



NYU  STERN

Salomon Center for the Study of Financial Institutions

Working Paper Series **CREDIT & DEBT MARKETS** Research Group

HOW RATING AGENCIES ACHIEVE RATING STABILITY

Edward I. Altman
Herbert A. Rijken

S-CDM-03-12

How rating agencies achieve rating stability

Edward I. Altman¹ and Herbert A. Rijken²

December 2003

JEL classification: G20, G33

Keywords: Rating agencies, "through-the-cycle" rating methodology, migration policy, credit scoring models

¹ NYU Salomon Center, Leonard N. Stern School of Business, New York University, 44 West 4th Street, New York, NY 10012, USA. email: ealtman@stern.nyu.edu

² Free University, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands. email: hrijken@feweb.vu.nl

Abstract

Surveys on the use of agency credit ratings reveal that most investors believe that rating agencies are relatively slow in adjusting their ratings. A well-accepted explanation for this perception on the timeliness of agency ratings is the "through-the-cycle" methodology, which agencies apply in their rating assessments, while investors have a "point-in-time" perception on the creditworthiness. The "through-the-cycle" methodology aims to suppress the sensitivity of the ratings to short-term fluctuations in credit quality. This article focuses on the migration policy of rating agencies as a second source of rating stability. In a benchmark study with credit scoring models we show that both the "through-the-cycle" methodology and the conservative migration policy are responsible for the investors' perception of the rigidity of agency ratings.

I Introduction

Credit ratings of Standard & Poor's, Moody's and Fitch, play a key role in the pricing of credit risk and in the delineation of investment strategies. The future role of these agency ratings will be even further expanded with the implementation of the Basle II accord, which establishes rating criteria for capital allocations of banks in 2007. Given the enormous increase in defaults in the 2001 – 2002 period, following the rather sudden meltdown in Asian countries and corporations, the timeliness of agency ratings has come under closer scrutiny and criticism. A recent survey conducted by the Association for Financial Professionals (2002) reveals that most respondents believe that the agency ratings are too slow in responding to changes in corporate credit quality¹. Investors prefer more timely rating changes, even at a cost of an increase in the number of reversal rating migrations, provided that rating migrations are beyond marginal. Surveys by Ellis (1998) and Baker et al. (2001) reveal the same finding. Recent failures of investment-graded companies reinforce these opinions.

A well-accepted explanation for this perception on the timeliness of agency ratings is the "through-the-cycle" rating methodology, which agencies apply in their rating assessments, while investors have a "point-in-time" perception on the creditworthiness. In their corporate ratings criteria document, Standard & Poor's (2003) take the position that "the value of its products is greatest when its ratings focus on the long term and do not fluctuate with near term performance"². In order to preserve the notion of long-term credit risk measurement, the rating agencies suppresses in their rating assessment the influence of short-term fluctuations in credit quality as much as possible.

Little details are known on the "through-the-cycle" methodology itself. Carey and Hrycay (2001) describe the "through-the-cycle" methodology as a rating assignment based on a stress scenario. When firms are consequently rated in the bottom of the credit quality cycle, agency ratings are insensitive to the credit quality cycle and long term focused. Löffler (2002) concludes, with caution, that the "through-the-cycle" methodology results in a higher rating stability at a cost of a lower capability to predict default. Still, how the "through-the-cycle" rating methodology is put into practice and to what extent it influences the rating stability, is not completely clear, at least not in a quantitative manner. In addition to the corporate issuer credit ratings, Standard and Poor's and Moody's provide information on the likelihood a rating might change in the near future (Standard & Poor's Credit Watch and Moody's Outlooks and Watch list). These outlooks provide information on a possible change in the long-term credit risk profile and do *not* provide information on the position within the credit quality cycle, i.e. the default profitability on short term.

In practice, investors judge the rating timeliness from a “point-in-time” perspective. In contrast to the “through-the-cycle” methodology, the “point-in-time” methodology measures the current position of a firm in the credit quality cycle. Depending on the time horizon it measures both the short-term and long-term default probability. By and large, banks apply a “point-in-time” methodology and a time horizon of 3 – 7 in their credit quality assessments (see Basle Committee on Banking Supervision, 2001 and Treacy and Carey, 2000). Presumably a vast majority of the investors, consciously or not, judge credit quality on the basis of a “point-in-time” estimate.

A second source of rating stability is the migration policy³ of the rating agencies. A delay and a spread in time of the necessary rating migrations can also increase the rating stability. In January 2002 Moody’s announced a number of initiatives to enhance the quality and timeliness of their ratings. One of the measures under consideration were “more aggressive ratings changes – such as downgrading a rating by several notches immediately in reaction to adverse news rather than slowly reducing the rating over a period of time – as well as shortening the rating review cycle to a period of weeks from the current period of several months“ (Financial Times, 19 January 2002). A few months later, after consultation, Moody’s concluded that the market participants strongly oppose to these changes because they generally desire rating stability, and they believe such changes would increase rating volatility (Moody’s, 2002a). Regardless of the outcome, this discussion provides practical evidence that the migration policy is a source of rating stability. To the best of our knowledge the influence of the migration policy has not been studied in academic literature so far.

The aim of this paper is twofold. First, to determine the relative importance of the two sources of rating stability. We find that both the “through-the-cycle” methodology and the prudent migration policy have an equal contribution to the rating stability. Secondly we characterized the migration policy by two parameters: a threshold parameter and an adjustment factor parameter. A rating migration is triggered when the credit quality exceeds a threshold level of 1.25 notch steps compared to the mean credit quality in a particular rating class. If triggered, ratings are only partly adjusted, by 75%, to a target rating level, which is set about 0.3 - 0.6 notch steps below the actual “through-the-cycle” credit quality level. Target rating adjustments are split and executed on different times. Presumably agencies follow a wait and see policy. Although these migration policy parameters appear not to be too conservative, the agency rating dynamics turn out to be very sensitive to these parameters.

