Some Elements of Rating-Based Credit Risk Modeling [1]

David Lando
Department of Operations Research
University of Copenhagen
Universitetsparken 5
DK-2100 Copenhagen Ø.
Denmark
fax: +45 35 32 06 78
e-mail: dlando@math.ku.dk

First version: November 17, 1998
This revision: February 24, 1999

[1] I am grateful to Brian Huge for helpful discussions.

1 Introduction

Default risk must be one of the oldest sources of financial risk that one can think of and yet there has been a surge of interest in the area in the 1990s. It is difficult to say why this surge did not come earlier, but now that it is there one can look for some reasons:

- Risk management of financial institutions is becoming increasingly complex and relies more and more on quantitative methods. The quantitative revolution in pricing and hedging 'market risk' is beginning to spill over into credit risk management and it is likely that banks will use increasingly sophisticated quantitative credit risk models to allocate capital internally and to convince regulators that capital adequacy standards are being met.
- Credit derivatives are becoming more popular, perhaps as part of a general tendency towards 'derivatizing' wider classes of risk and definitely as a low transaction cost means of diversifying loan portfolios. The idea behind such derivatives is typically to allow a bank to diversify without having to give up its core customer base which may have a strong regional concentration or a large concentration in a specific industry. The pricing and hedging of credit derivatives is impossible without a solid modeling framework.

The models we build to handle credit risk are centered around modeling credit events (e.g. defaults or downgrades) and payments on contracts made at such events. When choosing a modeling framework it is important to keep in mind what the purpose of the model is. Different models have advantages in different areas and an attempt to single out a single framework as generally superior is in my opinion misguided. What are the key issues that credit risk models have to address?

1. Pricing bank loans and corporate bonds: The fact that these types of loans exist is of course reason enough to study their pricing. But they serve a special role of basic building blocks for modeling credit risk in many applications, in much the same way that zero coupon bonds are used to set up a term structure of default free bonds from which term structure derivatives can be priced. At a more fundamental theoretical level, the pricing of corporate bonds is intimately linked

with questions of optimal capital structure and the study of conflicting interest between different creditors in a firm.

2. Risk management: Our models should be able to quantify risks associated with having credit risky instruments in a portfolio. The difficult question here is not only describing the stochastic behavior of prices but also to describe and measure correlation between price movements.

3. Pricing credit derivatives: Credit derivatives are contracts whose payout depend either on certain underlying credit events (such as default or downgrade) or measures of credit quality such as for example the yield spread between an A-rated bond and a treasury bond with similar payment structure. In this area, implied modeling is key: The ability to calibrate a model against observed data and price complicated structures off the model is the primary objective.

The main focus of this chapter is on understanding the idea of calibrating a rating based model using a low-dimensional modification of a known transition (or generator) matrix. This chapter is not intended as a survey but to put the chapter in a context, it does provide an introduction to a few basic ideas in credit risk modeling. The list of references is far from complete. A much more complete view of the literature would be obtained by consulting the references in such papers as Anderson and Sundaresan (1997) Cooper and Martin (1996) Duffie and Singleton (1997) Jarrow and Turnbull (1995) Lando (1997), Leland (1994) and Schönbucher (1997).

2 Connections between default probability and spreads

A first intuition which is important when modeling credit risk to know the connections between spreads in forward rates and default probabilities. First recall a little interest rate mathematics: Throughout, we let $B(t,T)$ denote the price at time t of zero coupon bond maturing at year T. The yield using discrete time (annual) compounding is

$$y(0,t) = \left(\frac{1}{B(0,t)}\right)^{\frac{1}{t}} - 1,$$

and the yield using continuous compounding is

$$y_c(0,t) = -\frac{1}{t}\log B(0,t)$$

The forward rate at time t seen from 0 (discrete time) is

$$f(0,t) = \frac{B(0,t)}{B(0,t+1)} - 1$$

and in continuous time the instantaneous forward rate at time t seen from 0 is

$$f(0,t) = -\frac{\partial}{\partial t}\log B(0,t)$$

Of course, the exact same quantities can be defined for a defaultable zero coupon bond. Let $v^i(0,t)$ denote the price of zero coupon bond maturing at t issued by a firm i. The credit spreads are obtained by taking differences: The yield spread is then defined as

$$ys^i(0,t) = y^i(0,t) - y(0,t)$$

and the forward spread is

$$fs^i(0,t) = f^i(0,t) - f(0,t),$$

and the continuous time versions may be defined this way as well.

