The Credit Rating Process and Estimation of Transition Probabilities: A Bayesian Approach

Catalina Stefanescu [a,*] , Radu Tunaru [b], Stuart Turnbull [c]

[a] *Management Science and Operations, London Business School, London, UK*

[b] *Cass Business School, London, UK*

[c] *Bauer College of Business, University of Houston, US*

Abstract

The Basel II Accord requires banks to establish rigorous statistical procedures for the estimation and validation of default and ratings transition probabilities. This raises great technical challenges when sufficient default data are not available, as is the case for low default portfolios. We develop a new model that describes the typical internal credit rating process used by banks. The model captures patterns of obligor heterogeneity and ratings migration dependence through unobserved systematic macroeconomic shocks. We describe a Bayesian hierarchical framework for model calibration from historical rating transition data, and show how the predictive performance of the model can be assessed, even with sparse event data. Finally, we analyze a rating transition data set from Standard and Poor's during 1981–2007. Our results have implications for the current Basel II policy debate on the magnitude of default probabilities assigned to low risk assets.

JEL classification: G21; G28; G32; C11; C13; C52

Key words: Ratings transitions, Bayesian inference, Latent factors, Markov Chain Monte Carlo

1 Introduction

The internal ratings based (IRB) approach in the New Basel Capital Accord (Basel II) allows banks to use their own internal credit ratings. Banks need to estimate the entire matrix of transition probabilities between rating

* Corresponding author. Management Science and Operations, London Business School, Regent's Park, London NW1 4SA, United Kingdom.

Email address: `cstefanescu@london.edu` (Catalina Stefanescu).

classes, and the Accord stresses that these probabilities must play an essential role in the calculation of regulatory capital, credit approval, risk management, internal capital allocation, and corporate governance functions of banks (Basel Committee on Banking Supervision, 2006b). For regulatory purposes, the Accord requires financial institutions to establish rigorous procedures for the validation of statistical models for internal ratings (Basel Committee on Banking Supervision, 2005). These procedures include out-of-sample tests, and they must make use of historical data over as long a time period as possible.

These requirements present a great technical challenge for many financial institutions that have a large number of high quality business lines for which extensive default data are not available. Such low default portfolios typically include exposures to sovereigns, large corporations, or financial institutions such as banks (in developed nations) and insurance companies, where very few defaults have been observed over long horizons. The scale of the issue is significant — in a joint industry survey of seven UK firms having nearly US $3 trillion in total gross assets, over 50% of total wholesale exposures had insufficient default data (British Bankers Association et al., 2004). Regulators expect that low default portfolios still follow minimum internal ratings based (IRB) standards for accuracy and conservatism of probabilities of default estimates, despite the data limitations. For low default portfolios, however, estimates of risk parameters based on simple historical averages or judgmental considerations alone, may underestimate capital requirements, raising the concern that financial institutions may not be able to apply the IRB approach for the many asset classes that have low number of defaults (British Bankers Association et al., 2004).

There are two main technical challenges related to low-default portfolios. The first issue is the estimation of default probabilities when no historical defaults have been recorded. Hamilton et al. (2007, Exhibit 21) report that over the period 1980–2006 there are sixteen years when there were no defaults for investment grade issues. However, none of these assets are default free, hence any reasonable model should assign a positive default probability. The second technical challenge for low default portfolios is the assessment of a model's predictive performance. The usual out-of-sample testing procedures (Shumway, 2001) cannot easily be applied in this case, since the zero realized default frequencies do not constitute a reasonable benchmark against which to compare the model's predictions.

This paper addresses these issues and makes three contributions. Our first contribution consists in developing a new model that describes the typical credit rating process that most major banks employ. In general, an obligor is assigned a credit rating based on an assessment of its current credit worthiness, which depends on many systematic and firm specific variables. The model includes the effects of a shared unobserved macroeconomic shock which induces dependence among transition probabilities for different credit classes in any given period. The model specification also allows for auto-correlation across time of transition probabilities from any credit class. Lastly, the model takes into account the heterogeneity in the credit worthiness of obligors in the same credit class., which can have significant effects on credit risk diversification (Hanson, Pesaran and Schuermann, 2008). It is difficult to calibrate our model with data from low default portfolios in a classical frequentist estimation framework, because the sparsity of data often leads to unrealistic transition probabilities. Therefore, we use instead a Bayesian hierarchical framework for model calibration, based on Markov Chain Monte Carlo (MCMC) techniques. The MCMC approach produces the inferred distribution for all parameters of interest, including credible intervals for transition probabilities, and it thus allows users of the methodology to directly address regulators' concerns about out-of-sample model testing (Basel Committee on Banking Supervision 2006a, §502). Within the Bayesian framework it is straightforward to model non-Markovian dynamics in ratings migrations, for which considerable empirical evidence exists (Altman and Kao, 1992; Nickell, Perraudin and Varotto, 2000; Bangia et al., 2002; Frydman and Schuermann, 2005). The Bayesian framework also offers a formal approach for taking into account expert opinion through the use of subjective prior distributions for the model parameters (Kiefer, 2007). This feature is important in the credit rating process where there is a large amount of non-quantifiable subjective information involved — senior credit risk officers often express opinions about the relative importance of certain inputs. Expert opinion gains even more weight in the case of low default portfolios, where there is a lack of objective historical transition data.

Our second contribution consists in addressing the difficult issue of assessing the predictive performance of a model when event data are sparse, such as is the case in low default portfolios. We employ two approaches to examine the predictive ability of a model. The first approach is based on a Bayesian measure of predictive power, the Deviance Information Criterion (DIC) developed by Spiegelhalter et al. (2002). When comparing the performance of several models, the model with the smallest DIC value is estimated to give the best predictions for a data set of the same structure as the data actually observed. The DIC measure has the advantage that it does not require the models to be nested for the purpose of comparison. Our second approach to investigating the predictive performance of a model is a variant of out-of-sample testing, taking into account not only the estimated transition probabilities but also the corresponding 95% credible intervals. For speculative grade classes where there is usually sufficient historical transition data, we compare observed default rates with the estimated default probabilities and their 95% credible intervals computed from the model. For investment grade classes where there are often no historical defaults, we can no longer use observed default rates as a benchmark. We thus compare instead the observed rates of staying in the same credit class with the estimated probabilities of no transitions and their 95% credible intervals computed from the model.

Our third contribution consists in applying our methodology to the analysis of an aggregate rating transitions data set from Standard and Poor's between 1981–2007, and deriving insights relevant for the current policy debate arising from Basel II. We calibrate different specifications of the credit rating process model and show that the estimated transition probabilities exhibit non-Markovian behavior, consistent with previous empirical evidence. We also find that the ratings transition matrix depends on the state of the economy, which can be partially described by two macroeconomic covariates — the return on the S&P 500 index and the Chicago Fed National Activities Index (CFNAI). As the CFNAI and S&P 500 return increase, the estimated default probabilities decrease for all credit classes as expected, and the effect of the macroeconomic conditions is generally larger for speculative grade than for investment grade classes. We find that the two macroeconomic variables, however, are not sufficient to capture the entire dynamics of the ratings transitions. The inclusion of a random unobserved macroeconomic shock significantly improves the predictive power of the credit rating process model, and at the same time accounts for the observed dependence among transition probabilities for different credit classes in any given period. We find that the estimated transition probability matrices are consistent with the monotonicity property, and that there is a potentially large heterogeneity among firms in the same credit class. We also find that the AAA rating class has very different dynamics than the other rating classes, and in particular that it is not sensitive to macroeconomic shocks. Finally, even in the absence of historical default data for top investment grade ratings, the credit rating process model always leads to positive estimated default probabilities in all credit classes, as required.

There is a considerable literature devoted to modelling and estimation of rating transition matrices. Lando and Skodeberg (2002) give a review of different approaches for estimating migration probabilities, which are extensively compared in Jafry and Schuermann (2004). Lando and Skodeberg (2002), and Christensen, Hansen and Lando (2004) address the issue of computing point and interval estimates for default probabilities with rare events, using a continuous time homogeneous Markov chain transition matrix. Christensen et al. (2004), Truck and Rachev (2005), and Hanson and Schuermann (2006) show that bootstrapped intervals for duration based estimates are relatively tight, however they are unable to distinguish default probabilities for investment grades. Pluto and Tasche (2005) address the issue of estimating the probability of default for low default portfolios by first specifying a confidence interval for the probability of default and then determining the maximum probability of default to be consistent with this confidence interval. Figlewski, Frydman and Liang (2008) investigate specifically conditioning on macroeconomic variables when modeling credit rating transitions and corporate default. None of the papers in this area addresses the issue of assessing the predictive performance of rating transition models, or comparing the models through out-of-sample testing for prediction purposes. In particular, the question of deciding whether any given model is appropriate for use in low default portfolios has, to our knowledge, never been addressed.

Our paper is perhaps closest to McNeil and Wendin (2006, 2007), who test several threshold Bayesian models with fixed and random effects using the traditional latent factor formulation for estimating transition probabilities. [1] There are two crucial differences between this paper and McNeil and Wendin (2006, 2007). First, the models described by McNeil and Wendin are variations of the traditional latent factor model commonly used in risk management. By contrast, the credit rating process model that we develop here is able to capture rating transition patterns that cannot be accounted for in the latent factor model framework. In particular, the probability of a transition in our model depends on the initial credit score assigned by a loan's officer. The credit rating process model thus naturally accounts for obligor heterogeneity within a credit class, since obligors have different initial credit scores even though they may have the same credit rating. Second, McNeil and Wendin (2006, 2007) do not address the issue of comparing models' predictive performance, which is crucial in light of the Basel II regulations. We describe in this paper a formal approach for assessing models' performance, using both out-of-sample testing and a Bayesian measure of predictive power. In particular, our approach addresses the issue of assessing whether a model is reasonable for the analysis of data from low default portfolios. For the Standard and Poor's data, we compare different specifications of the credit rating process model and of the traditional latent factor model using this approach, and find that the credit rating process model generally has superior predictive performance than the latent factor model.

When obligor level data are available, the credit rating process model and the MCMC Bayesian methodology are ideally suited to study credit rating momentum (Christensen, Hansen and Lando, 2004), and can be extended to investigate how Watchlist classifications affect rating behavior (Hamilton and Cantor, 2004). Many portfolio managers have mandates to track specified bond indices, such as the Lehman aggregate index. If an obligor is dropped from the index after a downgrade from an investment to a speculative rating, this triggers selling by portfolio managers, with resulting downward pressure on price. In an attempt to anticipate such events, managers often try to predict which obligors will be downgraded (upgraded). Our methodology can address this type of problem, as it allows the estimation of upgrade and downgrade probabilities at obligor level.

The paper is structured as follows. In Section 2 we develop the credit rating process model, we briefly discuss for comparison a traditional latent factor model for rating transitions, and we outline the Bayesian estimation framework. Section 3 contains a description of the Standard and Poor's rating transitions data set, and presents the results of our analyses. Section 4 concludes the paper with a discussion of some of the practical implications of our findings and with a summary of the results.

[1] Other applications of Bayesian techniques to credit issues can be found in Das, Fan and Geng (2002), Gössl (2005), Dwyer (2006), Kadam and Lenk (2007), and Farnsworth and Li (2007).

2 Model Specification and Estimation

In this section, we develop a new class of models that directly capture the credit rating process. These models can be used to describe the credit assessment changes for any individual obligor, provided that obligor specific data are available. For purposes of comparison, we then briefly discuss a second class of models arising from the traditional latent factor approach — see McNeil and Wendin (2007). Finally, we show how the parameters of both these classes of models are estimated in a Bayesian framework using Markov Chain Monte Carlo (MCMC) techniques.