These findings are obtained by benchmarking the agency ratings by well-established credit scoring models i.e. default-prediction models. If agency ratings are biased or delayed for whatever reason, an “objective” measurement of credit quality by a default-prediction model should be able to reveal that. In various case studies default-prediction models have showed up to

provide useful *additional* information to agency ratings. For example, a case study of Enron and Worldcom show that Altman Z-scores and EDF scores (KMV model) were signaling one year before default a strong downward trend in credit quality, while agency ratings were kept on an investment graded status (see McQuown, 1993, and Saunders and Allen, 2002).

The paper proceeds as follows. In section 2 the benchmark credit-scoring models are described and motivated. Section 3 examines the development of actual credit quality levels, as proxied by credit-scoring models, conditional on agency-rating migrations. Section 4 and 5 report on the consequences of the rating methodology and the migration policy on the agency rating dynamics, with a special focus on the rating-migration probabilities and the rating drift properties.

II Benchmark credit scoring models

2.1 Benchmark research setup

In the benchmark study the corporate issuer credit ratings of agencies are examined. These ratings reflect "the obligor's ability and willingness to meet its financial commitments on a timely basis" (see Standard's and Poor's, 2003) and "it generally indicates the likelihood of default regarding all financial obligations of the firm" (see Moody's, 2002b). According to these definitions the corporate issuer credit ratings of agencies ⁴ are measures of default probability.

Key in the benchmark study is the formulation of two benchmark credit scoring models:

1. The default-prediction model (DP-model) proxies for the investor's "point-in-time" perspective on credit quality. The DP-model is in fact a hazard rate model, which models the time a firm averts a default event. As Shumway (2001) pointed out, hazard rate models are preferable over static models (single-period classification models) in modeling default probability. First, hazard-rate models control for the different periods wherein financial troubled firms stand before they default. Second the estimation of hazard-rate models takes into account the changing default risk over time. Shumway has proven that the estimation of a discrete-time hazard-rate model is similar to the estimation of a multiperiod logit model. Therefore the DP-model is in fact a logit model, which models the one-year default probability p_i as follows

$$DP_i = \alpha + \beta X_i + \varepsilon_i \quad (2.1)$$

$$\log\left(\frac{p_i}{1-p_i}\right) = DP_i \quad (2.2)$$

where X_i is the set of model-variables for firm-year observation i , ε_i is normal distributed, and ε_i and ε_j ($i \neq j$) are independent. In reality these conditions to the error term are violated. To obtain the correct statistics we have used the Huber-White standard errors which control for firm dependence. The parameters α , β are estimated by a maximum likelihood procedure.

The DP-score is directly related to the one-year default probability p_i , which is a proxy for the "point-in-time" credit quality. A true proxy for the investor's "point-in-time" perception is difficult to achieve as no precise reference and no theoretical framework exists for the investor's perspective on credit quality. Therefore extensive attention is paid to formulate a

robust default-prediction model with a default prediction performance at least as good as agency ratings ⁵.

Strictly, the investor's perspective on credit quality includes both the short-term default probability *and* the long-term default probability. However, the *one-year* default-prediction model is slightly more sensitive to the actual position in the credit quality cycle, which emphasizes the consequences of suppressing the credit quality cycle information ⁶.

2. The agency-rating model (AR-model) reflects the *actual* “through-the-cycle” credit quality. The AR-model is in fact an ordered logit regression model, which models the discrete agency-rating scale y . In this model the AR-score, AR_i , is an unobservable variable

$$AR_i = \alpha + \beta X_i + \varepsilon_i \quad (2.3)$$

where X_i is the set of model-variables for observation i , ε_i is normal distributed, and ε_i and ε_j ($i \neq j$) are independent. It is assumed that AR_i is related to the agency rating y_i as follows

$$y_i = k \quad \text{if} \quad B_{k-1} < AR_i \leq B_k \quad (2.4)$$

where k is one of the agency-rating classes $\{1, 2, 3, \dots, 16\}$ ⁷, B_k is the upper boundary for the AR-score in rating class k , $B_0 = -\infty$ and $B_{16} = \infty$. In the ordered logit model the probability $y_i = k$ is specified by

$$P(y_i = k) = F(B_k - AR_i) - F(B_{k-1} - AR_i) \quad (2.5)$$

where F is the cumulative logistic function ⁸. The parameters α , β and B_k are estimated by a maximum likelihood procedure.

The AR-score reflects the actual “through-the-cycle” credit quality. It immediately follows the changes in the model-variables, i.e. the credit fundamentals and is obviously insensitive to the migration policy. We assume that the migration policy does not affect the estimated parameters α and β . If the amount of overstated ratings and the amount of understated ratings, due to the prudent migration policy, are equal, than the migration policy does not affect the estimation of the parameters α , β , but only widens the distribution of the error term ε .

Bond ratings are modeled mostly for the purpose of forecasting agency-rating migrations (see for example Ederington, 1985, Kaplan and Urwitz, 1979, Blume et al., 1998 and Kamstra et al, 2001). Applied statistical methodologies are ordinary least square analysis, ordered probit regression analysis and multinomial discriminant analysis. In order to be consistent with the statistical methodology of the default-prediction model, we have modeled the agency ratings by an ordered logit model, which is an alternative to the more popular ordered probit model.

In figure I the basic concept of the benchmark study is shown. The AR-score bridges the gap between de agency ratings and the DP-scores (see Figure I) and enables to reveal unambiguously the influence of the "through-the-cycle" methodology and the prudent migration policy on rating stability. Differences in dynamics between agency ratings and AR-scores exclusively reflect the influence of the prudent migration policy, and differences in dynamics between AR-scores and DP-scores exclusively reflect the influence of the credit quality measurement approach: "point-in-time" vs "through-the-cycle" ⁹.

