low-grade debt that occurred in the late 1980s and the associated increases in default rates. After controlling for "current liabilities of business failures" and for "corporate profits as a percentage of GNP," Jonsson and Fridson (1996) found a significant correlation between the quantity of unseasoned bottom-tier issues and default. This finding raises the possibility that the aging effects found by McDonald and Van de Gucht are spuriously affected by their omission of key cyclical factors. Further exploration of these issues seems warranted.

4 Ratings Transitions: Historical Patterns and Statistical Models

THE POSSIBILITY THAT major ratings agencies will change the credit rating of a bond issue is an important source of credit risk, above and beyond its implications for the direct risk of default. Changes in credit ratings may have an immediate effect on the values of defaultable bond portfolios, may mean that certain bonds are no longer admissible for investors who are subject to restrictions on the ratings of bonds in their portfolios, and may even lead to mandatory termination of some financial contracts. Certain *step-up* corporate bonds have coupon rates linked explicitly to credit rating. Proposed changes to the BIS accord, reviewed in Section 2.5.2, will determine the capital requirements of regulated banks based in part on the credit ratings of the debt instruments that they hold. These ratings-based regulatory capital charges will presumably be reflected in the market yield spreads charged for a given rating. Credit derivatives, and other contracts, sometimes have provisions for contingent payments based on changes in credit ratings. For these and related reasons, a model of the risk of ratings changes is a key ingredient for credit-risk management and for the valuation of many credit-sensitive instruments. This chapter reviews some of the historical patterns in ratings transitions as well as alternative approaches to modeling ratings transition risk.

4.1. Average Transition Frequencies

Major ratings agencies report the historical average incidence of transitions among credit ratings and into default in the form of a matrix of average transition frequencies. An example from Moody's is given in Table 4.1. Each row corresponds to the rating at the beginning of a year; the column heading gives the end-of-year rating. For example, of firms rated Baa at the beginning of a year, on average over this sample, 81% retained this rating,

Table 4.1. ***Moody's All-Corporate Average Transition Frequencies for 1980–2000***

	Rating at year end (percent)								
Initial rating	Aaa	Aa	A	Baa	Ba	B	Caa-C	Default	WR
Aaa	86.17	9.45	1.02	0.00	0.03	0.00	0.00	0.00	3.33
Aa	1.10	86.05	8.93	0.31	0.11	0.01	0.00	0.03	3.46
A	0.06	2.85	86.75	5.58	0.66	0.17	0.01	0.01	3.91
Baa	0.06	0.34	6.64	81.00	5.52	0.97	0.08	0.16	5.23
Ba	0.03	0.06	0.54	5.46	75.50	8.18	0.53	1.32	8.38
B	0.01	0.04	0.20	0.56	5.92	75.94	3.03	6.41	7.90
Caa-C	0.00	0.00	0.00	0.87	2.61	5.62	57.02	25.31	8.58

Source: Moody's Investor Services.

while approximately 5.5% made a transition to Ba. The fractions of transitions shown to WR correspond to withdrawn ratings. Although there may be some implications for credit quality for the event of becoming unrated, this effect is often ignored by normalizing each transition frequency by the total fraction of bonds that do not have a withdrawn rating.[1] The resulting normalized transition matrix is shown in Table 4.2.

It is not uncommon in industry practice to treat an annual average transition frequency matrix of the sort shown in Table 4.2 as though it is a matrix π of transition probabilities, with π_{ij} denoting the probability that a firm rated i at the beginning of the year is rated j at the end of the year. This implicitly assumes that transition probabilities are constant over time and that the sole determinant of an issuer's credit risk (including transition risk) is the issuer's current rating. This strong assumption would allow one to treat the rating of a firm as a Markov chain. As the number of years of data gets larger and larger, the historical transition frequency matrix would, under these assumptions, converge to the actual probability transition matrix π.