Let $\{1, ..., \Omega\}$ be the set of non default rating classes represented in ascending order of credit worthiness. That is, credit class 1 is the lowest credit quality and credit class Ω the highest credit quality. The state of default is denoted by 0, and is assumed to be absorbing. Let $A_\zeta(t)$ denote the set of obligors in credit class ζ at time t. For both the credit rating process model and the latent factor model, we assume that there exists a latent variable representing the credit worthiness of each obligor at time t, such that the obligor is assigned to a particular credit class if the latent variable lies within a certain interval. Let $\gamma_{\zeta,0} < \gamma_{\zeta,1} < .... < \gamma_{\zeta,\Omega-1} < \gamma_{\zeta,\Omega} = \infty$ denote unobserved critical thresholds specific to each credit class. [2] An obligor who is in credit class ζ at time t will default at time $t+1$ if the latent variable representing its credit worthiness is less than $\gamma_{\zeta,0}$. The obligor will move to credit class ψ at time $t+1$ if the latent variable lies in the interval $[\gamma_{\zeta,\psi-1}, \gamma_{\zeta,\psi})$, for $\zeta, \psi = 1, ..., \Omega$. We assume that $\gamma_{\zeta,0} = 0$ for all ζ, to ensure identifiability. The lengths of the risk category intervals need not be equal, and it is expected that the obligors in a given risk category will exhibit roughly the same default risk.

2.1 The Credit Rating Process Model

In a typical internal credit rating system, many quantitative and qualitative variables are combined in order to form an assessment of the credit worthiness of an obligor over some defined horizon, which may be one year or through the credit cycle. Examples of quantitative variables include cash flow, liquidity, and leverage measures, while the qualitative variables include competitive advantages and disadvantages, industry risk and trends, management quality, legal and financial structure. A summary of the variables considered in a typical rating process is given in Crouhy, Galai and Mark (2000). For each obligor, a loan officer assigns a score to the different factors. Given the magnitude of the final score and using professional judgment, the loan officer

[2] The assumption that the critical thresholds $\{\gamma_{\zeta,\psi}\}$ are credit class specific is necessary to accommodate non-Markovian behavior in the credit rating process. In practice, it is well known that the rating process is not Markovian; the loan officers assign the obligor to a particular credit class based not only on its current state and on expectations about its future credit worthiness, but also on its rating history.

reaches an overall assessment of the obligor's credit worthiness and assigns a credit rating. The rating is then reported to the bank's risk management system used to compute the value-at-risk number and the economic capital. The use of a coarse rating system implies that a wide range of obligor scores receive the same rating, and there is obligor heterogeneity in any given rating class. The median score within any rating class will vary over the credit cycle.

The outside regulators are mandated to assess the underlying methodology and the reasonableness of ratings assigned to different obligors. In practice, the regulators often observe only the final outcome given by the assigned credit rating, rather than the relative importance attached to each variable and the obligor specific information used by the loan officers, which is usually confidential.[3] The credit ratings correspond to certain default probabilities in each credit class, as well as transition probabilities between credit classes. The task facing a risk manager assessing the properties of the internal rating process, is to justify that the assigned default and transition probabilities are reasonable, and to present this evidence to the regulators.

We consider an obligor j belonging to credit class ζ. The bank's rating methodology combines systematic and obligor specific variables to reach an assessment of the credit worthiness of the obligor. Formally, we model this process by a continuous unobserved random variable denoted by $D_{jt} \in \Re$, which represents the loan officer's credit assessment[4] for obligor j at time t. Over the next period $t+1$, the systematic and obligor specific variables will change and this will affect the assessment of the obligor's credit worthiness. Some of these changes will be expected, while others will be unanticipated. We assume that the credit assessment for obligor j at time $t+1$ is given by

$$D_{jt+1} = \mu_{jt} + D_{jt} + \beta_j' w_{t+1} + e_{jt+1}. \tag{1}$$

Here μ_{jt} is a drift term representing the change in the credit assessment caused by expected changes in the systematic and idiosyncratic variables; w_t is a n-dimensional vector of zero mean systematic random shocks, $\beta_j \in \Re^n$ is a vector of coefficients, and e_{jt+1} is a m-dimensional vector of zero mean idiosyncratic random variables.[5]

The score D_{jt+1} at time $t+1$ depends on the score D_{jt} at time t, inducing auto-correlation in the latent credit worthiness of the obligor. For any obligor

[3] A bank is required to fully document the rating methodology and to describe its actual use. Regulators can request to see the raw obligor scores. However, given the number of obligors in a typical loan portfolio (usually measured in thousands), it is rare for regulators to see all the scores.

[4] The quantity reported to the bank's risk management system is not D_{jt}, but the credit rating assigned to the obligor. For this reason, we treat D_{jt} as unobserved.

[5] At the individual obligor level, expression (1) arises naturally from the assumption that the loan officer considers systematic variables, say F, and obligor specific variables, say f, for credit assessment purposes, so that $D = D(F, f)$. Under the assumption that these variables are described by diffusion processes, we have

$j \in A_\zeta(t)$, the obligor's rating at time $t+1$ depends on the position of its latent score D_{jt+1} relative to the set of thresholds $\gamma_{\zeta,0} < \gamma_{\zeta,1} < \ldots < \gamma_{\zeta,\Omega}$. For a typical bank there are many thousands of rated obligors and loan facilities that are part of the credit portfolio. Many obligors with very different characteristics may therefore receive the same rating, particularly when the bank uses a coarse rating system, leading to obligor heteroegeneity. Expression (1) implicitly accounts for heterogeneity within each credit class, since the credit assessment D_{jt+1} of obligor j at time $t+1$ depends on the assessment D_{jt} at time t, reflecting the fact that obligors with different credit scores at time t may belong to the same credit class.

Expression (1) describes the general form of the credit rating process model. Different distributional assumptions and different specifications for the structure of the systematic random shocks $\{w_t\}$ are possible, and this greatly increases the versatility of the model. In particular, the systematic random shocks induce dependence among rating transitions; this can be dependence among transitions taking place in the same time period (if w_t are independent), or dependence across time periods (if w_t are serially correlated and follow, for example, an autoregressive process). Within the multivariate structure of the systematic random shocks, it is also possible to account for sector, industry, or credit class effects by making some components of w_t sector specific or credit class specific. This flexibility enhances the model's ability to describe different patterns of heterogeneity and dependence. We discuss several special cases of the model later in this section, and we empirically test them in Section 3.2.

Representative Obligor

There are many situations where obligor specific information is unavailable. For example, when examining the properties of transition matrices published by rating agencies, the identities of the obligors in the data set are unknown. Also, since the credit risk management function in a bank is separate from the loans department, specific obligor information is often unavailable to the risk management group when testing the relevance of an internally generated transition matrix, even though the credit rating and the obligor identity may

---

$$
\begin{aligned}
dF &= \mu_F dt + \sigma_F dW_F, \\
df &= \mu_f dt + \sigma_f dW_f,
\end{aligned}
$$

where $\mu_F = \mu_F(F,T)$, $\sigma_F = \sigma_F(F,T)$, $\mu_f = \mu_f(f,T)$, $\sigma_f = \sigma_f(f,T)$, and W_F and W_f are vectors of independent Brownian motions. More complicated processes could also be assumed. Ito's lemma then implies that

$$
dD = \mu_D dt + \frac{\partial D}{\partial F}\sigma_F dW_F + \frac{\partial D}{\partial f}\sigma_f dW_f,
$$

using vector notation. Integrating the above expression, we derive an expression for the stochastic process describing the change in the credit assessment.

be known. This is particularly the case for private firms that constitute the majority of obligors for retail banks. In cases where obligor specific information is unavailable, it is necessary to impose restrictions on expression (1). The simplest way is first to assume that all obligors within a credit class are identically distributed — see McNeil and Wendin (2007). This implies that $\mu_{jt} = \mu_{\zeta t}$ for all obligors j in credit class ζ. Second, we assume that $\mu_{\zeta t}$ only depends on current macroeconomic conditions, since no obligor specific information can be observed. Thus $\mu_{\zeta t} = \alpha'_\zeta X_t$, where X_t is a n-dimensional vector of observable macroeconomic covariates (e.g., credit spread, term spread, or percentage change in GDP), and $\alpha_\zeta \in \Re^n$ is a vector of coefficients. Third, we assume that $\beta_j = \beta_\zeta$, for all obligors j in credit class ζ. Fourth, we assume that there is only one obligor specific idiosyncratic random variable, e_{jt+1}. Under these assumptions, for any obligor j in credit class ζ we have

$$D_{jt+1} = \mu_{\zeta t} + D_{jt} + \beta'_\zeta w_{t+1} + e_{jt+1}. \tag{2}$$

We cannot observe the obligor specific score, or the distribution of scores within the assigned credit class. We address this issue by introducing the notion of a representative obligor in credit class ζ at time t. Let $R_{\zeta t}$ be the score of the representative obligor [6] at time t. We replace D_{jt} with $R_{\zeta t}$, so that equation (2) becomes

$$D_{jt+1} = \mu_{\zeta t} + R_{\zeta t} + \beta'_\zeta w_{t+1} + e_{jt+1}. \tag{4}$$

Between time t and $t+1$, the scores of obligors in credit class ζ change, and ratings are reassessed at time $t+1$ when credit class membership may change. Define $R^-_{\zeta t+1}$ to be the unobserved value of the representative obligor's score just before credit ratings are reassessed at time $t+1$. This will be a function of $R_{\zeta t}$, the representative obligor's score at time t, and of changes in the macroeconomic conditions, so that we have [7]

$$R^-_{\zeta t+1} = \mu_{\zeta t} + R_{\zeta t} + \beta'_\zeta w_{t+1}. \tag{5}$$

The macroeconomic changes in the right-hand side of (5) are either antici-

[6] For instance, the score of the representative obligor may be defined as a weighted average

$$R_{\zeta t} = \sum_{j \in A_\zeta(t)} \nu_{j,t} D_{jt}, \tag{3}$$

where the weights sum to unity, $\sum_{j \in A_\zeta(t)} \nu_{j,t} = 1$ for all t and ζ.

[7] This expression can be derived as follows. From the definition of $R^-_{\zeta t+1}$, we can write

$$R^-_{\zeta t+1} = \sum_{j \in A_\zeta(t)} \nu_{j,t} D_{jt+1}.$$

From expressions (2) and (3) and the fact that the weights sum to unity, it follows that pated or unanticipated. As before, the term $\mu_{\zeta t} = \alpha'_\zeta X_t$ captures the anticipated changes modelled through the effect of observed covariates X_t. The term $\beta'_\zeta w_{t+1}$ represents the unanticipated macroeconomic changes. The macroeconomic shocks $\{w_t\}$ account for unobserved systematic risk, and therefore model heterogeneity beyond that which can be described with the observed covariates X_t.

The score at time $t+1$ of an obligor who is in credit class ζ at time t depends on effects of idiosyncratic events and on the impact of macroeconomic changes upon credit class ζ obligors. Therefore, we relate the score of the individual obligor to the representative obligor through the equation

$$D_{jt+1} = R^-_{\zeta t+1} + e_{jt+1}. \tag{6}$$

Substituting (5) in (6) we obtain equation (4).

