Bayesian Migration in Credit Ratings Based on Probabilities of Default

Sanjiv R. Das, Rong Fan, and Gary Geng

Sanjiv R. Das
is an associate professor of finance at the Leavey School of Business at Santa Clara University in Santa Clara, CA.
srdas@scu.edu

Rong Fan
is with Gifford Fong Associates in Lafayette, CA.
rong@gfong.com

Gary Geng
is with Gifford Fong Associates in Lafayette, CA.
gary@gfong.com

We develop a Bayesian model of rating changes based on the joint stochastic process of the probabilities of default of many issuers. Since changes in PDs are related to changes in ratings, we employ a modified (i.e., "biased") Bayesian model to calibrate the historical time series of PD changes to historical rating transition matrices.

Rating transitions punctuate changes in the prices of securities issued by firms. Firms such as Moody's, Standard & Poor's, and Fitch announce rating changes during periodic reviews of the creditworthiness of firms. There is now a vast history of rating transition data summarized into rating transition matrices (*see Exhibit 1 for an example*).

Rating transitions are important for market players for many reasons. First, they signal real changes in the value of firms, resulting in a series of repricings of issued securities.

Second, they impact investment portfolios subject to rating-based restrictions. For example, money market funds are not permitted to hold more than a small fraction of low-grade paper.

Third, securities that are indexed to ratings are impacted. Credit-sensitive notes, for example, are bonds whose coupons are indexed to rating levels. Fourth, credit portfolio risk is simulated according to rating transitions. Hence, ratings are important in all aspects of the credit markets.

Rating transition matrices have been used in the pricing of credit risk-sensitive securities. A framework for this form of modeling was first developed by Jarrow, Lando, and Turnbull [1997]. Extensions to this work have followed in rapid succession, as in Das and Tufano [1996], Kijima [1998], and Kijima and Komoribayashi [1998].

While these approaches use fixed transition matrices (albeit with time-dependent risk premiums), other models have used stochastic transitions (see Lando [1994, 1998], Nakazato [1997], Li [1998], and Arvanitis, Gregory, and Laurent [1999]). Applications to swap pricing are developed in Huge and Lando [1999].

Myriad approaches have been used to estimate rating transition matrices. A simple approach in continuous time is presented in Lando and Skodeberg [2002]. A detailed empirical examination of a range of state variables in the estimation of rating transition matrices has been carried out by Kavvathas [2001]. Bangia, Diebold, and Schuermann [1999] demonstrate the influence of business cycle effects on the ratings migration process.

Kiesel, Perraudin, and Taylor [2001] investigate the interaction of spread volatility with ratings. Since ratings are a coarser (and slower-moving) measure, they show that, for high-quality debt, spread volatility is a severe risk, even in the absence of ratings transition risk.

EXHIBIT 1
Sample Empirical Rating Transition Matrix

Rating	Rating Category						
	1	2	3	4	5	6	Default (D)
1	0.9081	0.0833	0.0068	0.0006	0.0012	0.0000	0.0000
2	0.0070	0.9065	0.0779	0.0064	0.0006	0.0014	0.0000
3	0.0009	0.0227	0.9105	0.0552	0.0074	0.0026	0.0006
4	0.0004	0.0035	0.0597	0.8695	0.0532	0.0119	0.0020
5	0.0020	0.0031	0.0084	0.0790	0.8070	0.0901	0.0123
6	0.0068	0.0079	0.0092	0.0111	0.0716	0.8414	0.0588
D	0	0	0	0	0	0	1

Based on data from Moody's.

The growing evidence on the dynamics of ratings may be distilled into some general conclusions. While rating transition matrices have traditionally been presented as constant, it is widely accepted that they are not time-homogeneous. Various reasons have been offered for this, mostly based on underlying macroeconomic state variables.

Also known is the feature of ratings drift, in that high-quality credits tend to drift lower in their ratings, and low-quality firms tend to drift higher. Hence mean reversion in ratings is also present.

Ratings are a slower-moving variable than the underlying state variables, partly on account of the human process involved, and naturally too because they are discrete measures as compared to the continuous variables on which they are based. Finally, many market participants see ratings as changing too infrequently, and being ex post; i.e., the information revealed in a ratings change has often already been reflected in the spreads for an issuer.

