $v \in \mathbb{C}_-^3$ and $0 \leq t \leq T$, where the \mathbb{C}_- -valued functions $\phi^{\text{co}} = \phi^{\text{co}}(t, v)$ and $\psi_i^{\text{co}} = \psi_i^{\text{co}}(t, v)$ solve the generalized Riccati equations

$$\begin{aligned} \partial_t^{\text{co}} \phi &= b_1 \psi_1^{\text{co}} + b_2 \psi_2^{\text{co}} - \ell(-\psi_2^{\text{co}})^\theta + c \left(e^{\psi_3^{\text{co}}} - 1 \right) \\ \phi(0, v) &= 0 \\ \partial_t \psi_1^{\text{co}} &= \alpha_1 (\psi_1^{\text{co}})^2 + \beta_1 \psi_1^{\text{co}} + \beta_{21} \psi_2^{\text{co}} - \lambda_1 (-\psi_2^{\text{co}})^\theta + \gamma_1 \left(e^{\psi_3^{\text{co}}} - 1 \right) - 1 \\ \psi_1^{\text{co}}(0, v) &= v_1 \\ \partial_t \psi_2^{\text{co}} &= \alpha_2 (\psi_2^{\text{co}})^2 + \beta_{22} \psi_2^{\text{co}} - \lambda_2 (-\psi_2^{\text{co}})^\theta + \gamma_2 \left(e^{\psi_3^{\text{co}}} - 1 \right) \\ \psi_2^{\text{co}}(0, v) &= v_2 \\ \partial_t \psi_3^{\text{co}} &= 0 \\ \psi_3^{\text{co}}(0, v) &= v_3 \end{aligned} \quad (19)$$

(only the equation for ψ_1^{co} differs from the original GREs (7)). The price of a zero-coupon corporate bond with zero recovery and maturity T is therefore

$$\begin{aligned}
P^{\text{co}}(t, T) &= \mathbb{E} \left[e^{-\int_t^T r_s ds} 1_{\{T < T_D\}} \mid \mathcal{F}_t \right] \\
&= \lim_{k \rightarrow \infty} \mathbb{E} \left[e^{-\int_t^T \Pi(Y_u) du} e^{-kY_T^3} \mid \mathcal{F}_t \right] \\
&= \lim_{k \rightarrow \infty} e^{\phi^{\text{co}}(T-t, -ke_3) + \langle \psi^{\text{co}}(T-t, -ke_3), Y_t \rangle} \\
&= e^{\tilde{\phi}^{\text{co}}(T-t, 0) + \tilde{\psi}_1^{\text{co}}(T-t, 0)Y_t^1 + \tilde{\psi}_2^{\text{co}}(T-t, 0)Y_t^2} 1_{\{t < T_D\}}
\end{aligned} \tag{20}$$

where $\tilde{\phi}^{\text{co}} = \tilde{\phi}^{\text{co}}(t, v)$ and $\tilde{\psi}_i^{\text{co}} = \tilde{\psi}_i^{\text{co}}(t, v)$ solve the GREs

$$\begin{aligned}
\partial_t \tilde{\phi}^{\text{co}} &= b_1 \tilde{\psi}_1^{\text{co}} + b_2 \tilde{\psi}_2^{\text{co}} - \ell(-\tilde{\psi}_2^{\text{co}})^\theta - c \\
\tilde{\phi}^{\text{co}}(0, v) &= 0 \\
\partial_t \tilde{\psi}_1^{\text{co}} &= \alpha_1 (\tilde{\psi}_1^{\text{co}})^2 + \beta_1 \tilde{\psi}_1^{\text{co}} + \beta_{21} \tilde{\psi}_2^{\text{co}} - \lambda_1 (-\tilde{\psi}_2^{\text{co}})^\theta - \gamma_1 - 1 \\
\tilde{\psi}_1^{\text{co}}(0, v) &= v_1 \\
\partial_t \tilde{\psi}_2^{\text{co}} &= \alpha_2 (\tilde{\psi}_2^{\text{co}})^2 + \beta_{22} \tilde{\psi}_2^{\text{co}} - \lambda_2 (-\tilde{\psi}_2^{\text{co}})^\theta - \gamma_2 \\
\tilde{\psi}_2^{\text{co}}(0, v) &= v_2,
\end{aligned} \tag{21}$$

which follows as (14). Again, $\tilde{\phi}^{\text{co}}(t, 0)$ and $\tilde{\psi}^{\text{co}}(t, 0)$ are given according to (10).

To summarize, we obtain an explicit affine expression for the zero-recovery yield spread curves.

Lemma 4.2. *The zero-recovery yield spread curve is*

$$\begin{aligned}
\Delta y(t, T) &= -\frac{1}{T-t} (\log P^{\text{co}}(T) - \log P^{\text{tr}}(t, T)) \\
&= -\frac{1}{T-t} \left(\tilde{\phi}^{\text{co}}(T-t, 0) - \phi^{\text{tr}}(T-t) + \left(\tilde{\psi}_1^{\text{co}}(T-t, 0) - \psi^{\text{tr}}(T-t) \right) Y_t^1 \right) \\
&\quad - \frac{1}{T-t} \tilde{\psi}_2^{\text{co}}(T-t, 0) Y_t^2.
\end{aligned} \tag{22}$$

In particular, in the limit $T \downarrow t$ we obtain

$$\Delta y(t, t) = c + \gamma_1 Y_t^1 + \gamma_2 Y_t^2. \tag{23}$$

Expression (23) shows that the zero time to maturity yield spread is strictly positive if $c + \gamma_1 + \gamma_2 > 0$ in general. This is a desirable feature as pointed out in [11].

4.2.2 Partial Recovery at Maturity

From the preceding results we can easily derive the time t -price $P_\delta^{\text{co}}(t, T)$ of a zero-coupon corporate bond which pays a (constant) fraction $\delta \in (0, 1)$ of face-value 1 at maturity $T \geq t$ in case of default. Indeed, the payoff at T is

$$1_{\{T < T_D\}} + \delta 1_{\{T \geq T_D\}} = (1 - \delta) 1_{\{T < T_D\}} + \delta.$$

Hence

$$P_\delta^{\text{co}}(t, T) = (1 - \delta)P^{\text{co}}(t, T) + \delta P^{\text{tr}}(t, T), \quad (24)$$

where $P^{\text{tr}}(t, T)$ and $P^{\text{co}}(t, T)$ are defined in (17) and (20), respectively.