Transition Probabilities

Obligor $j \in A_\zeta(t)$ defaults at time $t+1$ if $D_{jt+1} < \gamma_{\zeta,0}$. From expression (4) it follows that the conditional probability of default over the next period given $R_{\zeta t}$ and w_{t+1} is

$$\begin{aligned} P(D_{jt+1} < \gamma_{\zeta,0} \,|\, R_{\zeta t}, w_{t+1}) \\ &= P(e_{jt+1} < \gamma_{\zeta,0} - \mu_{\zeta t} - R_{\zeta t} - \beta'_\zeta w_{t+1} \,|\, R_{\zeta t}, w_{t+1}) \\ &= g(\gamma_{\zeta,0} - \mu_{\zeta t} - R_{\zeta t} - \beta'_\zeta w_{t+1}), \end{aligned}$$

where $g(\cdot)$ is a link function that depends on the distribution of e_{jt}. For example, $g(\cdot)$ is the logit link if e_{jt} has an Extreme Value distribution, and $g(\cdot)$ is the probit link if e_{jt} has a Normal distribution. Note that we do not observe $R_{\zeta t}$ and w_{t+1}. The unconditional probability of default for obligor j over the next period is

$$P(D_{jt+1} < \gamma_{\zeta,0}) = E_t[P(D_{jt+1} < \gamma_{\zeta,0} \,|\, R_{\zeta t}, w_{t+1})],$$

where the expectation is taken with respect to the distributions of $R_{\zeta t}$ and w_{t+1}. Thus the probability of default assigned to credit class ζ at time t is

$$p_t(\zeta, 0) = E_t[g(\gamma_{\zeta,0} - \mu_{\zeta t} - R_{\zeta t} - \beta'_\zeta w_{t+1})]. \tag{7}$$

The conditional probability of a transition from credit class ζ at time t to credit class ψ at time $t+1$ is given by

---

$$\begin{aligned} R^-_{\zeta t+1} &= \sum_{j \in A_\zeta(t)} \nu_{j,t}[\mu_{\zeta t} + D_{jt} + \beta'_\zeta w_{t+1} + e_{jt+1}] \\ &= \mu_{\zeta t} + R_{\zeta t} + \beta'_\zeta w_{t+1}. \end{aligned}$$

Here we neglect the term $\sum_{j \in A_\zeta(t)} \nu_{j,t} e_{jt+1}$, which is justified by appealing to the law of large numbers. See Gordy (2003).

$$
\begin{aligned}
&P(\gamma_{\zeta,\psi-1} \leq D_{jt+1} < \gamma_{\zeta,\psi} \mid R_{\zeta t}, w_{t+1}) \\
&\quad = P(\gamma_{\zeta,\psi-1} \leq \mu_{\zeta t} + R_{\zeta t} + \beta'_{\zeta} w_{t+1} + e_{jt+1} < \gamma_{\zeta,\psi} \mid R_{\zeta t}, w_{t+1}) \\
&\quad = g(\gamma_{\zeta,\psi} - \mu_{\zeta t} - R_{\zeta t} - \beta'_{\zeta} w_{t+1}) - g(\gamma_{\zeta,\psi-1} - \mu_{\zeta t} - R_{\zeta t} - \beta'_{\zeta} w_{t+1}),
\end{aligned}
$$

and the unconditional transition probability is then

$$
p_t(\zeta, \psi) = E_t[g(\gamma_{\zeta,\psi} - \mu_{\zeta t} - R_{\zeta t} - \beta'_{\zeta} w_{t+1})] - E_t[g(\gamma_{\zeta,\psi-1} - \mu_{\zeta t} - R_{\zeta t} - \beta'_{\zeta} w_{t+1})]. \tag{8}
$$

Note that the right-hand side of expression (8) is bounded between zero and one, since $\gamma_{\zeta,\psi} > \gamma_{\zeta,\psi-1}$ and $g(.)$ is a cumulative distribution function.

Special Cases

We next investigate several special cases of the credit rating process model that depend on the specifications of the latent shocks w_t and on the parameter constraints on α_ζ and β_ζ. We empirically fit these models in Section 3.2.

Mixed model

We assume that the macroeconomic shocks $\{w_t\}$ are independent and identically normally distributed. In general, $w_t = (w_{t1}, \ldots, w_{tn})$ is an n-dimensional vector, where n is the number of observable macroeconomic covariates in the model, and the corresponding vector of effects on credit class ζ is $\beta_\zeta = (\beta_{\zeta 1}, \ldots, \beta_{\zeta n})$. However, since the shocks are unobservable, not all their effects are identifiable on the basis of observed data. If the probability of default in credit class ζ increases, this could be due to the fact that $\beta_{\zeta 1} w_{t1} + \beta_{\zeta 2} w_{t2} + \ldots + \beta_{\zeta n} w_{tn}$ decreases, but it is not possible to identify the individual effects of $w_{t1}, w_{t2}, \ldots, w_{tn}$. Therefore, we shall impose the constraint that the effects of all macroeconomic shocks for the same credit class are equal, $\beta_{\zeta 1} = \ldots = \beta_{\zeta n}$, which is equivalent to assuming that the shock w_t is univariate. The model thus becomes:

$$
(R1): \quad
\begin{aligned}
&D_{jt+1} = \textstyle\sum_{i=1}^{n} \alpha_{\zeta i} X_{ti} + R_{\zeta t} + \beta_\zeta w_{t+1} + e_{jt+1}, \text{ for all } j \in A_\zeta(t) \\
&\qquad\qquad w_{t+1} \sim N(0, 1).
\end{aligned}
$$

Model (R1) implies dependence among default probabilities in different credit classes in the same time period through the shared macroeconomic shock w_{t+1}. The model also accounts for dependence of default probabilities for the same credit class across time periods, since the credit assessment at time $t+1$ depends on the assessment at time t.

Covariates only model

We wish to test whether the observed dynamics of ratings transitions can be completely explained by anticipated changes in the macroeconomic conditions. In other words, are the observed covariates X_t sufficient for explaining the ratings transitions, and are the unobserved macroeconomic shocks w_t redundant? This is equivalent to the constraint $\beta_\zeta = 0$ for all $\zeta = 1, \ldots, \Omega$, implying that the current values X_t of the macroeconomic covariates may affect the expected credit assessment, but unanticipated changes in w_t have no impact on the credit score. The model thus becomes:

$$(R2): \quad D_{jt} = \sum_{i=1}^{n} \alpha_{\zeta i} X_{ti} + R_{\zeta t} + e_{jt+1}, \text{ for all } j \in A_\zeta(t).$$

Note that in the absence of a shared macroeconomic shock w_t, model (R2) cannot account for dependence of default probabilities in different credit classes in the same time period. If model (R2) provides a better fit to the observed transition data than model (R1), this would support the hypothesis that ratings transitions can be completely explained by anticipated changes in the economy.

Independent shock only model

Conversely, we wish to test whether the dynamics of ratings transitions can be explained at least in part by expected changes in the macroeconomic conditions, or whether it is only unexpected changes in the economy that determine the transition probabilities. In other words, are the unobserved macroeconomic shocks $\{w_t\}$ sufficient for explaining the ratings transitions, and are the observed covariates X_t redundant? This is equivalent to the constraint $\alpha_\zeta = 0$ for all $\zeta = 1, \ldots, \Omega$, implying that the observed macroeconomic covariates have no additional impact on the credit score after the unobserved shock has been taken into account. Again, we assume that w_t is univariate, independent and identically normally distributed, so that the model becomes:

$$(R3): \quad \begin{aligned} D_{jt+1} &= R_{\zeta t} + \beta_\zeta w_{t+1} + e_{jt+1}, \text{ for all } j \in A_\zeta(t) \\ w_{t+1} &\sim N(0,1). \end{aligned}$$

If model (R3) provides a better fit to the observed transition data than model (R1), this would support the hypothesis that ratings transitions are entirely driven by unanticipated changes in the economy.

Autoregressive shock only model

Note that models (R1)–(R3) only allow for dependence of default probabilities in the same credit class across different time periods. They do not allow for dependence of default probabilities in different credit classes across different periods, for example, default probabilities in credit class ζ_1 at time t_1 and in credit class ζ_2 at time t_2, with $t_1 \neq t_2$. These patterns of dependence are induced by the shared unobserved shocks $\{w_t\}$, hence in order to model them across time we need to specify an autoregressive AR(1) structure for w_t. Again, in order to simplify the model we assume that there is no effect of observed macroeconomic covariates ($\alpha_\zeta = 0$ for all ζ), so that the model becomes

$$(R4):\quad \begin{aligned} D_{jt+1} &= R_{\zeta t} + \beta_\zeta w_{t+1} + e_{jt+1}, \text{ for all } j \in A_\zeta(t) \\ w_{t+1} &= a_w w_t + \varepsilon_{t+1}, \end{aligned}$$

where $a_w \in (-1,\, 1)$ to ensure stationarity, and $\varepsilon_{t+1} \sim N(0,1)$ are independent. Model (R4) captures dependence of rating transitions across time periods, induced by the serial correlation of the macroeconomic shocks w_t.

2.2 The Latent Factor Model

A second class of models for rating transitions relies on the traditional linear latent factor approach that is used both for risk management and for pricing credit derivatives (Schönbucher, 2003; Gagliardini and Gourieroux, 2005; McNeil, Frey and Embrechts, 2005, Chapter 8; McNeil and Wendin, 2006, 2007). We briefly describe this approach here.

Denote by $L_{\zeta t}$ a latent variable representing the credit worthiness of the representative obligor in credit class ζ at time t. We assume that at time $t+1$ the latent $L_{\zeta t+1}$ is related to a vector of common factors z_{t+1} through the expression

$$L_{\zeta t+1} = \mu_{\zeta t} + \beta'_\zeta z_{t+1}, \tag{9}$$

where $\mu_{\zeta t}$ is a drift term representing the effect of observed macroeconomic covariates, and $\beta_\zeta \in \Re^n$ is the sensitivity to the common latent factors z_{t+1}. We assume that for any obligor j in credit class ζ, the latent credit assessment L_{jt+1} is related to the latent credit assessment $L_{\zeta t+1}$ for the representative obligor through the expression

$$L_{jt+1} = L_{\zeta t+1} + e_{jt+1} = \mu_{\zeta t} + \beta'_\zeta z_{t+1} + e_{jt+1}, \tag{10}$$

where e_{jt+1} is an unobservable idiosyncratic random effect.[8] As in the previous subsection, the model can be calibrated at individual obligor level rather than credit class level if obligor specific data are available.

The conditional probability of default over the next period assigned to credit class ζ at time t is given by

$$\begin{aligned} P(L_{jt+1} < \gamma_{\zeta,0} \,|\, z_{t+1}) &= P(\mu_{\zeta t} + \beta'_\zeta z_{t+1} + e_{jt+1} < \gamma_{\zeta,0} \,|\, z_{t+1}) \\ &= g(\gamma_{\zeta,0} - \mu_{\zeta t} - \beta'_\zeta z_{t+1}), \end{aligned}$$

where $g(\cdot)$ is, as before, a link function that depends on the distribution of

[8] The model discussed by Heitfield (2005) is an extension of expression (10), including obligor specific covariates.

e_{jt+1}. The unconditional probability of default is then

$$p_t(\zeta, 0) = E_t[g(\gamma_{\zeta,0} - \mu_{\zeta t} - \beta'_\zeta z_{t+1})], \tag{11}$$

where the expectation is taken with respect to the distribution of z_{t+1}. The probability of a transition from credit class ζ at time t to credit class ψ at time $t+1$ is given by

$$p_t(\zeta, \psi) = E_t[g(\gamma_{\zeta,\psi} - \mu_{\zeta t} - \beta'_\zeta z_{t+1})] - E_t[g(\gamma_{\zeta,\psi-1} - \mu_{\zeta t} - \beta'_\zeta z_{t+1})]. \tag{12}$$

Note that the latent factor model captures the impact of economic factors (observed or latent) on the absolute level of the credit worthiness variable, while the credit rating process model captures the impact of economic factors on the relative change in the credit worthiness from one period to the next.

Similar to the credit rating process model, the value of the latent credit worthiness L_{jt+1} in the latent factor model depends on a drift term $\mu_{\zeta t}$, on changes in the macroeconomic factors z_{t+1}, and on the idiosyncratic random effect e_{jt+1}. However, unlike in the credit rating process model, the latent factor model does not account for the fact that obligors have different initial credit scores even though they may have the same credit rating, and therefore it does not directly reflect obligor heterogeneity within a credit class.

Special Cases

We investigate here two special cases of the latent factor model which we empirically test in Section 3.2.

Mixed latent factor model

A special case of the latent factor model is obtained under the assumption that the factor z_{t+1} is univariate, independent and identically normally distributed, so that the model becomes:

$$(L1): \quad \begin{array}{c} L_{jt+1} = \sum_{i=1}^{n} \alpha_{\zeta i} X_{ti} + \beta_\zeta z_{t+1} + e_{jt+1}, \text{ for all } j \in A_\zeta(t) \\ z_{t+1} \sim N(0,1). \end{array}$$

Model (L1) can account for dependence of rating transitions in the same time period through the latent factor z_{t+1} shared by all credit classes. However, it cannot accommodate dependence of rating transitions across time periods.

Factor only model

Another special case of the latent factor model is obtained under the assumption that there is no effect of observed macroeconomic covariates X_t, so that $\alpha_{\zeta i} = 0$ for all $\zeta = 1, \ldots, \Omega$ and $i = 1, \ldots, n$. The model thus becomes:

$$(L2): \quad \begin{aligned} L_{jt+1} &= \beta_\zeta z_{t+1} + e_{jt+1}, \text{ for all } j \in A_\zeta(t) \\ z_{t+1} &\sim N(0,1). \end{aligned}$$

We test the fit of models (L1) and (L2) to the observed transition data in Section 3.2.