Therefore, there is a need for a process of rating dynamics that is not as fast-changing as various state variables, yet is sufficiently timely. We show that this may be achieved in a simple Bayesian model of rating changes based on probabilities of default that incorporates the characteristics described.

A recent development in the credit markets has been the use of model-based probabilities of default (PD) for credit analysis. Various models based on Merton [1974] are now in vogue, such as those of KMV (see Crosbie [1999]) and RiskMetrics (see Finkelstein et al. [2002]). These models use the equity process of the firm as the basis for computing and updating PDs, as frequently as hourly. Other approaches, termed hybrid ones like Moody's, mix a market-based equity approach with balance sheet-based accounting information, calibrated to historical default experience, to compute PDs (see Sobehart et al. [2000]).

In either modeling approach, PDs proxy for a large number of state variables that drive changes in firms' credit quality, proxied by ratings. Thus, PDs can be used to refine the process of estimating the relationship between credit-related state variables and the stochastic process of rating transitions.

Estimating the process of rating transitions at the simplest level involves counting. Under the assumption of time-homogeneity of the transition matrix, the estimator of each cell in the transition matrix (i, j) is derived from the count of transitions from rating i to rating j, denoted $N_{ij}/\Sigma_j N_{ij}$, $\forall i$. If we assume that the transition matrix is not time-homogeneous, then the values in the cells of the matrix may be defined as functions of a set of state variables, as implemented by Kavvathas [2001]. Hence, estimation of rating transition matrices has been based on transition data and other state variables, in either a time-homogeneous or inhomogeneous framework.

The approach we take differs from other research in some important ways. First, it uses PDs as a sufficient statistic for other state variables. This has the advantage of making the model parsimonious. Second, the Bayesian approach enables calibration of the speed of rating transitions at a rate between that of the state variable and that of human judgment, i.e., in a more timely fashion. The approach also incorporates the human judgment factors that may be present in the ratings updates, by calibrating the Bayesian system to observed transition matrices.

The model we develop is useful for many reasons. First, it is simple and intuitive, requiring little more than a Bayesian updating rule. Second, the results mimic the true process of rating changes well. The transitions in the model are neither too fast, a problem with PD changes, nor too slow, an aspect of the human process for ratings. Third, rating agencies may use the model to propose rating changes, and market participants may find the approach useful for forecasting rating changes.

Finally, as we will show, the model replicates historical transitions, using only PD data, lending support to the argument that PDs may be sufficient statistics for rating changes.

I. THE MODEL

Some basic notation sets up the model. We assume there are K credit rating classes indexed by $k = 1, \ldots, K$. The rating classes are arranged in descending order of credit quality; i.e., rating class 1 is the best, and rating class K is the worst.

There are N issuers indexed by $i = 1, \ldots, N$, in these K classes. The rating class for each issuer is denoted R_{it}, where t indexes time.

Associated with each rating class is a mean probability of default (PD), denoted μ_{kt}, which is the mean PD for rating category k at time t. The stochastic process for the rating level PD is as follows:

$$\mu_{k,t+1} = \mu_{kt} + M_k[\theta_k - \mu_{kt}]\Delta t + \sigma_k(\mu_{kt})\varepsilon_{kt}, \quad \forall k \qquad (1)$$

$\sigma_k(\mu_{kt})$ is an arbitrary although positive and bounded function on μ_{kt}. M_k is the rate of mean reversion, and θ_k is the long-run mean of the process. The driving random shock is ε_{kt}. The only restriction imposed on the choice of stochastic process is that it be positive.

Associated with each issuer is a probability of default (PD), denoted λ_{it}, which denotes the likelihood of issuer i defaulting within one year of time t. The PDs follow a stochastic process.

For example, a simple version of the dynamics is

$$\lambda_{i,t+1} = \lambda_{it} + G_i[\mu_{k,t+1} - \lambda_{it}]\Delta t + \eta_i(\lambda_{it})\varepsilon_{it} \qquad (2)$$

$\eta_i(\lambda_{it})$ is an arbitrary although bounded function on λ_{it}. G_i is the rate of mean reversion, and μ_{kt} is the current mean of the process. The driving random shock is ε_{it}. Again, a positive stochastic process is assumed.