Such a Markov-chain assumption for ratings transitions is the basis for the popular CreditMetrics model (J. P. Morgan, 1997), although CreditMetrics does not require that the user's transition probability matrix π is equal to a historical average 1-year transition frequency matrix. The Markov-chain assumption greatly simplifies multiyear ratings transition and default probability calculations. For example, the n-year probability transition matrix is π^n, the n-fold product of the 1-year transition matrix. That is, for a firm initially rated i, the probability of being rated j after n years is the (i, j) element of the matrix π^n. Because 1-year periods are too "coarse-grained" for many applications, later in this chapter we extend this annual-period

[1] Roundoff errors imply that the rows of Table 4.2 do not add precisely to 100.

Table 4.2. ***Moody's Average Ratings Transition Frequency Matrix for 1980–2000, Normalized from Table 4.1 for Withdrawn Ratings***

	Rating at year end (percent)							
Initial rating	Aaa	Aa	A	Baa	Ba	B	Caa-C	Default
Aaa	89.14	9.78	1.06	0.00	0.03	0.00	0.00	0.00
Aa	1.14	89.13	9.25	0.32	0.11	0.01	0.00	0.03
A	0.06	2.97	90.28	5.81	0.69	0.18	0.01	0.01
Baa	0.06	0.36	7.01	85.47	5.82	1.02	0.08	0.17
Ba	0.03	0.07	0.59	5.96	82.41	8.93	0.58	1.44
B	0.01	0.04	0.22	0.61	6.43	82.44	3.29	6.96
Caa-C	0.00	0.00	0.00	0.95	2.85	6.15	62.36	27.68

Markov chain into a continuous-time setting, in which ratings changes can happen at any time during the year.

One should beware of the potential traps associated with this elegant and simple treatment of the rating of an issuer as a Markov chain. The reported transition frequencies are only averages and do not condition on all available information. In particular:

- The reported transition matrix represents the average of 1-year transitions over a long history, during which economic conditions change.
- Even ignoring cyclical considerations, 1-year transition rates may not be fixed over the age of the bond, the age of the firm, the previous rating, or time in current rating category. As with default risk, there is dependence of transition probabilities on duration in a rating category or age, often called an *aging effect,* documented by Carty and Fons (1994), Lando and Skødeberg (2000), and Kavvathas (2001). Moreover, Behar and Nagpal (1999), Lando and Skødeberg (2000), and Kavvathas (2001) all find that, for firms of certain ratings, the prior rating is an important determinant of the likelihood of a downgrade, versus that of an upgrade, over a given time horizon. There is indeed apparent *momentum* in ratings transitions data.
- Generally, different firms of the same rating have different credit qualities, and a given firm of a fixed rating has a credit quality that changes over time.

4.2. Ratings Risk and the Business Cycle

In order to assess the relationships between rating transitions and the business cycle, we computed four-quarter moving averages of the ratio of the total number of Moody's ratings upgrades to the the total number of downgrades (the *U/D* ratio). The *U/D* ratio was computed separately for the investment and speculative-grade (below Baa) ratings categories. These ratios were then correlated with the four-quarter moving average of U.S. gross domestic product (GDP) growth over various sample periods. The sample was split in 1983, because Moody's introduced a new alpha-numeric rating system in April 1982. Inspection of ratings transition activity during 1982 shows clearly that, with this new system, a large number of firms were reclassified to new coarse (letter) ratings. For instance, during the quarter in which Moody's introduced their alpha-numeric system, nearly 50% of Baa-rated U.S. industrial firms transitioned to new letter ratings. Such a large transition rate is well out of line with typical transition rates. The analysis of Kliger and Sarig (1997) suggests that the information release implicit in the new classification system was economically significant, in that there was a significant reaction of corporate bond yield spreads to the ratings refinements.

We report the sample correlations in Table 4.3, and show the underlying (standardized) time series of *U/D* ratios and GDP growth in Figure 4.1 for the postrevision sample period 1983–1997. The correlation is low during the first sample period but positive during the later sample period. The trends in upgrades relative to downgrades for speculative-grade issues, in particular, are strikingly well correlated with GDP growth during the second half of the sample. In Chapter 3, for example, we reported a consistent large negative correlation between default rates and GDP growth.