2.2 Model variables in the benchmark credit scoring models

The DP-model and AR-model employ both the same model-variables and are both based a logit regression methodology. This allows an unambiguous comparison of the dynamics of AR-scores and DP-scores. The DP-score (equation 2.1) and the AR-score (equation 2.3) are calculated on the basis of the following set of six model-variables

$$credit\ score = \alpha + \beta_1 \frac{WK}{TA} + \beta_2 \frac{RE}{TA} + \beta_3 \frac{EBIT}{TA} + \beta_4 \frac{ME}{BL} + \beta_5 Size + \beta_6 Age \quad (2.6)$$

where WK is the net working capital and equals to the total current assets minus the total current liabilities, RE are the retained earnings, TA are the total assets, EBIT are the earnings before interest and taxes, ME is the market value of equity and BL is the book value of total liabilities. Size equals to the total liabilities, normalized by the total value of the US equity market Mkt and log-transformed: $\ln(BL/Mkt)$. Age is the number of years since a firm has been first rated by a rating agency ¹⁰. In order to increase the effectiveness of the RE/TA-, EBIT/TA-, and ME/BL-variables in the logit model estimate, these variables are log-transformed as follows: RE/TA $\rightarrow -\ln(1 - RE/TA)$, EBIT/TA $\rightarrow -\ln(1 - EBIT/TA)$ and ME/BL $\rightarrow 1 + \ln(ME/BL)$ ¹¹.

The choice of the model-variables is inspired by the Z-score model (Altman, 1968) ^{12 13}. The WK/TA-variable is a proxy for the short-term liquidity of a firm. The RE/TA-, EBIT/TA-, and ME/BL-variables proxy for respectively the historic-, current- and future-profitability. The ME/BL-variable also proxies for the market leverage which can be interpreted as the current

willingness of the financial market to invest in a particular firm. Multiple interpretations are possible for the ME/BL-variable, as the market value of equity is a catchall variable of actual information regarding future earnings, confidence of investors etc. The empirical evidence of a "too-big-to-fail" default protection ¹⁴, and the empirical evidence of a strong negative relationship between Age and the default rate, for Age-values below 10 ¹⁵, motivates the inclusion of the Size-variable and the Age-variable in the credit-scoring models.

We did not try to find an optimal set of model-variables in the logit model. First, it would be beyond the scope of the paper to replicate the numerous studies on finding the optimal set of model-variables ¹⁶. Second, the Z-score variables have a good track record. Third, we believe that a variation in proxies for profitability and leverage would only marginally improve the performance of the credit-scoring models.

2.3 Parameter estimates of the credit-scoring models

Data on agency ratings is obtained from the Standard & Poor's CREDITPRO database ¹⁷. Less than half of the data in CREDITPRO could be linked with COMPUSTAT accounting data ¹⁸. Main part of this data reduction in observations is ascribed to the requirement of stock price data availability, which imposes a selection of public firms. Additionally, only non-financial US firms are selected.

The panel dataset includes the time series of 1772 obligors with a duration of 1 - 21 years. It contains 11890 firm-year observations with a known S&P rating and 1828 firm-year observations with a non-rated S&P status ¹⁹. Each firm-year observation consists of the S&P rating at the end of June and the values of the model-variables known to the public at that date. The market equity values are based on the stock price and total shares outstanding at the end of June. To ensure the accounting information is publicly available, all accounting information refers to the four most recent fiscal quarters ending at least six months before the end of June ²⁰. The six-month lagging condition for accounting information might be somewhat conservative, as most of the accounting data is available in the first few months after the end of a fiscal year/quarter. However, financial information of troubled firms is, in general, typically slower in reaching the financial community.

The panel observations are split into surviving observations and defaulting observations. For the 13447 surviving observations stock price data is available, both at the end of June in current year and at the end of June in subsequent year ²¹. For the 271 defaulting observations a default event happens in subsequent year ²². The dependent binary variable p_i in the logit-regression model (see equation 2.2) equals 1 for surviving observations and equals 0 for defaulting observations.

Table I provides the mean- and median-values for the model-variables after truncation of their most extreme values ²³. The log-transformations of the ME/BL-, RE/TA- and EBIT/TA-variables considerably reduce the skewness in their panel distributions. The panel distribution of the ME/BL-variable becomes even a normal distribution.

The parameters of the DP-model are estimated for the period 1981 - 1999 (see first column of Table II). The signs of all estimated parameters match the expectations. The ME/BL-variable turns out to be the dominant variable in the DP-model ²⁴. This is consistent with the success of the Moody's KMV structural model, in which market equity and total liabilities play a key role. Although the ME/BL-variable is the most important variable, accounting information, and in particular the obligor characteristics Size and Age, add substantially to the explanation of default incidence.

Because of the arbitrary nature of the DP-model, the robustness of the estimated DP-parameters has been extensively tested.

- For two sub-periods, 1981 - 1990 and 1990 - 1999, the DP-parameters largely agree with each other, which demonstrates the time stability of the DP-model for the entire 1981 - 1999 period. Observations in the period 2000 – 2001 are excluded from the estimation of the DP-model. In this period the DP-parameters differ significantly from the period 1981 - 1999 (see Table II). The EBIT/TA-variable has become more informative on credit quality. Furthermore, the "too-big-too-fail" default protection has disappeared. Instead, firms with large Size-values have experienced a higher default rate; 90 firms with liabilities greater than 1 billion defaulted over the 30 month period January 2001 to June 2003 (see Altman and Bana, 2003). We must wait to determine whether these abrupt changes in DP-parameters represent a regime change or are to be ascribed to temporarily exceptional circumstances. For example, a large amount of large liability failures have occurred in the telecommunications sector.
- The DP-parameters only slightly change, at maximum 20%, when controlling the model-variables for industry sector differences ²⁵. When estimating the DP-model for the industry sectors separately, the DP-parameters are comparable ²⁶.
- To ensure that the DP-parameters are not unique for this particular dataset and the Standard & Poor's definition of default, the DP-model is re-estimated for all bankruptcies as reported by COMPUSTAT, omitting only the Age-variable ²⁷. The new bankruptcy dataset covers the 1970 – 1998 period and contains 118154 surviving observations and 755 bankruptcy observations, analogously defined as the surviving observations and defaulting observations in the S&P corporate bond dataset. Only a very small percentage of these bankruptcy observations overlap with the defaulting observations in the S&P corporate bond dataset.

The relative weight of the DP-parameters are robust to a choice in the dataset and robust to the definition of the default event ²⁸. When using the bankruptcy dataset, the relative weight of the DP-parameters is stable over time, varying by at most 20% between the two sub-periods 1970 - 1980 and 1981 - 1998. This allows the DP-model to be considered as an out-of-sample model for the entire period 1981 - 2001.

