A Markov Model for the Term Structure of Credit Risk Spreads

Robert A. Jarrow
Cornell University

David Lando
University of Copenhagen

Stuart M. Turnbull
Queen's University

This article provides a Markov model for the term structure of credit risk spreads. The model is based on Jarrow and Turnbull (1995), with the bankruptcy process following a discrete state space Markov chain in credit ratings. The parameters of this process are easily estimated using observable data. This model is useful for pricing and hedging corporate debt with imbedded options, for pricing and hedging OTC derivatives with counterparty risk, for pricing and hedging (foreign) government bonds subject to default risk (e.g., municipal bonds), for pricing and hedging credit derivatives, and for risk management.

This article presents a simple model for valuing risky debt that explicitly incorporates a firm's credit rating as an indicator of the likelihood of default. As such, this article presents an arbitrage-free model for the term structure of credit risk spreads and their evolution through time. This model will prove useful for the pricing and hedging of corporate debt with

---

We would like to thank John Tierney of Lehman Brothers for providing the bond index price data, and Tal Schwartz for computational assistance. We would also like to acknowledge helpful comments received from an anonymous referee. Send all correspondence to Robert A. Jarrow, Johnson Graduate School of Management, Cornell University, Ithaca, NY 14853.

The Review of Financial Studies Summer 1997 Vol. 10, No. 2, pp. 481–523

imbedded options, for the pricing and hedging of OTC derivatives with counterparty risk, for the pricing and hedging of (foreign) government bonds subject to default risk (e.g., municipal bonds), and for the pricing and hedging of credit derivatives (e.g. credit sensitive notes and spread adjusted notes). This model can also be used for risk management purposes, as it is possible to calculate the expected credit exposure profile over the life of a contract or a portfolio of contracts. To our knowledge, this is the first contingent claims model to explicitly incorporate credit rating information into the valuation methodology. Our model is an extension and a refinement of that contained in Jarrow and Turnbull (1995).

Previous models for the pricing of risky debt can be subdivided into three classes. The first class of models views the firm's liabilities as contingent claims issued against the firm's underlying assets, with the payoffs to all the firm's liabilities in bankruptcy completely specified. Bankruptcy is determined via the evolution of the firm's assets in conjunction with the various debt covenants [e.g., Black and Cox (1976), Chance (1990), Merton (1974), and Shimko, Tejima, and van Deventer (1993)]. This approach is difficult to implement in practice because all of the firm's assets are not tradeable nor observable. Secondly, to utilize this technique, the complex priority structure of the payoffs to all of the firm's liabilities need to be specified and included in the valuation procedure. This is a difficult task [see Jones, Mason, and Rosenfeld (1984)]. Furthermore, since this approach does not use credit rating information, it cannot be used to price credit derivatives whose payoffs depend directly on the credit rating [e.g., credit sensitive notes and spread adjusted notes, see Das and Tufano (1995) for additional elaboration].

The second class of models views risky debt as paying off an exogenously given fraction of each promised dollar in the event of bankruptcy. Bankruptcy is determined when the value of the firm's underlying assets hits some exogenously specified boundary [e.g., Hull and White (1991), Longstaff and Schwartz (1992), and Nielsen, Saá-Requejo and Santa-Clara (1993)]. This approach simplifies the first class of models by both exogenously specifying the cash flows to risky debt in the event of bankruptcy and in simplifying the bankruptcy process. Although this approach simplifies computation by avoiding the need to understand the complex priority structure of payoffs to all of the firm's liabilities in bankruptcy, it still requires estimates for the parameters of the firm's asset value, which is nonobservable; and it still cannot handle various credit derivatives whose payouts depend on the credit rating of the debt issue.

The third class of models avoids these last two problems. This approach, like the second, also views risky debt as paying off a fraction of each promised dollar in the event of bankruptcy; but the time of bankruptcy is now given as an exogenous process [e.g., Jarrow and Turnbull (1995) and Litterman and Iben (1991)]. This bankruptcy process is specified exogenously and does not explicitly depend on the firm's underlying assets. The advantage of this approach is that it allows exogenous assumptions to be imposed only on observables. All three classes of models are arbitrage-free.[1] In fact, this modeling approach includes the previous two as special cases where the payoff process and/or the bankruptcy process are endogenously derived. Also, the third approach can easily be modified to include credit rating information in the bankruptcy process, and therefore it can be used to price credit derivatives dependent on the credit rating. This provides the motivation underlying the choice of the model in this article.

This article takes the Jarrow and Turnbull (1995) model, and characterizes the bankruptcy process as a finite state Markov process in the firm's credit ratings. This new credit risk model has the following characteristics:

1. Different seniority debt for a particular firm can be incorporated via different recovery rates in the event of default.
2. It can be combined with any desired term-structure model for default-free debt [e.g., Black, Derman, and Toy (1990), Cox, Ingersoll, and Ross (1985), or Heath, Jarrow, and Morton (1992)].
3. It utilizes historical transition probabilities for the various credit rating classes to determine the pseudo-probabilities (martingale, risk adjusted) used in valuation.
4. It can be utilized, as shown in Jarrow and Turnbull (1995), to price and hedge options on risky debt or credit derivatives. The pricing and hedging of vulnerable options is a special case of this analysis.

For implementation, we impose one simplifying assumption on the interaction between the default-free term structure and the firm's bankruptcy process. We assume that the two processes are statistically independent under the pseudo-probabilities. Alternatively stated, the Markov process for credit ratings is independent of the level of spot interest rates (under the pseudo-probabilities). For investment grade debt, this appears to be a reasonable first approximation in the historical probabilities, though for speculative grade debt, the accuracy

[1] An analogy can be made between this class of models and those used to price equity options, that is, Jarrow and Turnbull's (1995) exogenous assumption on observables is equivalent to Black and Scholes (1973), whereas Merton's (1974) assumption on firm values is equivalent to the approach of Geske (1979).

of the approximation deteriorates.[2] The merit of this assumption is an outstanding empirical issue. It is imposed as a simplifying assumption to facilitate empirical investigation. If needed, it can be easily relaxed, as shown in Jarrow and Turnbull (1995). The extension of our credit class model in continuous time to include spot rate–dependent probabilities can be found in Lando (1994).

An important application of this model is in the area of risk management. In the recent (1993) Group of Thirty report,[3] two of the recommendations address the issue of measuring current and potential credit exposure, and aggregating credit exposures. The model can be used to define the current credit exposure and to generate the distribution of credit exposure over the life of a contract. Two commonly used statistics can be computed: the maximum exposure and expected exposure time profiles. That is, starting from a particular credit class, one can compute the probability of being in a given credit class after a fixed time interval. For pricing purposes, it is necessary to use the "risk-neutral" probabilities. For risk management purposes, however, it is necessary to use both the "risk-neutral" and the empirical probabilities. This model can also be extended to portfolios of contracts such as interest rate and foreign currency swaps.

An outline of this article is as follows. Section 1 describes the relevant features of the Jarrow and Turnbull model. Section 2 presents the discrete time model. Section 3 presents the continuous time model. Section 4 concludes the article.

1. The Jarrow–Turnbull Model

The model we use is based on Jarrow and Turnbull (1995). We consider a frictionless economy with a finite horizon $[0, \tau]$. Trading can be discrete or continuous (both cases are studied below). The underlying uncertainty is represented by a filtered probability space $(\Omega, Q, \mathcal{F}_\tau, (\mathcal{F}_t)_{0\leq t\leq\tau})$. The details of this filtered probability space are specified later. Traded are default-free zero-coupon bonds of all maturities, a default-free money market account, and risky zero-coupon bonds of all maturities.

[2] Moody's Special Report (1992), Figures 15–18, show that 1-year default rates for investment grade bonds have had little variation over the time period 1970–1991. Although speculative grade 1-year default rates exhibit more variation, on an absolute level, it is not large. The standard deviation of 1-year default rates is highest for B-rated firms, and there it is only 5.04%. Longstaff and Schwartz (1992) provide some evidence inconsistent with this hypothesis.

[3] The Group of Thirty is an international financial policy organization whose members include representatives of central banks, international banks, and securities firms. They issued a report entitled "Derivatives: Practice and Principles" in July 1993 to improve the supervisory and capital requirements of the off balance sheet risks related to derivatives trading activity.

We assume *that there exists a unique equivalent martingale measure $\tilde{Q}$ making all the default-free and risky zero-coupon bond prices martingales, after normalization by the money market account.* This assumption is equivalent to the statement that the markets for default-free and risky debt are complete and arbitrage-free [see Harrison and Pliska (1981)]. Sufficient conditions for the satisfaction of this assumption can be found in Jarrow and Turnbull (1995). Given information at time t, we denote conditional expectation and conditional probability statements with respect to the equivalent probability measure by $\tilde{E}_t(\bullet)$ and $\tilde{Q}_t(\bullet)$, respectively.

Let $p(t, T)$ be the time t price of a default-free zero-coupon bond paying a sure dollar at time T where $0 \leq t \leq T \leq t$. We assume forward rates of all maturities exist and that they are defined (1) in the discrete time case by $f(t, T) \equiv -\log(p(t, T+1)/p(t, T))$ and (2) in the continuous time case by $f(t, T) \equiv \frac{-\partial}{\partial T} \log p(t, T)$. The default-free spot rate, denoted $r(t)$, is defined by $r(t) \equiv f(t, t)$. We do not specify any particular stochastic process for spot rates. They can be modeled directly as in Cox et al. (1985), or indirectly via forward rates as in Heath et al. (1992). The money market account accumulates returns at the spot rate and is denoted as either

$$B(t) = \exp\left(\sum_{i=0}^{t-1} r(i)\right) \qquad \text{in the discrete time case,}$$

$$\text{or} \qquad B(t) = \exp\left(\int_0^t r(s)\, ds\right) \qquad \text{in the continuous time case.}$$

Under the maintained assumption of arbitrage-free and complete markets, we can write default-free bond prices as the expected, discounted value of a sure dollar received at time T, that is,

$$p(t, T) = \tilde{E}_t\left(\frac{B(t)}{B(T)}\right). \tag{1}$$

Let $v(t, T)$ be the time t price of a risky zero-coupon bond promising to pay a dollar at time t where $t \leq T \leq \tau$. This promised dollar may not be paid in full if the firm is bankrupt at time T.[4] If bankrupt, the firm pays only $\delta < 1$ dollars. The fraction δ, called the recovery rate, can depend on the priority (seniority) of the risky zero-coupon debt relative to the other liabilities of the firm.

[4] The term "bankrupt" should not be interpreted too literally. In our context it covers any case of financial distress that results in the bondholders receiving less than the promised payment.

The recovery rate δ is taken to be an exogenously given constant. This constancy is imposed for simplicity of estimation. As shown in Jarrow and Turnbull (1995), it implies that the stochastic structure of credit spreads will be independent of the recovery rate, and dependent only on the stochastic structure of spot interest rates and the bankruptcy process. This assumption can easily be relaxed [see Das and Tufano (1995)].[5]

Let τ^* represent the random time at which bankruptcy occurs. Then,

$$v(t, T) = \tilde{E}_t \left(\frac{B(t)}{B(T)} (\delta 1_{\{\tau^* \leq T\}} + 1_{\{\tau^* > T\}}) \right) \tag{2}$$

where $1_{\{\tau^* \leq T\}}$ is the indicator function of the event $\{\tau^* \leq T\}$. The risky zero-coupon bond's price is seen to be the expected, discounted value of a "risky" dollar received at time T. Note that if bankruptcy has occurred prior to time t, it is assumed that claimholders will receive δ for sure at the maturity of the contract. This implies that the risky term structure simplifies considerably as $v(t, T) = \delta \tilde{E}_t(\frac{B(t)}{B(T)}) = \delta p(t, T)$. In other words, in bankruptcy, the term structure of the risky debt collapses to that of the default-free bonds.

Next, we assume *that the stochastic process for default-free spot rates* $\{r(t)_{0 \leq t \leq \tau}\}$ *and the bankruptcy process, as represented by* τ^*, *are statistically independent under* $\tilde{Q}$. This assumption is imposed for simplicity of implementation. This assumption implies that the bankruptcy process (under the pseudo-probabilities) is uncorrelated with default-free spot interest rates. Under the additional structure imposed below, it will also imply that the bankruptcy process (under the empirical probabilities) is uncorrelated with default-free spot interest rates. The reasonableness of this assumption is an outstanding empirical issue. The relaxation of this assumption is discussed in Jarrow and Turnbull (1995) and implemented in Lando (1994). Whether or not this generalization is required awaits empirical testing of the simpler model.

Under this assumption, Equation (2) simplifies to

$$\begin{aligned} v(t, T) &= \tilde{E}_t \left(\frac{B(t)}{B(T)} \right) \tilde{E}_t (\delta 1_{\{\tau^* \leq T\}} + 1_{\{\tau^* > T\}}) \\ &= p(t, T)(\delta + (1 - \delta) \tilde{Q}_t(\tau^* > T)), \end{aligned} \tag{3}$$

where $\tilde{Q}_t(\tau^* > T)$ is the probability under $\tilde{Q}$ that default occurs after date T. The risky zero-coupon bond's price is the default-free zero-

[5] Random recovery rates $\delta(x_T)$, where x_T is the state variable, are also easily fitted into the valuation model, as in Chapter 3 of Lando (1994).

coupon bond's value multiplied by the expected payoff (in dollars) at time T. As revealed in Equation (3), the evolution of the term structure of risky debt is uniquely determined via specifying a distribution for the time of bankruptcy, $\tilde{Q}_t(\tau^* > T)$, under the martingale probabilities. The contribution of our article is an explicit modeling of this distribution as the first hitting time of a Markov chain, with credit ratings and default being the relevant states. This modeling exercise is the content of the remaining sections.