In order to analyze the impact of business cycles on ratings transitions probabilities, Nickell et al. (2000), extending work by Wilson (1997a, b), estimated an ordered probit model, a class of statistical models explained in Section 4.4, for the likelihood of a given transition, using: (1) domicile in Section 4.4, for the likelihood of a given transition, using: (1) domicile (Japan, U.S., U.K., and non-U.K. European domiciles); (2) ten industry categories; and (3) the business cycle (peak, normal, and trough) as explanatory variables.

Nickell et al. (2000) allocated the years of their sample period, 1970–1997, into "peak," "normal," and "trough" categories, depending on whether GDP growth was in the upper, middle, or lower third of realized growth rates over the sample period. For U.S. industrial and banking Aa-rated firms (treated separately), they found only a few statistically significant differences between the transition probabilities for peak and trough

Table 4.3. ***Sample Correlations of GDP Growth and U/D Ratios for Investment and Speculative-Grade Issues, as rated by Moody's***

Sample period	Investment grade	Speculative grade
1971–1982	0.076	−0.012
1983–1997	0.198	0.652
1971–1997	0.068	0.230

Figure 4.1. *Upgrade/downgrade ratios for 1983–1997 (standardized for 1971–1997).*

periods. There were significant differences, however, between peaks and troughs for lower-rated firms, between U.S. banks and industrial firms, and between U.S.- and foreign-domiciled firms.

Such a model and estimation procedure implicitly assume independence across issuers within a year, given the business cycle. Correlation is induced only by business cycle changes (which affect all issuers the same way). A selection of the results of Nickell et al. (2000) is given in the estimated peak-conditional and trough-conditional transition probabilities shown in Table 4.4, where "A↑" represents the subset of firms rated A or higher. The entries shown in bold type are the percentage transition probabilities estimated for troughs of the business cycle. Those shown in standard type are associated with the peaks. One notes the obvious adverse influence on ratings of troughs relative to peaks.

In order to extend our Markov-chain setting for transitions so as to allow for the influence of the business cycle (or some other state variable), we can let $\Pi_{ij}(x)$ denote the probability of a transition from rating i to rating j over 1 year if the underlying driving state variable (such as the business cycle) is at x. For example, with three business cycle states, as in Nickell et al. (2000), x is one of {peak, normal, trough}.

Then, in order to deal with multiyear transition probabilities, we can make the additional simplifying assumption that the time series of underlying covariate vectors, $X_1, X_2, \ldots, X_t, \ldots$, forms a Markov process. Moreover, conditional on the path taken by this Markov state process X, the probability that a given issuer makes a transition from rating i at year t to j at year $s > t$ is the (i, j) element of the product $\Pi(X_t)\Pi(X_{t+1}) \cdots \Pi(X_{s-1})$ of 1-year conditional transition matrices. The unconditional multiyear transition probabilities can be calculated, for example, by simulating the paths of X and averaging the conditional transition probabilities over independently generated paths for X. This is called a doubly stochastic transition model. More computationally tractable doubly stochastic ratings-transition models, in continuous time settings, are introduced in Section 4.5.2.

Table 4.4. Comparison of Peak and Trough (bold) Estimated Transition Probabilities (percent)

	A↑	Baa	Ba	B	Caa↓	D
A↑	(**93.9**, 95.7)	(**5.6**, 3.9)	(**0.4**, 0.3)	(**0.0**, 0.1)	(**0.0**, 0.0)	(**0.0**, 0.0)
Baa	(**6.9**, 4.6)	(**86.8**, 92.2)	(**5.6**, 2.8)	(**0.4**, 0.3)	(**0.2**, 0.1)	(**0.1**, 0.1)
Ba	(**0.6**, 0.6)	(**5.9**, 4.8)	(**83.1**, 88.5)	(**8.4**, 5.0)	(**0.3**, 0.3)	(**1.7**, 0.7)
B	(**0.3**, 0.1)	(**0.8**, 0.3)	(**6.6**, 7.2)	(**79.6**, 85.8)	(**3.2**, 2.1)	(**9.4**, 4.5)
Caa	(**0.0**, 0.0)	(**0.0**, 0.9)	(**1.9**, 2.7)	(**9.3**, 5.4)	(**64.9**, 77.5)	(**23.1**, 12.6)

Source: Based on results from Nickell et al. (2000).