In summary, for the period 1981 – 1999, the DP-model is stable over time and robust to the definition of default and robust to the dataset choice. The DP-model is applicable for a large range of obligors in different industry sectors and different Size. This emphasizes the universal character of the DP-model, which makes the DP-model a suitable benchmark for agency ratings (excluding the financial sector).

The parameters of the AR-model are estimated for the period 1981 - 1999 (see third column of Table II). All parameter estimates have the expected sign ²⁹. Like the DP-parameters the AR-parameters are robust to a split in sample period: 1981 – 1990 and 1991 – 1999. Also the observations in 2000 and 2001 are excluded. In this period the AR-parameters differ strongly for the EBIT/TA- and ME/BL-variable (see Table II). Notable is the robustness of the AR-parameters to a split of the observations into non-investment graded (BB+ and below) firms and investment graded (BBB- and above) firms ³⁰. The AR-parameters do not depend significantly on the rating level, which enables to model the entire agency-rating scale with one parameter set.

2.4 Comparison of the DP- and AR-parameters

The relative weights of the model-variables in the DP-model and AR-model are compared with each other by the following formula

$$RW_i = \frac{|\beta_i| \sigma_i}{\sum_{j=1}^6 |\beta_j| \sigma_j}$$

where RW_i is the relative weight of model-variable i , β_i is the parameter estimate for model-variable i , and σ_i is the standard deviation of the model-variable i in period 1981 – 1999.

The ME/BL-variable dominates in the DP-model with a RW-value of 41,0 %. The ME/TL-, Size- and Age-variables account for 85.8% of the variation in the DP-score. The WK/TA-, RE/TA- and EBIT/TA-variable play only a minor role. The AR-model gives most weight to the Size-, RE/TA- and ME/BL-variables, in this order of sequence. These three variables account for 78.2% of the variation in the AR-score. Paradoxically, the ignorance of short-term fluctuations and the

exclusive focus on long term credit quality, results in a heavier weight on historic-earnings (RE/TA) and less weight on future-earnings expectations (ME/BL). Presumably it is impossible to forecast beyond 1 - 2 years. Instead, one has to rely on historic earnings track records (RE/TA) and on the proven "too-big-to-fail" default protection (Size).

III Timeliness of agency ratings

3.1 Relationship between CM-scores and agency-rating scale

For each agency-rating class y the average AR-score and average DP-score are calculated. In Figure II these average values are plotted against a numerical agency-rating scale y : CCC/CC/C = 1, B- = 2, B = 3, ..., AA- = 14, AA = 15, and AA+/AAA = 16. This numerical rating scale is an arbitrarily choice, but most obvious.

The relationship between the average CM-scores and the agency ratings y is close to linear. Apparently the CM-scores are sufficiently and nearly equally dispersed over the entire agency-rating scale. On a more detailed level two groups of rating classes with an almost perfect linear relationship can be distinguished: non-investment rating classes 2 – 7 with a slope of 0.405 between the DP-score and y and the investment rating classes 8 – 15 with a slope of 0.307 between the DP-score and y ³¹. At the edge of the numerical agency-rating scale ($y = 1$ and $y = 16$) the DP-scores are deviating from these linear relationships. These rating classes have respectively no lower and upper boundary to the credit quality. The same picture appears for AR-scores, with slope values of 0.690 for $y \in [1, \dots, 7]$ and 0.472 for $y \in [8, \dots, 15]$.

3.2 Definition relative credit-model scores

Within an agency-rating class the credit quality of firms may vary considerably and even significantly overlap with neighboring agency-rating classes. A measure for the divergence in credit quality with respect to the average credit quality in a particular rating class, is the relative credit-model score, ΔCM -score. Each year t , for each firm i in a particular-rating class y , the difference in CM-score $CM_i(y, t)$ and the average CM-score of all firms in that particular rating class y $\overline{CM}(y, t)$ is computed

$$\Delta CM_i(y, t) = CM_i(y, t) - \overline{CM}(y, t) \quad (3.1)$$

Subsequently we convert the ΔCM -scores to the agency-rating notch scale, by the following formula

$$\Delta CM_{i,t} \longrightarrow \frac{\Delta CM_i(y, t)}{\gamma(y, t)} \quad (3.2)$$

where $\gamma(r,t)$ is the slope in the relationship between the credit score CM and the rating class y : $CM = \alpha + \gamma y$ ³² (see Figure II). Since the slope γ depends on time t and rating class y , the correction factor $1/\gamma$ enables a comparison of ΔCM -scores among all rating classes and across time.

The time dependency of γ is illustrated in Table III. For both the DP-scores and the AR-scores the table presents the mean scores in the period 1981 – 1990 and 1991 – 2001. The results suggest a slight increase in the credit quality dispersion in the agency-rating scale. On the upper side of the agency-rating scale (rating classes A and above) the mean CM-scores have increased over time. Blume et al (1998) who reveal the same findings, ascribe this increase in credit quality to the more stringent rating standards set by agencies in their rating assessment. This explanation is consistent with the decrease in number of obligors in the upper side of the agency-rating scale (see Table III). In the lower side of the agency-rating scale (rating classes BBB and below) the mean CM-scores have decreased over time. When assuming a constant interpretation of the CM-scores, in terms of absolute credit quality levels, this would suggest a deterioration in credit quality for lower rating classes. This suggestion is supported by the increase in default rates for the lower rating classes in the last three decades, as reported by Zhou et al. (2001).

The tractable linear relationship between CM-scores and the numerical agency-rating scale is a further support for the ability of CM-scores to benchmark consistently the entire agency-rating scale. Moreover the accuracy of the CM-scores, are comparable for the lower and upper part of the agency-rating scale. An indication for the accuracy level of the CM-scores is the standard deviation in CM-scores within a particular rating class y (see Table IV). After controlling for γ , the standard deviation in CM-scores varies only up to 25%, leaving aside the agency-rating classes 1 and 16.

3.3 Relative credit model-scores conditional on the agency-rating migration event

Average ΔCM -scores are computed, conditional on the agency-rating migration event. These conditional ΔCM -scores, in the years before and after the migration event, provide insight the timing of the agency-rating migrations relative to the “actual” credit quality level.