5 Measure Change

In this section we provide an equivalent change of measure which preserves the form (2) of the generator of Y , and which therefore links the above real-world model (Section 3) with the risk-neutral model (Section 4). We consider the affine processes Y and r as at the beginning of Section 4, where \mathbb{P} now denotes the real-world measure, say.

We change the drift of the short rate process r ($= Y^1$ on $[0, T_\Delta)$), which will indirectly change the characteristics of the credit index process Y^2 . Changing the parameter θ of the Y^2 -jump characteristic, see (1), by an equivalent change of measure seems to be difficult if not impossible (the candidate integrand for the logarithm of the density process, $\psi(x, \xi) = \xi^{\tilde{\theta} - \theta}$, does not satisfy the sufficient integrability conditions of the main theorem in [5]). On the other hand, it has been shown in [5] that the mean reversion rate, $-\beta_1$, of r can be transformed into any $-\tilde{\beta}_1 \in \mathbb{R}$ under an equivalent change of measure. The same holds true for $b_1 \rightsquigarrow \tilde{b}_1$ only for $\tilde{b}_1 \geq \alpha_1$ and only if $b_1 \geq \alpha_1$. The latter property is equivalent to $r > 0$ a.s. if $r_0 > 0$.

We let therefore $\tilde{b}_1 \geq \alpha_1$ if $b_1 \geq \alpha_1$ and set $\tilde{b}_1 = b_1$ if $b_1 < \alpha_1$. Let $\tilde{\beta}_1 \in \mathbb{R}$ and define the continuous function $\Lambda : (0, \infty) \rightarrow \mathbb{R}$ by

$$\Lambda(r) := \frac{\tilde{b}_1 - b_1}{2\alpha_1 r} + \frac{\tilde{\beta}_1 - \beta_1}{2\alpha_1}.$$

Then $\Lambda(r) \equiv (\tilde{\beta}_1 - \beta_1)/(2\alpha_1)$ is simply constant if $b_1 < \alpha_1$. Let

$$r_t^c = r_t - r_0 - \int_0^t (b_1 + \beta_1 r_s) ds$$

denote the (continuous) martingale part of r and write $X \bullet Z$ for the stochastic integral $\int X dZ$.

It is shown in [5] that

$$D = \mathcal{E}(\Lambda(r) \bullet r^c)$$

is a strictly positive martingale with $\mathbb{E}[D_t] = 1$. Hence for every $t \geq 0$ we can define an equivalent probability measure $\mathbb{Q}^t \sim \mathbb{P}$ on \mathcal{F}_t by

$$\frac{d\mathbb{Q}^t}{d\mathbb{P}} = D_t.$$

To simplify the exposure we now assume that there exists a probability measure \mathbb{Q} on \mathcal{F} such that $\mathbb{Q}|_{\mathcal{F}_t} = \mathbb{Q}^t$ for all $t \geq 0$ (the existence of \mathbb{Q} follows by the Daniell–Kolmogorov extension theorem if Ω is the space of càdlàg paths

in $\mathbb{R}_+ \times \overline{\mathbb{R}_+^3}$ and (\mathcal{F}_t) the canonical right-continuous filtration, see [22, Theorem IV.38.9]). It follows from [5] that r is a Markov process under \mathbb{Q} with generator

$$\tilde{\mathcal{A}}_1 f(r) = \alpha_1 r f''(r) + (\tilde{b}_1 + \tilde{\beta}_1 r) f'(r).$$

Lemma 5.1. *Y is a regular affine process under \mathbb{Q} with generator*

$$\begin{aligned} \tilde{\mathcal{A}}f(y) &= \alpha_1 y_1 \partial_{y_1}^2 f(y) + (\tilde{b}_1 + \tilde{\beta}_1 y_1) \partial_{y_1} f(y) \\ &\quad + \alpha_2 y_2 \partial_{y_2}^2 f(y) + (b_2 + \beta_{21} y_1 + \beta_{22} y_2) \partial_{y_2} f(y) \\ &\quad + \int_{\mathbb{R}_{++}} (f(y + \xi e_2) - f(y)) (\ell + \lambda_1 y_1 + \lambda_2 y_2) \mu_\theta(d\xi) \\ &\quad + (f(y + e_3) - f(y)) (c + \gamma_1 y_1 + \gamma_2 y_2). \end{aligned} \quad (25)$$

Proof. In view of [10], $\tilde{\mathcal{A}}$ is the generator of a unique (in distribution) regular affine process. Hence uniqueness holds for the local martingale problem for $(\tilde{\mathcal{A}}, Y_0, T_n)$ for all $n \geq 1$, see [13]. We can assume that Y follows an $\tilde{\mathcal{A}}$ -regular affine process under some probability measure $\tilde{\mathbb{P}}$ on $(\Omega, \mathcal{F}, (\mathcal{F}_t))$. We then have to show that $\tilde{\mathbb{P}} = \mathbb{Q}$.

We first show that $\tilde{M}_t^{f, T_n} = \tilde{M}_{t \wedge T_n}^f$ is a \mathbb{Q} -martingale for all $n \geq 1$ and $f \in C_c^2(\mathbb{R}_+^3)$, where

$$\tilde{M}_t^f := f(Y_t) - f(Y_0) - \int_0^t \tilde{\mathcal{A}}f(Y_s) ds.$$

This holds if and only if $D\tilde{M}^{f, T_n}$ is a \mathbb{P} -martingale. Integration by parts yields (see (3) for M^{f, T_n})

$$\begin{aligned} D\tilde{M}^{f, T_n} &= D_- \bullet M^{f, T_n} + D_- \bullet (\tilde{M}^{f, T_n} - M^{f, T_n}) + \tilde{M}_-^{f, T_n} \bullet D + [D, \tilde{M}^{f, T_n}] \\ &\sim D_- \bullet (\tilde{M}^{f, T_n} - M^{f, T_n}) + [D, \tilde{M}^{f, T_n}], \end{aligned}$$