To establish a connection between spreads and default probabilities we must have a model for the pricing of risky bonds. The cleanest picture is obtained in the case of the following very simple model for pricing the risky bond at time 0: Let the (random) default time of the firm be denoted τ. Now consider a defaultable zero coupon bond with maturity t, whose recovery in the event of default is equal to a constant δ, received at the time of maturity. This is equivalent to receiving at the default time a fraction δ of a treasury bond maturing at the same time as the defaulted bond. We let the probability of the firm surviving past t be denoted $S^i(0,t)$, i.e. if our underlying probability measure is P, we have

$$S^i(0,t) = P(\tau > t).$$

A simple pricing formula for the risky bond is then given as

$$v^i(0,t) = \delta B(0,t) + (1-\delta)B(0,t)S^i(0,t) \tag{1}$$

with the obvious interpretation that the bondholder always gets at least the discounted value of the recovery and - if the firm survives - the remaining value is received as well. The assumptions underlying this expression could be that P is a risk neutral measure obtained in some underlying arbitrage-free model and that under this risk neutral measure the default event is independent of the evolution of the variables governing the pricing of riskless bonds. One could also merely assume risk neutrality. What is important is that this simple pricing formula has a structure general enough to capture the essential structure of many model setups.

In the case where recovery is 0, the price simplifies to the expression

$$v^i(0,t) = B(0,t)S^i(0,t).$$

The issue of stripping out such a price is a delicate one and information from credit swaps may have to be added in practice to get enough pricing information to use for the calibration. For more on decomposing an observed spread into default probability and recovery rate, see Das and Sundaram (1998). The one year conditional default probability between time t and $t+1$, i.e. the probability, computed at time 0, of default before $t+1$ given survival up to and including time t is given by

$$\frac{S^i(0,t) - S^i(0,t+1)}{S^i(0,t)}$$

and rewriting this using our pricing formula gives us

$$\begin{aligned}
\frac{S^i(0,t) - S^i(0,t+1)}{S^i(0,t)} &= 1 - \frac{S^i(0,t+1)}{S^i(0,t)} \\
&= 1 - \frac{v(0,t+1)}{v(0,t)} \frac{B(0,t)}{B(0,t+1)} \\
&= 1 - \frac{1+f(0,t)}{1+f_v(0,t)} \\
&\approx f_v(0,t) - f(0,t)
\end{aligned}$$

Hence we see that forward spreads in discrete time are (almost) conditional default probabilities. To get an exact relationship, consider the continuous time case. Before computing the spread recall that any distribution on the positive reals which has a density can be expressed as a function of its hazard rate h defined as

$$h(t) := \frac{f(t)}{1 - F(t)}$$

where $f(t)$ is the density of the distribution and $F(t)$ is the distribution function. From this it follows that

$$S^i(t) = \exp\left(-\int_0^t h(s)ds\right)$$

Now consider the forward spread:

$$\begin{aligned} -\frac{\partial}{\partial t}\log v^i(0,t) - f(0,t) &= -\frac{\partial}{\partial t}\log\left(P(\tau \geq t)\right) \\ &= \frac{-\frac{\partial}{\partial t}P(\tau \geq t)}{P(\tau \geq t)} \\ &= \frac{f}{1-F} \end{aligned}$$

which is precisely the hazard rate of τ. The hazard rate gives conditional default intensities in the following sense:

$$P(\tau \in (t, t+\Delta t] \,|\, \tau \geq t) \approx h(t)\Delta t$$

Note the conditioning information: We are looking at the conditional default probability given only the information of survival up to time t. The central element of *reduced form* models (or *intensity* models) is the conditional default probability given survival *and* additional information which may have been collected at time t.

3 Some notes on reduced form modeling

Rating based models are typically formulated as a special case of reduced form models. The fundamental object of modeling in these models is the stochastically varying instantaneous rate of default, i.e. the intensity λ. Let information at time t be represented by $\mathcal{F}_t$. Then, heuristically, λ is $\mathcal{F}_t-$adapted process such that[1]

$$P(\tau \in (t, t+\Delta t] \,|\mathcal{F}_t) \approx \lambda(t)1_{\{\tau > t\}}\Delta t$$

[1] For precise definitions of this and a review of some of the results, see for example Duffie and Lando (1998).

Under some technical conditions (see Duffie (1998)) that need not worry us here it is the case that the survival probability over a given horizon T is given by

$$S(0,T) = E\exp\left(-\int_0^T \lambda(s)\,ds\right)$$

Note that this is of the same functional form as the formula for a bond price computed from the spot rate process r :

$$B(0,T) = E\exp\left(-\int_t^T r(s)\,ds\right)$$

Expressing the bond price in terms of the (instantaneous) forward rate

$$B(0,T) = \exp\left(-\int_0^T f(0,s)\,ds\right)$$

and comparing with the form of the survival probability in terms of its hazard rate