2. Credit Ratings and Default-Probabilities: The Discrete Time Case

This section models the distribution of default time in a discrete trading economy. Discrete trading is illustrated first for two reasons. First, because its mathematical simplicity facilitates understanding. Second, because the discrete time model, parameterized in terms of its continuous limit (the next section), is the model formulation actually implemented on a computer. The theorems in this section parallel those developed in the continuous time section.

2.1 Valuation

The distribution for the default time is modeled via a discrete time, time-homogeneous Markov chain on a finite state space $S = \{1, \ldots, K\}$. The state space S represents the possible credit classes, with 1 being the highest (Aaa in Moody's rankings) and $K - 1$ being the lowest (C in Moody's rankings). The last state, K, represents bankruptcy.

The discrete time, time-homogeneous finite state space Markov chain $\{\eta_t\colon 0 \le t \le \tau\}$ is specified by a $K \times K$ transition matrix[6]

$$Q = \begin{pmatrix} q_{11} & q_{12} & \cdots & q_{1K} \\ q_{21} & q_{22} & \cdots & q_{2K} \\ \vdots & & & \\ q_{K-1,1} & q_{K-1,2} & \cdots & q_{K-1,K} \\ 0 & 0 & \cdots & 1 \end{pmatrix} \tag{4}$$

where $q_{ij} \ge 0$ for all i, j, $i \ne j$, and $q_{ii} \equiv 1 - \sum_{\substack{j=1 \\ j\ne i}}^{K} q_{ij}$ for all i. The (i, j)th entry q_{ij} represents the actual probability of going from state i to state j in one time step. We assume for simplicity that bankruptcy (state K) is an absorbing state, so that $q_{Ki} = 0$ for $i = 1, \ldots, K - 1$

[6] With a slight abuse of notation, we denote the probability measure and the transition matrix as Q. No confusion should arise because the meaning is made clear by the context.

and $q_{KK} = 1$. This explains the last row in the transition matrix. In Jarrow and Turnbull (1995) this assumption is relaxed.

The time homogeneity of this transition matrix is imposed for simplicity of estimation. Based on the evidence contained in Moody's Special Report (1992), it appears to be more reasonable for investment grade bonds, than it is for speculative grade bonds. Again, the validity of this assumption is an outstanding empirical issue.

Let $q_{ij}(0, n)$ denote the n-step transition probability of going from state i at time 0 to state j at time n. It is well known that the n-step $K \times K$ transition matrix, $Q_{0,n}$ whose (i, j)th entry is $q_{ij}(0, n)$ satisfies $Q_{0,n} = Q^n$, the n-fold matrix product of Q.

Estimates of this transition matrix, with a time step of 1 year, can be obtained from Moody's Special Report (1992) [see Lucas and Lonski (1992, Table 6) or Standard & Poor's Credit Review (1993)]. The nonzero entries of the historical matrix tend to be concentrated around the diagonal, with movements of two credit ratings (or more) in a year being rare or nonexistent.

Under the maintained assumption of complete markets and no arbitrage opportunities, the transition matrix from time t to time $t + 1$ under the equivalent martingale probability is given by

$$\tilde{Q}_{t,t+1} = \begin{pmatrix} \tilde{q}_{11}(t, t+1) & \tilde{q}_{12}(t, t+1) & \cdots & \tilde{q}_{1K}(t, t+1) \\ \tilde{q}_{21}(t, t+1) & \tilde{q}_{22}(t, t+1) & \cdots & \tilde{q}_{2K}(t, t+1) \\ \vdots & & & \\ \tilde{q}_{K-1,1}(t, t+1) & \tilde{q}_{K-1,2}(t, t+1) & \cdots & \tilde{Q}_{K-1,K}(t, t+1) \\ 0 & 0 & \cdots & 1 \end{pmatrix} \tag{5}$$

where $\tilde{q}_{ij}(t, t+1) \geq 0$, for all $i, j, i \neq j$, $\tilde{q}_{ii}(t, t+1) \equiv 1 - \sum_{\substack{j=1 \\ j \neq i}}^{K} \tilde{q}_{ij}(t, t+1)$, and $\tilde{q}_{ij}(t, t+1) > 0$ if and only if $q_{ij} > 0$ for $0 \leq t \leq \tau - 1$. Without additional restrictions, the martingale probabilities $\tilde{q}_{ij}(t, t+1)$ can depend on the entire history of the process up to time t. Hence, under the martingale probabilities, the process need not be Markov. To facilitate empirical implementation, it is desirable to impose more structure on these probabilities. In particular, we assume that *the risk premia adjustments are such that the credit rating process under the martingale probabilities satisfy*[7]

$$\tilde{q}_{ij}(t, t+1) = \pi_i(t) q_{ij} \text{ for all } i, j, i \neq j \text{ where} \tag{6}$$

[7] One can always write $\tilde{q}_{ij}(t, t+1) = \pi_{ij}(t) q_{ij}$ where $\pi_{ij}(t)$ can depend on the entire history of the process up to time t. The restriction in Equation (6) is twofold. First, $\pi_{ij}(t)$ is independent of j, and second, it is a deterministic function. The independence of j implies that moving from i to 1 receives the same risk premium as moving from i to K. This restriction may not be true in practice.

$\pi_i(t)$ is a deterministic function of time such that
$\tilde{q}_{ij}(t, t+1) \geq 0$ for all $i, j, i \neq j$ and
$\sum_{\substack{j=1 \\ j \neq i}}^{K} \tilde{q}_{ij}(t, t+1) \leq 1$ for $i = 1, \ldots, K$.

In matrix form, we can write this as

$$\tilde{Q}_{t,t+1} - I = \Pi(t)[Q - I]$$

where I is the $K \times K$ identity matrix and $\Pi(t) = \text{diag}(\pi_1(t), \ldots, \pi_{K-1}(t), 1)$ is a $K \times K$ diagonal matrix.

The last row in the transition matrix for $\tilde{Q}_{t,t+1}$ in conjunction with Equation (6) implies that $\pi_K(t) \equiv 1$ for all t. The proportionality adjustments $\pi_i(t)$ have the interpretation of being risk premiums. These transform the actual probabilities to those used in valuation. For later use, it is important to note that the restrictions imposed on the martingale probabilities $\tilde{q}_{ij}(t, t+1)$ in Equation (5) imply that $\pi_i(t) \geq 0$ for all i and t.

This assumption is imposed to facilitate statistical estimation since the historical transition matrix Q can be utilized in the inference process. The actual parameterization estimated is based on the continuous time limit, and it is discussed in the next section. The implication of this assumption (combined with the previous structure) on the evolution of credit spreads is discussed later in this section.

Given this structure, we can now compute the probability of default occurring after date T, that is, $\tilde{Q}_t(\tau^* > T)$. Let $\tilde{q}_{ij}(0, n)$ denote the n-step transition probabilities of going from state i at time 0 to state j at time n. It is well known that the n-step $K \times K$ transition matrix, $\tilde{Q}_{0,n}$, whose (i, j)th entry is $\tilde{q}_{ij}(0, n)$, satisfies

$$\tilde{Q}_{0,n} = \tilde{Q}_{0,1}\tilde{Q}_{1,2}, \ldots, \tilde{Q}_{n-1,n}. \tag{7}$$

Lemma 1 (Probability of Solvency in Terms of $\tilde{Q}$). *Let the firm be in state i at time t, denoted by $\eta_t = i$ and define $\tau^* = \inf\{s \geq t \colon \eta_s = K\}$, which represents the first time of bankruptcy. Then, the probability that default occurs after time T is*

$$\tilde{Q}_t^i(\tau^* > T) = \sum_{j \neq K} \tilde{q}_{ij}(t, T) = 1 - \tilde{q}_{iK}(t, T).$$

Proof. Since K is absorbing, the event $(\tau* > T)$ is equivalent to η_t not being in state K at time T, when starting in state i at time t. Using the $\tilde{Q}_{t,T}$ matrix gives the result. ∎

In Lemma 1 we make explicit the credit class dependence via a superscript on the conditional probability $\tilde{Q}_t^i$. Lemma 1 shows how to compute the probability of solvency at any future date T, starting from credit class i at time t. This solution requires the computation of an n-fold matrix product.[8]

We can use these characterizations of the probability of solvency to rewrite the risky zero-coupon bond's valuation. Let $v^i(t, T)$ be the value of a zero-coupon bond issued by a firm in credit class i at time t. Then[9]

$$v^i(t, T) = p(t, T)(\delta + (1 - \delta)\tilde{Q}_t^i(\tau^* > T)), \tag{8}$$

where $\tilde{Q}_t^i(\tau^* > T)$ is obtainable from Lemma 1.

Equation (8) provides a characterization of the credit risk spread. Given that the forward rate for the risky zero-coupon bond in credit class i is defined by

$$f^i(t, T) \equiv -\log(v^i(t, T+1)/v^i(t, T)),$$

Equation (8) yields

$$f^i(t, T) = f(t, T) + 1_{\{\tau^* > t\}} \log\left(\frac{[\delta + (1-\delta)\tilde{Q}_t^i(\tau^* > T)]}{[\delta + (1-\delta)\tilde{Q}_t^i(\tau^* > T+1)]}\right). \tag{9}$$

This gives an explicit representation for the credit risk spread $(f^i(t, T) - f(t, T))$ in terms of the recovery rate (δ) and the transition matrix for credit classes $\tilde{Q}$ (as given in Lemma 1). Equation (9) shows that credit spreads are always strictly positive in this model, except in bankruptcy. In bankruptcy, as in the discussion after Equation (2), $f^K(t, T) = f(t, T)$. Equation (9) also shows that the previous assumptions, in conjunction, imply that the credit spreads $(f^i(t, T) - f(t, T))$ and $(f^i(t, T) - f^j(t - T))$ are constants, except for random changes in credit ratings. Although this implication appears contrary to some preliminary evidence [see Das and Tufano (1995)], its rejection awaits a more thorough empirical investigation.

To get the spot rate, set $T = t$ in Equation (9) and simplify

$$r^i(t) = r(t) + 1_{\{\tau^* > t\}} \log(1/[1 - (1-\delta)\tilde{q}_{iK}(t, t+1)]). \tag{10}$$

In bankruptcy, $r^K(t) = r(t)$.

---

[8] In the special case that $\Pi(t)$ is a constant matrix, independent of t, $\tilde{Q}_{0,n} = \tilde{Q}^n$, where $\tilde{Q}^n$ is the n-fold matrix product. Then, $\tilde{Q}_t^i(\tau^* > T) = \tilde{Q}_0^i(\tau^* > T - t)$, as the process is time homogeneous.

[9] We can rewrite Equation (8) as

$$v^i(t, T) = p(t, T)[1 - (1-\delta)\tilde{Q}_t^i(\tau^* \leq T)].$$

The right side is the value of a default-free bond less the present value of the loss if default occurs.

2.2 Options and hedging

For practical applications, hedging jumps in credit ratings is essential. This section discusses how to hedge such jumps in both risky bonds and in (vulnerable) options on the term structure of credit risk spreads. To do so we must first extend our valuation methodology to apply to (vulnerable) options on the term structure of credit risk spreads. Afterwards, we analyze hedging these jumps in credit ratings.

As shown in Jarrow and Turnbull (1995), valuing (vulnerable) options on the term structure of credit risk spreads is a straightforward application of the martingale pricing technology. Given the random payoff to a credit-risky claim, say C_T at time T, its value at time t, denoted C_t, is

$$C_t = \tilde{E}_t(C_T/B(T))B(t). \tag{11}$$

For example, a European call option with exercise date T and strike price X on the risky firm's zero-coupon bond, maturing at time $M \geq T$, would have the random payoff $C_T = \max[v(T, M) - X, 0]$, with time t value given by Equation (11).

Valuation is thus transformed into an expected value calculation. Utilizing Lemma 1, simple formulas for risky coupon bonds, futures on risky bonds, options on risky debt, and vulnerable options can be obtained (via direct substitution). Expanding the traded securities to include common stock, this method also allows us to price equity options and convertible debt under the above scenario [see Jarrow and Turnbull (1995) for details].

To hedge credit rating changes (including bankruptcy), as shown in Jarrow and Turnbull (1995), the standard option hedging techniques are applicable. Sufficient information to apply these techniques is a careful description of the evolution of forward rates for the risky and riskless zero-coupon bonds.