More generally, we can treat the pair (X_t, C_t) as a Markov chain. Suppose, for instance, that X_t has three outcomes, as in the example above with peak, normal, and trough. With, say, seven ratings, we can now treat the entire system as a Markov chain with $3 \times 7 = 21$ states. If we follow the doubly stochastic assumption, we can let $p_{x,y}$ denote the conditional probability that $X_{t+1} = y$, given that $X_t = x$. In this case, the one-period transition probability from (x, i) to (y, j) is given by $p_{x,y} \times \Pi_{i,j}(x)$. Multiperiod ratings transition probabilities are now easy, following the matrix-product rule.

4.3. Ratings Transitions and Aging

There is evidence that the transition rates to new ratings depend on the length of time that an issue has held its current rating and also on an *aging effect*. These effects are explored by Carty and Fons (1994), who assumed that the probability distribution for the time spent continually in a given credit rating was described by the Weibull duration model introduced in Chapter 3 in the context of default time. The Weibull duration model has a hazard rate for transition out of the current rating of

$$h(t) = \gamma p(\gamma t)^{p-1}, \tag{4.1}$$

where p and γ are constants, and the time t should be interpreted as time since entering a given rating category. If $p > 1$, then $h(t)$ is increasing in t, so the hazard function $h(\cdot)$ is said to exhibit *positive duration dependence*. If $p = 1$, then $h(t)$ does not depend on t, whereas $p < 1$ implies that there is negative duration dependence. Positive duration dependence implies that the longer a bond has a given credit rating, the more likely it is that there will be a change in rating in the near term. The Weibull distribution implies only monotonically increasing or decreasing hazard rates for nonzero p. A "humped" pattern, for instance, cannot be captured by the Weibull model.

Carty and Fons estimated positive duration dependence for Aaa-rated bonds, approximately no duration dependence for Baa-rated bonds, and negative duration dependence for B-rated and Caa-rated bonds (see Figure 4.2). For the lowest credit rating, the negative duration dependence suggests that the intensity of default declines with the duration of survival time from issuance. However, these forward survival probabilities are not conditioned on information about the state of the economy or other credit covariates.

The study by Carty and Fons (1994) also examined the presence of ratings *drift* and *momentum*. They found that, since the mid-1970s, the differences between the percentage of firms upgraded and the percentage of firms downgraded generally declines with later calendar years. This was

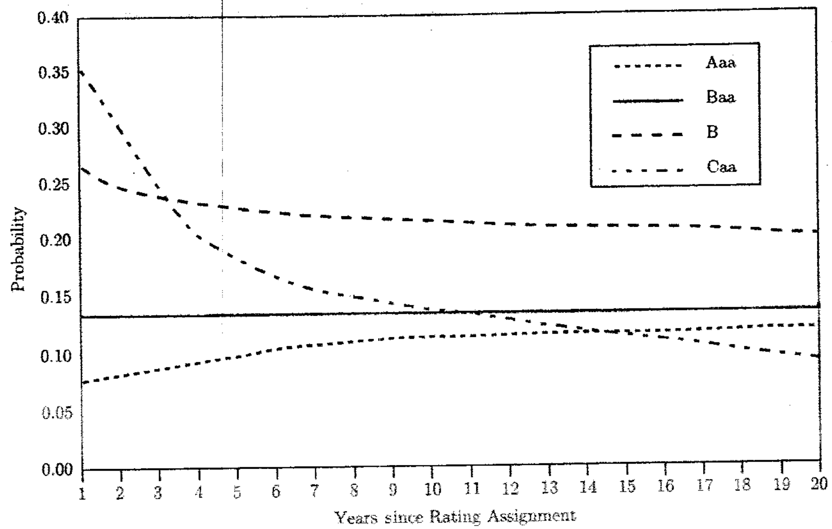

Figure 4.2. Hazard functions for selected long-term ratings. (Source: Carty and Fons, 1994.)