We may allow for arbitrary correlation among all the random shocks.

Equation (1) may be estimated from a time series of the average of the PDs for each rating class. Equation (2) is estimated from the time series of PDs for each issuer, according to:

$$\lambda_{i,t+1} = \lambda_{it} + G_i[\mu_{i,t+1} - \lambda_{it}]\Delta t + \eta_i(\lambda_{it})\varepsilon_{it} \qquad (3)$$

where

$$\mu_{i,t+1} = \mu_{k,t+1} = \mu_{R_{it},t+1}$$

Hence data from the rating level series is introduced to proxy for the latent mean of the process.

Bayesian Changes in Rating

Ratings and PDs are positively related. That is, as PDs increase, we expect ratings to rise (remember, the quality drops as the rating level increases from 1 to K). We posit a conditional mapping from PDs to ratings; i.e., there is a function $h_i | C_i$, such that $h_i\!: \lambda_i | C_i \to k$. C_i is a general function of the history of PDs.

Because this is a many-to-few type of mapping, the relationship is coarse—each rating class includes a range of PDs. Across rating classes, these PD ranges overlap as well, so they are not distinctly defined. Therefore, our techniques are more useful and relevant.

Ratings do not change as frequently as PDs. There are two reasons for this. First, the coarse mapping means that every time a PD changes, it need not result in a rating change.

Second, even when the PD changes enough to populate the range of another rating class, this may simply be on account of a wild overreaction in the equity markets, which the rating agencies do not necessarily respond to. In fact, if they did overreact, and the overreaction is corrected the following day, this would mean a rating reversal within a two-day period. Therefore actual ratings cannot be mapped to PDs too directly, as any model that does so would induce excessive ratings volatility, not observed empirically.[1]

Yet it is also known that rating changes are often slow, and the rating agencies are criticized for not updating their ratings frequently enough. Thus, there is a need for ratings to track PDs somewhat more closely than they do at present, but not too closely as to be too volatile. We present a Bayesian approach to achieving this desired middle ground.

Using time series data of PDs (λ_{it}) for each issuer and ratings (R_{it}), we calibrate a Bayesian model of rating changes.

At every time t, there is a known prior probability of the issuer being in every rating class. This prior probability is denoted

$$Pr[R_{it} = k], \quad \forall i$$

This probability is also the posterior probability of the rating class from the previous period ($t - 1$).

From time t to time $t + 1$, there is an observed change in the PD, i.e., a move from λ_{it} to $\lambda_{i,t+1}$. The model then determines whether this change in PD amounts to a change in rating as well.

Using the estimated stochastic processes in Equations (1) and (2), we can determine the conditional probability density of the innovation in λ_{it}. This density function is denoted

$$f[\lambda_{i,t+1} \mid \lambda_{it}, R_{it} = k]$$

This density is also conditioned on the current rating level, which is necessary to parameterize Equation (2).

A simple application of Bayes' theorem enables the computation of posterior probabilities:

$$Pr[R_{i,t+1} = k \mid \lambda_{it}] = \frac{f[\lambda_{i,t+1} \mid \lambda_{it}, R_{it} = k].Pr[R_{it} = k]}{\sum_{k=1}^{K} f[\lambda_{i,t+1} \mid \lambda_{it}, R_{it} = k].Pr[R_{it} = k]}, \quad \forall i, k \tag{4}$$

Given these posterior probabilities, we determine the rating as follows:

$$R_{i,t+1} = \text{argmax}_k \, Pr[R_{i,t+1} = k \mid \lambda_{it}]$$

These posteriors are then used in the next round of updating.

This approach has some useful features that are consistent with the empirical reality of ratings transitions. First, transitions are dynamic. Second, the model makes rating changes persistent, which matches the empirical stickiness of ratings. Third, the passage to default is progressive and not sudden. Hence, the Bayesian approach formalizes the fact that sufficient evidence is required of an improvement or degradation in credit quality before a rating change is announced. Finally, because the rating changes are based on the stochastic process for PDs, they are time- and state-dependent.

Initializing the Transition System

Given a time series of PD and rating data, we can establish the unconditional probability density for each rating category. At the simplest level, this may be thought of as the histogram of PDs within a rating class. If the unconditional density function is denoted $g(\lambda_i | k)$, we can initialize the system for each issuer at one of two possible values.