where we write $A \sim B$ if $A - B$ is a local martingale. Notice that

$$\tilde{M}_t^{f, T_n} - M_t^{f, T_n} = \int_0^{t \wedge T_n} (A - \tilde{\mathcal{A}})f(Y_s) ds$$

is continuous. Moreover, if we write A^c for the continuous martingale part of a semimartingale A ,

$$D^{T_n} = D_-^{T_n} \bullet (\Lambda(Y_-^1) \bullet Y^{1, T_n, c})$$

and

$$M^{f, T_n, c} = \nabla f(Y_-) \bullet Y^{T_n, c}.$$

Hence

$$[D, \tilde{M}^{f, T_n}] = D_-^{T_n} \bullet ((\Lambda(Y_-^1) \partial_{y_1} f(Y_-)) \bullet [Y^{1, T_n, c}, Y^{1, T_n, c}])$$

because $[Y^{1, T_n, c}, Y^{2, T_n, c}] = 0$ (there is no $\partial_{y_1} \partial_{y_2}$ -term in (2)). Notice that

$$[Y^{1, T_n, c}, Y^{1, T_n, c}]_t = 2\alpha_1 \int_0^{t \wedge T_n} Y_s^{1, T_n} ds$$

and, by (25),

$$\tilde{\mathcal{A}}f(y) = \mathcal{A}f(y) + 2\alpha_1 y_1 \Lambda(y_1) \partial_{y_1} f(y)$$

for $y \in \mathbb{R}_+^3$ with $y_1 > 0$ if $b_1 \geq \alpha_1$ and for all $y \in \mathbb{R}_+^3$ otherwise. Since in the former case ($b_1 \geq \alpha_1$) we have that $Y^{1,T_n} = r^{T_n} > 0$ a.s., we conclude that in both cases $D\tilde{M}^{f,T_n} \sim 0$, whence \tilde{M}^{f,T_n} is a \mathbb{Q} -martingale.

By the uniqueness of the local martingale problem for $(\tilde{\mathcal{A}}, Y_0, T_n)$ we conclude that $\mathbb{Q} = \tilde{\mathbb{P}}$ on $\mathcal{G} := \sigma(\cup_n \mathcal{F}_{T_n}^Y)$ (see [13]).

We now claim that

$$\mathcal{F}_t^Y \subset \mathcal{G} \quad \forall t \geq 0, \quad (26)$$

which yields the lemma. The proof of (26) follows from the observation that, for any Borel set $B \subset \mathbb{R}_+^3$, we have the almost sure equalities

$$\{Y_t \in B\} = \{Y_t \in B\} \cap \{T_\Delta > t\} = \bigcup_{n \geq 1} \underbrace{\{Y_t \in B\} \cap \{T_n > t\}}_{\in \mathcal{F}_{T_n}^Y} \in \mathcal{G}$$

and

$$\{Y_t = \Delta\} = \bigcap_{n \geq 1} \{T_n \leq t\} \in \mathcal{G}.$$

□

6 Empirical Testing

In this section, we empirically examine the cross-sectional fitting abilities of the affine model proposed in the previous sections. For comparison, we consider the two extreme cases of

- a purely structural (PS) model with $\theta = 0.75$ and

$$\alpha_2 = b_2 = \beta_{21} = c = \gamma_1 = \gamma_2 = 0, \quad \lambda_2 = 1, \quad \text{and hence } T_{\mathcal{D}} = T_\Delta;$$

- a purely reduced-form (RF) model; i.e., it is assumed that

$$\ell = \lambda_1 = \lambda_2 = 0, \quad \gamma_2 = 1, \quad \text{and hence } T_{\mathcal{D}} = T_J;$$

and the mixture of both:

- a mixed (MX) model; i.e., it is assumed that

$$\theta = 0.75, \quad \alpha_2 = \beta_{21} = c = 0, \quad \lambda_2 = 1, \quad \text{and hence } T_{\mathcal{D}} = T_\Delta \wedge T_J.$$

It is worth noting that the parameters in the three models are not fully identifiable, since the state variable Y^2 is subject to an arbitrary scale. Therefore, without loss of generality, we fix $\lambda_2 = 1$ for the PS and MX models, and set $\gamma_2 = 1$ for the RF model, and then the remaining parameters are identifiable.

The data we used, including both Treasury and corporate bond price quotes, have been downloaded from Bondpage.com. It consists of one-time observations of 50 Treasury note and bond prices and more than 600 month-end quoted prices of corporate bonds issued by investment-grade firms with ratings between Baa and Aaa. All bonds are non-callable with at least half a year remaining to maturity and share the same settlement date.

6.1 Estimation Strategy

A simple nonlinear least squares algorithm is applied to estimating the parameters to fit the term structure of both Treasury rates and the corporate bond spreads using a snapshot of the market data.

However, two problems come up. First, although the data include 600 non-callable corporate bond prices, no individual firm has more than 10 observations. Hence the credit index estimation for each individual firm is subject to substantial uncertainty. Duffee [9] encountered similar problems when estimating the default intensity of each firm. A way to overcome this difficulty is to form four rating groups, Aaa, Aa, A, Baa, and estimate a typical credit index value, Y_{Aaa}^2 , Y_{Aa}^2 , Y_A^2 , Y_{Baa}^2 , for each of these groups, respectively.

The second problem is the difficulty of estimating θ , which determines the jump characteristics of Y^2 . The parameter θ turns out to be dominant over the other parameters. Changing the value of θ results in significant value changes of the other parameters, but the differences between measurement errors are rather small, which implies that estimating the parameter θ by minimizing the mean square error is infeasible. Therefore, we fix θ equal to 0.75 instead for both the PS and MX model when implementing the optimization algorithm.