2.3 Bayesian Estimation

We describe in this section a framework for Bayesian estimation of the parameters of the credit rating process model discussed in Section 2.1. This framework can be easily adapted for estimation of the parameters of the latent factor model from Section 2.2 as well.

Let us denote by θ the vector of parameters of the model, which includes $\{\alpha_\zeta\}$, $\{\beta_\zeta\}$, and the parameters of the distribution of w_t. Standard likelihood based inference for the credit rating process model specified in (4) is difficult to achieve, because the multivariate structure and serial dependence of the unobserved macroeconomic shocks lead to joint migrations distributions in the form of multivariate integrals lacking closed form expressions. As an alternative to maximum likelihood inference, we use a Bayesian approach to estimating the parameters of the models from sample data, which can be implemented in a Markov chain Monte Carlo (MCMC) framework. An introduction to MCMC techniques is given by Gilks, Richardson and Spiegelhalter (1996).

A Bayesian specification requires prior distributions to be chosen for θ. Let $p(\theta)$ be the probability density of the prior distribution of θ, and let $\mathcal{Y}$ be the sample data. The joint posterior density $p(\theta \mid \mathcal{Y})$ of all parameters given the observed data is proportional to the product of the likelihood function $p(\mathcal{Y} \mid \theta)$ and the prior density:

$$p(\theta \mid \mathcal{Y}) \propto p(\mathcal{Y} \mid \theta) \cdot p(\theta). \tag{13}$$

We make the standard assumption that the model parameters are a priori independent, so that the prior density $p(\theta)$ is a product of prior densities for each parameter. With little external information available, we generally would like to specify non–informative priors $p(\cdot)$ for the components of θ. For instance, in the application described in Section 3 we specify normal priors with large variances for the parameters α_ζ and β_ζ, and gamma priors with large variances for the threshold increments $\gamma_{\zeta,\psi+1} - \gamma_{\zeta,\psi}$. However, if expert opinion is available, it can be incorporated into the analysis by specifying more concentrated priors. For example, a downturn in the economy that may lead to a general increase in the defaults for all rating categories, can be accounted for in the model by changing the priors for β_ζ to plausible ranges.

In a Bayesian framework, inference on model parameters is based on their marginal posterior density obtained by integrating out the other parameters from the joint posterior density given by (13). This is difficult to achieve analytically, therefore we propose the use of Gibbs sampling (Geman and Geman, 1984) for generation of the marginal posterior distributions. The Gibbs sampler is an iterative algorithm for generation of samples from a multivariate distribution. It proceeds by updating each variable by sampling from its conditional distribution given current values of all other variables. After a sufficiently large number of iterations, under mild conditions it can be proven that the values of the updated variables so obtained form a sample from the joint posterior distribution.

After convergence of the Gibbs sampler, 95% credible intervals for all model parameters can be computed from the samples of observations generated from the posterior densities, and these can be then used to test specific hypotheses about the parameters. In the application described in Section 3, we use the Deviance Information Criterion (DIC) to choose among different models fitted to the same data set, following Spiegelhalter et al. (2002). This criterion is a Bayesian alternative to Akaike's Information Criterion (AIC), and it is an estimate for the expected predictive deviance which has been suggested as a measure of model fit when the goal is to choose a model with best out-of-sample predictive power.

The DIC is defined as follows. Consider first the deviance defined as usually through

$$dev = -2\log(\text{likelihood}) = -2\log p(\mathcal{Y} \mid \theta). \tag{14}$$

Denote by $\overline{dev}$ the posterior mean of the deviance, and by $\widehat{dev}$ the point estimate of the deviance computed by substituting the posterior means $\widehat{\theta}$ of theta in (14). Thus $\widehat{dev} = -2\log p(\mathcal{Y} \mid \widehat{\theta})$. Denote by pD the *effective number of parameters* defined as the difference between the posterior mean of the deviance and the deviance of the posterior means:

$$pD = \overline{dev} - \widehat{dev}.$$

Then the Deviance Information Criterion is given by

$$DIC = \overline{dev} + pD.$$

Note that DIC is defined similarly with the Akaike Information Criterion. The model with the smallest DIC value is estimated to be the model that would best predict a data set of the same structure as the data actually observed. An advantage of the DIC measure is that it can be used to compare models that are not nested, as we exemplify in the next section. It is, however, a difficult task to define what constitutes an important difference in DIC values between models; the distribution of DIC is not known and thus no formal hypothesis testing can be done. Spiegelhalter et al. (2002) propose the following rule of thumb: if the difference in DIC is greater than 10, then the model with a larger DIC value has considerably less support than the model with a smaller DIC value.

3 Analysis of Standard and Poor's Data

In this section, we first describe the data set used in the study. These data are not obligor specific, therefore we use the representative obligor forms (4) and (10) of the credit rating and latent factor models. We test different specifications of these models and discuss the results and implications of our analysis.

3.1 Data Description

The source of the data set that we analyze here is an aggregated version of the CreditPro 7.72 database of long–term local currency issuer credit ratings. The database pooled the information from 13,162 companies that were rated by Standard and Poor's as of 31 December 1980, or that were first rated between that date and 31 December 2007. These rated issuers include industrials, utilities, financial institutions, and insurance companies around the world. [9] Public information ratings as well as ratings based on the guarantee of another company were not taken into consideration. The data also did not include structured finance vehicles, public–sector issuers, subsidiaries whose debt is fully guaranteed by a parent or whose default risk is considered identical to that of their parents, as well as sovereign issuers.

The data set contains information on seven rating categories: AAA, AA, A, BBB, BB, B, and CCC/C. The data consists of static pool one-year transition matrices, with the number of issuers and the number of transitions between each pair of rating categories (including default) available for every year during the 27 year horizon. [10] The data available for this study records transitions at the aggregate credit class level rather than at obligor level. As such, it is not possible to identify individual obligor's transitions during the horizon. For this reason, the results discussed in this section are only based on calibrating the representative obligor form of the credit rating process model based on equation (4), rather than the individual obligor form given by equation (1).

In this data set, a default has been recorded by Standard and Poor's upon the first occurrence of a payment default on any financial obligation, rated or

[9] Since we use two US macroeconomic variables as explanatory covariates, as we describe later in this section, the fact that the ratings transitions include non-US obligors as well raises a problem. This may lower model performance and reduce the power for finding statistically significant covariate effects, to the extent that the credit worthiness of foreign obligors is less sensitive than that of US obligors to the US macroeconomic climate. Our results would be strengthened if individual obligors could be identified and perhaps additional relevant macroeconomic covariates were included. Our data, however, is at an aggregate level and does not allow the identification of individual obligors.

[10] The CreditPro 7.72 database records individual obligor transitions. However, for this study we only had access to an aggregated version of the database rather than the full database.

unrated, other than a financial obligation subject to a bona fide commercial dispute.[11] The classification of an issuer into a credit rating category reflects Standard and Poor's opinion of a company's overall capacity to pay its obligations, focusing on the obligor's ability and willingness to meet its financial commitments on a timely basis. The rating generally indicates the likelihood of default regarding all financial obligations of the firm. Note, however, that a company may not have rated debt but it may still be assigned a credit rating. A Standard and Poor's rating reflects a through-the-cycle assessment of the credit risk of an obligor. Moreover, the agencies assign ratings based on a "stress scenario" for the borrowers, therefore the estimate is close to the estimate of the borrower's default probability at the time of rating assignment only if the borrower already is in the stress scenario (Carey and Hrycay, 2001). This implies that Standard and Poor's will overestimate the credit risk of the obligor when the economy is in a good state, and underestimate it when the economy is in a bad state.

Standard and Poor's report acknowledges that their ongoing enhancement of the CreditPro database from which this data is extracted may lead to outcomes that differ to some extent from those reported in Standard and Poor's previous studies. The data set that we analyze here is the latest version of Standard and Poor's data.

In our analysis we investigated several macroeconomic covariates: the credit spread, the term spread, the three month Treasury yield, the growth in real GDP, the growth in personal income, the return on the S&P 500 index, and the Chicago Fed National Activities Index (CFNAI). Among these covariates, we found that the CFNAI and the return on the S&P 500 index had the highest explanatory power, therefore in the results described in this section we report only the insights based on these two macroeconomic covariates. The CFNAI is the index of national economic activity developed in Stock and Watson (1999), and constructed from the first principal component of 85 economic series. Stock and Watson (1999) found that this single index can be used to obtain good forecasts of inflation and of overall economic activity, and CFNAI has also been used as explanatory variable in McNeil and Wendin (2006, 2007). The return on the S&P 500 index is the cumulated monthly return on the index over the year, and it has been used as a covariate in Chava, Stefanescu and Turnbull (2006). Figure 1 gives the plot of the time series values of the two macroeconomic covariates. During the initial period 1981–1988 the covariates appear to be negatively correlated, while in the last period 1995–2007 they appear to be positively correlated. The plot thus suggests that CFNAI and the S&P 500 return capture, as expected, different characteristics of the economy's performance. On average, large values of CFNAI and of the S&P 500 return indicate that the economy is doing well, and the average credit worthiness of obligors should be high. Consequently, we expect positive coefficients for these

[11] This is not true when an interest payment missed on the due date is made within the grace period.

covariates in the credit rating models.

[Figure 1 about here.]

3.2 Empirical Results

We calibrate the models developed in Section 2 using the Standard and Poor's data and the Bayesian methodology described in Section 2.3. We test the four different specifications (R1)–(R4) of the credit rating process model, and the two specifications (L1)–(L2) of the latent factor model. Throughout the study, we choose $g(\cdot)$ to be the logit link function corresponding to the logistic distribution for the idiosyncratic terms $e_{\zeta t}$. Other link functions (e.g. probit or log-log) could be chosen, but our investigations showed that this would have a minimal impact on the results of the analysis.

We used WinBUGS version 1.4 (Spiegelhalter et al., 2003) for model calibration. We specified diffuse but proper priors for all parameters, however other priors are also possible if specific prior information on some parameters is available. We chose $N(0, 10^3)$ priors for parameters α_1, α_2 and β_ζ representing the effects of the macroeconomic covariates (CFNAI and S&P 500 return) and of the macroeconomic shock w_t. To assess the impact of the choice of prior variance, we carried out a sensitivity analysis by investigating different prior normal distributions with variances ranging between 10^3 and 10^6. The results were broadly similar, hence here we report only the summary statistics based on the $N(0, 10^3)$ priors. We also specified a gamma prior with large variance $\Gamma(.001, .001)$ for the positive threshold increments $\gamma_{\zeta,\psi+1} - \gamma_{\zeta,\psi}$. For the autoregressive parameter a_w in model R4 we chose a uniform prior on $(-1, 1)$. These are standard choices of non-informative prior distributions.

For each model we ran two parallel Markov chains started with different sets of initial values. The Gibbs sampler ran for 10,000 iterations, with the first 5,000 iterations discarded as a burn–in period. Gelman and Rubin's diagnostic (Gelman et al., 1995) indicated satisfactory convergence of all chains. After convergence, inference on the parameters of interest was based on the pooled sample iterations of both chains.

In-Sample Model Comparison

The first aim of the analysis is to compare measures of model fit and decide which of the models best captures the different patterns of heterogeneity and migration dependence observed in the transition data. Table 1 reports the values of the Deviance Information Criterion (DIC) for each model. Model R1 has the lowest DIC value, implying that this model, with two observed macroeconomic covariates and the independent macroeconomic shock w_t, would best predict a data set of the same structure as Standard and Poor's data.

[Table 1 about here.]

The fit of R1 as measured by DIC is significantly better [12] than that of the latent factor model L1 with independent factors. The fit of R1 is also better than the fit of R4, which includes an AR(1) specification for the macroeconomic shock w_t. Model R1 has better predictive power because the time dependence of transition probabilities is inherently taken into account in the credit rating model, without the necessity of specifying an autoregressive structure for w_t. For the credit rating process model, the right hand side of equation (4) includes $R_{\zeta t}$, hence the credit worthiness at time $t+1$ directly depends on the value at time t of the unobserved credit worthiness of the representative obligor. Therefore the credit rating process model implicitly accounts for the auto-correlation in the credit worthiness variable across time. By contrast, in the latent factor model specified by (10), the credit worthiness at time $t+1$ depends only on values at time $t+1$ of the covariates and latent factors. Hence in this case the only solution for capturing time dependence of migrations is to include serial dependence in the latent factor z_t.