Define the function

$$\phi(t, T, i) = \begin{cases} \log\left(\dfrac{[\delta+(1-\delta)\tilde{Q}_t^i(\tau^* > T)]}{[\delta+(1-\delta)\tilde{Q}_t^i(\tau^* > T+1)]}\right) & \text{for } i = 1, \ldots, K-1, \\ 0 & \text{for } i = K. \end{cases} \tag{12}$$

The change in the firm's forward rate over $[t, t+1]$ is easily deduced from Equation (9). Let the firm's time t credit rating be $\eta_t = i$, then

$$\begin{aligned} f^{\eta_{t+1}}(t+1, T) - f^i(t, T) &= [f(t+1, T) - f(t, T)] \\ &\quad + [\phi(t+1, T, \eta_{t+1}) - \phi(t, T, i)] \end{aligned} \tag{13}$$

for $\eta_{t+1} \in \{1, 2, \ldots, K\}$ with pseudo-probabilities $\tilde{q}_{i\eta_{t+1}}(t, t+1)$. The first component of the change in the risky firm's forward rates is due to changes in the default-free forward rate structure and a shortening of time to maturity. Any desired term structure model can be applied [e.g., Cox, Ingersoll, and Ross (1985) or Heath, Jarrow, and Morton (1992)]. This component of the risk is hedged in the same manner as employed in the literature on interest rate options.

The second component of the change in the risky firm's forward rates is due to changes in the default probability arising from an unpredictable change in credit class and a (predictable) shortening of the time to maturity. This risk has at most K different outcomes (if $\tilde{q}_{i\eta_{t+1}}(t, t+1) > 0$ for all possible η_{t+1}). To hedge this credit class risk, in general, one needs at most K of this firm's credit risky securities. As the entire term structure of the risky firm's zero-coupon bonds trade, enough securities are available to implement this trading strategy. The hedging procedure is analogous to that used in any multinomial (but Markov) process, and a description of this technique can be found in Jarrow and Turnbull (1995).

2.3 Fitting the credit class zero-curves

Given estimates of the empirical transition matrix Q, and the risk premium $(\pi_1(t), \ldots, \pi_{K-1}(t)$ for $0 \le t \le \tau - 1)$, Equation (8) provides a theoretical pricing formula for risky zero-coupon bonds. This formula can be utilized to identify arbitrage opportunities across the credit-risky term structures.

Alternatively, this methodology can be utilized to price and hedge options on the risky debt (Section 2.2). This section presents a recursive procedure for selecting the risk premium $(\pi_1(t), \ldots, \pi_{K-1}(t)$ for all $t)$ such that the theoretical pricing formula exactly matches a given initial term structure of credit-risky zeros (or credit class spreads).

From Equation (8), given estimates of the risky zero coupon bonds $(v^i(t, T))$, the default-free zeros $(p(t, T))$, and the recovery rate (δ), the initial (time 0) credit class zero-curves will be matched if $\tilde{Q}_0^i(\tau^* \le T)$ is selected such that

$$\tilde{Q}_0^i(\tau^* \le T) = [p(0, T) - v^i(0, T)]/[p(0, T)(1 - \delta)] \qquad (14)$$
$$\text{for } i = 1, \ldots, K \text{ and } T = 1, 2, \ldots, \tau.$$

We next show how to select $(\Pi(t)\colon t \in 0, 1, \ldots, \tau - 1)$ such that Equation (14) is satisfied. The procedure is recursive.

Given the empirical transition matrix Q, we have

$$\tilde{Q}_{0,1} = I + \Pi(0)(Q - I).$$

From this matrix, and Lemma 1, we get

$$\tilde{Q}_0^i(\tau^* \le 1) = \pi_i(0) q_{iK}.$$

Substitution into Equation (14) and algebra gives

$$\pi_i(0) = [p(0,1) - v^i(0,1)]/p(0,1)(1-\delta)q_{iK} \text{ for } i = 1, \ldots, K-1. \quad (15)$$

Given that we have computed these risk premium, we can now return to Equation (6) and compute $\tilde{Q}_{0,1}$.

Given the empirical transition matrix Q and $\tilde{Q}_{0,t}$ (calculated from the previous step), we have

$$\tilde{Q}_{0,t+1} = \tilde{Q}_{0,t}[I + \Pi(t)(Q - I)].$$

From this matrix, and Lemma 1, we get

$$\tilde{Q}_0^i(\tau^* \leq t+1) = \sum_{j=1}^{K} \tilde{q}_{ij}(0,t)\pi_j(t)q_{jK}.$$

In matrix form, after substitution into Equation (14) and algebra gives

$$\tilde{Q}_{0,t}\begin{bmatrix} \pi_1(t)q_{1K} \\ \vdots \\ \pi_K(t)q_{KK} \end{bmatrix} = \begin{bmatrix} [p(0,t+1) - v^1(0,t+1)]/p(0,t+1)(1-\delta) \\ \vdots \\ [p(0,t+1) - v^K(0,t+1)]/p(0,t+1)(1-\delta) \end{bmatrix}.$$

Assuming that $\tilde{Q}_{0,t}^{-1}$ exists (denote its entries by $\tilde{q}_{ij}^{-1}(0,t)$), the solution to the matrix equation is

$$\pi_i(t) = \sum_{j=1}^{K} \tilde{q}_{ij}^{-1}(0,t)[p(0,t+1) - v^i(0,t+1)]/[p(0,t+1)(1-\delta)q_{iK}]. \quad (16)$$

Given these risk premium, we can now compute $\tilde{Q}_{0,t+1}$ via Equation (6).

In practical applications one should, of course, check that the resulting risk premium is nonnegative and the resulting matrix is indeed a probability transition matrix. If they are not, this indicates an inconsistency of the data with the model. An illustrative computation in Section 3.4.4 clarifies this issue and provides an alternative estimation procedure when an inconsistency is observed.

This completes the recursive procedure. Equations (15) and (16) will also prove useful for estimating the parameters of the continuous time Markov chain model introduced in subsequent sections.

2.4 Discussion

This completes our analysis of the discrete time case. Some additional results can now be derived with respect to the shape and behavior of credit class spreads, but to avoid redundancy, these are only provided for the continuous time case. The purpose of this section was to introduce the fundamental insights in a simple fashion and to introduce a structure suitable for computer implementation. Having accomplished this task, we quickly move on to the continuous time analysis.

3. Credit Ratings and Default Probabilities: The Continuous Time Case

This section models the distribution for default time in a continuous trading economy. The continuous time setting facilitates the derivation of additional insights due to the availability of stochastic calculus. It also provides the parameterization of the bankruptcy process most suitable for estimation.

3.1 Valuation

The distribution for default time is modeled via a continuous time, time-homogeneous Markov chain on a finite state space $S = \{1, 2, \ldots, K\}$. As in the discrete time setting, these states represent the various credit classes, with state 1 being the highest and state K being bankruptcy.

A continuous time, time-homogeneous Markov chain $\{\eta_t \colon 0 \leq t \leq \tau\}$ is specified in terms of its $K \times K$ generator matrix

$$\Lambda = \begin{pmatrix} \lambda_1 & \lambda_{12} & \lambda_{13} & \ldots \lambda_{1,K-1} & \lambda_{1K} \\ \lambda_{21} & \lambda_2 & \lambda_{23} & \ldots \lambda_{2,K-1} & \lambda_{2K} \\ \vdots & & & & \\ \lambda_{K-1,1} & \lambda_{K-1,2} & \lambda_{K-1,2} & \ldots \lambda_{K-1} & \lambda_{K-1,K} \\ 0 & 0 & 0 & \ldots 0 & 0 \end{pmatrix} \tag{17}$$

where $\lambda_{ij} \geq 0$ for all i, j, and

$$\lambda_i = -\sum_{\substack{j=1 \\ j \neq i}}^{K} \lambda_{ij} \text{ for } i = 1, \ldots, K.$$

The off-diagonal terms of the generator matrix λ_{ij} represent the transition rates of jumping from credit class i to credit class j. The last row of zeros implies that bankruptcy (state K) is absorbing.

The $K \times K$ t-period probability transition matrix for η is given by[10]

$$Q(t) = \exp(t\Lambda) = \sum_{k=0}^{\infty} (t\Lambda)^k / k!, \tag{18}$$

[see Karlin and Taylor (1975, p. 152)]. For the discrete time approximation to this system, let $[0, \tau]$ be divided into τ steps of equal length. Then the one-step transition matrix of Equation (4) is given by $Q(1)$ via Equation (18). Denote the (i, j)th entry of the $Q(t)$ matrix by $q_{ij}(t)$.

By construction, state 1 is viewed as the best credit rating and $K-1$ is the worst credit rating prior to the state of default. To be sure that the Markov chain used to model credit rating changes reflects the fact that "lower credit classes are riskier," there exists a simple condition on the generator matrix that one can check.

Lemma 2 (Credit Ratings Versus Risk). *The following statements are equivalent:*

1. $\sum_{j \geq k} q_{ij}(t)$ *is a nondecreasing function of i for evey fixed k and* t.
2. $\sum_{j \geq k} \lambda_{ij} \leq \sum_{j \geq k} \lambda_{i+1,j}$ *for all i, k such that* $k \neq i + 1$.

Proof. This follows by combining Proposition 7.3.2, Theorem 7.3.4, and the remark on p. 251 in Anderson (1991). ■

Note that a generator matrix, in which the $(K-1)\times(K-1)$ submatrix of Λ has the structure of a birth-death chain and where the default intensities in the Kth column do not decrease as a function of row number, will satisfy the conditions of this lemma.

We assume that the generator matrix under the equivalent martingale probability is given by

$$\tilde{\Lambda}(t) \equiv U(t)\Lambda, \tag{19}$$

where $U(t) = \mathrm{diag}(\mu_1(t), \ldots, \mu_{K-1}(t), 1)$ is a $K \times K$ diagonal matrix whose first $K - 1$ entries are strictly positive deterministic functions of t that satisfy

$$\int_0^T \mu_i(t)\, dt < +\infty \text{ for } i = 1, \ldots, K-1.$$

[10] To get Equation (18) as the limit of the discrete time case, first divide 1 into n equal periods of length $(1/n)$. Define $[I + \Lambda/n]$ to be the transition matrix over each subperiod. Then, $Q = [I + \Lambda/n]^n$. As $n \to \infty$, $Q \to \exp(\Lambda)$.

For the more general case, the U matrix could be both history and time dependent. This is the analogous assumption to that preceding Equation (6) in the discrete time case. The entries $(\mu_1(t), \ldots, \mu_{K-1}(t), 1)$ are interpreted as risk premia, that is, the adjustments for risk that transform the actual probabilities into the pseudo-probabilities suitable for valuation purposes.

The $K \times K$ probability transition matrix from time t to time T for η under the equivalent martingale measure is given as the solution to the Kolmogorov differential equations [see Cox and Miller (1972, p. 181)]:

$$\frac{\partial \tilde{Q}(t, T)}{\partial t} = -\tilde{\Lambda}(t)\tilde{Q}(t, T) \text{ and} \tag{20a}$$

$$\frac{\partial \tilde{Q}(t, T)}{\partial T} = \tilde{Q}(t, T)\tilde{\Lambda}(T) \text{ with the initial condition} \tag{20b}$$

$$\tilde{Q}(t, t) = I. \tag{20c}$$

Let the (i, j)th entry of the $\tilde{Q}(t, T)$ matrix be denoted $\tilde{q}_{ij}(t, T)$.

Under the assumption in Equation (19), the credit rating process is still Markov, but it is now inhomogeneous. In the selection of this restriction, we faced a trade-off between computability of the transition matrix $\tilde{Q}(t, T)$ and the ability to match any given initial term structure of credit risk spreads. Equation (19) was our compromise. A special case illustrates this reasoning. When

$$\tilde{\Lambda} = \operatorname{diag}(\mu_1, \ldots, \mu_{K-1}, 1)\Lambda$$

for strictly positive constants $\mu_1, \ldots, \mu_{K-1}$, the Markov process is time homogeneous. Here, the solution to Equation (20) is easily computed as

$$\tilde{Q}(t, T) = \exp(\operatorname{diag}(\mu_1, \ldots, \mu_{K-1}, 1)\Lambda(T - t)).$$

But this restriction does not enable the model to match any given initial term structure of credit risk spreads. Equation (19) is the simplest extension, which satisfies both requirements.

Analogous to Lemma 1 for the discrete time case we have

Lemma 3 (Probability of Solvency in Terms of $\tilde{Q}$). *Let the firm be in state i at time t, that is, $\eta_t = i$ and define $\tau^* = \inf\{s \geq t \colon \eta_s = K\}$. Then*

$$\tilde{Q}_t^i(\tau^* > T) = \sum_{j \neq K} \tilde{q}_{ij}(t, T) = 1 - \tilde{q}_{iK}(t, T).$$

The proof of this lemma is identical to that used in Lemma 1.

Conditional survival probabilities can be obtained via the expression

$$\tilde{Q}_t^i(\tau^* > T+1 | \tau^* > T) = \tilde{Q}_t^i(\tau^* > T+1)/\tilde{Q}_t^i(\tau^* > T). \qquad (21)$$

Given this lemma we can value the risky firm's zero-coupon bonds as in Equation (8), that is,

$$v^i(t, T) = p(t, T)(\delta + (1-\delta)\tilde{Q}_t^i(\tau^* > T)), \qquad (22)$$

where $\tilde{Q}_t^i(\tau^* > T)$ is obtainable from Lemma 3.

Given that the forward rate for the risky zero-coupon bond in credit class i is defined by $f^i(t, T) \equiv \frac{-\partial}{\partial T} \log v^i(t, T)$, Equation (22) yields

$$f^i(t, T) = f(t, T) - 1_{\{\tau^* > t\}} \left[\frac{(1-\delta)\dfrac{\partial}{\partial T}\tilde{Q}_t^i(\tau^* > T)}{\delta + (1-\delta)\tilde{Q}_t^i(\tau^* > T)} \right]. \qquad (23)$$

In bankruptcy, $f^i(t, T) = f(t, T)$. We define the credit risk spread to be $[f^i(t, T) - f(t, T)]$.