For each year T , at the end of June, firms are classified into three samples, conditional on the agency-rating migration ΔN in the previous annual period ($T - 1, T$): one sample of firms with an upgrade ($\Delta N > 0$), one sample of firms with a downgrade ($\Delta N < 0$) and one sample of firms with no migration ($\Delta N = 0$). The mean ΔCM -score of these three samples is computed for the 3 years before the migration event and for the 4 years after the migration event, resulting in respectively $\Delta CM_{T,t}(+)$, $\Delta CM_{T,t}(-)$ and $\Delta CM_{T,t}(0)$ with $t \in (T-3, T+3)$. Subsequently these figures are averaged

over the sample period $T \in (1981, 2001)$, resulting in $\Delta CM_t(+)$, $\Delta CM_t(-)$ and $\Delta CM_t(0)$ with $t \in (-3, 3)$.

The composition of the three conditional samples differs for $t \in (-3, 3)$, because of missing observations due to defaults, mergers, delistings etc. For example, firms defaulting in the period $(-1, 0)$ are missing in the mean calculations of ΔCM_0 , ΔCM_1 , ΔCM_2 and ΔCM_3 , which include only the surviving, healthier firms. For this reason the ΔCM -scores are positively biased for $t \in (0, 3)$. However of most interest are the differences between $\Delta CM_t(+)$ and $\Delta CM_t(0)$ and the differences between $\Delta CM_t(-)$ and $\Delta CM_t(0)$, not the absolute levels of the ΔCM -scores. Therefore $\Delta CM_t(+)$ and $\Delta CM_t(-)$ are scaled to $\Delta CM_t(0)$: $\Delta CM_t(+)$ - $\Delta CM_t(0) \rightarrow \Delta CM_t(+)$ and $\Delta CM_t(-)$ - $\Delta CM_t(0) \rightarrow \Delta CM_t(-)$. Table IV presents the $\Delta CM_t(+)$ -scores and $\Delta CM_t(-)$ -scores for $t \in (-3, 3)$.

The average agency rating-migration steps ΔN_t in period $(t, t-1)$, surrounding the migration event at $t = 0$, are computed in a similar way. At the end of June in year T , for the three conditional samples of firms (see above), the average agency-rating migration ΔN is computed for 3 annual periods before the migration event and 4 annual periods after the migration event, resulting in respectively $\Delta N_{T,t}(+)$, $\Delta N_{T,t}(-)$ and $\Delta N_{T,t}(0)$ with $t \in (T-3, T+3)$. Subsequently these figures are averaged over the sample period $T \in (1981, 2001)$, resulting in $\Delta N_t(+)$, $\Delta N_t(-)$ and $\Delta N_t(0)$ with $t \in (-3, 3)$. Just like the ΔCM -scores $\Delta N_t(+)$ and $\Delta N_t(-)$ are scaled to $\Delta N_t(0)$: $\Delta N_t(+)$ - $\Delta N_t(0) \rightarrow \Delta N_t(+)$ and $\Delta N_t(-)$ - $\Delta N_t(0) \rightarrow \Delta N_t(-)$. Table IV presents the $\Delta N_t(+)$ -scores and $\Delta N_t(-)$ -scores for $t \in (-3, 3)$.

At the end of June in year -1, on average one half year before the upgrade, $\Delta AR_{-1}(+)$ equals 1.03. At the end of June in year 0, on average one half year after the upgrade, $\Delta AR_0(+)$ equals 0.41. Apparently the magnitude of the upgrade is not sufficient to reduce $\Delta AR(+)$ to zero values. The new rating level is kept below the actual “through-the-cycle” credit quality. In the years after the upgrade $\Delta AR(+)$ slightly decreases, but remains positive. Just before the upgrade, the actual “through-the-cycle” credit quality deviates by 1.39 notch steps from the average credit quality in the particular rating class ($= \frac{1}{2}\Delta AR_{-1}(+) + \frac{1}{2}\{\Delta AR_0(+) + \Delta N\}$). In case of a downgrade a comparable, but slightly different picture shows up. Just before the downgrade, the actual “through-the-cycle” credit quality deviates by - 0.91 notch steps from the average credit quality in the particular rating class. After the downgrade $\Delta AR_0(-)$ is positive (0.29), which suggests an overreaction in the rating adjustment. In subsequent years $\Delta AR(-)$ even increases to a level of 0.65. These findings point to a prudent migration policy. First, a downgrade is triggered at a lower the threshold than an upgrade. Second, new ratings, after a migration event, are kept on safe side, regardless whether the rating is upgraded or downgraded.

Differences between ΔAR -scores and ΔDP -scores are best demonstrated by a cumulative ΔCM -score

$$\text{Cumulative } \Delta CM_t = \Delta CM_t + \sum_{k=-3}^t \Delta N_k \quad (3.3)$$

The cumulative ΔCM -score reflects the change in actual credit quality relative to the base date $t = -4$, conditional on the rating migration event in period $(-1,0)$. Figure III presents the time-series of the cumulative $\Delta CM_t(+)$ -score and cumulative $\Delta CM_t(-)$ -scores for $t \in (-3,3)$. The cumulative ΔCM -scores are close to zero at $t = -3$ and changes by about two notch steps in the subsequent six year period, both for DP-scores and AR-scores and both for upgrades and downgrades.

Especially for $t = -1, 0$ and 1 the cumulative ΔDP -scores clearly differs from the cumulative ΔAR -scores. Apparently a change in long-term credit quality perspectives coincides with a temporary extra deterioration / rise in short term credit quality. In contrast, as expected, the cumulative ΔAR -scores, which have an exclusive focus on long term, do not show this temporary divergence in credit quality. One possible cause for the difference between cumulative ΔDP -scores and cumulative ΔAR -scores is the relative heavy load of the ME/TL-variable in the DP-scores see section 2.4). A change in long-term credit quality perspectives is associated with a strong temporary “overreaction” in market equity value. It is not yet clear whether this is an overreaction effect or a real temporary change in (short-term) credit quality.