$$S(0,T) = \exp\left(-\int h(s)ds\right)$$

we note the analogy between, on one hand, spot rates and forward rates in bond pricing and, on the other hand, intensities and hazard rates in survival probability modeling. This analogy is fundamental in the approach developed by Duffie, Schroder and Skiadas (1996), Duffie and Singleton (1997) and Lando (1994,1998) which is able to handle dependence between the stochastic intensity and the spot rates very easily and essentially reduce the computation of prices of defaultable claims to the same form as that of default-free bonds. This makes it technically appealing from the viewpoint of computation and the advantages are perhaps more clear when computing prices of credit derivatives, see for example Schönbucher (1997). But there are empirical reasons as well for taking this path: Given the observation that spreads in forward rate curves are (in the case of zero recovery) the same as the hazard rate and that in particular the spread in the spot rates is equal to the intensity, it is clear that the spreads in the short end on yields will be strictly positive as long as there is a positive intensity of default. In the classical approach, taken in the works by Black and Scholes (1973) and Merton (1974), the spread in the short end goes to zero. To see this, note that if we assume that the value of a firm's equity starts out at a positive level S_0 and if default is defined as the first time the equity hits zero, then one has the following limiting result:

$$\lim_{h \to 0} \frac{P(\tau \leq h)}{h} = \lim_{h \to 0} \frac{\log P(\tau \geq h)}{h} = 0 \tag{2}$$

This fact makes short spreads go to zero in the short end.[2] To see this, think of the case with constant r and zero recovery:

$$v(0,h) = B(0,h)P(\tau \geq h)$$

where v is price of defaultable zero coupon bond. The yield spread is

$$-\frac{\log v(0,h)}{h} = y(0,h) - \frac{\log P(\tau \geq h)}{h}$$

and inserting the result from (2) shows that spreads go to 0. The empirical difficulties of diffusion models for firm value in explaining spreads on corporate bonds, as noted for example in Jones, Mason and Rosenfeld (1984), might be attributed to this fact. Introducing small fixes such as stochastic interest rates will not change this by much. More drastic features such as introduction of strategic debt service by the equity holders and absolute priority violations may partially remedy this (see for example Anderson and Sundaresan (1996)). Another immediate fix is to apply a jump-diffusion model for the asset value of the firm. This is done for example in Zhou (1997).

It may seem that there is a huge difference in economic content between, on one hand, modeling default as the first time the issuer's assets fall below a certain level and on the other hand modeling default as a Poisson-like process in which the default event is not specified directly as a condition on the firm's assets (or cash flows)[3]. Results of Duffie and Lando (1998) show however, that using a classical approach in a model in which there is imperfect observation of the firm's assets, assumed to follow a diffusion process, default becomes a Poisson-like process whose intensity process may be characterized and computed explicitly in certain cases. The yield spread

[2]In the Black-Scholes-Merton formulation, one has to assume that the firm is has a value larger than the face value of debt to get this result. Otherwise the yield goes to infinity.

[3]Note however, that the intensity of default may well depend on the firm's assets in a reduced-form model

in the short end then becomes strictly positive and the size of the spread reflects not only asset value but also the quality of information on the firm that bondholders have.

4 Approaches using ratings

One of the simplest ways of having a stochastically varying default intensity is to let a finite state space Markov chain - representing for example a rating - 'modulate' the intensity. We will from now on focus primarily on using ratings and for illustrational purposes we will look at a pricing model which uses *only* ratings, as in Jarrow, Lando and Turnbull (1997). We will explain the idea behind calibration of such a model and give three examples based on modifying an underlying generator matrix of a continuous-time Markov chain. For more on calibration of rating based models, see Das and Tufano (1996), Kijima (1998) and Kijima and Komoribayashi (1998). Rating based models which include stochastically varying transition intensities were introduced in Lando (1994, 1998) and recent contributions include Arvanitis, Gregory and Laurent (1998), T. Li (1997) and Nakazato (1997).

We think now a rating system for pricing corporate debt or perhaps indices of corporate debt. The example chosen here is one where only the rating category matters. In practice, more advanced methods and more explanatory variables will have to be added, but the calibration method discussed below is important in these frameworks as well.