interpreted as a downward drift in the Moody's ratings. Additionally, they examined whether prior ratings changes have predictive power for future ratings changes, that is, whether there is *ratings momentum*. Using data from January 1938 through June 1993, Carty and Fons (1994) found that, for all ratings between Aa and B, there was a significantly higher probability of a downgrade within a year given that the previous rating change had been a downgrade. They found essentially no evidence of a comparable upgrade momentum. Extending from this work, Behar and Nagpal (1999), Lando and Skødeberg (2000), and Kavvathas (2001) all find that for firms of a given rating, the prior rating is an important determinant of the likelihood of a downgrade, versus that of an upgrade, over a given time horizon.

4.4. Ordered Probits of Ratings

Most of the statistical analyses linking credit ratings to observable credit-related covariates are special cases of qualitative-response models, such as ordered probit. As opposed to ratings-transition probabilities, these studies focus on an explanation of how ratings are assigned at a given time as a function of currently observable variables related to credit quality.

An ordered qualitative-response model is obtained by assuming that ratings are linked to a credit-quality index Z_t that depends on observable covariates measured in a vector X_t, in that

$$Z_t = \alpha + \beta \cdot X_t + \epsilon_t, \qquad (4.2)$$

for an intercept coefficient α and a vector β of slope coefficients to be estimated, where ϵ_t is a *noise* variable that is generally assumed to be independent and identically distributed across observations. We let $C(t)$ denote the credit rating of a bond at date t and suppose that there are K credit ratings, as well as default $(K+1)$. The ratings of bonds are then assumed to be determined by

$$C(t) = j, \qquad \text{if} \quad z_{j-1} \leq Z_t < z_j, \qquad (4.3)$$

for boundary coefficients $z_1, z_2, \ldots, z_K$, which are also to be estimated from the data on the covariates X_t and ratings $C(t)$ of issuers. [In (4.3), we take $z_0 = -\infty$ and $z_{K+1} = +\infty$.] In other words, the probability that the rating of this bond is j is the probability that the credit index Z_t is in the interval $[z_{j-1}, z_j)$.

Ordered probit models were estimated by Kaplan and Urwitz (1979), Ederington and Yawitz (1987), Cheung (1996),[2] and Blume et al. (1998).

Blume et al. (1998) examined a data set of over 14,000 observations covering the period 1973–1992. The covariate vector X_t included balance-sheet information about the firm as well as information about the firm's stock-return beta. Blume et al. also allowed the intercept coefficient α in (4.2) to depend on the year. They found that these estimated time-varying intercepts became increasingly negative over time, especially after 1980, and interpreted this finding to mean that the standards used in assigning ratings became more stringent over time. This conclusion comes from the implication that a firm with a given credit rating in one year is more likely to have a lower credit rating in subsequent years, holding the conditioning variables (X_t) fixed. These results may reflect the same ratings-drift phenomenon documented by Carty and Fons (1994). They, however, interpreted ratings drift as being due to "a prolonged deterioration in overall credit quality that started in 1980," and not to changes in the stringency of the criteria used by Moody's to rate firms. Blume et al. (1998) also omit business cycle effects on both the likelihood that a given firm will default and possible cyclical changes in the ranges of characteristics that define the credit ratings (the boundary coefficients z_j).

[2] Cheung (1996) focused on provincial debt ratings in Canada and thus considered a quite different set of explanatory variables.

4.5. Ratings as Markov Chains

We now extend the notion of rating as a Markov chain by using the intensity approach of Chapter 3, in this case to allow for ratings transition that may occur at any time within a year, as follows.