To get the spot rate, let $T \to t$ in Equation (23). One obtains

$$r^i(t) = r(t) + (1-\delta)\lambda_{iK}\mu_i(t) 1_{\{\tau^* > t\}}, \qquad (24)$$

since $\partial \tilde{Q}_t^i(\tau^* > T)/\partial T = -\partial \tilde{Q}_t^i(\tau^* \leq T)/\partial T = -\lambda_{iK}\mu_i(t)$. In bankruptcy, $r_i(t) = r(t)$.

The spot rate on risky debt is seen to exceed the default-free spot rate by a credit risk spread of $(1-\delta)\lambda_{iK}\mu_i(t)$, where δ is the recover rate and $\lambda_{iK}\mu_i(t)$ is the pseudo-probability of default. This expression is analogous to that given in the discrete time case [Equation (10)], and it can also be found in Jarrow and Turnbull (1995).

3.2 Options and hedging

Analogous to the discrete time case, option valuation is a straightforward application of the martingale pricing technology and corresponds to computing an expectation as in Equation (11). Consequently we need not repeat that discussion.

Hedging credit rating changes is also analogous to the discrete time case. Similar to that case, given a specification of the stochastic process for the risky forward rates, hedging follows the standard procedures [details are in Jarrow and Turnbull (1995)]. All that remains to be done, then, is to derive the stochastic process for changes in forward rates. This is the content of this section.

Let $N_{ij}(t)$ where $i = 1, \ldots, K-1$ and $j = 1, \ldots, K$ be independent Poisson processes under the probability measure Q with intensities

λ_{ij}. The counting process $N_{ij}(t)$ represents a credit class jump from state i to state j at time t if $N_{ij}(t) - N_{ij}(t-) = 1$ where $N_{ij}(t-) = \lim_{\varepsilon \to 0} N_{ij}(t-\varepsilon)$.

Under the martingale measure $\tilde{Q}$, $N_{ij}(t)$ has intensities $\lambda_{ij}\mu_i(t)$ for $i = 1, \ldots, K-1$ and $j = 1, \ldots, K$.

We can represent the Markov chain process $\{\eta_t\colon 0 \leq t \leq \tau\}$ as

$$\eta_t = \eta_0 + \sum_{i=1}^{K-1}\sum_{j=1}^{K}\int_0^t (j-i)1_{\{\eta_{t-}=i, \eta_t \neq K\}}\, dN_{ij}(t). \tag{25}$$

Define the function

$$\phi(t, T, i) = \begin{cases} \dfrac{-(1-\delta)\partial \tilde{Q}_t^i(\tau^* > T)/\partial T}{\delta + (1-\delta)\tilde{Q}_t^i(\tau^* > T)} & \text{for } i = 1, \ldots, K-1 \\ 0 & \text{for } i = K. \end{cases} \tag{26}$$

Then, letting $\eta_0 = i_0$ be the firm's credit rating at time 0, we have

$$\begin{aligned} f^{\eta_t}(t, T) - f^{i_0}(0, T) = {} & [f, (t, T) - f(0, T)] \\ & + \int_0^t \phi'(s, T, \eta_{s-})\, ds + \sum_{i=1}^{K-1}\sum_{j=1}^{K} \\ & \times \int_0^t (\phi(s, T, j) - \phi(s, T, i))1_{\{\eta_{s-}=i\}}\, dN_{ij}(s), \end{aligned} \tag{27}$$

where $\phi'(s, T, \eta_{s-}) = d\phi(s, T, \eta_s)/ds$.

The first component of the change in the risky debt's forward rates is due to changes in the default-free forward rates. Any desired default-free term structure model can be applied to hedge this risk as is done in the interest rate option literature [e.g., Heath, Jarrow, and Morton (1992)].

The second component is a smoothly varying term arising from the shortening of time to maturity. If the Markov chain is time homogeneous, this is always negative.

The third component of the change in the risky debt's forward rates is due to jumps in the credit rankings. At any time s in state $\eta_{s-} = i$, there are at most K jump processes needed to be hedged, that is, $N_{i1}(s), \ldots, N_{iK}(s)$. These can be hedged in the standard manner using at most K risky zero-coupon bonds (or any other traded claims subject to this risk). A description of these hedging procedures can be found in Jarrow and Turnbull (1995).

3.3 Examples

This section illustrates some important insights related to credit risk spreads via three examples. The first two examples illustrate the importance of taking into account the entire structure of the generator matrix when calculating default probabilities and hence credit risk spreads. The third shows how to generate the model used in Jarrow and Turnbull (1995) as a special case of this article's analysis.

Example 1. Consider the generator matrix

$$\Lambda = \begin{pmatrix} -0.11 & 0.10 & 0.01 \\ 0.05 & -0.15 & 0.1 \\ 0 & 0 & 0 \end{pmatrix}$$

In this example, there are three possible states. Examining the first row, the probability of staying in the first credit class over a small period of time Δt is approximately $1 - 0.11\Delta t$. The transition rate from the first credit class to the second credit class is 0.10, and the rate of default from the first credit class is 0.01. One could then make the inappropriate estimate that a firm that is in this credit class will have a probability equal to $1 - \exp(-0.01t)$ of defaulting before time t since the holding time is exponential. This, of course, does not take into account the possibility of downgrading (occurring with rate 0.1) and subsequent default (occurring with rate 0.1 in the lower class). In Figure 1 we illustrate the difference by graphing the true default probability as a function of time (the top curve) and the one based on the rate of default directly from the highest credit rating.

It is important to estimate risks of downgrading and risks of default in lower rating classes to get an accurate estimate of default rates for the top-rated firms. In Figure 2 we show the spreads [see Equation (23)] under the assumption that $\delta = 0$, and it is clear that it is quite possible for the spread to be decreasing over time. This will typically happen to very low credit classes. Recall that the spread is really a survival contingent spread: it reflects forward rates obtained if no default occurs before the maturity date. Hence, if a low-rated firm has survived long, it may have a lower spread simply because survival increases the probability that it has reached a higher credit rating.

Example 2. Now consider the generator matrix

$$\Lambda = \begin{pmatrix} -0.13 & 0.10 & 0.02 & 0.01 \\ 0.05 & -0.16 & 0.1 & 0.01 \\ 0 & 0.05 & -0.15 & 0.1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

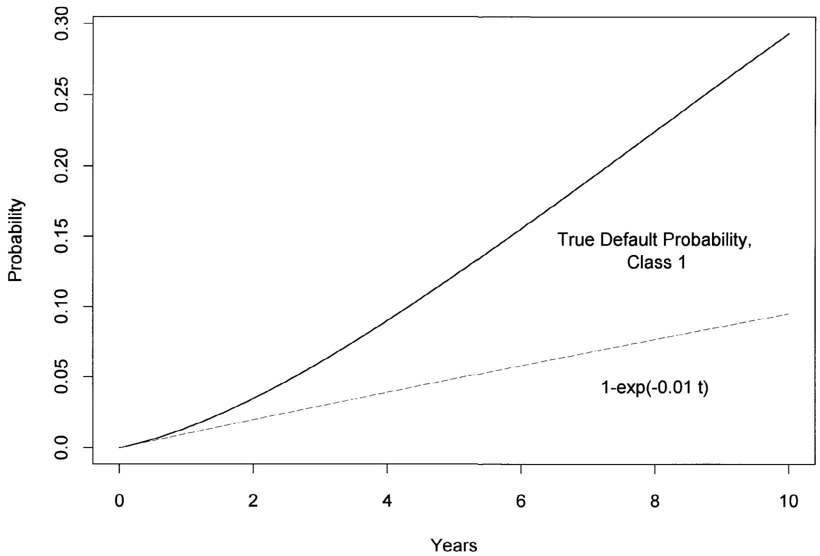

Figure 1
A comparison of the true default probability as a function of time for credit class 1 and an exponential distribution based on the one-step default probability for credit class 1. This comparison is for Example 1.

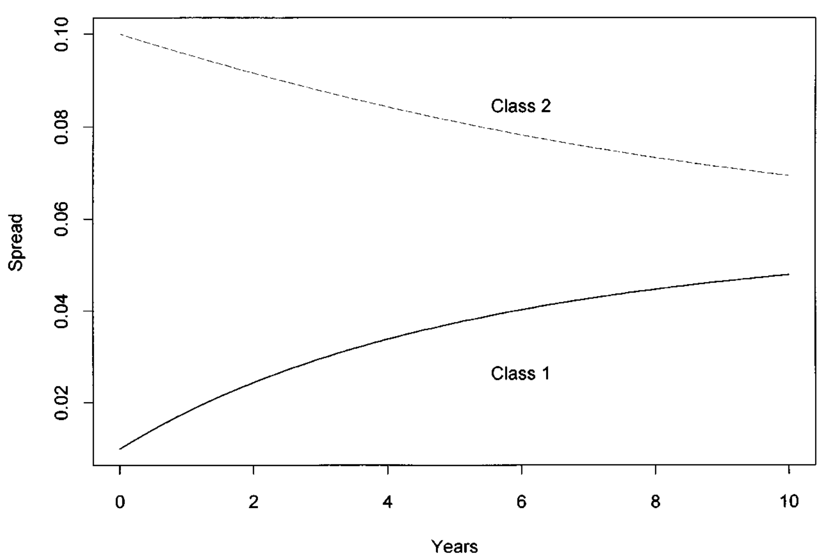

Figure 2
The spreads (forward rate credit class i less the default-free forward rate) as a function of maturity are graphed for credit classes 1 and 2. These spreads are for Example 1. A zero recovery rate is assumed.

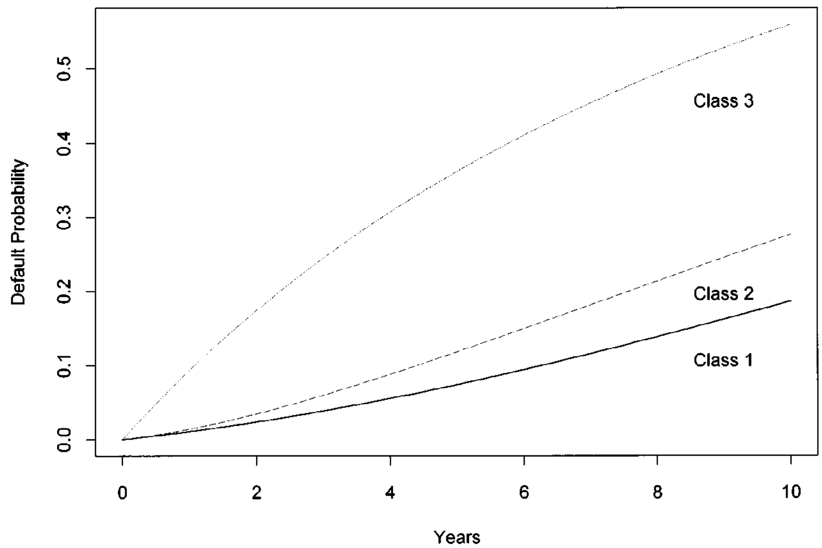

Figure 3
Default probabilities as a function of time for credit classes 1, 2, and 3. These default probabilities are for Example 2.

In this example there are four possible states. From the first row, the rate of default is 0.01. Note in this example that the two top credit classes have the same risk of direct default (the rate is 0.01 for direct default), but as Figure 3 illustrates, the probability of default for the second class (middle graph) is higher than that of the first class, simply because the second class is closer to the risky third class, whose high default probabilities are the top graph in Figure 3.

Example 3. If we change the last row of the generator matrix in Example 2 so that the rate of default is 0.01 (and change the diagonal element correspondingly to -0.06), we see that the default probabilities, and therefore the spread structure, becomes independent of credit class. The lower graph of Figure 1 illustrates the default probability as a function of time (an exponential distribution). It is equivalent to a bankruptcy process independent of credit classes, which is the model analyzed in Jarrow and Turnbull (1995).

3.4 Parameter estimation

To utilize this model to price and hedge credit-risky bonds and options, we need to estimate the parameters of the stochastic process given in Equation (23). This estimation procedure can be decomposed into two parts. The first part estimates the parameters generating the default-free forward rates, $f(t, T)$; the second part estimates the parameters generating the credit risk spread [the second term in Equation (23)].

3.4.1 Estimation of default-free parameters. The parameter estimation problem for the stochastic process generating default-free forward rates is a well-studied issue in the realm of interest rate options [see Amin and Morton (1993) and Heath et al. (1992)]. Consequently we refer the reader to that literature for the relevant insights.

3.4.2 Estimation of the bankruptcy process parameters. The parameter estimation problem for the stochastic process generating the credit risk spread involves estimating the recovery rate (δ), which can depend on the seniority of the debt, and estimating the generator matrix $\tilde{\Lambda}$ defined in Equation (19).

3.4.3 Estimating the recovery rate. An estimate of the recovery rate (δ) can be obtained historically from past defaults, or implicitly via market prices.

Table 1 reproduces the historical recovery rates from Moody's Special Report (1992). As can be seen, these recovery rates increase as the seniority of the debt increases. These numbers vary across time.

An implicit procedure generates an estimate for δ (given estimates for all the other parameters) by setting the theoretical value equal to a market price for some traded derivative security and then inverting (most often numerically). The implicit procedure could utilize zero-coupon bond prices [Equation (22)], spot rates [Equation (24)], coupon bonds, or options.