If the investor’s perception on the change in actual credit quality, conditional on an agency-rating migration, is reflected by the cumulative ΔDP -score, than the investor’s dissatisfaction on the timeliness of agency ratings has two reasons

1. The migration policy masks the true dynamics in credit quality. Agency ratings do not immediately respond to changes in the actual “through-the-cycle” credit quality. In chapter 5 the migration policy is characterized in detail.
2. An additional temporary decline / rise in short-term credit quality is perceived by investors, when the long-term perspectives on credit quality change. The “through-the-cycle” methodology ignores this effect.

The results and conclusions presented above are robust to rating level and sample period. The sensitivity of $\Delta CM(+)$ -scores and $\Delta CM(-)$ -scores to investment grade and sample period is tested³³. Only some modest differences appear between investment-graded firms and non-investment graded firms³⁴. This robustness has two important implications. First no major differences in migration policy appear between the high and low credit quality range and no major change in migration policy has been detected between the sample periods 1981 – 1990 and

1991 – 2001. Second, the CM-scores are just as effective in estimating the credit quality in the high credit quality spectrum as they are in the low credit quality spectrum.

IV Agency rating dynamics

A direct measure of rating stability is the rating-migration probability. Both the ignorance of short-term fluctuations in actual credit quality and the prudent migration policy reduce the rating-migration probability. To what extent the rating-migration probability is reduced and the relative importance of these two potential sources, is the aim of the benchmark study below. Earlier preliminary benchmark studies show a large increase in rating migration probabilities, when ratings are based on the EDF-scores of the KMV-model (Moody's, 2002b).

Additionally, the rating drift properties of agency ratings are benchmarked. Bond rating drift, a series of mild downgrades (or upgrades) was first described by Altman and Kao in 1992. Both the direction from which a rating class is reached and the duration of a stay in a particular class, are correlated with the following downgrade or upgrade intensity (see also Lando and Skødeberg, 2002).

4.1 Conversion of CM-scores to CM-ratings

The agency rating dynamics are benchmarked against the dynamics of CM-scores. For a correct comparison, CM scores are converted to CM-ratings, equivalent to agency ratings. Each year, at the end of June, all firms are ranked by their CM-score. On the basis of their ranking sixteen CM-ratings, AAA/AA+, AA, ..., B-, CCC/CC, equivalent to agency ratings, are assigned to firms. Each year, the number of firms within each CM-rating class is exactly equal to the number of firms in the equivalent agency-rating class. In order to be able to calculate average migration numbers, the agency ratings and the CM-ratings are converted to numbers: CCC/CC/C = 1, B- = 2, B = 3, ..., AA- = 14, AA = 15, and AA+/AAA = 16.

4.2 Unconditional rating migration

The average rating migration for all observations is the unconditional rating migration. The unconditional rating migration is evaluated for agency ratings and CM-ratings to make sure differences in rating dynamics are not imposed by the boundaries of the dataset. The unconditional rating migration equals -0.15 for agency ratings and AR-ratings and equals -0.13 for DP-ratings. From a technical perspective, this unconditional downward migration equals to the difference in rating level between firms entering the dataset and firms exiting the dataset, divided by the number of years of unbroken stay in the dataset (= on average 6.35 years). Firms enter the dataset (1) obviously at the beginning of the dataset in 1981, (2) when they are newly rated, (3) when the non-rated agency rating status is lifted up or (4) when COMPUSTAT data becomes available. Firms exit the dataset (1) obviously at the end of the dataset in 2001, (2) in

case of a default event, (3) when a rating changes to a non-rated status or (4) when COMPUSTAT data becomes unavailable ³⁵.

Table V presents the analysis of the unconditional migration of agency ratings, and CM-ratings. These ratings agree on the rating level when firms enter the dataset R_{EN} and exit the dataset R_{EX} . R_{EN} is about 7 (BB+) and R_{EX} is about 6 (BB), which results in an unconditional downward migration of about 1 notch for all firms during their stay within the dataset. Dividing this figure by the duration of 6.35 years of unbroken stay in the dataset, gives the unconditional annual rating migration of - 0.15. Grouping by reasons to exit and enter the dataset, reveals that 80% of this rating migration is due to financial troubled firms in the 2 – 3 years approaching the default event. When eliminating these firms from the analysis, R_{EX} gets close to R_{EN} (see Table V). Obviously, default migrations only occur at the exit and, when leaving everything equal, this creates by obviously an unconditional downward rating migration.

4.3 Rating-migration probability and rating drift

For agency ratings and CM-ratings the distribution of rating migrations M (in terms of notch steps) is determined: $M < -3$, $M = 3$, ..., up to $M = 2$, $M > 2$. The symmetric properties of this distribution are examined by the average rating migration figure and the ratio in number of downgrades D and the number of upgrades U (D/U). The two-sided sign test determines whether the number of downgrades deviates significantly from the number of upgrades.

The one year migration probability, of at least one notch step within one year, is 23.4% for agency ratings, 49% for AR-ratings and 60.7% for DP-ratings, including the migrations to default (see Table VI). The “through-the-cycle”-methodology and the migration policy both significantly add to the rating stability. This finding is robust to the rating migration sign (-,0,+) in previous annual period, robust to sub-samples of non-investment graded firms and investment graded firms and robust to two sample periods 1981 - 1990 and 1991 - 2001.

Conditional on a downgrade in previous annual period, the average agency-rating migration equals - 0.47 for agency ratings. The downward drift in agency ratings appears to be significant. Conditional on no rating migration in previous annual period, the average migration equals - 0.14, about equal to the unconditional rating average rating migration. Conditional on an upgrade in previous annual period, the average agency-rating migration equals +0.08. The upward drift in agency ratings is much smaller than downward drift in agency ratings. This asymmetric behavior in rating drift is reported by Altman and Kao (1992). When controlling for the unconditional rating migration, the rating drift becomes of equal magnitude both for the downside and de

upside, which implies that the underlying source of rating drift is equally effective in both directions.

In contradiction to agency ratings no significant rating drift is observed for DP-ratings and AR-ratings. Conditional on the rating migration in previous year, the average AR-rating migration and the average DP-rating migration do not differ significantly from the unconditional rating migration of respectively -0.15 and -0.13. Rating drift in agency ratings is apparently a migration policy effect.