The objective function can be written as

$$F(\varsigma) = \left\{ \sum_{i=1}^N (p_i - P_i(\varsigma, \vec{T}_i, c_i))^2 \right\},$$

where p_i denotes the observed price of bond i , and $P_i(\varsigma, \vec{T}_i, c_i)$ denotes the model implied bond price given the parameter set ς , the semiannual coupon rate c_i and coupon payment dates $\vec{T}_i = (T_{i,1}, T_{i,2}, \dots, T_{i,m_i})$. Here zero-recovery at default is assumed when calculating corporate bond prices. Therefore the estimator is given by

$$\varsigma^* = \arg \min \{F(\varsigma)\}.$$

There exist several standard nonlinear least squares algorithms. We employ the Levenberg–Marquardt method, a simple but robust nonlinear least squares algorithm. The basic idea is to approximate F with a simple linear function which reasonably reflects the behavior of F in a neighborhood of the initial point ς_0 , and thus we are able to attack the nonlinear least squares problem using the linear least squares algorithm. Since the Gauss-Newton method often encounters problems when the approximated Hessian matrix is singular, the Levenberg–Marquardt method is applied to overcome this difficulty by adding a typical

positive-definite diagonal matrix. Moreover, the step length is determined by a linear search. For more details about the Levenberg–Marquardt method, we refer to [19, 20]. For robustness of the estimation, thirty independent procedures (experiments) have been performed with different initial values for ς and for each of the three models: PS, RF and MX models.

6.2 Estimation Results

6.2.1 Parameter Estimation

Table 1 summarizes the estimation results of the parameters for each model based on the previously described nonlinear least squares algorithm.

[Table 1 about here.]

As shown in Table 1, for each of the models, the credit index process Y^2 is mean-reverting under the risk-neutral measure. Moreover, the non-zero values for λ_1 , β_{21} and γ_1 suggest that the risk-free rate does have a significant impact on the credit migration. This empirically supports the hypothesis of a stochastic dependence between risk-free rates and credit risk. Finally, we conclude that the mixed model (MX) outperforms the other two models (PS and RF) with regard to the smaller mean square error (MSE) of the optimization.

6.2.2 Estimation of Credit Indices

The estimates for the typical credit indices ($Y_{Aaa}^2, Y_{Aa}^2, Y_A^2, Y_{Baa}^2$) (standard error in parentheses) assigned to each class (Aaa, Aa, A, Baa) for each model (PS, RF, MX) are shown in Table 2. For each model we have $Y_{Aaa}^2 < Y_{Aa}^2 < Y_A^2 < Y_{Baa}^2$ as expected.

[Table 2 about here.]

It speaks for the quality of a model if the values for credit indices Y^2 do not vary too much for firms within one rating class. In view of the affine yield spread curve (22), this is equivalent to saying that the Y^2 -sensitive part, $-\frac{1}{T}\tilde{\psi}_2^{co}(T, 0)$, has an appropriate shape. It is therefore an interesting test to solve for the credit index Y^2 of every individual firm, after having fixed all the remaining parameters given by the preceding estimation.

For the PS case, there exists a significant downward drift of the Y^2 -value from short-term bonds to long-term bonds as shown in Figure 1. This means that $T \mapsto -\frac{1}{T}\tilde{\psi}_2^{co}(T, 0)$ is too steep, resulting in an overestimate of long term credit spreads which has to be compensated by smaller values of Y^2 . The zero yield spread at zero maturity (see (23)) also contributes to this phenomenon, see also [14].

[Figure 1 about here.]

For the RF model notice that quite a few Aaa-rated bonds imply negative Y^2 -values (shown in Figure 2), which is not allowed in our affine setup. This means that the fixed yield spread part (the first two summands in (22)) is too large and has to be compensated by subtracting the Y^2 -sensitive part.

[Figure 2 about here.]

The MX model clearly outperforms PS and RF models given its flat term structures of the credit indices across four credit rating classes as shown in Figures 3.

[Figure 3 about here.]

6.2.3 Spread Curves and Default Probabilities

Figure 4 compares the yield spread curves for the Baa-rated class for the three different models. We can see that the PS model has a zero yield spread at zero maturity, and how the MX model lies between the two extreme cases PS and RF.

[Figure 4 about here.]

We finally assume that the change from real-world measure \mathbb{P} to risk-neutral measure $\tilde{\mathbb{P}}$ is given according to Section 5 by $\tilde{b}_1 = b_1$ and $\tilde{\beta}_1 - \beta_1 = -0.02$ (larger mean reversion rate under \mathbb{P} : $|\tilde{\beta}_1| > |\beta_1|$). This estimate of the market price of risk is taken from [8]. Figure 5 shows the resulting default distributions for each model for rating class Baa. The main difference between the models is in the short end, where the PS model has a flat distribution function.

[Figure 5 about here.]

A Proof of Lemma 4.1

We first prove an auxiliary result.

Lemma A.1. Y_-^1 is continuous on $[0, T_\Delta]$ a.s. Hence $Y_{T_\Delta-}^1 = \lim_{t \uparrow T_\Delta} Y_t^1$ exists a.s.

Proof. In view of (3), $N_{t \wedge \tau}^g$ is a continuous local martingale for every stopping time $\tau < T_\Delta$ and $g \in C_b^2(\mathbb{R}_+)$, where

$$N_t^g := g(Y_t^1) - g(Y_0^1) - \int_0^t \mathcal{A}_1 g(Y_s^1) ds \quad (27)$$

(recall Remark 2.1 and (16)).

We now claim that there exists a universal constant $C_0 = C_0(t)$ such that

$$\mathbb{E} \left[\sup_{s < t \wedge T_\Delta} (Y_s^1)^2 \right] \leq C_0 (1 + (Y_0^1)^2) e^{C_0 t}. \quad (28)$$

Indeed, in view of (11) and (27)

$$Z_t^n := Y_{t \wedge T_n}^1 - Y_0^1 - \int_0^{t \wedge T_n} (b_1 + \beta_1 Y_s^1) ds$$

is a continuous local martingale with quadratic variation

$$\langle Z^n \rangle_t = \int_0^{t \wedge T_n} \alpha_1 Y_s^1 ds$$

for every $n \geq 1$. From Doob's maximal inequality we obtain

$$\mathbb{E}[\sup_{s \leq t} (Z_s^n)^2] \leq 4\mathbb{E}[(Z_t^n)^2] = 4\mathbb{E}[\langle Z^n \rangle_t] = 4\mathbb{E} \left[\int_0^{t \wedge T_n} \alpha_1 Y_s^1 ds \right].$$

Moreover,

$$\begin{aligned} (Y_{t \wedge T_n}^1)^2 &= \left(Y_0^1 + \int_0^{t \wedge T_n} (b_1 + \beta_1 Y_s^1) ds + Z_t^n \right)^2 \\ &\leq 3 \left((Y_0^1)^2 + t \int_0^{t \wedge T_n} (b_1 + \beta_1 Y_s^1)^2 ds + (Z_t^n)^2 \right). \end{aligned}$$