3.4.4 Estimating the generator matrix Λ. An estimate of the generator matrix $\tilde{\Lambda}$ can also be obtained either historically or implicitly. The implicit procedure estimates the generator matrix $\tilde{\Lambda}$ implicitly using market prices for the risky zero-coupon bonds. Implicit estimation of the generator matrix $\tilde{\Lambda}$ would be preferred to historical estimation if it is believed that the historical transition matrix does not represent the actual transition matrix. This could be due to either nonstationaries in the bankruptcy process or untimely changes in the credit ratings by the rating agency. For example, if credit rating changes lag the actual changes in the default probabilities, then the historical transition matrix of changes in credit ratings will not be the relevant probability matrix.

In an implicit estimation procedure for $\tilde{\Lambda}$, $(K-1) \times (K-1)$ parameters need to be estimated [see Equation (17)]. This is a large number

Table 1
Recovery prices and seniority of debt, 1974–1991

	1974–1991	1991	1990	1989	1988	1987	1986	1985	1984	
Senior secured	$67.13	$60.81	$33.72	$74.63	$76.29	$87.78	$64.30	$76.04	$59.02	
Senior unsecured	46.53	45.50	35.29	43.26	44.57	67.39	33.65	69.97	64.56	
Senior subordinated	32.10	33.68	26.47	40.27	39.85	47.73	28.84	32.87	45.69	
Subordinated	26.44	16.58	23.39	25.95	35.03	37.75	31.14	33.93	39.32	
Junior subordinated	17.95	3.63	18.86	17.50	55.00					
Weighted average	39.22	32.65	28.17	39.11	44.29	72.57	35.76	40.29	48.83	
	1983	1982	1981	1980	1979	1978	1977	1976	1975	1974
Senior secured		$72.50					$35.08		$32.84	
Senior unsecured	$67.73	38.98	$12.00	$44.02		$60.00	29.93		25.33	
Senior subordinated		44.19					40.50			
Subordinated	40.18	36.48	40.00	40.63	$34.50	67.00	35.00		$5.81	
Junior subordinated										
Weighted average	50.24	39.24	30.49	40.98	34.50	61.88	32.01		27.76	5.81

Reproduced from Moody's Special Report (1992).

of parameters to imply out from the data. We therefore suggest using a decomposition of the generator matrix under the martingale measure into a product of the empirical generator matrix, which may be readily estimated, and a low-dimensional vector of time-dependent risk premia. Prices observed in the markets and the empirical generator matrix (historical or predicted) may then be used to imply out the risk premia.

Alternative ways of simplifying the estimation of $\tilde{\Lambda}$ could be to restrict the number of strictly positive transition probabilities between states so that, for example, only jumps to the adjacent state can occur. This will reduce the number of parameters to $2(K-1)$. Under this restriction it is also possible to relax the assumption on the risk premia so that $\pi_i(t)$ becomes $\pi_{ij}(t)$. Another approach would be to reduce the number of states to three classes: investment grade, speculative grade, and default. However, if the true process of ratings with the K-dimensional state space is Markovian, it will typically not be the case that this process on the smaller state space is Markovian. Applying a function that is not one to one to a Markov chain destroys the Markov property.

We now concentrate on the historical approach. Our estimation procedure is based on the zero-coupon bond prices and utilizes the decomposition of the generator matrix $\tilde{\Lambda}$ into the product of the empirical generator matrix Λ and the risk premia $(\mu_1(t), \ldots, \mu_{K-1}(t))$ [see Equation (19)].

3.4.5 Estimation of the empirical generator matrix Λ. Estimates of the empirical generator matrix Λ can be obtained from past observations of credit rating changes [see Moody's Special Report (1992) or Standard & Poor's Credit Review (1993)]. From a statistical viewpoint, the Markov chain model we have specified is very convenient. If we observe over a time period $[0, T]$ the exact times of credit class transitions and defaults, a classical estimator of an off-diagonal element λ_{ij} of the generator matrix is given by

$$\lambda_{ij} = \frac{N_{ij}(T)}{\int_0^T Y_i(s)\, ds},$$

where

$N_{ij}(T) =$ total number of transitions from i to j over the period $[0, T]$

and

$$Y_i(s) = \text{ number of firms in class } i \text{ at time } s.$$

For a derivation of this result, and for extensions to estimation of transition rates that are time dependent, have external covariates, or are based on censored observations, see Andersen and Borgan (1985).

For this article we use data reported in Standard and Poor's Credit Review (1993). These data do not include the exact timing of transitions, but they do include estimates of 1-year transition probabilities that are obtained by observing the credit ratings of a fixed group of firms at the beginning and at the end of the year. We use these estimates to obtain an approximation to the generator matrix, and this, in turn, will be used to obtain transition probabilities $\hat{Q}(t)$ for every t.

Given the estimate of $\hat{Q}(1)$, we can estimate the generator matrix by assuming that the probability of making more than one transition per year is small. In other words, every firm is assumed to have made either zero or one transition throughout the year. Under this hypothesis, it can be shown given $\lambda_i \neq 0$ for $i = 1, \ldots, K-1$ that[11]

$$\exp(\Lambda) \approx \begin{bmatrix} e^{\lambda_1} & \frac{\lambda_{12}(e^{\lambda_1}-1)}{\lambda_1} & \cdots & \frac{\lambda_{1K}(e^{\lambda_1}-1)}{\lambda_1} \\ \frac{\lambda_{21}(e^{\lambda_2}-1)}{\lambda_2} & e^{\lambda_2} & \cdots & \frac{\lambda_{2K}(e^{\lambda_2}-1)}{\lambda_2} \\ \frac{\lambda_{K-1,1}(e^{\lambda_{K-1}}-1)}{\lambda_{K-1}} & \frac{\lambda_{K-1,2}(e^{\lambda_{K-1}}-1)}{\lambda_{K-1}} & \cdots & \frac{\lambda_{K-1,K}(e^{\lambda_{K-1}}-1)}{\lambda_{K-1}} \\ 0 & 0 & \cdots & 1 \end{bmatrix}. \tag{28}$$

To obtain our estimates of $\hat{\Lambda}$, we set $\hat{Q}(1)$ equal to the right side of Equation (28) to obtain[12]

$$\hat{q}_{ii} = e^{\hat{\lambda}_i} \text{ for } i = 1, \ldots, K-1 \text{ and} \tag{29a}$$

$$\hat{q}_{ij} = \hat{\lambda}_{ij}(e^{\hat{\lambda}_i} - 1)/\hat{\lambda}_i \text{ for } i \neq j \text{ and } i, j = 1, \ldots, K-1. \tag{29b}$$

The solution to this system is

$$\hat{\lambda}_i = \log(\hat{q}_{ii}) \text{ for } i = 1, \ldots, K-1 \text{ and} \tag{30a}$$

$$\hat{\lambda}_{ij} = \hat{q}_{ij} \log(\hat{q}_{ii})/[\hat{q}_{ii} - 1] \text{ for } i \neq j \text{ and } i, j, \ldots, K-1. \tag{30b}$$

A 1-year transition matrix is given in Standard and Poor's Credit Review (1993) and is reproduced in Table 2 in the Appendix. It poses one problem for our analysis. Transitions to the "Not rated" (NR) cate-

[11] See the Appendix for a proof. Since $Q(1) = \exp(\Lambda) \approx I + \Lambda$, we could have used the approximation $\hat{\Lambda} = \hat{Q}(1) - I$, but we choose Equation (28) since it provided a better fit of $\exp(\hat{\Lambda})$ to $\hat{Q}(1)$ in our data.

[12] The notation $\hat{q}_{ij}$ should not be confused with the pseudo-probabilities $\tilde{q}_{ij}$ defined in the discrete time section.

Table 2
Average 1-year transition probabilities, 1981–1991

Initial rating	Rating at the end of year AAA	AA	A	BBB	BB	B	CCC	D	NR
AAA	0.8746	0.0945	0.0077	0.0019	0.0029	0.0000	0.0000	0.0000	0.0183
AA	0.0084	0.8787	0.0729	0.0097	0.0028	0.0028	0.0000	0.0000	0.0246
A	0.0009	0.0282	0.8605	0.0628	0.0098	0.0044	0.0000	0.0009	0.0324
BBB	0.0006	0.0041	0.0620	0.7968	0.0609	0.0151	0.0017	0.0043	0.0545
BB	0.0004	0.0020	0.0071	0.0649	0.7012	0.0942	0.0115	0.0218	0.0970
B	0.0000	0.0017	0.0027	0.0058	0.0451	0.7196	0.0380	0.0598	0.1272
CCC	0.0000	0.0000	0.0102	0.0102	0.0179	0.0665	0.5729	0.2046	0.1176

Standard and Poor's Credit Review (1993), Table 10.

Table 3
Modified transition probabilities

Initial rating	Rating at the end of year AAA	AA	A	BBB	BB	B	CCC	D
AAA	0.8910	0.0963	0.0078	0.0019	0.0030	0.0000	0.0000	0.0000
AA	0.0086	0.9010	0.0747	0.0099	0.0029	0.0029	0.0000	0.0000
A	0.0009	0.0291	0.8894	0.0649	0.0101	0.0045	0.0000	0.0009
BBB	0.0006	0.0043	0.0656	0.8427	0.0644	0.0160	0.0018	0.0045
BB	0.0004	0.0022	0.0079	0.0719	0.7764	0.1043	0.0127	0.0241
B	0.0000	0.0019	0.0031	0.0066	0.0517	0.8246	0.0435	0.0685
CCC	0.0000	0.0000	0.0116	0.0116	0.0203	0.0754	0.6493	0.2319
D	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000

No-rating category eliminated.

gory are listed, but no estimates of subsequent defaults or reentrance to the rating categories are reported.[13] We choose to eliminate this part of the sample by redefining the transition probability from state i to j (both states different from NR) as being

$$q_{ij} = \frac{\text{fraction of firms going to } j \text{ from } i}{\text{fraction of firms going to a state different from NR}}. \tag{31}$$

This is the modified transition matrix reported in Table 3. Using the estimation procedure described above we get an estimate of the generator matrix $\hat{\Lambda}$, reported in Table 4.

3.4.6 Estimation of the risk premium. Estimates of the risk premia $(\mu_1(t), \ldots, \mu_{K-1}(t))$ can be obtained from the zero-coupon bond prices [Equation (22)].

[13] As noted in the Standard & Poors Credit Review (1993), the majority of the transitions to an NR originate from issuers repaying outstanding debt or bringing it below a limit of $25 million, but some are caused by insufficient information being provided by the issuer and subsequent defaults do occur.

Table 4
Generator matrix estimated from modified transition probabilities in Table 3

Initial rating	Rating at the end of year							
	AAA	AA	A	BBB	BB	B	CCC	D
AAA	−0.1154	0.1019	0.0083	0.0020	0.0031	0.0000	0.0000	0.0000
AA	0.0091	−0.1043	0.0787	0.0105	0.0030	0.0030	0.0000	0.0000
A	0.0010	0.0309	−0.1172	0.0688	0.0107	0.0048	0.0000	0.0010
BBB	0.0007	0.0047	0.0713	−0.1711	0.0701	0.0174	0.0020	0.0049
BB	0.0005	0.0025	0.0089	0.0813	−0.2530	0.1181	0.0144	0.0273
B	0.0000	0.0021	0.0034	0.0073	0.0568	−0.1929	0.0479	0.0753
CCC	0.0000	0.0000	0.0142	0.0142	0.0250	0.0928	−0.4318	0.2856
D	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Using an analytic approximation to the forward equation [Equation (20b)], we get that over a small time period Δt[14]

$$\tilde{Q}(t, t+\Delta t) \approx I + \tilde{\Lambda}(t)\Delta t = I + U(t)\Lambda\Delta t. \tag{32}$$

Letting $\Pi(t) \equiv U(t)$ and $Q - I \equiv \Lambda\Delta t$ in Equations (15) and (16), yields an estimation procedure for the risk premia $U(t)$. $U(t)$ in this approximation is assumed to be right continuous over $[t, t+\Delta t)$. This estimation procedure will be such that the model prices given via Equation (22) will exactly match the initial credit-risky zero-coupon bond price curves.

3.4.7 Survival probabilities and spreads under risk neutrality. We can illustrate some main features of the model by looking at the risk-neutral case, that is, the case where the pricing measure and the empirical measure are the same.

From the estimated generator matrix, we have calculated the survival probabilities for firms of the various credit ratings as a function of time. In other words, if a firm is in a given credit class today, and if its rating evolves according to the Markov chain specified by the empirical generator matrix, what is the probability of having no defaults within the next t years? The result is shown in Figure 4. Note that with the modified empirical transition matrix specified in Table 3, we could have obtained estimates for survival probabilities for integer years in the future by simply calculating powers of that matrix and reading off the entry in the default column. But we are interested in default probabilities for every time to maturity, and in particular for

[14] This approximation comes from the forward equation [Equation (20b)]. For small Δt,

$$\partial\tilde{Q}(t, T)/\partial T\Big|_{T=t+\Delta t} \approx [\tilde{Q}(t, t+\Delta t) - I]/\Delta t \text{ and } \tilde{Q}(t, t+\Delta t)\tilde{\Lambda}(t+\Delta t) \approx \tilde{Q}(t, t)\tilde{\Lambda}(t) = \tilde{\Lambda}(t).$$

Substitution into Equation (20b) gives Equation (32).