V Characterization of the migration policy

5.1 Migration policy parameters

The differences in rating-migration probability and rating drift properties between agency ratings and AR-ratings (reflecting the actual “through-the-cycle” credit quality) are ascribed to the migration policy (see Figure I). In order to understand the impact of the migration policy on rating dynamics, we propose a simple model to represent the migration policy of rating agencies. In this model the migration policy is characterized by two migration policy parameters:

- The threshold parameter TH specifies the size of a credit quality interval $[-TH, TH]$, centered to the average credit quality in a particular rating class. In this interval the credit quality is allowed to fluctuate without triggering a rating migration³⁶. This threshold prevents small credit quality fluctuations to trigger a rating migration and as a result the rating-migration probability is reduced.
- The adjustment factor AF reflects the non-full adjustment of agency ratings to a target rating level, when a rating migration is triggered. This partial adjustment of agency ratings, or spreading the target rating adjustment over time, is responsible for the rating drift.

The threshold level TH can be estimated from the time-series of the cumulative ΔAR -scores (see Figure III). The cumulative ΔAR -score at which a rating migration is triggered is about 1.50 notch steps for an upgrade and about -1.25 notch steps for a downgrade (a rating migration occurs on average at $t = -0.5$). So the threshold level TH is likely to be about 1.25 - 1.50 notch steps. A best guess for the adjustment fraction AF can be made as follows. The total agency-rating migration ΔN_{tot} in the period $(-4,3)$, conditional on a downgrade and upgrade in period $(-1,0)$, equals respectively - 2.4 and 1.8 notch steps. On average $2/3$ of this total migration ΔN_{tot} in period $(-4,-3)$ occurs in period $(-1,0)$. The other part occurs in the periods surrounding $(-1,0)$. Apparently, if triggered, ratings are, on average, only adjusted by a fraction $2/3$ to the target rating level. The remainder part of the rating adjustment is executed on a later date. If the intended rating adjustments are *not* spread over more years by the agencies, the average rating migration in the period surrounding $(-1,0)$ is expected to be close to zero (given the Markovian behavior of AR-ratings and the small net reversion rate over the entire rating scale).

5.2 Simulation of agency rating dynamics

A simulation experiment is carried out to estimate in an alternative way the two migration policy parameters TH and AF. The simulation experiment involves three steps. In the first step a modified AR-score is defined, AR^M , which represents a particular migration policy. In the second step the AR^M -scores are converted to $AR(TH,AF)$ -ratings. In the third step the migration policy

parameters TH and AF are determined by searching for a match in rating-migration probability and a match in rating drift properties between agency ratings and AR(TH,AF)-ratings.

Step 1: Modification of AR-scores

For each observation the AR-score is converted to a modified score AR^M , in such a way that it represents a specific migration policy, characterized by a threshold TH and an adjustment factor AF. When following the time-series of the AR_t -scores for a particular firm, the modified AR^M_t -scores are computed. At the beginning of the time-series of each firm, AR^M_0 is set equal to AR_0 . The AR^M_t -score is held constant as long as the AR_t -score stays within the threshold interval

$$AR^M_t = AR^M_{t-1}, \quad \text{if } \frac{|AR_t - AR^M_{t-1}|}{\gamma} < TH \quad (5.1)$$

where $t \in (0, t^{\max})$ and t^{\max} is the duration of a particular firm in the dataset. TH is expressed in notch steps, γ converts the AR-score a notch scale (see section 3.2). As soon as the AR_t -score exceeds the threshold interval, the AR^M_t -score is adjusted: if $AF = 1$, then the AR^M_t -score is fully adjusted to the actual “through-the-cycle” credit quality and set equal to AR_t , if $AF < 1$, then the AR^M_t -score is partially adjusted to the actual “through-the-cycle” credit quality as follows

$$AR^M_t = AF \times (AR_t - AR^M_{t-1}) + AR^M_{t-1}, \quad \text{if } \frac{|AR_t - AR^M_{t-1}|}{\gamma} \geq TH \quad (5.2)$$

Step 2: Conversion of AR^M -scores to AR(TH,AF)-ratings

AR^M -scores are converted to AR(TH,AF)-ratings, equivalent to agency ratings, by following the procedure as described in section 4.1. The time-series of AR^M -scores is an irregular pattern of up- and downward jumps. The duration between these jumps varies in between 1 and t^{\max} years. Crucial in the simulation experiment is the unambiguous conversion of these jumps into AR(TH,AF)-rating migrations. This unambiguous conversion is checked and safeguarded as follows:

- The minimum size of the jump in AR^M -scores is $AF \times TH$, which is sufficient to convert nearly all jumps in the modified AR^M -score into AR(TH,AF)-rating migrations.
- Additionally, the score-rating conversion procedure does not prevent a migration in AR(TH,AF)-ratings, when no jump occurs in the AR^M -score. To prevent these non-intended migrations, AR(TH,AF)-ratings are replaced by it’s lagged rating when AR^M_t equals AR^M_{t-1} . As a consequence the distribution of the AR(TH,AF)-ratings is slightly altered. The number

of observations in each rating class changes by 10% at most. This change in rating distribution does not seriously affect the comparability of the AR(TH,AF)-ratings with agency ratings.

Step 3: Determination of migration policy parameters

Table VII presents the rating-migration probability and rating drift properties of agency ratings and AR(TH,AF)-ratings. The rating drift properties are measured by the average rating migration, conditional on the rating migration in previous year M_{t-1} . Migrations to default are excluded from the simulation experiment, as these migrations are obviously not initiated by rating agencies ³⁷.

The migration policy parameters TH and AF are determined by searching for a match in rating-migration probability and a match in rating drift properties between agency ratings and AR(TH,AF)-ratings. A variation of the threshold TH in the simulation experiment shows that a threshold of 1.25 notch steps best matches the rating migration probabilities of agency ratings and AR(1.25,AF)-ratings. This match does depend on the value of the adjustment fraction AF. The two migration policy parameters appear to influence the rating dynamics nearly independent from each other: the threshold level exclusively affects the rating-migration probability, the adjustment fraction exclusively affects the rating drift properties.