1. Set the starting rating for each issuer, i.e., R_{i0} at the most likely value:

$$R_{i0} = \text{argmax}_k \, g(\lambda_{i0} \mid k)$$

2. Set the starting value at the mean PD for issuer i, i.e., equal to $\text{argmax}_k \, g(\bar{\lambda}_i | k)$. We then set the prior probabilities equal to:

$$Pr[R_{i0} = k] = g(\lambda_{i0} \mid k), \quad \forall i, k$$

Of course, the Bayesian system would not work if the prior were exactly zero, so we ensure that there is some small positive probability for each prior, even if the probability is zero.

Modified Bayesian Model

It is possible to generate time series data based on the parameters of the PD stochastic processes, and run the entire PD system forward using Monte Carlo simulation, keeping track of all rating transitions. A count of the transitions developed from this simulated data is summarized in a rating transition matrix (call this the *fitted* transition matrix). We may then compare this transition matrix to that developed from actual transitions over time (the *empirical* matrix), frequently published by rating agencies like Moody's and Standard & Poor's.

Of course, the two transition matrices need not coincide, unless the Bayesian system perfectly captures the behavior of the rating agencies in the way they process PD data. There are many reasons why differences may occur.

First, rating agencies use information beyond that in the PDs. Second, rating updates as the rating agencies see revised PDs need not weight the probabilities in the pure Bayesian form we have described.

This may happen because they update probabilities in a biased manner, i.e., not exactly as per Equation (4). By modifying the Bayesian model, we can capture the effects of these biases, so as to provide a more flexible system that better depicts the way ratings evolve with PDs.

We achieve this by introducing K new parameters (γ_k, $k = 1, \ldots, K$) into the updating rule:

$$Pr[R_{i,t+1} = k \mid \lambda_{it}] = \frac{\gamma_k f[\lambda_{i,t+1} \mid \lambda_{it}, R_{it} = k].Pr[R_{it} = k]}{\sum_{k=1}^{K} \gamma_k f[\lambda_{i,t+1} \mid \lambda_{it}, R_{it} = k].Pr[R_{it} = k]}, \quad \forall i, k \tag{5}$$

We may think of the new parameters γ_k, $k = 1, \ldots, K$ as coefficients that bias each conditional probability for each rating. For example, if the rating agency were biased in favor of rating $k = 3$, then γ_3 would be a large number relative to the other γ_k, $k \neq 3$ values.

There is another benefit to the flexibility introduced with these parameters. By optimally choosing them, we can ensure that the historical system generated from the data generates a fitted transition matrix that is as close as possible to the empirical one. Hence, the introduction of K parameters enables effective calibration. We demonstrate in an application that it is possible to carry out such a calibration, which may be used to support 1) the prediction of rating changes, and 2) automated changes in ratings.

Finally, after the K parameters have been fitted, the user of the methodology may modify the parameters to change the rate of migration between rating classes. For example, if the modeler believes that rating changes should occur faster than in the current model, a reduction in the bias parameters γ_k will result in quicker migrations.

Hence, the Bayesian approach to rating changes provides flexibility in the modeling of transitions. A further extension would make the bias parameters themselves functions of exogenous state variables, thereby allowing the rating transition matrix's time- and state-dependence to be based on factors other than PDs.

EXHIBIT 2

System State Variables

$\mu_{kt} \in \Re^K$	Mean of the PD for rating class k
$\lambda_{it} \in \Re^N$	PD for issuer i
$R_{it} \in \Re^N$	Rating of issuer i
$Pr[R_{it} = k] \in \Re^{N \times K}$	Posterior probability of rating

II. MONTE CARLO SIMULATION

The simulation of rating changes is easy to implement, once the stochastic processes for PDs have been calibrated to historical data.[2] The following are the three Monte Carlo steps at each time t:

1. Generate a sample of the rating mean PDs at time $(t + 1)$. Given the current μ_{kt}, draw $\mu_{k,t+1}$, $\forall k$, using parameterized Equation (1).
2. Generate a sample of the individual PDs at time $(t + 1)$. Given current λ_{it}, draw $\lambda_{i,t+1}$, $\forall i, k$, using the $\mu_{k,t+1}$ values from Step 1.
3. Update the rating category R_{it}, for each i, using the modified Bayesian rule. Here, we determine rating changes, if any. Then roll forward to the next period by repeating Step 1.