The rating process is modeled as a Markov chain whose basic ingredients are a *state space* - often labeled $\{1, 2, \ldots, K\}$, where one should think of 1 as the top rating (AAA, say) and K as default - and a *transition matrix*

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & & p_{1K} \\ p_{21} & p_{22} & \cdots & \cdots & p_{2K} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ p_{K1} & p_{K2} & \cdots & \cdots & p_{KK} \end{pmatrix}. \quad (3)$$

Here, p_{ij} is the probability that the Markov chain which is currently in state i will be in state j next period. All entries in the transition matrix are non-negative and the rows sum to one:

$$\sum_{j=1}^{K} p_{ij} = 1$$

Formally: Let η_n denote the rating at time n. Then

$$\begin{aligned} & Prob(\eta_n = j | \eta_0 = i_0, \eta_1 = i_1, \ldots, \eta_{n-1} = i) \\ = \; & Prob(\eta_n = j | \eta_{n-1} = i) = p_{ij} \end{aligned}$$

The key property to note (the Markov property) is that only the rating at time $n-1$ and *not* the entire history is relevant for determining the probability of being in rating j at time n. From the one-period transition matrix, we get transition probabilities over several periods by multiplying the transition matrix with itself:

$$Prob(\eta_n = j | \eta_0 = i) = p_{ij}^{(n)}$$

where $p_{ij}^{(n)}$ is the $ij-$th entry of P^n

There are two key assumptions underlying this many period transition phenomenon: Time-homogeneity and the Markov property. The Markov property has already been described above. Non-homogeneity occurs if a separate matrix is used for each period (but we still have the Markov property). The multiperiod matrix will then have to be indexed by both the starting date and the ending date of the period over which we consider transitions. The length of the period is not sufficient. In this case the Markov property would be captured by the semi-group property of transition matrices:

$$P(s, u) = P(s, t)P(t, u)$$

whenever $s < t < u$. This means that the state at a certain time is still sufficient for determining the transition probabilities in the future but the transition probabilities are changing over (calendar) time. An example where even the Markov property breaks down would be a case where the number of time periods spent in the current rating or the precise level of the previous rating entered into the probability. One can then regain the Markov property by enlarging the state space but this is costly in terms of new parameters that are introduced into the model.

However, we will look at the case where the Markov property holds and we are interested in trying to deduce the implied transition probabilities from market prices. In principle it is possible to imply out parameters from a homogeneous Markov chain implied by prices if we have known recovery rates (or known expected recovery rates) in our model for bond prices given by (1): To see this, note from (1) that for each starting rating category and each time to maturity t we have

$$S^i(0,t) = \frac{v^i(0,t) - \delta B(0,t)}{(1-\delta)B(0,t)}$$

Since (expected) recovery rates are assumed known, we can observe the implied survival probabilities. To get the discrete-time transition probabilities, first note that

$$p_{iK} = 1 - S^i(0,1)$$

Then note that

$$S^i(0,t) = \sum_{j=1}^{K-1} p_{ij} S^j(1,t) \quad (4)$$

$$= \sum_{j=1}^{K-1} p_{ij} S^j(0,t-1) \quad (5)$$

where we used time-homogeneity to replace $S^j(1,t)$ - the probability of surviving past t given rating j at time 1 - by $S^j(0,t-1)$. This gives one new equation in the transition probabilities from category i for each time to maturity, since for fixed i and $t \geq 2$ we get

$$S^i(0,t) = p_{i1}S^1(0,t-1) + \cdots p_{i,K-1}S^{K-1}(0,t-1)$$

Hence from K maturities we get all the parameters $p_{i1}, \ldots, p_{iK}$ from this system. This procedure is then repeated for every rating category.

When trying to do this in practice one could replace one equation by the restriction that the probabilities sum to one but even then there is no guarantee that the solution is positive as probabilities should be. This reflects the fact that time-homogeneity is much too restrictive. Why not then assume a general non-homogeneous structure? The problem is that the unknown parameters cumulate quickly. For each new maturity that we consider we have as many new equations as there are rating categories, i.e. (K-1), but we have $(K-1)^2$ new parameters to determine.

Therefore, the real world attempts of calibrating a model often introduce non-homogeneity by finding a low dimensional modification of a known homogeneous Markov chain such that we only add as many parameters for each maturity as we can actually determine from the data using a known base matrix.

To illustrate this approach, assume that a one-period transition matrix P is given as in (3). This matrix will be our 'base matrix'. In Jarrow, Lando and Turnbull (1997) it is based on empirically observed transitions but it need not be in a general approach. The key is that it is *known* and that implied default probabilities for different maturities will be obtained from this matrix by a 'low dimensional' modification. Throughout, we will assume that there exists a matrix

$$\Lambda = \begin{pmatrix} -\lambda_1 & \lambda_{12} & \cdots & & \lambda_{1K} \\ \lambda_{21} & -\lambda_2 & \cdots & \cdots & \lambda_{2K} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \lambda_{K1} & \lambda_{K2} & \cdots & \cdots & -\lambda_K \end{pmatrix}$$

with

$$\begin{aligned} \sum_{j \in \{1,\ldots,K\} \setminus \{i\}} \lambda_{ij} &= \lambda_i \\ \lambda_i &\geq 0, \quad i = 1, \ldots, K \end{aligned}$$

such that

$$exp(\Lambda) = P.$$

The goal is to create a family of transition matrices $(Q(0,t))_{t>1}$ such that the implied default probabilities for each maturity t are matched by the corresponding entries in the last column of $Q(0,t)$. The general procedure is as follows: Let implied survival probabilities for rating category i over a time horizon t be denoted $S^i(t)$, so that the implied default probability is $1 - S^i(t)$.