Note that models R3 and R4 are closest to R1 in terms of predictive performance among the credit rating class models. Among the latent factor models, the performance of L1 is superior to that of L2. However, the performance of R1 is considerably better than that of R3, R4 and L1. Therefore, for the remaining discussion of the analysis insights, we present the results only for model R1.

Model Estimation

Table 2 reports the posterior means and standard errors for model R1 parameters. [13] The covariate effects parameters α_1 and α_2 and the macroeconomic shock effect parameters β_ζ are positive and statistically significant for all credit classes, except for β_{AAA}. The effect β_{AAA} of the unobserved macroeconomic shock on credit class AAA is not statistically significant, implying that this credit class is not sensitive to unanticipated changes in the economy. McNeil and Wendin (2006) reach similar insights about the AAA rating class, consistent also with results reported in Altman and Kao (1992b) and Parnes (2007) on the stability of AAA ratings. The effect of the unobserved shock β_ζ is generally larger for speculative grade than for investment grade classes. This is not surprising, and we would expect lower credits to be more sensitive to unanticipated changes in the state of the economy.

The positive signs of α_1, α_2 and β_ζ show that the default probability does indeed decrease both with increasing values of CFNAI and of the S&P 500 return, and with an improving state of the economy, as expected. The fact

[12] The difference in DIC values between models R1 and L1 is 38 ($= 5,042 - 5,004$).
[13] We also fit the model with credit class specific covariate effects $\alpha_{1,\zeta}$ and $\alpha_{2,\zeta}$, but this model had a higher DIC value than R1, and the estimated posterior means for the covariate effects were not statistically significant. Hence in model R1 we have made the additional assumption that the effect of each covariate is common across all credit classes, so that $\alpha_{1,\zeta} = \alpha_1$ and $\alpha_{2,\zeta} = \alpha_2$ for all ζ.

that both CFNAI and the S&P 500 return have statistically significant effects α_1 and α_2, is consistent with the expectation that these two covariates capture different facets of economic conditions, reinforced by their history plotted in Figure 1.

[Table 2 about here.]

Figure 2 gives the time series values of the unobserved macroeconomic shock w_t estimated from model R1, together with the observed default rates (percentages) during 1981–2007. The upper graph plots the posterior means of w_t with 95% credible limits. The lower graph plots the observed overall default rates in our sample, as well as the observed default rates for investment grade and for speculative grade ratings during 1981–2007. A comparison of the upper and lower graphs of Figure 2 shows that default frequencies increased during 1985–1986, 1989–1991, and 1997–2001, while the values of w_t decreased during the same periods. This is consistent with the interpretation of w_t as an indicator of economic health.

[Figure 2 about here.]

A comparison between the history of w_t in Figure 2 and the history of the CFNAI and S&P 500 return covariates in Figure 1 shows a slight correlation between the time series of the observed covariates and of the macroeconomic shock. However, it is also apparent that the unobserved shock w_t captures some variability in transition rates that is left unexplained by the two observed covariates. Indeed, the correlations between the posterior means of the unobserved shock and the values of the two covariates across the 1981–2007 horizon are very small (less than 10%). This is also reflected in the model's predictive power, and explains why model R2 that includes only covariates and no macroeconomic shock has a much lower predictive power than model R1 that includes both covariates and the unobserved shock.

Table 3 reports the posterior means and standard errors for the threshold values $\gamma_{\zeta,\psi}$. Notice that, for each credit class ζ, the widest threshold interval is $(\gamma_{\zeta,\zeta-1}, \gamma_{\zeta,\zeta})$, corresponding to the probability of no transition from the credit class. This implies that there is potentially large heterogeneity among firms in the same credit class, which can have significant effects on credit risk diversification (Hanson, Pesaran and Schuermann, 2008).

[Table 3 about here.]

Transition probabilities between each pair of rating classes in every year are also estimated using model R1. As an illustration, Table 4 reports the matrix of posterior means for the estimated transition probabilities in 2007. Note that, as required, the model predicts positive default probabilities for all credit classes, even though no defaults have been observed in 2007 in the investment grade classes AAA, AA, A and BBB. As expected, the probability of default increases as the credit quality decreases. The transition matrix has high probability mass on the main diagonal since obligors are most likely to maintain their current rating, due to the rating agencies' desire for ratings stability (Altman and Rijken, 2004), or possibly to a signal extraction issue where the rating agency is reluctant to change the rating unless the signal is strong enough (Löffler, 2004). The probability of remaining within the current credit class overall decreases as the credit quality decreases, a phenomenon often observed in transition matrices (Standard and Poor's, 2007). Moreover, the probability of staying in the same credit class is higher for A and AA ratings than for the AAA rating, consistent with results in Altman and Kao (1992b) who found that bonds rated A and AA are more stable than AAA bonds.

[Table 4 about here.]

Note also that, conditional on the initial rating of an obligor, the second largest transition probabilities in Table 4 are in direct neighborhood to the diagonal, consistent with the monotonicity property. Obligors rated A and AA have greater probability of being downgraded than upgraded, whereas obligors rated BBB, BB and B are more likely to be upgraded than downgraded.

Figures 3 and 4 give the time series plots of posterior means and 95% credible limits for default probabilities between 1981–2007, in the investment grade and speculative grade rating classes respectively. The time series are comparable with the year-by-year estimates reported in Figure 3 from Hanson and Schuermann (2006). The plots emphasize the overall increase in the estimated default probabilities during 1985–1986, 1989–1991 and 1997–2001. This is consistent with the observed high default rates for speculative grade bonds in these periods — see Figure 2. The time series show some degree of co-movement, partially explained by the common effect of the state of the economy on both series of estimates. This is to be expected given the results in Table 2, where both the macroeconomic covariate coefficients and the unobserved shock coefficients are positive and statistically significant. The notable exception are default probabilities for credit class AAA which do not show almost any sensitivity to macroeconomic conditions. This is again consistent with results in Table 2, where the effect β_{AAA} of the macroeconomic shock on credit class AAA is not statistically significant.

[Figure 3 about here.]

[Figure 4 about here.]

Figures 5 and 6 give the time series plots of posterior means for upgrade and downgrade probabilities for each credit class between 1981–2007, with investment grade rating classes in the upper graphs and speculative grade rating classes in the lower graphs. The probability of an upgrade increases with decreasing credit quality in all years, with the exception of credit class BB in a few years. Conversely, the probability of a downgrade decreases with decreasing credit quality in all years, again with the exception of credit classes BB and AAA. It is apparent that the yearly evolutions of upgrade and downgrade probabilities exhibit opposite patterns, and there is a high degree of correlation between the time series of probabilities across credit classes. Moreover, the patterns are similar for investment and speculative grade rating classes, as a result of the strong impact of common economic conditions. The notable exception are again the downgrade probabilities for AAA ratings, which reinforces the earlier finding that the AAA class behaves differently than the other rating classes and that it is much less sensitive to macroeconomic influences.

[Figure 5 about here.]

[Figure 6 about here.]

Note that in several years credit class AAA had the largest downgrade probability among investment grade ratings. In particular, the downgrade probability for AAA in 2005 was larger than the downgrade probability for any other rating class except CCC. One potential explanation for the propensity for downgrades of AAA rated companies, lies in the conservative structural standards required by the rating agencies for assigning a AAA rating.

Out-of-Sample Model Comparison

For internal risk management and capital allocation purposes, banks require predicted probabilities of default for different rating classes, and associated transition probabilities based on current information. A common test for accuracy is to examine the model's ability to correctly identify out-of-sample obligors that default over the horizon (Shumway, 2001). This type of test can be performed for low quality investment grade and speculative grade obligors, for which sufficient default events exist. For high investment grade obligors, however, this type of test is not feasible due to the lack of default events which implies that the observed default rates in these rating classes are not the appropriate benchmark against which to assess a model's predictive performance. A reasonable model should predict a nonzero probability of default as economic theory requires, although ex post there may be no defaults over many periods.

Standard and Poor's (2007) note that no internal risk system or methodology can be considered best in a low default portfolio, and that consequently the issue of methodological comparison is not a question of one model being superior to another. Therefore, we do not attempt here to reach an absolute assessment of the superior predictive performance of the credit rating model, compared to models from the latent factor class. Instead, we investigate the out-of-sample predictions from two models within the Bayesian estimation framework, with the aim of gaining insights on the relative advantages and disadvantages of each model.

To this end, we focus on the out-of-sample predictions of models R1 and L1, the two best performing models from the credit rating process class and the latent factor class, as measured by the DIC values. We estimate the parameters of the models on the transition data from years 1981–2004, then generate the posterior distributions of the out-of-sample transition probabilities for 2005. We then update the models' parameters using data for 1981–2005, and generate the posterior distributions of the out-of-sample transition probabilities for 2006. Finally, a similar exercise is repeated for 2007.

First, we investigate predictions of default probabilities. Tables 5–6 report the posterior means, medians, standard errors, the endpoints and the width of 95% credible intervals for the out-of-sample default probabilities in 2005–2007, as well as the observed default rates for all credit classes. The results for model R1 are given in Table 5, and those for model L1 in Table 6.

[Table 5 about here.]

[Table 6 about here.]

The inherent difficulty of comparing estimates of default probability with actual outcomes is apparent from Tables 5–6. There were no observed defaults between 2005 and 2007 in credit classes AAA, AA, and A, and no defaults observed for 2006 and 2007 for credit class BBB, yet a benchmark of zero is clearly *not* appropriate for assessing predictive default probabilities in these rating classes. Both models R1 and L1 generate positive default probabilities for all rating classes, even those with little or no historical default data. In particular, there have been no defaults for credit class AAA over the entire period 1981–2006, yet both models predict a positive, although small, default probability for AAA obligors in 2007. In general, model R1 predicts smaller default probabilities than model L1 for all credit classes, and the predicted default probabilities from both models are larger than the observed default rates. One of the major strengths of the Bayesian estimation approach is the ability to easily generate 95% credible intervals for predicted default probabilities, using their posterior distributions. The credible intervals from model R1 are generally narrower than the ones from model L1 during 2006 and 2007.

As a further form of out-of-sample testing, we compare the probabilities of staying in the same credit class during the next year, for years 2005–2007. Tables 7–8 report the posterior means, medians, standard errors, the endpoints and the width of 95% credible intervals for the out-of-sample probabilities of no transition in 2005–2007, for all credit classes. The tables also report the actual realized percentages of firms staying in the same credit class during the year. The results for model R1 are given in Table 7, and those for model L1 in Table 8.

[Table 7 about here.]

[Table 8 about here.]

The observed rates of no-transition in all credit classes are close to the posterior means of the no-transition probabilities predicted by both models. The credible interval for model R1 contains the observed rate of no-transition more often than model L1. For example, in year 2005, the credible interval for model R1 contains the observed rate of no-transition for all credit classes, while this is not the case for credit class AA for model L1.

4 Conclusions

In this paper, we develop a credit rating process model that describes a typical form of internal credit assessment used by financial institutions. The loan officer reaches an assessment of the credit worthiness of an obligor and assigns a certain rating based on the range within which this assessment falls. This implies that there is heterogeneity in the credit worthiness of the obligors within any rating class, given the typical large number of obligors and small number of credit classes. The credit rating process model takes this heterogeneity into account, while extant latent factor models completely ignore it. We propose a Bayesian estimation methodology that jointly uses all available transition data, and thus overcomes the technical challenges related to the estimation of default probabilities and to the assessment of predictive performance in low default portfolios. Consequently, this paper directly addresses some of the issues raised by regulators and industry groups pertaining to low default portfolios.

Our empirical study of the Standard and Poor's ratings transition data shows that, among the models we tested, the best predictive credit rating process model includes two observed macroeconomic covariates and a random unobserved macroeconomic shock which accounts for dependence of transition probabilities for different credit classes in any given period. The corresponding implied rating transition matrix depends thus on the state of the economy, and we find that the effect of macroeconomic conditions is generally larger for speculative grade than for investment grade classes. In particular, we also find that the AAA class is much less sensitive to macroeconomic shocks, and in general has very different dynamics than the other rating classes. In terms of prediction, the performance of the best credit rating process model as assessed by the DIC measure of predictive power, is superior to that of the latent factor models that we tested. For out-of-sample prediction, we showed that the credit rating process model can be used to easily generate transition probabilities, and in particular default probabilities, and corresponding 95% credible intervals. As required by economic theory, all the predicted default probabilities are non-negative, even for those credit classes with few or no historical defaults.