We allow for K nondefault ratings as well as the default rating, numbered $K+1$. For any current rating i, there is a transition intensity for a jump to any other rating j given at time t of $\Lambda_{ij}(t)$. For example, any firm currently rated i has a default intensity of $\Lambda_{i,K+1}(t)$. We suppose that there is no transition out of default, so that $\Lambda_{K+1,i}(t) = 0$ for all $i < K+1$. This is a technical convenience and is not to be taken as a statement of whether a firm can be reorganized after default. For now, we can suppose that these transition intensities are deterministic and constant. Later, we consider deterministic time variation, and then stochastic variation, in intensities.

For example, we could take the $K = 6$ nondefault ratings, Aaa, Aa, A, Baa, Ba, B, and let default be rating number 7, labeled D. The total intensity of a change in rating (including default) by a B-rated issuer is then

$$\Lambda_{B,Aaa} + \Lambda_{B,Aa} + \Lambda_{B,A} + \Lambda_{B,Baa} + \Lambda_{B,Ba} + \Lambda_{B,D}.$$

Similarly, for each rating i, the total intensity of a change in rating is $-\Lambda_{ii}$, where

$$\Lambda_{ii} = -(\Lambda_{i,1} + \Lambda_{i,2} + \cdots + \Lambda_{i,i-1} + \Lambda_{i,i+1} + \cdots + \Lambda_{i,7}). \tag{4.4}$$

This implies that, starting in rating i at time t, the probability of remaining within rating i continually until some time $s > t$ is $e^{(s-t)\Lambda_{ii}}$, analogous to the modeling in Chapter 3 of survival probabilities based on a constant default intensity.

The *generator* for the Markov chain is the 7×7 matrix Λ, with the ith diagonal element Λ_{ii}, given by

$$\Lambda = \begin{pmatrix} \Lambda_{11} & \Lambda_{12} & \Lambda_{13} & \cdots & \Lambda_{1,6} & \Lambda_{1,D} \\ \Lambda_{21} & \Lambda_{22} & \Lambda_{23} & \cdots & \Lambda_{2,6} & \Lambda_{2,D} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \Lambda_{61} & \Lambda_{62} & \Lambda_{63} & \cdots & \Lambda_{6,6} & \Lambda_{6,D} \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}. \tag{4.5}$$

The zero bottom row of the generator Λ in (4.5) reflects our standing assumption that default is a *trapping state*.

For example, extrapolating from estimates in work by Jarrow et al. (1997), a constant generator that is "close" in some sense to that implied by the average transition frequencies of Figure 4.1 is given by

	A↑	BBB	BB	B	CCC	D
A↑	−0.086	0.069	0.011	0.005	0.000	0.001
BBB	0.077	−0.171	0.070	0.017	0.002	0.005
BB	0.012	0.081	−0.252	0.118	0.014	0.027
B	0.005	0.007	0.057	−0.192	0.048	0.075
CCC	0.014	0.014	0.025	0.093	−0.432	0.286
D	0	0	0	0	0	0

after collapsing AAA, AA, and A into a single rating denoted A↑.

Israel et al. (2001) provide alternative techniques for the "calibration" of a generator to a 1-year transition probability matrix (or a transition matrix of any discrete period length). It is well known, however, that for certain 1-year transition matrices, there need not be any generator.[3] This would mean that the 1-year transition matrix could not be that of a continuous-time Markov chain. Moreover, as pointed out by Israel et al. (2001), there can be distinctly different generators consistent with a given 1-year transition matrix, and these different generators imply different transition probabilities for time horizons other than 1 year.