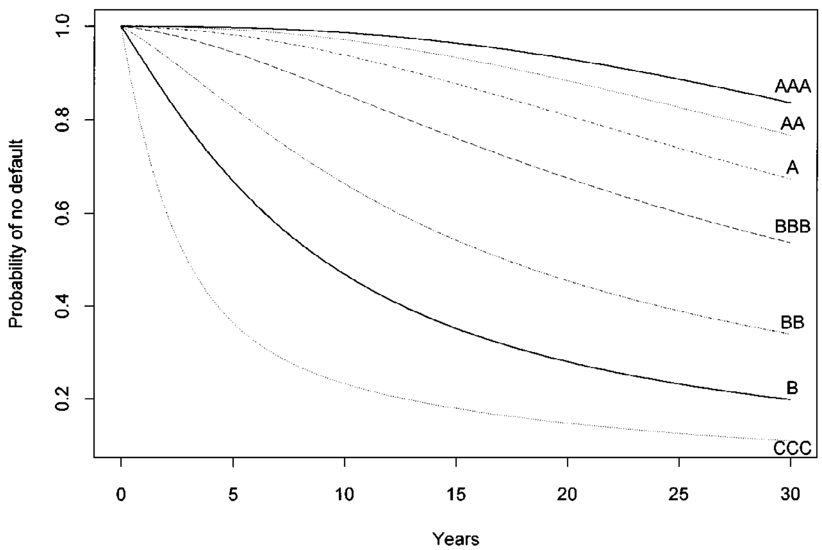

Figure 4
The survival probabilities (no defaults in the next t years) for the different credit classes as a function of time. These graphs are based on the generator matrix given in Table 4.

periods shorter than 1 year. Our estimate of the generator matrix provides us with estimates of these probabilities. As expected, the default probabilities for every fixed time are lower for the better rated firms.

In Figures 5 and 6 we have graphed forward spreads as defined in Equation (23) for investment grade and speculative grade bonds, respectively. We have set the recovery rate equal to zero in order to give the graphs an alternative interpretation as hazard rates: The hazard rate corresponding to credit class i at time t is the rate of default at time t for a firm that is in class i at time 0 and which has not defaulted up to time t. The spread at time 0 is the credit spread on the spot rate. If we set the recovery rate to $\delta > 0$, these spot rate credit spreads are multiplied by $1 - \delta$. For a value of δ equal to 0.67, this would bring the extremely high spreads on speculative rates down in a more reasonable range, but it would also lower the already very low spreads on the investment grade bonds.

3.4.8 An illustrative estimation of the risk premia. This section illustrates the computation of the risk premia $(\mu_1(t), \ldots, \mu_{K-1}(t))$ for $t = 0, 1, \ldots, \tau$ utilizing the procedure described above. This example demonstrates the feasibility of performing the described computations

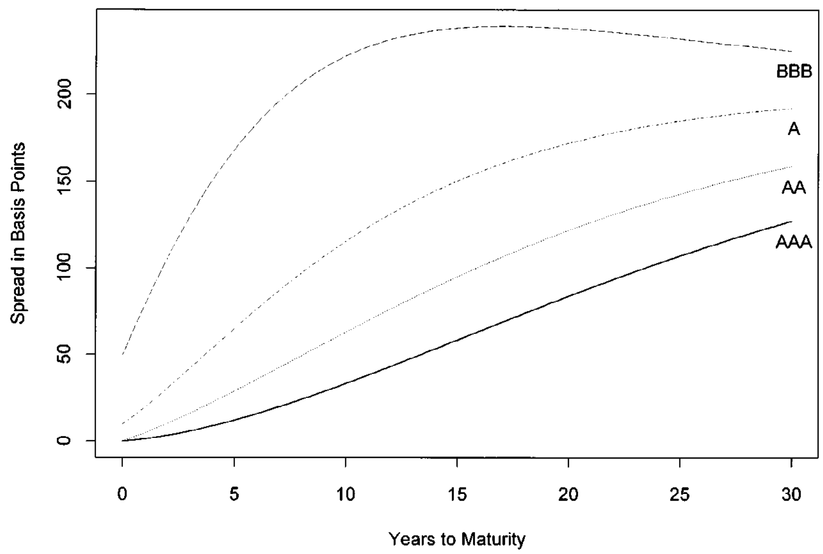

Figure 5
The spreads (forward rate credit class *i* less the default-free forward rate) in basis points as a function of maturity for credit classes AAA, AA, A, and BBB. These spreads assume risk neutrality and a zero recovery rate. They are based on the generator matrix given in Table 4.

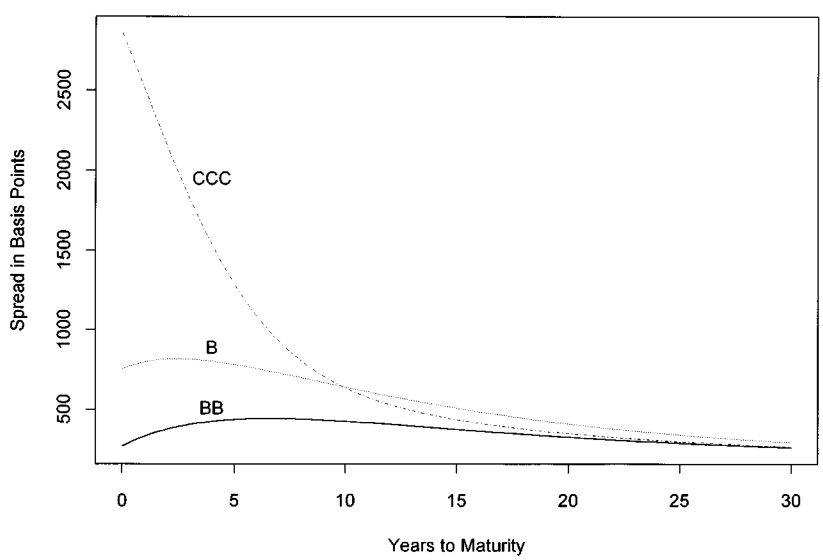

Figure 6
The spreads (forward rate credit class *i* less the default-free forward rate) in basis points as a function of maturity for credit classes BB, B, CCC. These spreads assume risk neutrality and a zero recovery rate. They are based on the generator matrix given in Table 4.

and provides some preliminary insights useful for designing a more thorough empirical investigation of the proposed model.

The computations were performed using bond index price data for Friday, December 31, 1993. The corporate bond price data was kindly provided by Lehman Brothers, and it is included in Table 5.

Table 5 contains a matrix whose columns correspond to various maturity ranges, and whose rows correspond to credit classes. Within each cell of the matrix is given (1) the number of bond issues represented in the cell, (2) the market value weighted coupon payment for the cell, and (3) the (bid) "yield to worst" (market value weighted) for the cell. The yield to worst is the bond's yield, except when a bond is callable. When callable, the yield to worst is the minimum of the following two numbers: (1) the yield computed based on the bond's maturity date, and (2) the yield computed based on the bond's first call date. Obviously this is an incomplete adjustment to incorporate a bond's call provision into the calculation of the bond's yield. As such it can introduce significant error into the following procedure. In essence the computation of yield to worst can place a callable bond's yield in the wrong maturity cell. This observation error is important to keep in mind when evaluating the subsequent results.

The first step in the computation of the risk premium from this corporate bond data is to strip out the risky zero-coupon bond prices. Our procedure for doing this is described in the Appendix. The results are provided in Table 6. Also included in Table 6 are default-free zero-coupon bond prices. These are the U.S. Treasury strip prices as reported in the *Wall Street Journal* for December 31, 1993. Figures 7 and 8 contain the graphs of these zero-coupon bond prices.

Occasional "mispricings" can be observed in these various zero-coupon bond price curves. Indeed, whenever the credit class term structures cross, potential arbitrage opportunities exist. For example, at the maturity of 1 year, the AA zero is priced above the AAA zero-coupon bond, and the B zero-coupon bond is priced above the BA-zero coupon bond. These crossings are inconsistent with the previous model structure. For the purposes of this section, we attribute these mispricings to noise in the data (most likely due to the yield-to-worst issue previously discussed). The noise in the data is particularly bothersome when the number of issues in a cell are small, such as in the case of the B-rated bonds in the first two maturity classes (only one issue each).

These mispricings will generate risk premium $(\mu_1(t), \ldots, \mu_{K-1}(t))$ for $t = 0, \ldots, \tau$ that do not satisfy the conditions imposed in Equation (6) [using $\Pi(t) = U(t)$ from Equation (32)], which guarantee that $\tilde{Q}$ is

Table 5
Lehman Brothers corporate bond index data for December 31, 1993

		Maturity								
Quality		0–1.75	1.75–2.25	2.75–3.25	3.25–4.75	4.75–5.25	5.25–6.5	6.5–7.5	9.5–10.5	10.5–18
AAA	1.	14	4	13	20	3	14	12	3	28
	2.	8.415	7.592	7.459	9.009	9.383	8.076	8.990	7.369	8.138
	3.	4.351	4.508	4.974	5.493	5.034	5.843	5.841	6.277	6.240
AA	1.	13	10	17	46	14	35	24	22	66
	2.	7.086	6.463	7.250	6.753	8.419	7.267	8.476	8.133	8.740
	3.	4.233	4.555	4.731	5.394	5.347	5.670	5.989	6.381	6.144
A	1.	70	58	61	153	39	139	81	71	179
	2.	7.791	7.289	7.879	7.898	8.012	7.726	9.041	6.743	8.191
	3.	4.286	4.689	4.944	5.403	5.523	5.899	6.187	6.480	6.719
BAA1	1.	22	13	37	87	29	92	51	56	89
	2.	8.357	7.628	8.409	8.585	8.635	8.734	9.095	7.961	8.847
	3.	4.870	5.006	5.181	5.944	6.182	6.514	6.701	7.102	7.037
BA	1.	2	1	7	15	8	36	22	28	58
	2.	10.975	9.450	8.885	8.808	11.194	9.288	7.593	9.272	9.636
	3.	7.291	6.023	6.224	7.494	7.433	7.531	9.155	8.455	8.579
B	1.	1	1	4	21	11	46	34	61	50
	2.	10.000	9.500	9.987	7.763	11.551	10.309	10.715	9.434	8.186
	3.	5.428	7.807	9.981	11.608	9.418	9.289	9.676	9.712	9.969
CAA	1.	1	0	1	2	2	3	3	5	1
	2.	15.000	0.000	11.375	2.475	9.886	11.976	13.129	2.823	11.000
	3.	8.571	0.000	11.170	15.085	15.210	13.425	11.806	10.276	10.407

Within each cell is listed:
1. Number of issues
2. Coupon market weighted
3. Yield to worst

For the BA, B, CAA quality, the 10.5–18 column is really 10.5+ with no terminal date provided.

Table 6
Stripped zero-coupon bond prices for December 31, 1993

	Maturity													
Quality	1.00	2.00	3.00	4.00	5.00	6.00	7.00	8.00	9.00	10.00	11.00	12.00	13.00	14.00
GOVT	96.969	92.656	87.844	82.719	77.719	73.000	68.250	63.781	60.176	56.570	52.965	49.359	45.754	42.148
AAA	95.830*	91.549	86.353*	80.420*	78.240*	70.427*	66.497	62.076	57.655	53.234	50.412	47.589*	44.766*	41.944*
AA	95.939	91.459	87.001	80.789	76.804	71.357	65.660	61.275	56.891	52.507	50.128	47.749	45.370	42.991
A	95.890	91.218	86.448	80.780	76.133	70.322	64.701	60.581	56.460	52.340	48.898	45.455	42.013	38.571
BAA1	95.356	90.684	85.894	79.033	73.619	67.742	62.570	58.060	53.550	49.040	46.194	43.347	40.500	37.654
BA	93.204	89.057	83.473	74.439	69.396	64.232	52.237	49.377	46.517	43.658	40.395	37.133	33.871	30.608
B	94.851*	85.060	74.525	63.482*	64.000	59.000	52.190	47.904	43.617	39.331	35.905	32.478*	29.052*	25.626*
CAA	92.106	82.266	72.426	56.553	47.551	46.663	47.626*	44.633	41.641	38.648	35.885	33.122	30.358	27.595

*Indicates a potential mispricing.