1. Let $Q(0,0) = I$.
2. Given $Q(0,t)$. Choose $Q(t,t+1)$ such that $Q(0,t)Q(t,t+1)$ satisfies $(Q(0,t+1))_{iK} = 1 - S^i(t+1) \qquad \text{for } i = 1, \ldots, K-1.$
3. Go to step 2

What distinguishes the methods is the way in which step 2 is performed. Throughout we will consider an example with a rating system consisting of 4 categories A,B,C,D with D denoting default. For our example (which is purely for illustration) we let P be given as

P	A	B	C	D
A	0.95	0.03	0.01	0.01
B	0.1	0.7	0.1	0.1
C	0.1	0.2	0.4	0.3
D	0	0	0	1

and the associated generator is

Λ	A	B	C	D
A	−0.0539	0.0350	0.0125	0.0064
B	0.1126	−0.3889	0.19037	0.0859
C	0.1369	0.3795	−0.9612	0.4448
D	0	0	0	0

For comparison we also list the two period transition matrix obtained by squaring P :

$P(0,2)$	A	B	C	D
A	0.9065	0.0515	0.0165	0.0255
B	0.175	0.513	0.111	0.201
C	0.155	0.223	0.181	0.441
D	0	0	0	1

Assume that from prices of bonds we have deduced the following implied default probabilities:

	A	B	C
$1 - S^i(0,1)$	0.02	0.12	0.35
$1 - S^i(0,2)$	0.045	0.215	0.490

We now show different ways of modifying P (based on modifications of Λ) such that the implied default probabilities are matched. Note that the procedure does not match transition probabilities other than those to default. If further contracts are available - such as default swaps protection against downgrades - a higher dimensional modification of P can be used to fit not only implied default probabilities but also implied transitions probabilities between non-default categories. In each procedure we let

$$
\begin{aligned}
Q(0,1) &= \exp(\Lambda(0)) \\
Q(1,2) &= \exp(\Lambda(1))
\end{aligned}
$$

where $\Lambda(0)$ is a modification of Λ depending on $\pi_1 = (\pi_{11}, \pi_{12}, \pi_{13})$ and $\Lambda(2)$ is a modification of Λ depending on $\pi_2 = (\pi_{21}, \pi_{22}, \pi_{23})$. and we determine the parameter such that

$$\begin{aligned}
(Q(0,1))_{i,D} &= 1 - S^i(0,1) \qquad i = A, B, C \\
(Q(0,2))_{i,D} &= (Q(0,1)Q(1,2))_{i,D} = 1 - S^i(0,2), \qquad i = A, B, C
\end{aligned}$$

The extension to several periods is obvious. In each case we need to check whether the modified generator still is non-negative in its off-diagonal elements, so that we still have a generator matrix. Once this is checked, the matrix exponential will automatically give us a transition probability matrix.
Method 1: Modifying default intensities.
In this method we modify the default column of the generator matrix by letting

$$\begin{aligned}
\lambda_{AD}(0) &= \pi_{11} \cdot 0.0064 \\
\lambda_{BD}(0) &= \pi_{12} \cdot 0.0859 \\
\lambda_{CD}(0) &= \pi_{13} \cdot 0.4448
\end{aligned}$$

where $\pi_{11}, \pi_{12}, \pi_{13}$ are all strictly positive. The diagonal elements need modification too so that the matrix remains a generator with rows summing to zero:

$$\begin{aligned}
\lambda_{AA}(0) &= -0.0539 - (\pi_{11} - 1) \cdot 0.0064 \\
\lambda_{BB}(0) &= -0.3889 - (\pi_{12} - 1) \cdot 0.0859 \\
\lambda_{CC}(0) &= -0.9612 - (\pi_{13} - 1) \cdot 0.4448.
\end{aligned}$$

$\Lambda(1)$ is computed similarly using $(\pi_{21}, \pi_{22}, \pi_{23})$ instead of $(\pi_{11}, \pi_{12}, \pi_{13})$. Solving the matching conditions numerically, we obtain the following values:

$$\begin{aligned}
(\pi_{11}, \pi_{12}, \pi_{13}) &= (2.4998, 1.2158, 1.2116) \\
(\pi_{21}, \pi_{22}, \pi_{23}) &= (2.6725, 0.7884, 1.1486)
\end{aligned}$$

The associated implied transition probability matrices are

$Q(0,1)$	A	B	C	D
A	0.940879	0.0295479	0.00957321	0.02
B	0.098418	0.68669	0.0948917	0.12
C	0.0956735	0.189793	0.364534	0.35
D	0	0	0	1