Combining this we get

$$\begin{aligned} g_n(t) &:= \mathbb{E}[\sup_{s \leq t} (Y_{s \wedge T_n}^1)^2] \\ &\leq C_1 \left((Y_0^1)^2 + \mathbb{E} \left[\int_0^{t \wedge T_n} ((b_1 + \beta_1 Y_s^1)^2 + \alpha_1 Y_s^1) ds \right] \right), \end{aligned}$$

which implies

$$g_n(t) \leq C_2 \left(1 + (Y_0^1)^2 + \int_0^t g_n(s) ds \right),$$

where the constants C_1, C_2 depend only on t, α_1, b_1 and β_1 . Gronwall's inequality yields

$$g_n(t) \leq C_0(1 + (Y_0^1)^2)e^{C_0 t}$$

with $C_0 = C_0(t, C_2)$. Monotone convergence for $n \rightarrow \infty$ yields (28).

On the other hand, we have by the same arguments as above that

$$\sup_{s \leq t} (Y_{s \wedge T_n}^1 - Y_{s \wedge T_m}^1)^2 \leq 2 \left(t \int_{t \wedge T_m}^{t \wedge T_n} (b_1 + \beta_1 Y_u^1)^2 du + \sup_{s \leq t} (Z_s^n - Z_s^m)^2 \right)$$

and hence

$$\begin{aligned} \mathbb{E} \left[\sup_{s \leq t} (Y_{s \wedge T_n}^1 - Y_{s \wedge T_m}^1)^2 \right] &\leq 2\mathbb{E} \left[t \int_{t \wedge T_m}^{t \wedge T_n} (b_1 + \beta_1 Y_u^1)^2 du + \int_{t \wedge T_m}^{t \wedge T_n} \alpha_1 Y_s^1 ds \right] \\ &\leq C_3 \mathbb{E} \left[\left(1 + \sup_{s < t \wedge T_\Delta} (Y_s^1)^2 \right) (t \wedge T_n - t \wedge T_m) \right] \end{aligned}$$

for all $n \geq m$, where C_3 does not depend on m, n . Using (28) and dominated convergence we conclude that

$$\lim_{m, n \rightarrow \infty} \mathbb{E} \left[\sup_{s \leq t} (Y_{s \wedge T_n}^1 - Y_{s \wedge T_m}^1)^2 \right] = 0.$$

Hence $Y_{t \wedge T_n}^1 = Y_{t \wedge T_n -}^1$ converges uniformly in t on compacts in probability to $Y_{t \wedge T_\Delta -}^1$, which proves the lemma. \square

Proof of Lemma 4.1 Recall that the stochastic basis is rich enough to carry an (\mathcal{F}_t) -Brownian motion W . Define the $(\mathcal{F}_{T_\Delta+t})_{t \geq 0}$ -Brownian motion

$$W_t^\Delta := W_{T_\Delta+t} - W_{T_\Delta}, \quad t \geq 0$$

and consider the stochastic differential equation

$$\begin{aligned} dR_t &= (b_1 + \beta_1 R_t) dt + \sqrt{2\alpha_1 R_t} dW_t^\Delta \\ R_0 &= Y_{T_\Delta -}^1. \end{aligned} \tag{29}$$

It is well known that a unique continuous $(\mathcal{F}_{T_\Delta+t})_{t \geq 0}$ -adapted strong solution R exists. Notice that

$$\{(t - T_\Delta)^+ \leq c\} = \{T_\Delta \geq t\} \cup (\{T_\Delta < t\} \cap \{T_\Delta \geq t - c\}) \in \mathcal{F}_t \quad \forall c \geq 0,$$

hence $R_{(t-T_\Delta)^+} 1_{\{t \geq T_\Delta\}}$ is (\mathcal{F}_t) -adapted.

We then define the continuous adapted process

$$r_t := Y_t^1 1_{\{t < T_\Delta\}} + R_{t-T_\Delta} 1_{\{t \geq T_\Delta\}}.$$

Let $g \in C_c^2(R_+)$. By Itô's formula and since $\int_{T_\Delta}^{T_\Delta+t} \phi_u dW_u = \int_0^t \phi_{T_\Delta+u} dW_u^\Delta$ we have

$$\begin{aligned} g(r_t) - g(r_{t \wedge T_\Delta}) &= g(R_{(t-T_\Delta)^+}) - g(R_0) \\ &= \int_0^{(t-T_\Delta)^+} \mathcal{A}_1 g(R_s) ds + \int_0^{(t-T_\Delta)^+} \sqrt{2\alpha_1 R_s} g'(R_s) dW_s^\Delta \\ &= \int_{t \wedge T_\Delta}^t \mathcal{A}_1 g(r_s) ds + \int_{t \wedge T_\Delta}^t \sqrt{2\alpha_1 r_s} g'(r_s) dW_s. \end{aligned}$$

Hence

$$\begin{aligned} N_t^g &= g(r_t) - g(r_0) - \int_0^t \mathcal{A}_1 g(r_s) ds \\ &= \int_{t \wedge T_\Delta}^t \sqrt{2\alpha_1 r_s} g'(r_s) dW_s \\ &\quad + g(r_{t \wedge T_\Delta}) - g(r_{t \wedge T_n}) - \int_{t \wedge T_n}^{t \wedge T_\Delta} \mathcal{A}_1 g(r_s) ds \\ &\quad + g(Y_{t \wedge T_n}^1) - g(Y_0^1) - \int_0^{t \wedge T_n} \mathcal{A}_1 g(Y_s^1) ds \end{aligned}$$

satisfies for $s \leq t$

$$\begin{aligned}
\mathbb{E}[N_t^g \mid \mathcal{F}_s] &= \int_{s \wedge T_\Delta}^s \sqrt{2\alpha_1 r_u} g'(r_u) dW_u \\
&+ \mathbb{E} \left[g(r_{t \wedge T_\Delta}) - g(r_{t \wedge T_n}) - \int_{t \wedge T_n}^{t \wedge T_\Delta} \mathcal{A}_1 g(r_u) du \mid \mathcal{F}_s \right] \\
&+ g(Y_{s \wedge T_n}^1) - g(Y_0^1) - \int_0^{s \wedge T_n} \mathcal{A}_1 g(Y_u^1) du \\
&= N_s^g - \left(g(r_{s \wedge T_\Delta}) - g(r_{s \wedge T_n}) - \int_{s \wedge T_n}^{s \wedge T_\Delta} \mathcal{A}_1 g(r_u) du \right) \\
&+ \mathbb{E} \left[g(r_{t \wedge T_\Delta}) - g(r_{t \wedge T_n}) - \int_{t \wedge T_n}^{t \wedge T_\Delta} \mathcal{A}_1 g(r_u) du \mid \mathcal{F}_s \right].
\end{aligned}$$