The results reported here have implications for the current policy debate arising from Basel II. The Accord imposes a floor of 3 basis points on any probability of default estimate (Basel Committee on Banking Supervision, 2006a, §285), yet there is little evidence on whether this floor is realistic or not. From Table 5, the predicted out-of-sample default probabilities in 2007 for credit classes AAA, AA and A are respectively 0.4, 1.0 and 3.2 basis points. Our results show that the threshold of 3 basis points falls between ratings A and AA, hence it is overly conservative for AAA and AA rated corporations. The threshold implies that firms rated AA and AAA would be treated the same from a regulatory perspective, potentially distorting lending decisions by banks subject to Basel II regulations.

Finally, we have focused in this paper on the application of our methodology to the analysis of rating transition data from a public rating agency at aggregate level. The power and the flexibility of this methodological framework, however, make it ideally suitable also for the analysis of obligor-level data, and in particular as part of the assessment of any typical internal rating system employed by banks subject to Basel II regulations.

Acknowledgements

The authors are grateful to Robert Cowell, Mathias Drehmann, Bruce Hardie, Soosung Hwang, Alexander McNeil, Stephen Schaefer, Til Schuermann, Raman Uppal, Richard Verrall and Ingrid Weston for useful comments at several stages of preparation of this article. We also appreciate comments and suggestions by participants at seminars at the Bank of England, Cambridge University, Heriot-Watt University, Cass Business School, the CREDIT 2007 conference, and the INFORMS 2006 general meeting, as well as from two anonymous referees. Financial support from an RAMD grant from London Business School is gratefully acknowledged.

References

Altman, E., Kao, D.L., 1992a. Rating drift in high yield bonds. Journal of Fixed Income 2, 15–20.

Altman, E., Kao, D.L., 1992b. The implications of corporate bond rating drift. Financial Analysts Journal 48, 64–75.

Altman, E., Rijken, H.A., 2004. How rating agencies achieve rating stability. Journal of Banking and Finance 28, 2679–2714.

Bangia, A., Diebold, F., Kronimus, A., Schagen, C., Schuermann, T., 2002. Rating migration and the business cycle, with application to credit portfolio stress testing. Journal of Banking and Finance 26, 445–474.

Basel Committee on Banking Supervision, 2005. Studies on the Validation of the Internal Rating Systems. Working paper No.14, Bank for International Settlements.

Basel Committee on Banking Supervision, 2006a. International Convergence of Capital Measurement and Capital Standards: A Revised Framework. Bank for International Settlements.

Basel Committee on Banking Supervision, 2006b. The IRB Use Test: Background and Implementation. Basel Committee Newsletter No.9, Bank for International Settlements.

British Bankers' Association, London Investment Banking Association and the International Swaps and Derivatives Association, 2004. Introductory Paper on Low Default Portfolios. Discussion Paper 080904.

Carey, M., Hrycay, M., 2001. Parameterizing credit risk models with rating data. Journal of Banking and Finance 25, 197–270.

Chava, S., Stefanescu, C., Turnbull, S., 2006. Modeling expected loss. Working paper, Bauer College of Business, University of Houston.

Christensen, J.H.E., Hansen, E., Lando, D., 2004. Confidence sets for continuous time rating transition probabilities. Journal of Banking and Finance 28, 2575–2602.

Crouhy, M., Galai, D., Mark, R., 2000. Risk Management. New York: McGraw-Hill.

Das, S.R., Fan, R., Geng, G., 2002. Bayesian migration in credit ratings based on probabilities of default. Journal of Fixed Income 12, 17–23.

Dwyer, D.W., 2006. The Distribution of Defaults and Bayesian Model Validation. Moody KMV.

Farnsworth, H., Li, T., 2007. The dynamics of credit spreads and rating migrations. Journal of Financial and Quantitative Analysis 42, 595-620.

Figlewski, S., Frydman, H., Liang, W., 2008. Modeling the effect of macroeconomic factors on corporate default and credit rating transitions. NYU Stern Business School working paper.

Frydman, H., Schuermann, T., 2008. Credit rating dynamics and Markov mixture models. Journal of Banking and Finance 32, 1062–1075.

Gagliardini, P., Gourieroux, C., 2005. Stochastic migration models with application to corporate risk. Journal of Financial Econometrics 3, 188–226.

Gelman, A., Carlin, J.C., Stern, H., Rubin, D.B., 1995. Bayesian Data Analysis. New York: Chapman & Hall.

Geman, S., Geman, D., 1984. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. IEEE Transactions on Pattern Analysis and Machine Intelligence 6, 721–741.

Gilks, W.R., Richardson, S., Spiegelhalter, D.J., 1996. Markov Chain Monte Carlo in Practice. London: Chapman & Hall.

Gordy, M., 2003. A risk factor model for ratings based bank capital rules. Journal of Financial Intermediation 12, 199–232.

Gössl, C., 2005. Predictions Based on Certain Uncertainties — A Bayesian Credit Portfolio Approach. Working Paper, Hypo Vereinsbank AG, London.

Hamilton, D.T., Cantor, R., 2004. Rating Transitions and Defaults Conditional on Watch, Outlook and Rating History. Moody's Investment Service, Special Comment.

Hamilton, D.T., Ou, S., Kim, F., Cantor, R., 2007. Corporate Default and Recovery Rates, 1920–2006. Moody's Investor Service.

Hanson, S., Schuermann, T., 2006. Confidence intervals for probabilities of default. Journal of Banking and Finance 30, 2281–2301.

Hanson, S., Pesaran, M.H., Schuermann, T., 2008. Firm heterogeneity and credit risk diversification. Journal of Empirical Finance 15, 583–612.

Heitfield, E., 2005. An Illustrative Model of Rating System Dynamics and PD Quantification. Appendix A, Basel Committee on Banking Supervision. Studies on the Validation of the Internal Rating Systems. Working paper No.14.

Jafry, Y., Schuermann, T., 2004. Measurement, estimation and comparison of credit migration matrices. Journal of Banking and Finance 28, 2603–2639.

Kadam, A., Lenk, P., 2007. Bayesian inference for issuer heterogeneity in credit ratings migration. Cass Business School Working Paper.

Kiefer, N., 2007. The probability approach to default estimation. Risk, July, 146–150.

Lando, D., Skødeberg, T., 2002. Analyzing rating transitions and rating drift with continuous observations. Journal of Banking and Finance 26, 423–444.

Löffler, G., 2004. An anatomy of rating through the cycle. Journal of Banking and Finance 28, 695–720.

McNeil, A., Frey, R., Embrechts, P., 2005. Quantitative Risk Management. Princeton: Princeton University Press.

McNeil, A., Wendin, J., 2006. Dependent credit migrations. Journal of Credit Risk 2, 87–114.

McNeil, A., Wendin, J., 2007. Bayesian inference for generalized linear mixed models of portfolio credit risk. Journal of Empirical Finance 14, 131–149.

Nickell, P., Perraudin, W., Varotto, S., 2000. Stability of rating transitions. Journal of Banking and Finance 24, 203–227.

Parnes, D., 2007. A density dependent model for credit ratings migration dynamics. Journal of Fixed Income 17, 26-37.

Pluto, K., Tasche, D., 2005. Thinking positively. Risk, August, 72-78.

Schönbucher, P.J., 2003. Credit Derivatives Pricing Models. New Jersey: John Wiley & Sons.

Shumway, T., 2001. Forecasting bankruptcy more accurately: a simple hazard model. Journal of Business 74, 101–124.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P., van der Linde, A., 2002. Bayesian measures of model complexity and fit. Journal of Royal Statistical Society, Series B 64, 583–640.

Spiegelhalter, D.J., Thomas, A., Best, N.G., Lunn, D., 2003. WinBUGS Version 1.4 User Manual. MRC Biostatistics Unit, Cambridge.

Standard and Poor's, 2007. Annual Global Corporate Default Study.

Stock, J., Watson, M., 1999. Inflation forecasting. Journal of Monetary Economics 44, 293–335.

Trück, S., Rachev, S.T., 2005. Credit portfolio risk and probability of default confidence sets through the business cycle. Journal of Credit Risk 1, 61–88.

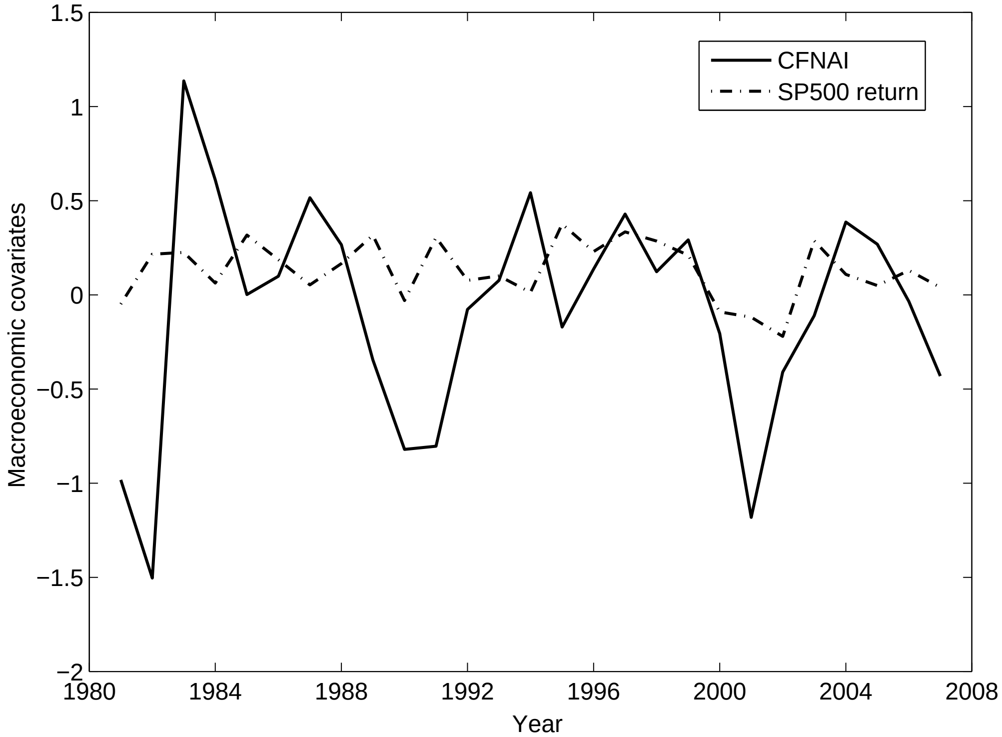


Fig. 1. History of the macroeconomic covariates CFNAI and SP500 return

Table 1
Values of the Deviance Information Criterion (DIC)

Model	R1	R2	R3	R4	L1	L2
DIC	5004.39	5969.01	5037.80	5040.10	5042.29	5061.62

Note: DIC values for the credit rating process models R1–R4, and for the latent factor models L1 and L2 estimated from the Standard and Poor's transition rating data.