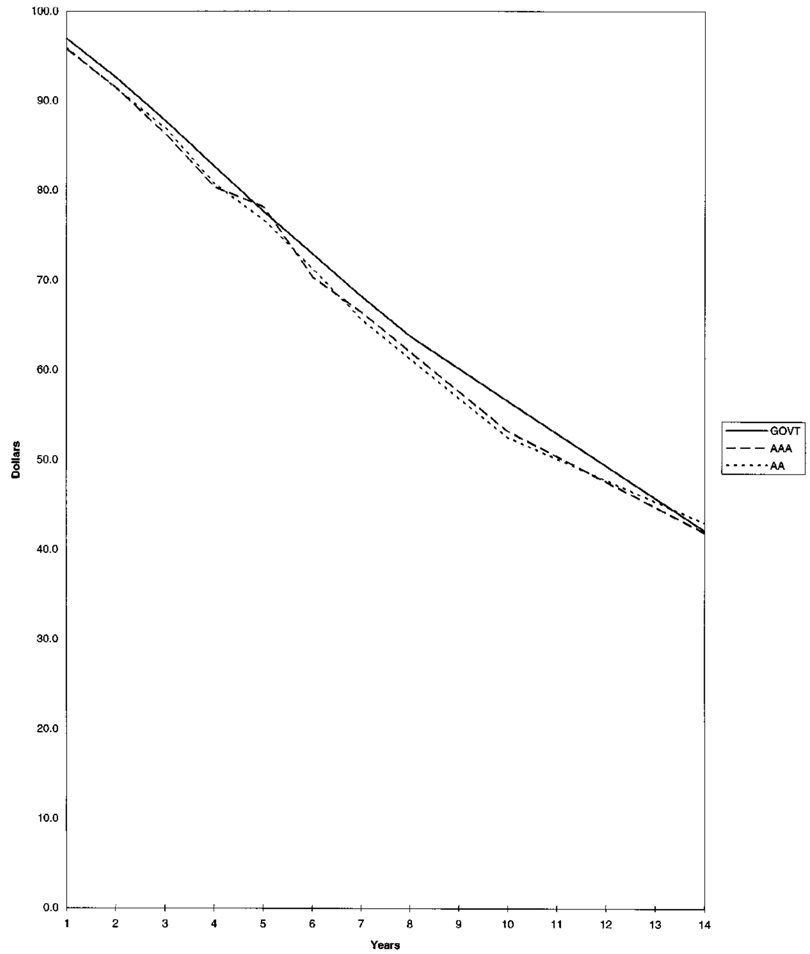

Figure 7
Stripped zero-coupon bond prices on December 31, 1993, versus time to maturity for Treasuries, AAA, and AA rated debt. A face value of $100 is assumed. Treasuries are ——, AAA are – – –, and AA are - - - - - -.

a probability matrix. The relevant condition is that the risk premium is nonnegative, $\mu_i(t) \geq 0$, for all credit classes i and times t.[15]

To illustrate this outcome we compute the risk premia using the (unconstrained) procedure as described in Section 2.3. Needed for this

[15] Strictly speaking, the condition is that $\mu_i(t) > 0$ for all credit classes and times (where the generator matrix has strictly positive values). However, imposing strict positivity will not allow us to solve the minimization problem. Therefore we do not impose this strict positivity condition (see footnote 16).

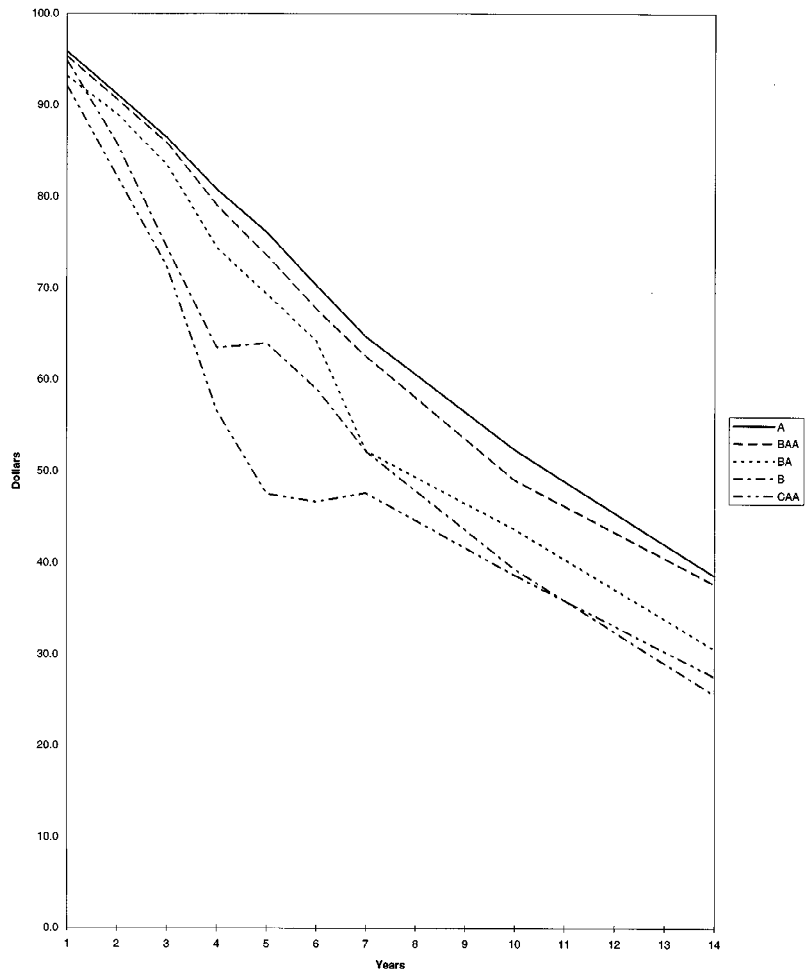

Figure 8
Stripped zero-coupon bond prices on December 31, 1993, versus time to maturity for A, BAA, BA, B, and CAA rated debt. A face value of the \$100 dollars is assumed. A are ——, BAA are – – –, BA are - - - - - -, B are – - – - –, and CAA – - - – - - –.

computation are an estimate of the recovery rate (δ) and an estimate of the Q matrix. We set $\hat{\delta} = 0.3265$, which is the weighted average value for 1991 as given in Table 1. We set $\hat{Q}$ equal to the generator matrix given in Table 4 [as required by Equation (32)], with one small modification. In order that the risk premia $\mu_i(0)$ for $i = 1, \ldots, K-1$ are well-defined for credit class i via Equation (15), we need to guarantee that the one-step default probability $q_{iK} \neq 0$. This is not satisfied by the default probabilities in the first two rows of Table 4. For this reason we replace the zero default probabilities q_{iK} for these two rows with 0.0001 (and reduce the diagonal element in row i by an equal amount). This is the smallest significant number possible within Table 4.

The results of this computation are given in Table 7. As noted, many of the risk premia are negative, indicating the existence of arbitrage opportunities, data errors, or model rejection. We believe these mispricings are most likely due to the noise introduced into the computation of the zero-coupon bond prices due to the yield-to-worst issue previously discussed.

Given that an unconstrained estimation procedure for $(\mu_1(t), \ldots, \mu_{K-1}(t))$ for $t = 0, \ldots, 14)$ indicates mispricings, we next compute the best values for the risk premia consistent with no arbitrage, where "best" is defined to mean minimizing the sum of squared errors of the actual prices from the model prices. This procedure uses all the data, including the mispriced zero-coupon bonds.

The constrained estimation technique follows a recursive procedure similar to that described in Section 2.3. We now briefly describe this procedure.[16]

At step 0, we find such that $(\mu_1(0), \ldots, \mu_{K-1}(0))$ such that $\sum_{i=1}^{K-1}[v^i(0,1) - v^i(0,1)^{\text{observed}}]^2$ is minimized subject to $\mu_i(0) \geq 0$ for $i = 1, \ldots, K-1$, where $v^i(0,1)$ is determined from Equation (3) and $\tilde{Q}_0^i(\tau^* > 1)$ in Equation (3) is determined as Equation (15) in the text.

Given the optimized values for $(\mu_1(0), \ldots, \mu_{K-1}(0))$, to prepare for the next step in the procedure, we compute the one-step transition matrix $\tilde{Q}_{0,1}$, using Equation (6).

At step $t+1$, we enter with $\tilde{Q}_{0,t}$. We compute $(\mu_1(t+1), \ldots, \mu_{K-1}(t+1))$ such that $\sum_{i=1}^{K-1}[v^i(0,t+1) - v^i(0,t+1)^{\text{observed}}]^2$ is minimized subject to $\mu_i(t+1) \geq 0$ for $i = 1, \ldots, K-1$, where $v^i(0, t+1)$ is determined from Equation (3) and $\tilde{Q}_0^i(\tau^* > t+1)$ in Equation (3) is determined as Equation (16) in the text. In particular,

$$\tilde{Q}_0^i(\tau^* > t+1) = 1 - \sum_{j=1}^{K} \tilde{q}_{ij}(0,t)\mu_j(t)q_{jk}.$$

Given the optimized values for $(\mu_1(t+1), \ldots, \mu_{K-1}(t+1))$, we then compute $\tilde{Q}_{0,t+1}$ using Equation (6) in order to prepare for the next step.

[16] To satisfy the strict positivity condition in footnote 15, one could add the restriction that $\mu_i(0) \geq 0.0001$ for all i. The estimates obtained in this case would be similar to those subsequently reported.

Table 7
Risk premia computed using an unconstrained procedure on December 31, 1993

	Maturity													
Quality	0.00	1.00	2.00	3.00	4.00	5.00	6.00	7.00	8.00	9.00	10.00	11.00	12.00	13.00
AAA	174.29	419.87	−261.93	−429.98	554.71	4745.25	−25111.87	269.69	649.39	1076.34	5.44	−1318.54	5.93	−64.38
AA	157.68	429.97	−255.96	−428.99	466.93	4744.94	−24906.09	267.70	646.62	1071.26	8.55	834.32	16.90	33.24
A	16.51	44.69	−30.55	−50.20	55.26	492.50	−2546.54	27.57	65.84	109.38	−0.22	−3.38	0.60	−0.41
BAA1	5.04	5.44	−7.74	−10.18	15.54	94.54	531.55	5.73	13.60	22.61	−0.23	−0.97	−0.11	−0.21
BA	2.11	−2.31	−0.05	2.40	14.62	7.40	−106.12	1.00	2.40	3.93	−0.03	−0.16	0.00	−0.02
B	0.43	1.10	2.02	2.15	−7.93	0.10	0.90	−0.06	0.24	0.26	0.59	0.79	1.07	1.41
CAA	0.26	0.36	0.41	1.12	1.64	−2.40	−3.34	0.01	0.12	0.14	−0.02	−0.03	−0.03	−0.02
D	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

The recursive procedure stops when $t + 1 = 14$.

The constrained estimates for the risk premia are contained in Table 8. The theoretical risky zero-coupon bond prices consistent with these risk premia are contained in Table 9. The standard error of the estimate and the percentage error (standard error/average price in the column) are provided under each column in Table 6. The percentage error is seen to be the smallest for the shortest maturity zeros, and increases as time to maturity increases. This is to be expected as pricing errors compound in the computing procedure. The dollar pricing error for maturities less than 4 years is no more than \$1.0826 per \$100 face value bond. For maturities exceeding 4 years, the pricing error is at most \$3.7908 per \$100 face value bond. These risky zero-coupon bonds and the risk premia $(\mu_1(t), \ldots, \mu_{K-1}(t))$: $t = 1, \ldots, 14)$ in Tables 8 and 9 provide the necessary inputs for computing the values of options on risky debt and credit derivatives.

The previous tables demonstrate the simplicity of obtaining the parameter estimates for the Markov model. They also provide additional insights into numerous estimation issues that need to be addressed in a more thorough investigation. First, the data underlying the estimates for the risky zero-coupon bonds need to be filtered to remove all imbedded options. This filtering will remove the yield-to-worst bias present in the previous results. Second, to get a more homogeneous set of bond prices, partitioning by industry groups may be a good idea. Third, time-series observations are needed to (1) verify the time consistency of the estimations and (2) for testing trading strategies to see if the "mispricings" are due to arbitrage opportunities, data errors, or model misspecification.

4. Conclusion

This article develops a Markov model for the term structure of credit risk spreads, extending the work of Jarrow and Turnbull (1995). The bankruptcy process is modeled via a discrete state space, continuous time, time-homogeneous Markov chain in credit ratings. This bankruptcy processes' parameters can be estimated from observable data. An illustrative example is provided. The validation or rejection of this model awaits subsequent research.