This holds for any $n \geq 1$. Letting $n \rightarrow \infty$ we get by continuity of r and dominated convergence that

$$\mathbb{E}[N_t^g \mid \mathcal{F}_s] = N_s^g \quad \text{a.s.}$$

Hence N^g is a martingale and the proof of Lemma 4.1 is complete.

References

- [1] C. Albanese, J. Campolieti, O. Chen, and A. Zavidonov, *Credit barrier models*, Risk Magazine (June 2003), 109–113.
- [2] T. R. Bielecki and M. Rutkowski, *Credit risk: modelling, valuation and hedging*, Springer Finance, Springer-Verlag, Berlin, 2002.
- [3] T. Björk, *Arbitrage theory in continuous time*, Oxford University Press, 1998.
- [4] L. Chen and D. Filipović, *Credit derivatives in an affine framework*, working paper, Princeton University, 2004.
- [5] P. Cheridito, D. Filipović, and M. Yor, *Equivalent and absolutely continuous measure changes for jump-diffusion processes*, working paper, Princeton University (www.math.ethz.ch/~filipo/ftp/preprints/pubpre.html), 2003.
- [6] J. C. Cox, J. Ingersoll, and S. A. Ross, *A theory of the term structure of interest rates*, *Econometrica* **53** (1985), 385–408.
- [7] R. Douady and M. Jeanblanc, *A rating-based model for credit derivatives*, working paper, Evry University, 2002.

- [8] J. C. Duan and J. G. Simonato, *Estimating and testing exponential-affine term structure models by Kalman filter*, Review of Quantitative Finance and Accounting **13** (1999), 111–135.
- [9] G. R. Duffee, *Estimating the price of default risk*, The Review of Financial Studies **12** (1999), 41–59.
- [10] D. Duffie, D. Filipović, and W. Schachermayer, *Affine processes and applications in finance*, Ann. Appl. Probab. **13** (2003), no. 3, 984–1053.
- [11] D. Duffie and D. Lando, *Term structure of credit spreads with incomplete accounting information*, Econometrica **69** (2001), 633–664.
- [12] D. Duffie and K. Singleton, *Credit risk: Pricing, measurement, and management*, Princeton University Press, 2003.
- [13] S. N. Ethier and T. G. Kurtz, *Markov processes. Characterization and convergence*, John Wiley & Sons Inc., New York, 1986.
- [14] K. Giesecke, *Credit risk modeling and valuation: An introduction*, Working Paper, Cornell University, 2003.
- [15] J. Jacod and A. N. Shiryaev, *Limit theorems for stochastic processes*, Grundlehren der mathematischen Wissenschaften, vol. 288, Springer-Verlag, Berlin-Heidelberg-New York, 1987.
- [16] R. A. Jarrow, D. Lando, and S. M. Turnbull, *A Markov model for the term structure of credit risk spreads*, The Review of Financial Studies **10** (1997), 481–523.
- [17] R. A. Jarrow and S. M. Turnbull, *Pricing derivatives on financial securities subject to credit risk*, Journal of Finance **50** (1995), 53–85.
- [18] D. Lando, *On Cox processes and credit-risky securities*, Review of Derivatives Research **2** (1998), 99–120.
- [19] K. Levenberg, *A method for the solution of certain non-linear problems in least squares*, Quarterly Applied Mathematics **2** (1944), 164–168.
- [20] D. Marquardt, *An algorithm for least squares estimation of nonlinear parameters*, SIAM Journal of Applied Mathematics **11** (1963), 431–441.
- [21] D. Revuz and M. Yor, *Continuous martingales and Brownian motion*, Grundlehren der mathematischen Wissenschaften, vol. 293, Springer-Verlag, Berlin-Heidelberg-New York, 1994.
- [22] L. C. G. Rogers and D. Williams, *Diffusions, Markov processes, and martingales. Vol. 2*, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 2000, Itô calculus, Reprint of the second (1994) edition.

List of Figures

1	Credit Indices for the PS Model	19
2	Credit Indices for the RF Model	20
3	Credit Indices for the MX Model	21
4	Spread Curves for Baa-Rating Class	22
5	Default Distributions for the Baa-Rating Class	23