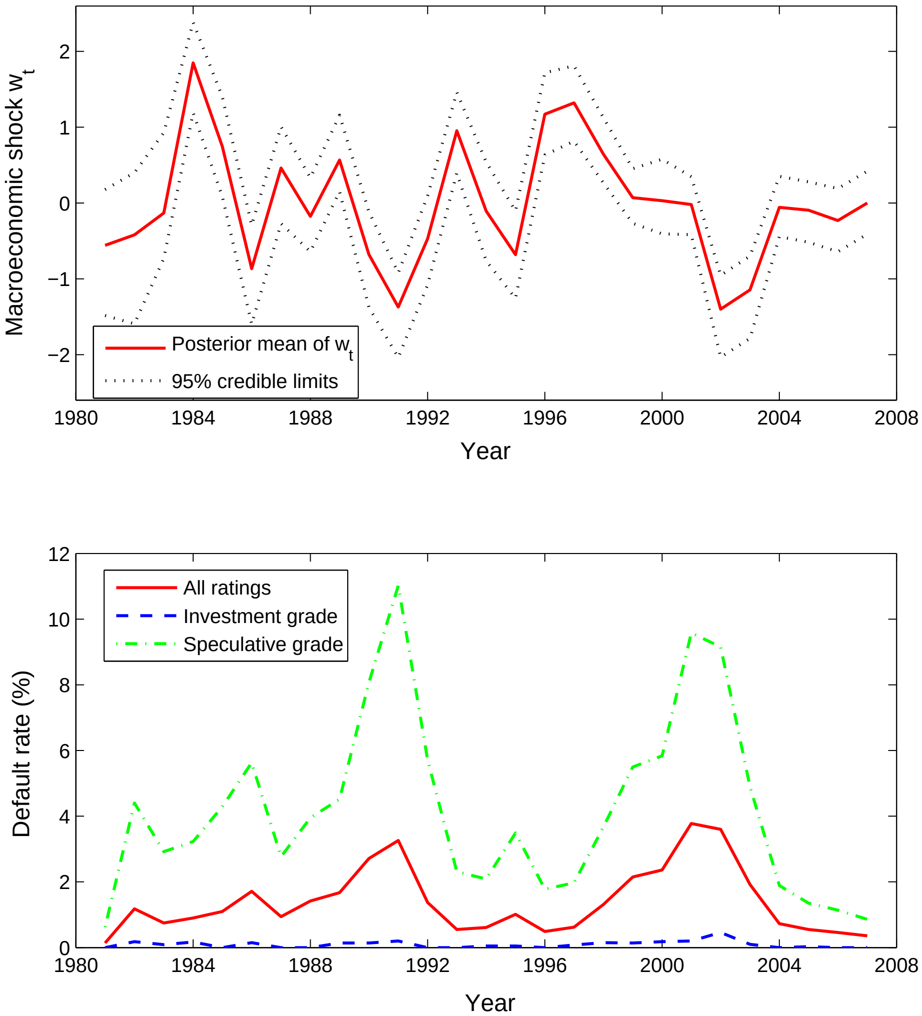


Fig. 2. Implied history of the macroeconomic shock w_t (upper graph) and historical default rates in percentages (lower graph).

Table 2
Bayesian estimates for model R1 parameters.

β_{AAA}	β_{AA}	β_{A}	β_{BBB}	β_{BB}	β_{B}	$\beta_{CCC/C}$	α_1	α_2
-0.048	0.345	0.305	0.254	0.247	0.449	0.510	0.243	0.167
(0.063)	(0.056)	(0.048)	(0.042)	(0.044)	(0.068)	(0.086)	(0.033)	(0.036)

Note: Posterior means of the parameters with standard errors in parentheses. The effect of the observed macroeconomic covariates CFNAI and SP500 return on all credit classes is given by α_1 and α_2 respectively, while the effect of the unobserved macroeconomic shock w_t on credit class ζ is given by β_ζ.

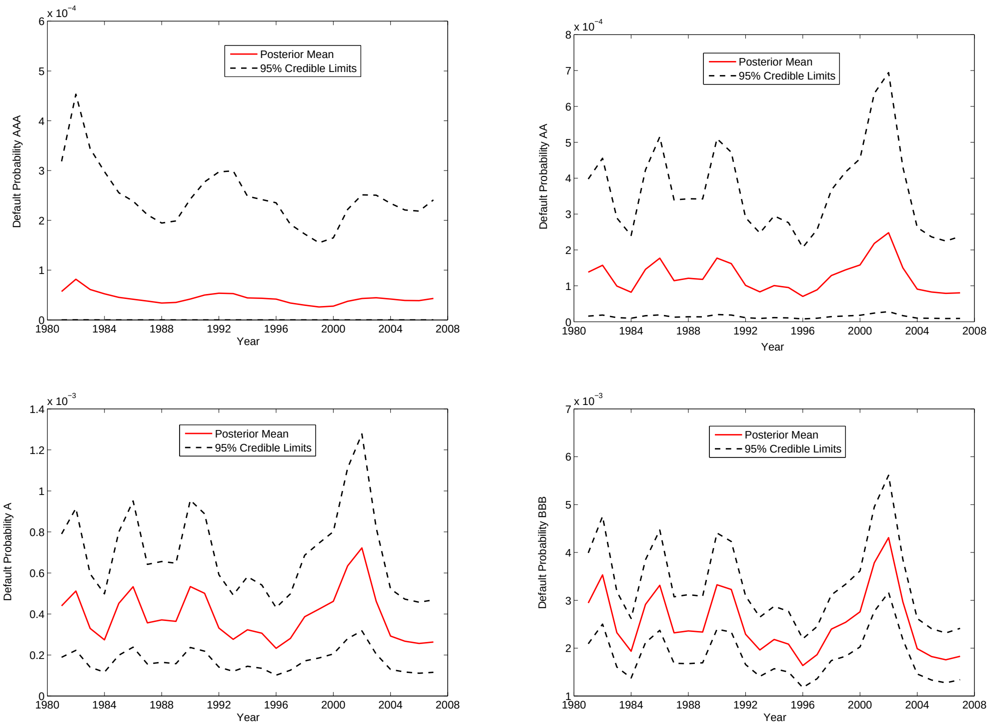


Fig. 3. Annual default probabilities (percentages) for investment grade ratings; posterior means with 95% credible intervals.

Table 3
Threshold parameters $\gamma_{\zeta,\psi}$ estimated from model R1.

	AA	A	BBB	BB	B	CCC/C
AAA	8.861	6.217	4.773	3.832	2.023	0.997
	(1.950)	(1.940)	(1.896)	(1.841)	(1.435)	(0.990)
AA	14.510	6.886	4.373	2.964	2.567	1.080
	(0.856)	(0.844)	(0.835)	(0.814)	(0.801)	(0.637)
A	15.540	11.750	5.184	2.857	1.814	0.655
	(0.464)	(0.350)	(0.346)	(0.335)	(0.314)	(0.231)
BBB	14.660	12.120	9.120	3.160	1.567	0.511
	(0.556)	(0.214)	(0.151)	(0.143)	(0.132)	(0.093)
BB	12.560	11.390	10.080	7.262	2.287	0.588
	(0.497)	(0.292)	(0.164)	(0.094)	(0.084)	(0.057)
B	13.200	10.450	9.010	8.187	5.612	0.672
	(1.297)	(0.370)	(0.186)	(0.129)	(0.056)	(0.029)
CCC/C	8.703	8.202	6.707	5.895	4.904	2.713
	(1.017)	(0.867)	(0.434)	(0.301)	(0.188)	(0.090)

Note: Posterior means of the $\gamma_{\zeta,\psi}$ parameters with standard errors in parentheses. The rows represent the credit classes at the beginning of the year, and the columns represent the credit classes at the end of the year. $\gamma_{\zeta,0} = 0$ for all ζ.

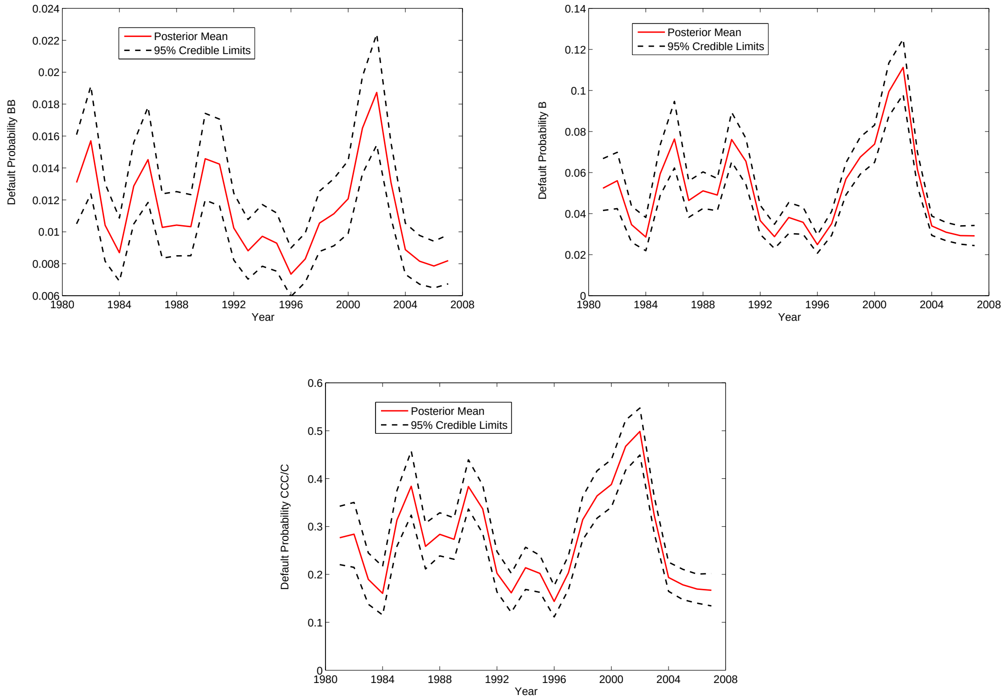
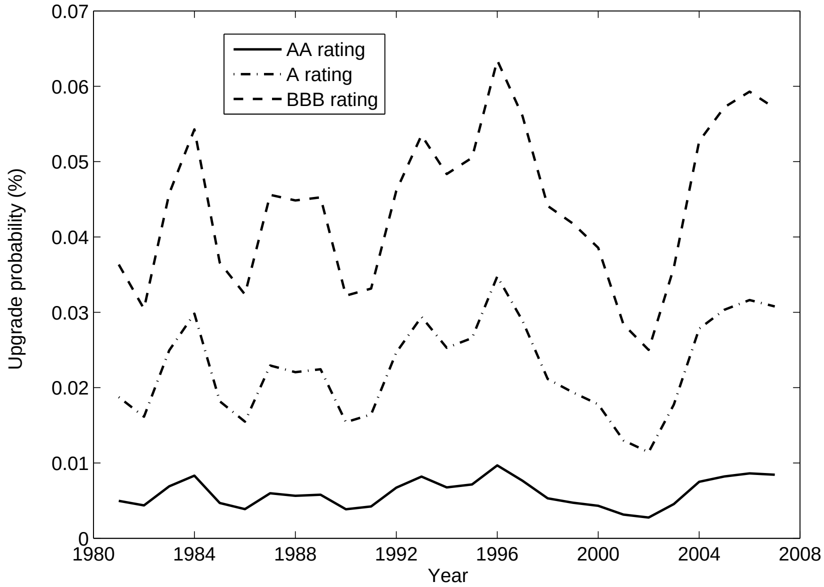


Fig. 4. Annual default probabilities (percentages) for speculative grade ratings; posterior means with 95% credible intervals.

Table 4
Transition probabilities for 2007 estimated from model R1.

	AAA	AA	A	BBB	BB	B	CCC/C	D
AAA	91.730	7.620	0.485	0.094	0.054	0.009	0.004	0.004
AA	0.844	93.641	5.041	0.355	0.038	0.061	0.012	0.008
A	0.075	3.002	92.680	3.814	0.277	0.104	0.022	0.026
BBB	0.027	0.277	5.404	90.191	3.236	0.561	0.121	0.183
BB	0.048	0.096	0.372	7.356	84.631	6.033	0.645	0.819
B	0.011	0.092	0.311	0.515	9.972	83.550	2.633	2.916
CCC/C	0.125	0.062	0.482	0.755	2.226	21.360	58.290	16.700

Note: Posterior means (percentages) of the transition probabilities. The rows represent the credit classes at beginning of the year, and the columns represent the credit classes at the end of the year.



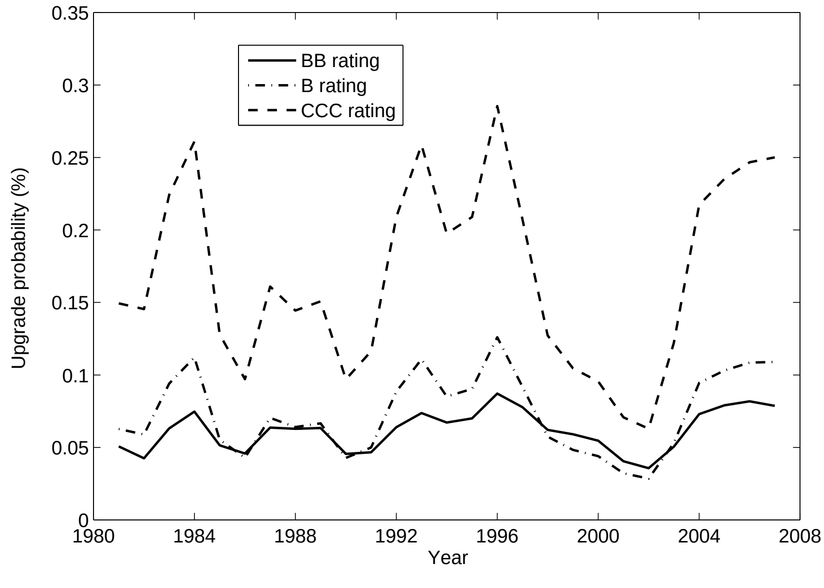


Fig. 5. Annual upgrade probabilities (percentages) for investment grade classes (upper graph) and speculative grade classes (lower graph).