$Q(0,2)$	A	B	C	D
A	0.888184	0.0512025	0.0156132	0.045
B	0.170443	0.510694	0.103864	0.215
C	0.144236	0.209551	0.156213	0.49
D	0	0	0	1

Method 2: Modifying rows of the generator.
This method is the one proposed in Jarrow, Lando and Turnbull (1997) (see equation (19) on page 495) but in there a numerical approximation of the transition matrix using the generator is used to solve for the adjustment parameters. The advantage of the numerical approximation is that it gives linear equations which of course can be solved analytically. But the procedure requires fairly small time intervals and the resulting matrix is not guaranteed to be a transition matrix. Here we do not use a linear approximation. Hence the transition matrix property is guaranteed, but we can only solve numerically: Let

$$\Lambda(0) = \begin{pmatrix} -0.0539 \cdot \pi_{11} & 0.0350 \cdot \pi_{11} & 0.0125 \cdot \pi_{11} & 0.0064 \cdot \pi_{11} \\ 0.1126 \cdot \pi_{12} & -0.3889 \cdot \pi_{12} & 0.19037 \cdot \pi_{12} & 0.0859 \cdot \pi_{12} \\ 0.1369 \cdot \pi_{13} & 0.3795 \cdot \pi_{13} & -0.9612 \cdot \pi_{13} & 0.4448 \cdot \pi_{13} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

for strictly positive values of $\pi_{11}, \pi_{12}, \pi_{13}$. $\Lambda(1)$ is computed similarly using $(\pi_{21}, \pi_{22}, \pi_{23})$ instead of $(\pi_{11}, \pi_{12}, \pi_{13})$. The resulting parameters are

$$\begin{aligned} (\pi_{11}, \pi_{12}, \pi_{13}) &= (1.8988, 1.1606, 1.2925) \\ (\pi_{21}, \pi_{22}, \pi_{23}) &= (1.4754, 0.7005, 1.6628) \end{aligned}$$

and the associated transition matrices are

$Q(0,1)$	A	B	C	D
A	0.908042	0.0547708	0.0171868	0.02
B	0.112348	0.667519	0.100133	0.12
C	0.115383	0.223701	0.310916	0.35
D	0	0	0	1

$Q(0,2)$	A	B	C	D
A	0.847867	0.090461	0.0166715	0.045
B	0.166842	0.556799	0.0613593	0.215
C	0.162213	0.266138	0.0816494	0.49
D	0	0	0	1

Method 3: Modifying eigenvalues of the transition probability matrix.
Finally we consider a case studied first in Lando (1994) and (1998). We assume that P (and therefore also Λ) are diagonalizable, we let B denote a matrix of eigenvectors of P, and we let D denote a diagonal matrix of eigenvalues of Λ. Hence we have

$$\Lambda = BDB^{-1}$$

and we now let

$$\Lambda(0) = B\Pi(0)DB^{-1}$$

where $\Pi(0)$ is a diagonal matrix with diagonal elements $(\pi_{11}, \pi_{12}, \pi_{13})$ and 0, where the value 0 should correspond to the eigenvalue of zero in the generator. The same procedure is used for $\Lambda(1)$. In our example this gives the following value of parameters:

$$\begin{aligned}(\pi_{11}, \pi_{12}, \pi_{13}) &= (1.4124, 1.18906, 1.3326) \\ (\pi_{21}, \pi_{22}, \pi_{23}) &= (1.2601, 0.9561, 2.8896)\end{aligned}$$

where the order of parameters here are chosen such that π_{11} modifies the (numerically) largest eigenvalue. The associated transition matrices are

$Q(0,1)$	A	B	C	D
A	0.935037	0.0336963	0.0112667	0.02
B	0.112148	0.652385	0.115467	0.12
C	0.113185	0.230881	0.305933	0.35
D	0	0	0	1

$Q(0,2)$	A	B	C	D
A	0.886296	0.0518704	0.0168333	0.045
B	0.175185	0.478481	0.131333	0.215
C	0.161481	0.263352	0.0851667	0.49
D	0	0	0	1

.