Figure 1: Credit Indices for the PS Model

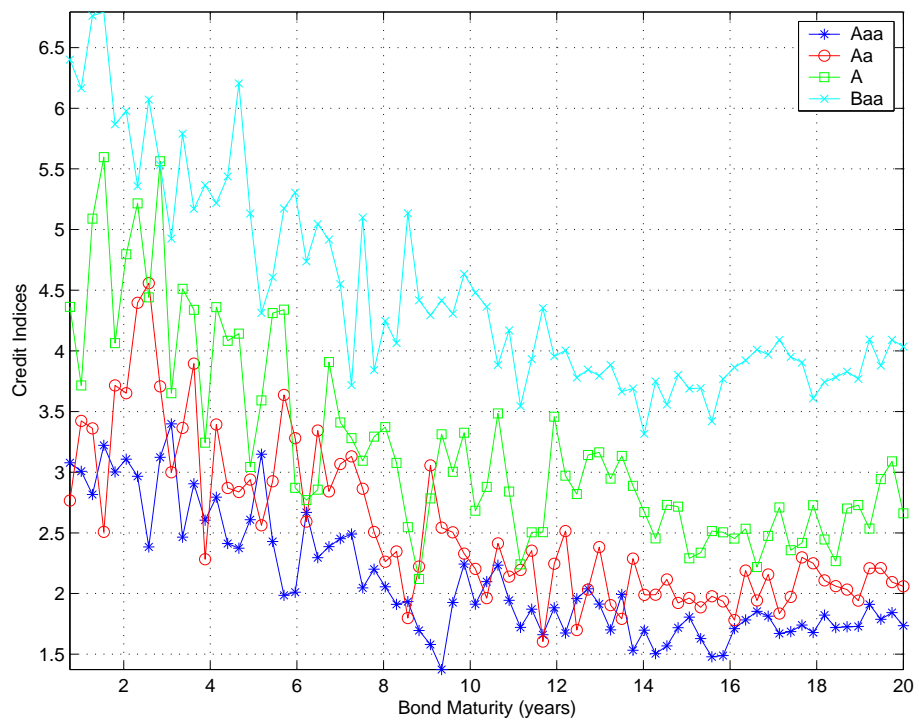


Figure 2: Credit Indices for the RF Model

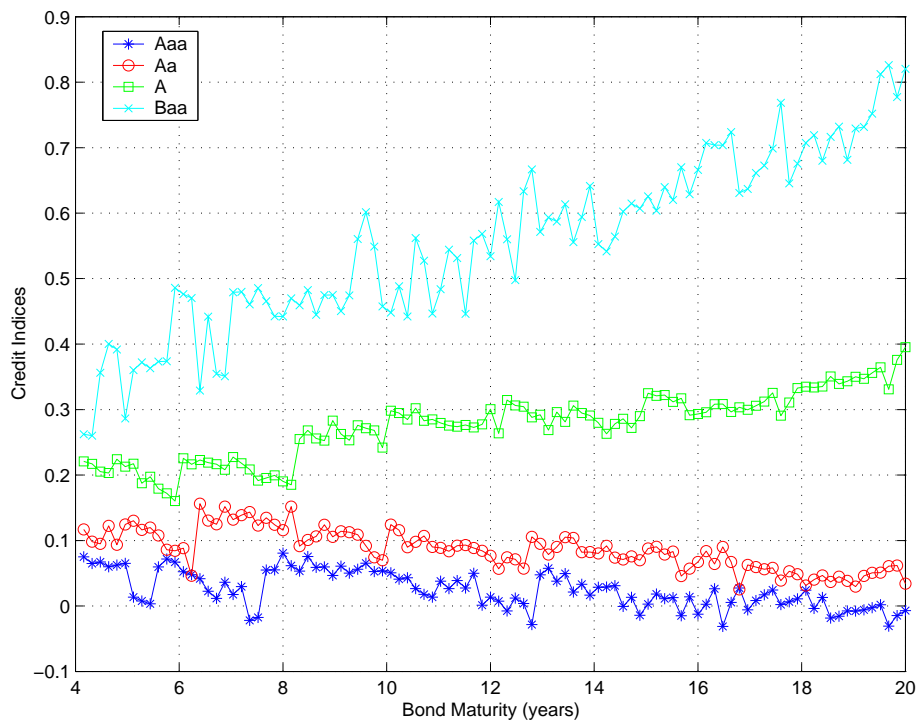


Figure 3: Credit Indices for the MX Model

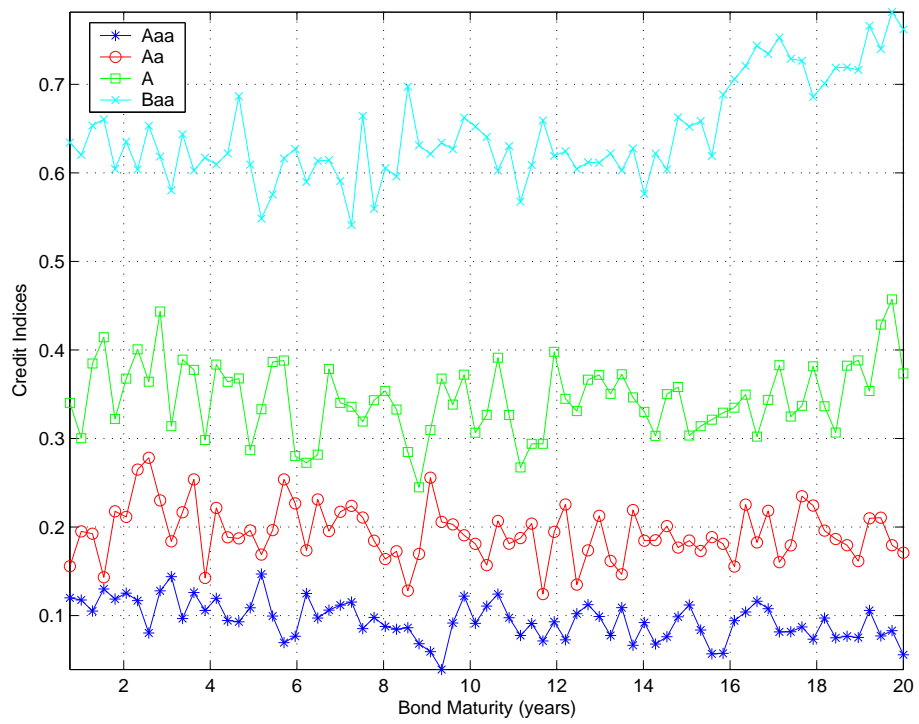


Figure 4: Spread Curves for Baa-Rating Class

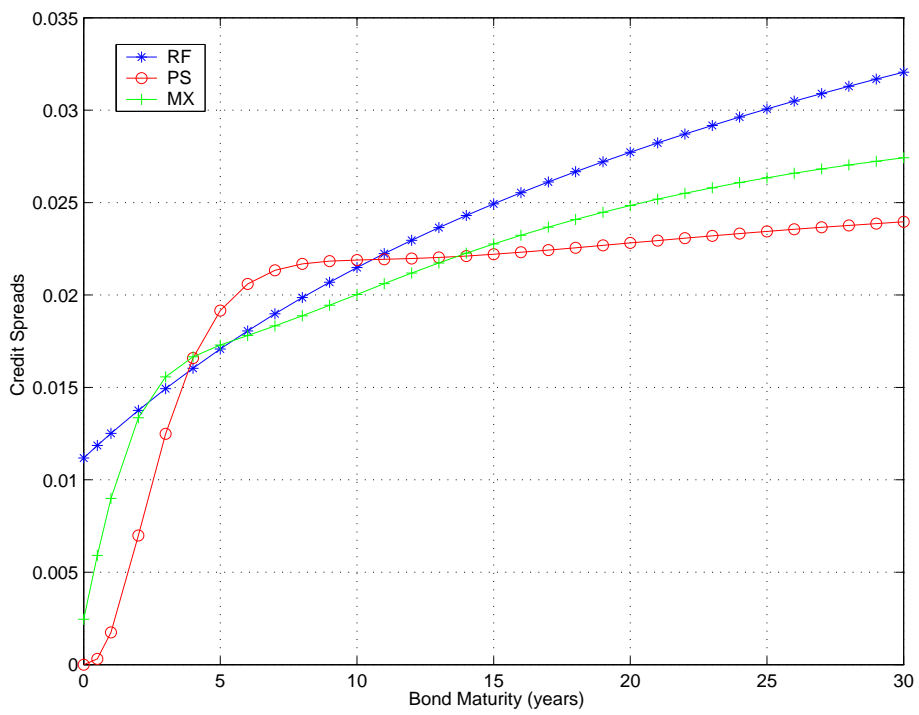
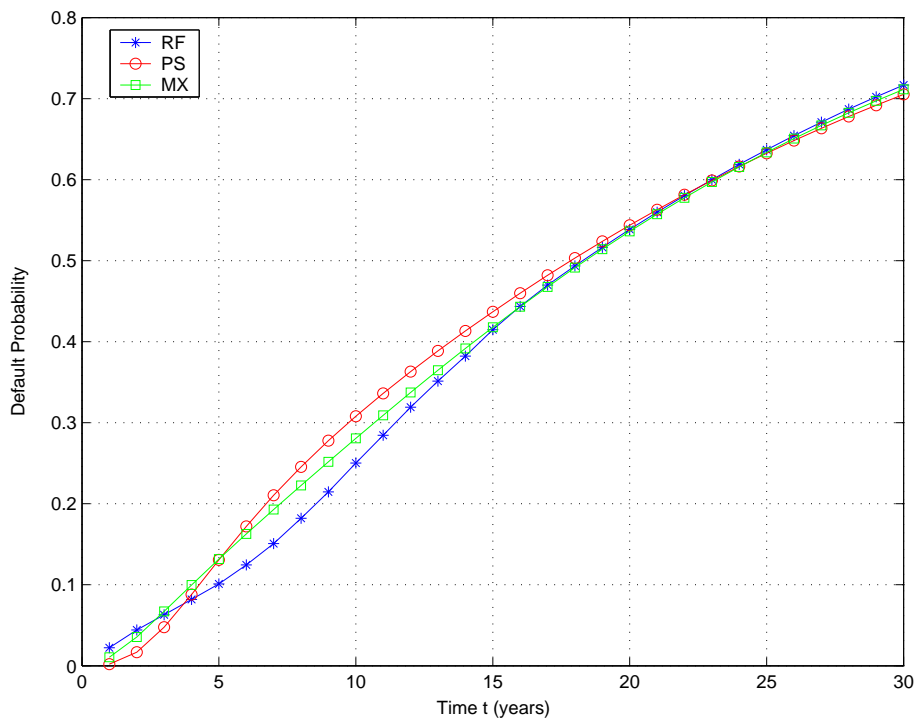


Figure 5: Default Distributions for the Baa-Rating Class



List of Tables

1	Parameter Estimates for Fitting the Current Term Structure of Treasury Rates and Credit Spreads	25
2	Summary Statistics of Credit Indices	26

Table 1: Parameter Estimates for Fitting the Current Term Structure of Treasury Rates and Credit Spreads

Parameter	PS	RF	MX
b_1	0.0107 (0.00018)	0.01167 (0.00024)	0.0112 (0.00032)
β_1	-0.141 (0.0049)	-0.155 (0.0050)	-0.141 (0.0061)
α_1	1.245e-5 (4.3e-5)	1.736e-5 (3.4e-5)	1.383e-5 (2.8e-5)
r	0.0107 (0.00031)	0.0105 (0.00041)	0.0117 (0.00035)
b_2	-	2.534e-6 (5.2e-5)	2.632e-6 (8.1e-6)
β_{22}	-1.571 (0.82)	-1.473 (1.03)	-2.253 (0.93)
β_{21}	-	0.0323 (0.00215)	-
α_2	-	0.7323 (0.53)	-
c	-	9.132e-7 (0.13e-6)	-
γ_1	-	0.2635 (0.00064)	0.103 (0.034)
γ_2	-	1	0.00186 (0.00032)