It is important during the simulation to account for the correlation between the various rating class shocks ε_{kt}, $k = 1, \ldots, K$, that may be computed from the residuals of Equation (1). Within a rating class, the correlation of issuer shocks ε_{it}, $i = 1, \ldots, K$, for a given k, is obtained from the residuals of Equation (2).

We summarize the set of system state variables during simulation in Exhibit 2. In each stage of the simulation, we keep track of the mean PD of each rating class k, the PD of each issuer i, and the rating of each issuer.

III. NUMERICAL FITTING OF SYSTEM TO HISTORICAL TRANSITIONS

Our data set comprises all rating transitions over the period from January 1993 through December 1997. There are 589 issuers in this data set. Following Moody's rating scheme, the issuers are divided into six rating categories. We begin with the initial rating and PD of each issuer, and the unconditional distributions of PDs for each rating class.

EXHIBIT 3

Fitted Rating Transition Matrix and Bias Parameters

Rating	Rating Category						
	1	2	3	4	5	6	Default (D)
1	0.9096	0.0864	0.0030	0.0009	0.0000	0.0000	0.0000
2	0.0058	0.9425	0.0507	0.0002	0.0006	0.0002	0.0000
3	0.0514	0.0042	0.9112	0.0002	0.0243	0.0084	0.0003
4	0.0723	0.0039	0.0008	0.8694	0.0003	0.0515	0.0018
5	0.0130	0.1397	0.0044	0.0017	0.7709	0.0638	0.0065
6	0.0026	0.0251	0.0253	0.0485	0.0003	0.8353	0.0628
D	0	0	0	0	0	0	1
γ_k	1.6789	4.1071	23.2983	23.4913	5.2505	2.2548	0.4065

From simulations based on PD data from Moody's.

Each month, we follow the change in PD for each issuer, and simulate the change in rating based on the modified Bayes rule in Equation (5). Over the entire period, the simulated rating changes are used to compute a rating transition matrix. We execute 50 simulation runs, and the fitted matrix is taken to be the average of transition matrices across all simulations.

This matrix depends on the set of bias parameters γ_k, $k = 1, \ldots, K$. We attempt to choose these parameters so that the *distance* between the empirical transition matrix and the simulated matrix is as small as possible. We define the distance to be the sum of the absolute difference in cell values between the two transition matrices. We undertake an intensive numerical search over the parameters to obtain the best fit transition matrix.

The fitted matrix is presented in Exhibit 3, and may be compared with the empirical one in Exhibit 1.

The bias parameters are also provided. Notice that for rating classes 1 to 6, the parameters $\gamma_k > 1$. Therefore, the parameters tend to bias the issuer toward remaining in the same rating category as far as possible, until sufficient updating has occurred to provide impetus for a rating change. This slows down the impact of drastic changes in the PDs, causing slower changes in ratings than in PDs.

Notice also that the γ_k are higher for the issuers in rating categories 3 and 4, i.e., the medium-quality ratings. This occurs because ratings are mean-reverting. Firms with high ratings tend to drift downward in quality, and firms with low ratings, conditional on survival, tend to drift higher in quality. Hence, firms in the middle of the quality spectrum are more likely to remain within their rating class, on a conditional basis from month to month.[3]

IV. CONCLUDING COMMENTS

The advent of models for computing probabilities of default (PD) has provided a supplementary measure of default likelihood in addition to credit ratings. Credit ratings are a coarser measure of default likelihood, and embed the same information as PDs plus a modicum of human judgment. Rating transitions tend to occur less frequently than PD changes, since the human judgment involved overrides temporary spikes in state variables driving PDs.

We have developed a Bayesian model based on PD changes to mimic rating changes. The free parameters in the model are tuned to historical data to fit the human judgment element in rating transitions.

The model is easy to implement. We generate a simulation-fitted transition matrix that mimics the historical empirical one closely. This lends support to the often-made argument that PDs may be used as sufficient statistics for rating changes. Rating agencies may use this model as a basis for proposing rating changes to credit analysts, and finally, portfolio managers may use the model to obtain forecasts of rating changes, based on the observed historical time series of firm PDs.