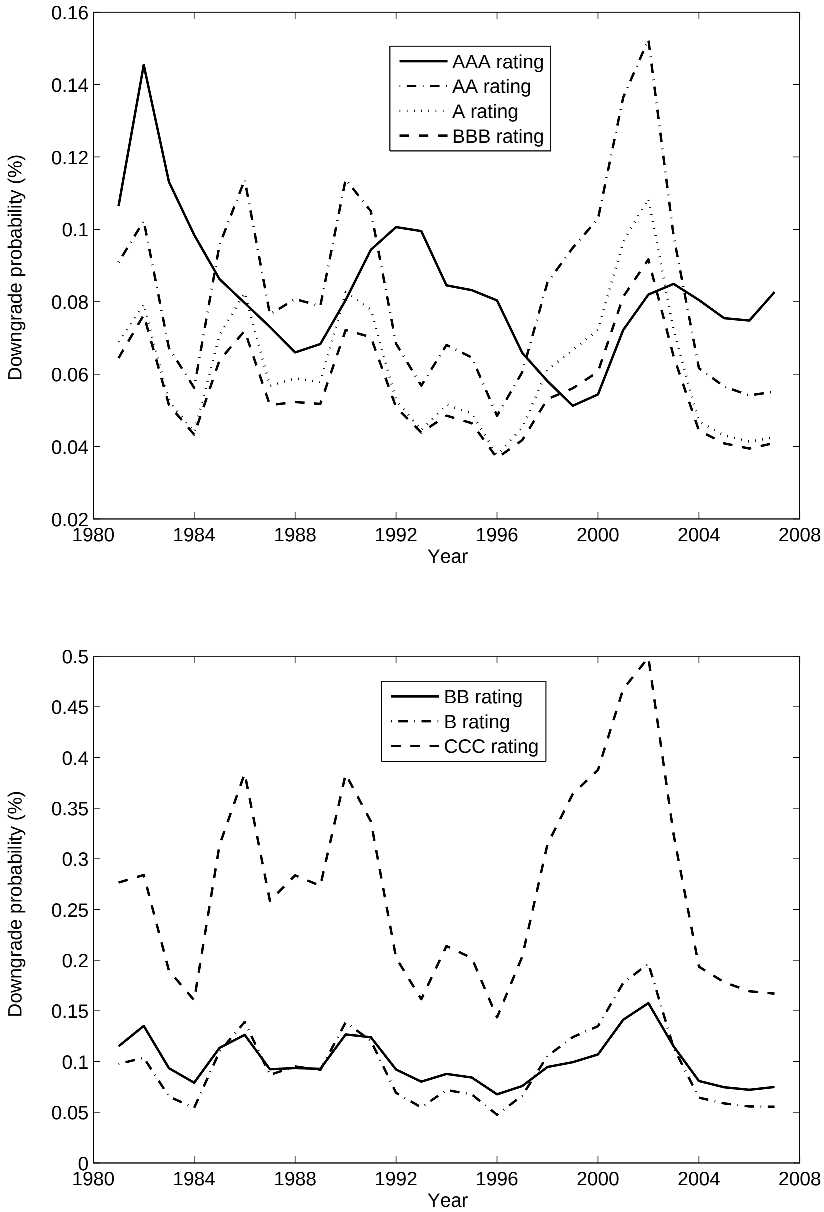


Fig. 6. Annual downgrade probabilities (percentages) for investment grade classes (upper graph) and speculative grade classes (lower graph).

Table 5
Out-of-sample default probabilities for 2005–2007, model R1.

Rating class	Mean	Median	Standard error	95% credible intervals	Observed percentage	Width of 95% credible intervals
Year 2005						
AAA	0.004	0.001	0.007	$(0, 0.023)$	0	0.023
AA	0.011	0.007	0.012	$(0, 0.040)$	0	0.040
A	0.035	0.029	0.024	$(0.008, 0.094)$	0	0.086
BBB	0.239	0.217	0.115	$(0.088, 0.533)$	0.07	0.445
BB	0.996	0.910	0.478	$(0.351, 2.186)$	0.219	1.835
B	4.477	3.669	3.441	$(0.874, 13.490)$	1.860	12.616
CCC/C	22.660	20.590	12.530	(5.109, 54.690)	10.577	49.581
Year 2006						
AAA	0.004	0.002	0.008	$(0, 0.025)$	0	0.025
AA	0.010	0.007	0.009	$(0, 0.035)$	0	0.035
A	0.030	0.027	0.017	$(0.009, 0.071)$	0	0.062
BBB	0.213	0.197	0.087	$(0.093, 0.421)$	0	0.328
BB	0.907	0.850	0.330	$(0.421, 1.721)$	0.380	1.300
B	3.793	3.230	2.323	$(1.104, 9.680)$	0.840	8.576
CCC/C	19.490	17.370	10.540	$(4.963, 45.680)$	14.630	40.717
Year 2007						
AAA	0.004	0.002	0.007	$(0, 0.022)$	0	0.022
AA	0.010	0.008	0.009	$(0.001, 0.035)$	0	0.034
A	0.032	0.028	0.016	$(0.011, 0.071)$	0	0.060
BBB	0.219	0.207	0.072	$(0.113, 0.398)$	0	0.285
BB	0.965	0.930	0.275	$(0.557, 1.621)$	0.280	1.064
B	3.820	3.458	1.820	$(1.459, 8.387)$	0.340	6.928
CCC/C	20.130	18.680	8.630	$(7.427, 40.540)$	18.460	33.113

Table 6
Out-of-sample default probabilities for 2005–2007, model L1.

Rating class	Mean	Median	Standard error	95% credible intervals	Observed percentage	Width of 95% credible intervals
Year 2005						
AAA	0.007	0.006	0.006	$(0, 0.020)$	0	0.020
AA	0.024	0.020	0.015	$(0.006, 0.061)$	0	0.055
A	0.043	0.037	0.025	$(0.013, 0.105)$	0	0.092
BBB	0.295	0.279	0.103	$(0.139, 0.542)$	0.07	0.403
BB	1.197	1.173	0.269	$(0.730, 1.803)$	0.219	1.073
B	5.912	5.547	2.407	$(2.317, 12.290)$	1.860	9.973
CCC/C	29.310	28.540	9.816	(12.010, 51.610)	10.577	39.600
Year 2006						
AAA	0.002	0.000	0.004	$(0, 0.015)$	0	0.015
AA	0.014	0.011	0.011	$(0.003, 0.039)$	0	0.036
A	0.044	0.038	0.026	$(0.017, 0.114)$	0	0.097
BBB	0.291	0.276	0.102	$(0.142, 0.536)$	0	0.394
BB	1.154	1.137	0.208	$(0.789, 1.616)$	0.380	0.827
B	5.610	5.243	2.149	$(2.498, 10.920)$	0.840	8.422
CCC/C	28.270	26.750	10.630	$(10.990, 53.130)$	14.630	42.140
Year 2007						
AAA	0.004	0.002	0.006	$(0, 0.026)$	0	0.026
AA	0.022	0.017	0.015	$(0.005, 0.066)$	0	0.061
A	0.051	0.047	0.023	$(0.020, 0.106)$	0	0.086
BBB	0.294	0.277	0.108	$(0.145, 0.548)$	0	0.403
BB	1.344	1.324	0.259	$(0.884, 1.911)$	0.280	1.027
B	6.604	6.172	2.918	$(2.635, 13.500)$	0.340	10.865
CCC/C	32.090	31.120	11.690	$(12.540, 57.470)$	18.460	44.930

Table 7
Out-of-sample probabilities of no transition for 2005–2007, model R1.

Rating class	Mean	Median	Standard error	95% credible intervals	Observed percentage	Width of 95% credible intervals
Year 2005						
AAA	92.420	92.620	1.403	(88.730, 94.510)	89.691	5.780
AA	92.100	92.890	3.195	(83.740, 95.420)	94.373	11.680
A	91.700	92.280	1.743	(86.860, 92.980)	92.551	6.120
BBB	89.210	89.690	1.480	(84.790, 90.401)	89.646	5.611
BB	83.000	83.660	2.174	(76.610, 84.730)	84.510	8.120
B	80.800	82.750	5.285	(64.640, 84.490)	82.907	19.850
CCC/C	53.830	56.200	7.060	(33.100, 60.570)	56.731	27.470
Year 2006						
AAA	92.040	92.180	1.209	(89.450, 93.930)	97.674	4.480
AA	92.660	93.180	2.374	(86.430, 95.430)	94.949	9.000
A	92.140	92.460	1.069	(89.310, 93.080)	91.966	3.770
BBB	89.340	89.650	1.060	(86.450, 90.400)	90.185	3.950
BB	83.600	83.920	1.206	(80.380, 84.850)	88.378	4.470
B	81.290	82.640	3.916	(70.560, 84.380)	83.835	13.820
CCC/C	53.930	55.770	6.103	(36.970, 60.450)	57.073	23.480
Year 2007						
AAA	91.470	91.550	1.026	(89.200, 93.350)	97.647	4.150
AA	92.580	92.990	2.028	(87.710, 95.250)	94.864	7.540
A	92.140	92.400	0.870	(89.800, 92.990)	92.296	3.100
BBB	89.860	89.990	0.581	(88.410, 90.580)	91.751	2.170
BB	84.200	84.360	0.746	(82.350, 85.150)	84.712	2.800
B	82.570	83.320	2.217	(76.380, 84.590)	83.842	8.210
CCC/C	56.440	57.410	3.952	(45.640, 61.320)	56.923	15.680

Table 8
Out-of-sample probabilities of no transition for 2005–2007, model L1.

Rating class	Mean	Median	Standard error	95% credible intervals	Observed percentage	Width of 95% credible intervals
Year 2005						
AAA	90.860	90.930	1.200	$(88.230, 93.220)$	89.691	4.990
AA	90.140	94.200	2.486	$(84.250, 94.200)$	94.373	9.950
A	91.230	91.670	1.612	$(86.930, 92.880)$	92.551	5.950
BBB	89.570	89.800	0.813	$(87.440, 90.440)$	89.646	3.000
BB	83.750	83.890	0.757	$(81.720, 84.810)$	84.510	2.690
B	82.510	83.180	2.112	$(76.240, 84.440)$	82.907	8.200
CCC/C	54.730	56.000	0.164	(41.850, 60.580)	56.731	18.730
Year 2006						
AAA	91.590	91.650	0.930	$(89.550, 93.420)$	97.674	3.870
AA	90.790	91.140	2.161	$(85.830, 94.110)$	94.949	8.280
A	91.470	91.820	1.325	$(87.950, 92.940)$	91.966	4.990
BBB	89.650	89.830	0.767	$(87.770, 90.440)$	90.185	2.670
BB	83.880	83.960	0.627	$(82.420, 84.810)$	88.378	2.390
B	82.750	83.320	1.805	$(77.680, 84.410)$	83.835	6.730
CCC/C	54.960	56.410	5.127	$(40.570, 60.670)$	57.073	20.100
Year 2007						
AAA	92.350	92.420	1.231	(89.850, 94.480)	97.650	4.630
AA	89.660	90.070	3.194	(82.270, 94.290)	95.860	12.020
A	90.950	91.380	1.754	(86.520, 92.840)	93.300	6.320
BBB	89.630	89.880	0.965	(87.340, 90.560)	91.750	3.220
BB	83.390	83.520	0.994	(81.030, 84.890)	84.710	3.860
B	82.010	82.970	2.873	(74.560, 84.560)	87.840	10.000
CCC/C	53.800	55.560	6.221	(37.760, 60.830)	56.920	23.070