One needs to check in this method that the modified generator is indeed a generator. The rows are guaranteed to sum to zero, but off-diagonal elements may become negative. This is a potential problem of the method. The advantage is the fact that the equation solving for the parameter vectors π_1 and π_2 are linear and can therefore be solved explicitly. To see this, let b_{jK}^{-1}denote the $(j, K)'th$ entry of B^{-1}, where we assume that the $K'th$

eigenvalue of the generator is 0. One may check that $b_{iK}b^{-1}_{KK} = 1$. Defining $\beta_{ij} = -b_{ij}b^{-1}_{jK}$ we can write the probability of no default before time 1 as

$$1 - P(0,1)_{i,K} = \sum_{j=1}^{K-1} \beta_{ij} \exp(d_j(1))$$

where $d_j(i) = d_1\pi_{1i}$ and since β_{ij} is known for every i, j this system of equations determines $(d_1(1), \ldots, d_{K-1}(1))$, and therefore π_1. Now consider the survival probability for two periods. Rewrite as before

$$S^i(0,2) = \sum_{j=1}^{K-1} P_{i,j}(0,1) S^j(1,2)$$

and use the fact that there are only $K-1$ new parameters to determine here, namely the constants $(d_1(2), \ldots, d_{K-1}(2))$ entering into the expression $S^j(1,2)$. Fortunately, there are also $K-1$ equations, which are linear in the variables $\exp(d_j(2)), \quad j = 1, \ldots, K-1$. The procedure is then repeated for maturity three years and so forth. The problem using this method in practice is finding a good base matrix. Empirical generators seem to have too small default intensities for high rated issuers and this leads to strange results and possibly no solutions to the calibration. The key assumption for the procedure to work is that there exists a common basis of eigenvectors for the implied transition matrices of all maturities. Arvanitis, Gregory and Laurent (1998) present results supporting this assumption. Fot the simplification and connection with affine models which are obtained in rating based models by using this representation in continuous time, see Lando (1994,1998).

5 Credit swaps

Given a calibrated intensity model of default, such as the rating based model of the previous section, one may proceed to price credit derivatives. One of the most popular credit derivatives is the credit default swap (or credit swap) whose basic idea is to protect the holder of the swap against default on an underlying *reference* security. As we will see below, with a known recovery rate the credit default swap can be priced just knowing the default probabilities and therefore the particular choice of calibration method is irrelevant. However, by changing the credit event to include a downgrade to a non-default category, the calibration method will matter. In technical terms, this is because we are only calibrating to a subset of securities which does not span all sources of uncertainty in our model.

In a credit default swap, the protection buyer pays a fixed coupon $c(T)$ every period (typically every 3 or 6 months.) T is the maturity date of the contract - the size of the constant payment naturally depends on the length of the swap contract. The payment continues until which ever comes first: default or maturity of the credit swap. If default occurs before the maturity of the swap, there is (in the case of cash settlement) a payment from the protection seller to the protection buyer usually equal to the difference between the notional amount of the bond and the recovery value (based for example on an average quote among dealers in the immediate post-default market). As with ordinary swaps, the fixed side payment is set so that the contract value at initiation is zero. In abstract terms, the cash flow of the contract is as follows: The payment at a coupon date i for the protection buyer is

$$c(T)1_{\{\tau>i\}}$$

The payment of the protection seller is equal to at time τ

$$(1-\delta)1_{\{\tau\leq T\}}$$

To find $c(T)$ such that the cash flows of the two sides are equal assuming known survival probabilities under the pricing measure, we find assuming a constant interest rate r :

$$c(T)=\frac{(1-\delta)E[e^{-r\tau}1_{\{\tau\leq T\}}]}{\sum_{i=1}^{n}e^{-ri}E[1_{\{\tau>i\}}]}$$

which can be readily calculated and is easy to generalize to cases with stochastic interest rates and dependence between interest rates and default intensities. To see the formula illustrated in the example of the previous chapter, consider a two period default swap entered into at time 0, and with a maturity at date 2. The protection buyer pays $c(2)$ at date 1 if there is no default and $c(2)$ again at date 2 if there is no default at or before time 2. If default happens at time 1 the protection buyer receives $1-\delta$ where δ is the recovery of face value and the same applies at date 2. If the default date is denotes τ then it is clear that the value of the cash flow paid by the protection buyer is

$$c(2)B(0,1)Q(\tau>1)+c(2)B(0,2)Q(\tau>2).$$

The value of the cash flow paid by the protection seller is

$$(1-\delta)B(0,1)Q(\tau=1)+(1-\delta)B(0,2)Q(\tau=2).$$

The swap premium is then the value of $c(2)$ for which these two cash flows have the same value. Note that we do not consider the default risk of the parties to the contract but focus on the reference security. We also ignore for simplicity the effects of accrued interest if default happens between coupon dates. For the example of the previous section, the premium $c(2)$ as a function of initial rating assuming a continuously compounded interest rate of 5%, a recovery of 50% and a notional amount of 100 is given below:

	A	B	C
$c(2)$	1.159	6.466	21.28

In practice, as swaps are becoming sufficiently liquid, the observed prices are used to generate their own implied term structures supplementing scarce data on defaultable bonds.