ENDNOTES

The authors are grateful for helpful discussions with Moody's Risk Management Services and Fitch Ratings. Das thanks a Breetwor Fellowship and the Dean Witter Foundation for support.

[1]Note that some models that generate rating changes depending on asset value changes still suffer from the deficiency of excessive ratings volatility.

[2]Based on square root diffusions, the means estimated for the six rating classes are: {0.0001, 0.0001, 0.0006, 0.0020, 0.0123, 0.0588}. The rates of mean reversion for each rating class are: {0.4813, 0.0018, 0.6680, 0.4811, 0.4305, 2.8350}. And the volatility parameters used are: {0.0054, 0.0102, 0.0353, 0.0241, 0.0307, 0.0778}. These are based on the data set available, and are illustrative of the implementation. Since there are close to 600 individual default intensities, we do not report the estimated parameters for Equation (2).

[3]This inference cannot be made directly from an inspection of the rating transition matrix, since it is the conditional matrix of one-year transitions, as opposed to conditional monthly transitions.

REFERENCES

Arvanitis, A., J. Gregory, and J-P. Laurent. "Building Models for Credit Spreads." *The Journal of Derivatives*, 6 (1999), pp. 27-43.

Bangia, Anil, Francis Diebold, and Til Schuermann. "Rating Migration and the Business Cycle, with Application to Credit Portfolio Stress Testing." Working paper, Oliver Wyman Institute, 1999.

Crosbie, P. "Modeling Default Risk." KMV Corporation, 1999.

Das, Sanjiv, and Peter Tufano. "Pricing Credit Sensitive Debt when Interest Rates, Credit Ratings and Credit Spreads Are Stochastic." *Journal of Financial Engineering*, 5 (2) (1996), pp. 161-198.

Finkelstein, Vladimir, Jean-Pierre Lardy, George Pan, Thomas Ta, and John Tierney. "CreditGrades Technical Document." RiskMetrics working paper, 2002.

Huge, Brian, and David Lando. "Swap Pricing with Two-Sided Default Risk in a Rating-Based Model." *European Finance Review*, 3 (1999), pp. 239-268.

Jarrow, Robert, David Lando, and Stuart Turnbull. "A Markov Model for the Term Structure of Credit Risk Spreads." *Review of Financial Studies*, 10 (2) (1997), pp. 481-523.

Kavvathas, Dimitrios. "Estimating Credit Rating Transition Probabilities for Corporate Bonds." Working paper, University of Chicago, 2001.

Kiesel, Rudiger, William Perraudin, and Alex Taylor. "The Structure of Credit Risk: Spread Volatility and Ratings Transitions." Working paper, Bank of England, 2001.

Kijima, Masaaki. "Monotonicities in a Markov Chain Model for Valuing Corporate Bonds Subject to Credit Risk." *Mathematical Finance*, 8 (3) (1998), pp. 229-247.

Kijima, Masaaki., and K. Komoribayashi. "A Markov Chain Model for Valuing Credit Risk Derivatives." *The Journal of Derivatives*, 6 (1) (1998), pp. 97-108.

Lando, David. "On Rating Transition Analysis and Correlation." *Credit Derivatives: Applications for Risk Management, Investment and Portfolio Optimization*. New York: Risk Publications, 1998, pp. 147-155.

——. "Three Essays on Contingent Claims Pricing." Ph.D. thesis, Cornell University, 1994.

Lando, David, and Torben Skødeberg. "Analyzing Rating Transitions and Rating Drift with Continuous Observations." *Journal of Banking and Finance*, 26 (2002), pp. 423-444.

Li, T. "A Model of Pricing Defaultable Bonds and Credit Ratings." Working paper, Olin School of Business, Washington University, 1998.

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

Nakazato, D. "Gaussian Term Structure Model with Rating Classes." Working paper, Industrial Bank of Japan, 1997.

Sobehart, J., R. Stein, V. Mikityanskaya, and L. Li. "Moody's Public Firm Risk Model: A Hybrid Approach To Modeling Short Term Default Risk." Moody's Investors Service, March 2000.