Appendix A

Proof of Equation (28). This matrix can be derived using a probabilistic argument. Under this hypothesis, we can approximate the evolution of the firm's credit rating over a year by assuming that if the firm leaves state i before the year is over for state j, it is absorbed in state j. To

Table 8
Risk premia computed using constrained optimization on December 31, 1993

	Maturity													
Quality	0	1	2	3	4	5	6	7	8	9	10	11	12	13
AAA	8.6580	8.6580	8.6580	8.6580	0.0000	8.1190	8.6580	8.6580	8.6580	8.6580	0.0000	0.0000	0.0000	0.0000
AA	9.5785	9.5785	0.0000	9.5785	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
A	8.5324	8.5324	8.5324	8.5324	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
BAA1	5.0404	1.8159	0.0000	1.5447	0.0000	0.0000	0.0000	0.0000	1.1770	5.8445	0.0000	0.0000	0.0000	0.0000
BA	2.1117	0.0000	0.0000	0.0000	0.0000	0.0000	3.1267	0.4607	3.9526	3.9526	0.0000	0.0000	0.0000	0.0000
B	0.4307	0.9865	1.8181	2.5875	0.0000	0.0000	0.0000	0.0000	0.0000	0.4456	0.0309	0.0201	0.0244	0.0244
CAA	0.2607	0.3530	0.4034	1.1068	0.6978	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Table 9
Arbitrage-free risky zero-coupon bond prices for December 31, 1993

	Maturity													
Quality	1	2	3	4	5	6	7	8	9	10	11	12	13	14
AAA	96.9215	92.5014	87.1225	81.2513	76.2492	70.8646	65.1637	60.8654	57.0747	52.7439	49.363	45.9904	42.6181	39.2471
	(1.0825)	(0.9524)	(0.7695)	(0.8313)	(−1.9908)	(0.4376)	(−1.3333)	(−1.2106)	(−0.5803)	(−0.4901)	(−1.049)	(−1.5986)	(−2.1479)	(−2.6969)
AA	96.9064	92.0011	86.7370	80.9128	75.9209	70.5794	64.9239	60.6421	56.8693	52.5555	49.187	45.8266	42.4665	39.1076
	(0.9674)	(0.5421)	(−0.264)	(0.1238)	(−0.8831)	(−0.7776)	(−0.7361)	(−0.6329)	(−0.0217)	(0.0485)	(−0.941)	(−1.9224)	(−2.9035)	(−3.8834)
A	96.4118	91.5915	86.1866	80.2087	75.2303	69.9297	63.9945	59.7643	55.9816	51.7065	48.3897	45.0821	41.7748	38.4689
	(0.5218)	(0.3735)	(0.2614)	(−0.5713)	(−0.9027)	(−0.3923)	(−0.7065)	(−0.8167)	(−0.4784)	(−0.6335)	(−0.5083)	(−0.3729)	(−0.2382)	(−0.1021)
BAA1	95.356	90.3754	84.8356	78.7447	73.7927	68.6292	62.1318	58.0049	54.2150	50.0835	46.8659	43.6595	40.4533	37.249
	(0.0000)	(−0.3086)	(−1.0584)	(−0.2883)	(0.1737)	(0.8872)	(−0.4382)	(−0.0551)	(0.6650)	(1.0435)	(0.6719)	(0.3125)	(0.0467)	(−0.405)
BA	93.204	87.6054	81.0866	73.9117	68.9893	63.7359	57.2526	53.4365	49.8739	46.1042	43.1381	40.1842	37.2303	34.2787
	(0.0000)	(−1.4516)	(−2.3864)	(−0.5273)	(−0.4067)	(−0.4961)	(5.0156)	(4.0595)	(3.3569)	(2.4462)	(2.7431)	(3.0512)	(3.3593)	(3.6707)
B	94.851	86.2495	75.3329	63.6358	58.5354	51.8234	47.3623	44.2297	41.4798	38.4922	36.0173	33.5518	31.0865	28.6229
	(0.0000)	(0.3895)	(0.8080)	(0.1538)	(−5.4646)	(−7.1766)	(−4.8277)	(−3.6743)	(−2.1372)	(−0.8388)	(0.1123)	(1.0738)	(2.0345)	(2.9969)
CAA	92.106	82.3102	72.5022	56.5784	49.6692	43.8038	40.6174	37.9481	35.7206	33.3455	31.2104	29.0795	26.9488	24.8187
	(0.0000)	(0.0442)	(0.0762)	(0.0254)	(2.1182)	(−2.8592)	(-7.0086)	(−6.6849)	(-5.9204)	(−5.3025)	(−4.6746)	(−4.0425)	(−3.4092)	(−2.7763)
Standard error	.5831	.7267	1.0826	.4501	2.3935	2.9680	3.7908	3.3210	2.7228	2.2846	2.1409	2.1809	2.3949	2.7436
Percent error	.0061	.0082	.0132	.0061	.0350	.0473	.0661	.0620	.0543	.0489	.0493	.0539	.0638	.0794

In parentheses under each cell is the pricing error $\varepsilon_i(T) \equiv [v^i(0,T) - v^i(0,T)^{\text{obs}}]$. The standard error is $\sqrt{\sum_{i=1}^{7} \frac{\varepsilon_i(T)^2}{7}}$. The percent error is the standard error divided by the average price in the column.

make our argument, we define τ_{ij} to be independent exponentially distributed random variables with intensities λ_{ij}. Let $\tau_i \equiv \min_j \tau_{ij}$, then

$$\tau_i \text{ is exponentially distributed with parameter } -\lambda_i \equiv \sum_{\substack{j=1 \\ j \neq i}}^{K} \lambda_{ij}$$

and $\text{prob}(\tau_i = \tau_{ij} \mid \tau_i \leq 1) = \text{prob}(\tau_i = \tau_{ij}) = \lambda_{ij}/-\lambda_i$ [see Resnick (1992; Exercise 4.45 and Proposition 5.1.1)]. Our hypothesis implies that

$$\begin{aligned} q_{ii} &= \text{prob}(\tau_i > 1) = e^{\lambda_i} \text{ and} \\ q_{ij} &= \text{prob(going to state } j \text{ before time 1)} \\ &= \text{prob}(\tau_i = \tau_{ij} \text{ and } \tau_i \leq 1) \\ &= \text{prob}(\tau_i = \tau_{ij} \mid \tau_i \leq 1)\text{prob}(\tau_i \leq 1) = \frac{\lambda_{ij}}{-\lambda_i}(1 - e^{\lambda_i}). \end{aligned}$$

This completes the probabilistic argument. We now provide a second proof.

The hypothesis, in analytic form, is that

(i) $\lambda_{ij}\lambda_{jk} \approx 0$ for $i \neq j$ and $j \neq k$, and
(ii) For each i, $\sum_{n=0}^{\infty}(\sum_{m=0}^{n} \lambda_i^m \lambda_j^{n-m})/n!$ is independent of j.

Roughly, in probabilistic terms, (i) states that starting in state i, it is unlikely to jump to j and then k, in a small time period; (ii) states that starting in state i, given that one jumps out, the likelihood of jumping to state j (in a small time period) is proportional to λ_{ij}.

Now, $\exp(\Lambda) = \sum_{n=0}^{\infty} \Lambda^n/n!$.
Performing the matrix multiplication term by term, and then summing, one gets that the diagonal elements are e^{λ_j} for $j = 1, \ldots, K-1$. This follows from (i) alone. The off-diagonal elements (i, j) for $i \neq j$ are

$$\lambda_{ij}\left[\sum_{n=0}^{\infty}\left(\sum_{m=0}^{n} \lambda_i^m \lambda_j^{n-m}\right) \Big/ n!\right].$$

This also follows from (i) alone. But since the rows of $\exp(\Lambda)$ sum to one, we get

$$e^{\lambda_i} + \sum_{\substack{j=1 \\ j \neq i}}^{K} \lambda_{ij}\left(\sum_{n=0}^{\infty}\left(\sum_{m=0}^{n} \lambda_i^m \lambda_j^{n-m}\right) \Big/ n!\right) = 1,$$

hence the result, since $\lambda_i \equiv -\sum_{\substack{j=1 \\ j \neq i}}^{K} \lambda_{ij}$. This completes the argument.

Appendix B

Procedure for stripping risky zero-coupon bond prices
We use the maturities $T = 1, 2, 3, 4, 5, 6, 7, 10$, and 14 years. These are the midpoints of the maturity ranges given in Table 5.

Given are

$$\begin{aligned} c &= \text{coupon}, \\ y_T^i &= \text{yield to worst for the } i\text{th credit class, and} \\ T &= \text{maturity}. \end{aligned}$$

Let $\mathcal{B}_T^i$ be the bond's price.

To strip out the zeros, we first compute $\mathcal{B}_T^i$ by

$$\mathcal{B}_T^i = \sum_{j=1}^{T} \frac{c}{(1+y_T)^j} + \frac{100}{(1+y_T)^T}.$$

This assumes the coupon payment is made once per year and the face value is 100. Then, using a triangular system of equations, we compute $v^i(T)$ recursively as follows:

$$\begin{aligned} \mathcal{B}_1^i &= (c+100)v^i(1) \text{ compute } v^i(1) \\ \mathcal{B}_2^i &= cv^i(1) + (c+100)v^i(2) \text{ compute } v^i(2) \\ &\vdots \\ \mathcal{B}_7^i &= cv^i(1) + cv^i(2) + \cdots + (c+100)v^i(7) \text{ compute } v^i(7), \end{aligned}$$

missing $T = 8, 9$. To determine these we used linear interpolation. Let

$$\begin{aligned} v^i(8) &\equiv (2/3)v^i(7) + (1/3)v^i(10) \\ v^i(9) &\equiv (1/3)v^i(7) + (2/3)v^i(10). \\ \mathcal{B}_{10}^i &= cv^i(1) + \cdots + cv^i(7) + c[(2/3)v^i(7) + (1/3)v^i(10)] \\ &\quad + c[(1/3)v^i(7) + (2/3)v^i(10)] \\ &\quad + [c+100]v^i(10) \text{ compute } v^i(10), \end{aligned}$$

missing $T = 11, 12, 13$. To determine these let

$$\begin{aligned} v^i(11) &\equiv (3/4)v^i(10) + (1/4)v^i(14) \\ v^i(12) &\equiv (1/2)v^i(10) + (1/2)v^i(14) \\ v^i(13) &\equiv (1/4)v^i(10) + (3/4)v^i(14). \end{aligned}$$

$$\begin{aligned}\mathcal{B}_{14}^{i} = \; & cv^{i}(1) + \cdots + cv^{i}(10) + c[(3/4)v^{i}(10) + (1/4)v^{i}(14)] \\ & + c[(1/2)v^{i}(10) + (1/2)v^{i}(14)] \\ & + c[(1/4)v^{i}(10) + (3/4)v^{i}(14)] \\ & + [c + 100]v^{i}(14) \text{ compute } v^{i}(14).\end{aligned}$$

References

Amin, K., and A. Morton, 1993, "Implied Volatility Functions in Arbitrage Free Term Structure Models," forthcoming in *Journal of Financial Economics*.

Andersen, P. K., and O. Borgan, 1985, "Counting Process Models for Life History Data: A Review," *Scandinavian Journal of Statistics*, 12, 97–158.

Anderson, W. J., 1991, *Continuous-Time Markov Chains*, Springer, New York.

Black, F., and J. Cox, 1976, "Valuing Corporate Securities: Some Effects of Bond Indenture Provisions," *Journal of Finance*, 31, 351–367.

Black, F., E. Derman, and W. Toy, 1990, "A One-Factor Model of Interest Rates and Its Application to Treasury Bond Options," *Financial Analysts Journal*, 46(1), 33–39.

Black, F., and M. Scholes, 1973, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economics*, May, 637–659.

Chance, D., 1990, "Default Risk and the Duration of Zero Coupon Bonds," *Journal of Finance*, 45(1), 265–274.

Cox, J., J. Ingersoll, and S. Ross, 1985, "A Theory of the Term Structure of Interest Rates," *Econometrica*, 53, 385–407.

Cox, D., and H. Miller, 1972, *The Theory of Stochastic Processes*, Chapman and Hall, London.

Das, S. R., and P. Tufano, 1995, "Pricing Credit-Sensitive Debt When Interest Rates, Credit Ratings, and Credit Spreads are Stochastic," unpublished manuscript, Harvard University.

Group of Thirty, 1993, *Derivatives: Practices and Principles*, report by the Global Derivatives Study Group, Group of Thirty, Washington, D.C.

Fons, J., and A. Kimball, 1991, "Corporate Bond Defaults and Default Rates 1970–1990," *Journal of Fixed Income*, June, 36–47.

Geske, R., 1979, "The Valuation of Compound Options," *Journal of Financial Economics*, March, 63–81.

Harrison, J. M., and S. Pliska, 1981, "Martingales and Stochastic Integrals in the Theory of Continuous Trading," *Stochastic Processes and Their Applications*, 11, 215–260.

Heath, D., R. Jarrow, and A. Morton, 1992, "Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation," *Econometrica*, 60, 77–105.

Hull, J., and A. White, 1991, "The Impact of Default Risk on Option Prices," working paper, University of Toronto.

Jarrow, R., and S. Turnbull, 1995, "Pricing Derivatives on Financial Securities Subject to Credit Risk," *Journal of Finance*, 50, 53–86.

Jones, E., S. Mason, and E. Rosenfeld, 1984, "Contingent Claims Analysis of Corporate Capital Structures: An Empirical Investigation," *Journal of Finance*, 39, 611–627.

Karlin, S., and H. Taylor, 1975, *A First Course in Stochastic Processes* (2nd ed.), Academic Press, San Diego.

Lando, D., 1994, *Three Essays on Contingent Claims Pricing*, Ph.D. thesis, Cornell University.

Litterman, R., and T. Iben, 1991, "Corporate Bond Valuation and the Term Structure of Credit Spreads," *Financial Analysts Journal*, Spring, 52–64.

Longstaff, F., and E. Schwartz, 1992, "Valuing Risky Debt: A New Approach," working paper, UCLA.

Lucas, D., and J. Lonski, 1992, "Changes in Corporate Credit Quality 1970–1990," *The Journal of Fixed Income*, March, 7–14.

Merton, R. C., 1974, "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates," *Journal of Finance*, 2, 449–470.

Moody's Special Report, 1992, *Corporate Bond Defaults and Default Rates*, Moody's Investors Service, New York.

Nielsen, L., J. Saa-Requejo, and P. Santa-Clara, 1993, "Default Risk and Interest Rate Risk: The Term Structure of Default Spreads," working paper, INSEAD.

Resnick, S., 1992, *Adventures in Stochastic Processes*, Birkhauser, Boston.

Seneta, E., 1973, *Non-negative Matrices*, Wiley, New York.

Shimko, D., N. Tejima, and D. vanDeventer, 1993, "The Pricing of Risky Debt When Interest Rates are Stochastic," *Journal of Fixed Income*, September, 58–66.

Standard & Poor's Credit Review, 1993, *Corporate Default, Rating Transition Study Updated*, McGraw Hill, New York.