It is clear that one could generalize to swaps insuring against a downgrade in addition to a default such that the protection seller would receive a specified amount in the event of a transition to a category lower than a certain boundary. In the procedure above, simply define the relevant credit event and modify the payment accordingly. In this case, prices will reflect the fact that transition probabilities between non-default categories depended on the calibration method, as can be seen by comparing the implied transition matrices. There is clearly significant variation in the probabilities of ending in category C over, say, a two year time period. This dependence is not surprising given that we only fitted default probabilities.

The intensity framework (and in particular the ratings-based approach) generalizes easily to a first-to-default type swap contract, in which there is a basket of reference securities. The loss on the first of the reference securities to default is then covered by the swap contract. The framework developed to handle such contracts is also relevant for taking credit risk of the parties to the credit swap into account. The advantage of reduced form models here is that intensities of baskets behave very nicely in that the sum of the intensities is the intensity of the first default. For more on this see Duffie (1998b).

References

[1] R. Anderson and S. Sundaresan (1996) "Design and Valuation of Debt Contracts," *Review of Financial Studies* **9:**37-68

[2] A. Arvanitis, J.K. Gregory, J-P. Laurent (1998), "Building Models for Credit Spreads," Working Paper, Paribas and CREST

[3] F. Black and M. Scholes (1973), "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy* **81:** 637-654.

[4] I. Cooper and M. Martin (1996), "Default risk and derivative products," *Applied Mathematical Finance* **3:** 53-74

[5] S. Das and R. Sundaram(1998), "A direct approach to arbitrage-free pricing of credit derivatives," Working Paper, Harvard University and NUY.

[6] S. Das and P. Tufano (1996), "Pricing Credit Sensitive debt when Interest rates and Credit Spreads are Stochastic," *Journal of Financial Engineering* **5:**161-198

[7] D. Duffie (1998a), "Credit Swap Valuation"' Working Paper, Graduate School of Business, Stanford University,

[8] D. Duffie (1998b), "First-to-Default Valuation," Working Paper, Graduate School of Business, Stanford University, and Institut de Finance, Universté de Paris, Dauphine.

[9] D. Duffie and D. Lando (1998), "Term Structures of Credit Spreads with Incomplete Accounting Information" Working Paper, Stanford University and University of Copenhagen.

[10] D. Duffie, M. Schroder, and C. Skiadas (1996), "Recursive Valuation of Defaultable Securities and the Timing of Resolution of Uncertainty," *Annals of Applied Probability*, **6:** 1075-1090.

[11] D. Duffie and K. Singleton (1997), "Modeling Term Structures of Defaultable Bonds," Forthcoming in: *Review of Financial Studies.*

[12] R. Jarrow, D. Lando, and S. Turnbull (1997), "A Markov Model for the Term Structure of Credit Spreads," *Review of Financial Studies* **10:** 481-523

[13] R. Jarrow and S. Turnbull (1995), "Pricing Options on Financial Securities Subject to Default Risk," *Journal of Finance* **50:** 53-86.

[14] E. Jones, S. Mason, and E. Rosenfeld (1984), "Contingent Claims Analysis of Corporate Capital Structures: An Empirical Investigation," *The Journal of Finance* **39:** 611-625.

[15] M. Kijima (1998), "Monotonicities in a Markov Chain Model for Valuing Corporate Bonds Subject to Credit Risk" *Mathematical Finance* **8:**229-247.

[16] M. Kijima and K. Komoribayashi (1998), "A Markov Chain Model for Valuing Credit Risk Derivatives" *Journal of Derivatives* Fall,pp. 97-108.

[17] D.Lando (1994), "Three Essays on Contingent Claims Pricing," Ph.D. Thesis. Cornell University.

[18] D. Lando (1997), "Modeling Bonds and Derivatives with Default Risk," *in* M. Dempster and S. Pliska (editors) *Mathematics of Derivative Securities*, pp. 369-393, Cambridge University Press.

[19] D. Lando (1998), "On Cox Processes and Credit Risky Securities," *Review of Derivatives Research* **2:**pp.99-120.

[20] H. Leland (1994), "Corporate Debt Value, Bond Covenants, and Optimal Capital Structure," *Journal of Finance* **49:** 1213-1252.

[21] R. Merton (1974), "On The Pricing of Corporate Debt: The Risk Structure of Interest Rates," *The Journal of Finance* **29:** 449-470.

[22] D. Nakazato (1997), "Gaussian Term Structure Model with Credit Rating Classes," Working Paper, The Industrial Bank of Japan.

[23] P. Schönbucher (1997), "The Term Structure of Defaultable Bond Prices," Forthcoming in *Review of Derivatives Research*.

[24] C. Zhou (1997) "A Jump Diffusion Approach to Modeling Credit Risk and Valuing Defaultable Securities," Working Paper, Federal Reserve Board.