In any case, assuming that we have a generator Λ in hand, we are ready to calculate the probability $\Pi_{ij}(t, s)$, beginning at time t in rating i, of being in rating j at time $s > t$. In particular, $\Pi_{i,K+1}(t, s)$ gives the probability, for an i-rated firm at time t, of having defaulted by time s. (This uses the assumption that default is a trapping state, for once having defaulted at any time before s, the firm is assured of being in default at time s.) This transition matrix is

$$\Pi(t, s) = e^{\Lambda(s-t)}, \tag{4.6}$$

where the exponential of any matrix A is defined by

$$e^A = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots.$$

4.5.1. Time-Varying Transition Intensities

If the ratings-transition generator $\Lambda(t)$ varies over time, then special methods are required to compute ratings-transition probabilities. It is *not*

[3] Israel et al. (2001) provide sufficient conditions on a transition matrix for the existence of a generator and conditions for the failure of the existence of a generator.

generally true that one can extend (4.6) to get $\Pi(t, s) = e^{\int_t^s \Lambda(u)\,du}$. Instead, for a time-varying but deterministic generator $\{\Lambda(t) : t \geq 0\}$, $\Pi(t, s)$ solves the linear ordinary differential equation (ODE)

$$\frac{\partial}{\partial s}\Pi(t, s) = -\Lambda(s)\Pi(t, s). \tag{4.7}$$

Without special structural assumptions on the generator Λ, (4.7) is solved numerically.

In one of the most comprehensive empirical studies of a time-varying generator, Kavvathas (2001) assumed that the ijth entry of $\Lambda(t)$ is given by

$$\Lambda_{ij}(t) = \exp\left[\eta_{ij} + \gamma_{ij}' X(t)\right]. \tag{4.8}$$

The parameters were estimated using historical ratings-transition information from Standard & Poor's and with various choices of $X(t)$, including bond, equity, and credit market variables. He found that his econometric model with state-dependent intensities out-performed, in out-of-sample forecasting, various reference models (including the constant-intensity model).

Computation of $\Pi(t, s)$ from a model such as Kavvathas's can be burdensome, however, so it is instructive to examine cases where the linear ODE can be solved more explicitly. One such case is when

$$\Lambda(s)\Lambda(t) = \Lambda(t)\Lambda(s), \tag{4.9}$$

for all s and t. This commutativity property obviously holds if the transition intensities are constant. More generally, commutativity applies if, for all t, $\Lambda(t)$ can be diagonalized in the form

$$\Lambda(t) = B\mu(t)B^{-1}, \tag{4.10}$$

where B is the matrix whose columns are the eigenvectors of $\Lambda(t)$, and $\mu(t)$ is the diagonal matrix whose ith diagonal element is the eigenvalue of $\Lambda(t)$ associated with the ith eigenvector. By assuming that B does not depend on t, while $\mu(t)$ may, we have commutativity (4.9). Under this commutativity condition, it follows from the theory of linear ODEs that

$$\Pi(t, s) = B\exp\left[\int_t^s \mu(u)\,du\right]B^{-1}. \tag{4.11}$$

For example, from (4.11), the probability, beginning in rating i at time t, of being in some other given rating k at time $s > t$, is

$$\begin{aligned}\Pi_{ik}(t, s) &= \sum_{j=1}^{K+1} B_{ij}\exp\left[\int_t^s \mu_j(u)\,du\right]B_{jk}^{-1}\\ &= \sum_{j=1}^{K+1} \beta_{ijk}\exp\left[\int_t^s \mu_j(u)\,du\right],\end{aligned} \tag{4.12}$$

where $\beta_{ijk} = B_{ij}B_{jk}^{-1}$.

4.5.2. Lando's Stochastic Transition-Intensity Model

In order to allow for stochastic transition intensities, which are natural given random fluctuations over time in the economic environment that determines default risk, but still maintain tractability, Lando (1998) provides the following insightful model. He maintains the special diagonalizability assumption (4.10) on the generator $\Lambda(t)$, now assumed to be a stochastic process, and moreover assumes that we have the affine dependence,

$$\mu_j(t) = \gamma_{j0} + \gamma_{j1}\cdot X_t, \tag{4.13}$$

of the eigenvalues of $\Lambda(t)$ on some n-dimensional affine state process X that reflects changing economic conditions determining ratings-transition intensities.