ℓ	0.163 (0.0032)	-	-
λ_1	5.682 (0.92)	-	2.353 (1.34)
λ_2	1	-	1
	-	-	-

Yield Maturity	RMSE (basis points)		
	PS Model	RF Model	MX Model
0-4 yr. Treasury	3.240	3.281	3.215
4-20 yr. Treasury	0.942	0.992	0.941
0-4 yr. Aaa rated bonds	24.493	14.689	5.032
4-20 yr. Aaa rated bonds	4.322	6.235	3.843
0-4 yr. Aa rated bonds	16.492	5.517	4.321
4-20 yr. Aa rated bonds	1.242	2.477	0.942
0-4 yr. A rated bonds	19.324	19.176	9.421
4-20 yr. A rated bonds	4.231	5.435	4.193
0-4 yr. Baa rated bonds	30.213	24.764	15.332
4-20 yr. Baa rated bonds	5.313	6.394	5.028

The parameters are estimated using the price data of more than 650 Treasury and corporate bonds. Thirty independent experiments are performed. The mean estimates together with the standard errors (in parentheses) are presented. The fitting errors (RMSE) are also presented with respect to different maturities and four rating classes.

Table 2: Summary Statistics of Credit Indices

Rating	Mean (Stdev)		
	PS	RF	MX
Y_{Aaa}^2	2.393 (0.429)	0.0250 (0.0282)	0.0951 (0.0184)
Y_{Aa}^2	2.848 (0.558)	0.0864 (0.0306)	0.193 (0.0265)
Y_A^2	4.002 (0.670)	0.275 (0.0504)	0.345 (0.0354)
Y_{Baa}^2	5.194 (0.730)	0.554 (0.131)	0.644 (0.0450)