Lando's model is doubly stochastic in the sense that, conditional on the entire future path of the state process X, the transition probabilities from rating to rating are given by (4.12). It follows from (4.12) that, for an issuer currently rated i, the conditional probability at time t, given X_t, of being in rating k at a future time s is

$$\begin{aligned}\pi_{ik}(t, s, X_t) &= E_t\left(\sum_{j=1}^{K+1} \beta_{ijk}\exp\left[\int_t^s \mu_j(u)\,du\right]\right)\\ &= \sum_{j=1}^{K+1} \beta_{ijk}E_t\left(\exp\left[\int_t^s (\gamma_{j0} + \gamma_{j1}\cdot X_u)\,du\right]\right)\\ &= \sum_{j=1}^{K+1} \beta_{ijk}\, e^{a_j(t,s)+b_j(t,s)\cdot X(t)},\end{aligned} \tag{4.14}$$

for coefficients $a_j(t, s)$ and $b_j(t, s)$, which can be easily calculated, as in any affine model setting, in the manner explained in Appendix A. For example, if the coordinates of X are CIR (square-root) diffusions, or mean reverting with jumps, or variants discussed in Appendix A, then we can use the explicit formulas for affine processes that were applied in Chapter 3 to calculate default probabilities.

In order to consider the special structure associated with Lando's generator-decomposition assumptions (4.10) and (4.13) for the correlations among the various transitions and default intensities, we momentarily specialize to the case of $K = 2$ nondefault ratings. We label rating 1 as I, for investment grade, and rating 2 as S, for speculative grade. The ratings-transition intensities $\Lambda_{IS}(t)$ and $\Lambda_{SI}(t)$ governing the transitions from rating I to S, and conversely, can be calculated from (4.10) as

$$\Lambda_{IS}(t) = \eta_1 [\mu_1(t) - \mu_2(t)]$$
$$\Lambda_{SI}(t) = \eta_2 [\mu_1(t) - \mu_2(t)],$$

where

$$\eta_1 = -\frac{B_{11}B_{12}}{B_{11}B_{22} - B_{12}B_{21}}$$
$$\eta_2 = \frac{B_{21}B_{22}}{B_{11}B_{22} - B_{12}B_{21}}.$$

It follows that the ratings-transition intensities are proportional to each other in this model and, hence, perfectly correlated. In particular, with $K = 2$, this model cannot capture asymmetric upgrade/downgrade patterns, such as the tendency for there to be more upgrades than downgrades in economic expansions and the opposite during contractions. We note that this conclusion is independent of assumptions such as (4.13) regarding the behavior of the eigenvalues $\mu_1(t)$ and $\mu_2(t)$.

Turning to the intensities for default, after computing $B\mu(t)B^{-1}$ and selecting the relevant elements of this matrix, we obtain[4] the default intensity from ratings I and S of

$$\Lambda_{ID}(t) = -[\Lambda_{IS}(t) + \omega\mu_1(t) + (1-\omega)\mu_2(t)]$$
$$\Lambda_{SD}(t) = -[\Lambda_{SI}(t) + (1-\omega)\mu_1(t) + \omega\mu_2(t)],$$

respectively, where $\omega = B_{11}B_{22}/(B_{11}B_{22} - B_{12}B_{21})$. As $\Lambda_{IS}(t)$ and $\Lambda_{SI}(t)$ are perfectly correlated, these default intensities are also perfectly correlated if $\omega = 0.5$.

More generally, by counting the equations restricting the transitions intensities with K nondefault ratings, Lando (1998) finds that the K^2 transition intensities vary in a $(K-1)$-dimensional subspace. As we have already found for $K = 2$, this implies that these intensities must be perfectly correlated.

In Chapter 6, we reinterpret these results in order to consider the effect of stochastic variation in transition intensities on the term structures of credit spreads for each rating.

[4] Here, we use the facts that

$$B^{-1}_{13} = \frac{B_{12} - B_{22}}{B_{11}B_{22} - B_{12}B_{21}}, \quad B^{-1}_{23} = \frac{B_{21} - B_{11}}{B_{11}B_{22} - B_{12}B_{21}}.$$