In case of a full adjustment ($AF = 1$) no significant rating drift is observed, regardless of the threshold level. The average rating migration conditional on M_{t-1} does not differ significantly between AR(1.25,1)-ratings and AR-ratings, which are in fact AR(0,1)-ratings ³⁸. This finding is consistent with the supposed Markovian behavior of AR-ratings (see section 4.3 and Table VI) ³⁹. In a further refinement of the simulation experiment the adjustment factor AF is varied in between 0.5 and 1. As expected the rating drift appears as soon as ratings are partly adjusted and the magnitude of the rating drift increases with lower adjustment factors. Best match in rating drift properties with agency ratings is obtained for AR(1.25,0.66)-ratings and AR(1.25,0.83)-ratings. Conclusion is that the rating agencies apply, on average, a threshold of 1.25 and partially adjust rating by a factor of 0.75. These migration policy parameters are not extreme, they point towards a reasonable prudent migration policy.

Table VII compares the one-year migration matrices of agency ratings and AR(1.25, 0.66)-ratings on a major rating level ⁴⁰. The elements in these migration look quite similar. The rating migration probabilities agree within 1%, apart from the CCC/CC rating class. The agency rating dynamics are almost perfectly simulated by AR(1.25, 0.66)-ratings.

In the simulation experiment the target level of the new ratings is set equal to the actual “through-the-cycle” credit quality. However, in reality, the target rating level is set to 0.30 - 0.65 notch steps above the actual “through-the-cycle” credit quality (see Table IV). However the dynamics of AR(TH,AF)-ratings are not sensitive to a choice in the target level.

VI Discussion: in search for the "true" credit quality migration matrix

Crucial input to rating-based credit-pricing models is the agency migration matrix, which contains the information on credit quality dynamics. These models, using the agency-rating migration matrix as input, predict too low credit spreads for corporate bonds. Elton et al (2001) suggest that a higher expected default rate (compensation for default loss) and/or a higher variation in the unexpected default rate (credit risk premium) should account for a relative large part of the credit risk spread. Both a higher expected default rate and a higher variation in the unexpected default rate points towards a more volatile migration matrix.

So far the academic literature on agency-rating migration matrices focuses mostly on the influence of the business cycle, bond rating age and industry (see Altman, 1998, Nickell et al, 2000, Hu et al , 2002 and Bangia et al, 2002). This paper aims to draw the attention on the volatility of the agency-rating migration matrix. As shown in previous sections the stability of agency ratings is significantly enhanced by the migration policy and "through-the-cycle" methodology. With the knowledge of the migration policy parameters and the ability to simulate the agency rating dynamics, we are now in a position to reveal how the agency-rating migration matrix would look like if rating agencies would relax their prudent migration policy and if rating agencies would incorporate more credit quality cycle information in their rating assessment.

Table VIII presents the one-year migration matrices for agency ratings, AR(1.25,0.66)-ratings, AR-ratings and DP-ratings on a major rating level. The influence of the migration policy of rating agencies is illustrated by the AR-ratings. Eliminating the impact of the migration policy, as reflected by AR(0,1)-ratings, increases the rating migration probabilities to neighboring rating classes by a factor 2. Switching to a "point-in-time" rating methodology, instead of a "through-the-cycle" rating methodology and eliminating the influence of the migration policy, as reflected by DP-ratings, increases the off diagonal migration probabilities by a factor 3.

Two dynamic properties of the one-year migration matrix are analyzed in more detail, per rating class: the average rating migration and the rating-migration probability in a one-year period. The rating migration $RM(y)$ for each of the 7 rating classes y ($AAA = 7$, $CCC = 1$) equals

$$RM(y) = \frac{1}{N(y)} \sum_{k=1}^7 M(y,k)(y-k) \quad (5.3)$$

where $N(y)$ is the amount of observations within rating class y and $M(y,k)$ is the amount of migrations from rating class y to rating class k ⁴¹. The rating-migration probability $RMP(y)$ for each of the 7 rating classes y equals

$$RMP(y) = \frac{1}{N(y)} \sum_{k=1}^7 M(y,k) , \quad k \neq y \quad (5.4)$$

RM(y) and RMP(y) are given in Table IX. In order to highlight the rating-migration reversion rates, the default migrations are excluded from the analysis. The default migration trigger is assumed to follow a complete different stochastic process. In absence of default migration events, ratings tend to migrate towards a mean investment grade (see also Altman et al, 1998). Part of this reverting behavior could be ascribed to the restricted number in possible upgrades and downgrades especially for the highest and lowest rating classes. Another part of the reverting behavior is more fundamental. Corporate credit quality, as measured by credit-scoring models, tend to revert towards a mean investment grade, with a rate depending on the current credit quality level. The exact characterization of this statistical process requires further study.

The rating reversion rate and the migration probability are strongly influenced by the "through-the-cycle" methodology and the migration policy. For DP-ratings, representing the investor's perspective on credit quality, the rating migration probabilities and rating reversion rate are roughly a factor 2 – 3 higher for all rating classes, compared to the agency ratings. To what extent a more volatile migration matrix alters the outcome of rating-based credit-pricing models is an interesting topic for follow-up research.

VII Conclusions

Rating stability is achieved by both the prudent migration policy and the “through-the-cycle” rating methodology. These two measures increase the rating stability, in terms of rating-migration probability, by a factor three. This conclusion is obtained by benchmarking the agency rating dynamics against two rather well-established credit scores, which proxy for the actual “through-the cycle” credit quality and the “point-in-time” perspective of investors.

In a simulation experiment the migration policy is characterized by two parameters. A rating migration is triggered, when the actual “through-the-cycle” credit quality exceeds a threshold of 1.25 notch steps, relative to the average “through-the-cycle” credit quality in a particular rating class. If triggered, the rating migration only closes 75% of the gap between the current rating level and the target rating level, which is set on the safe side, 0.3 – 0.6 notch steps below the actual “through-the-cycle” credit quality. Although these parameters seem not to point to an extreme prudent migration policy, they are sufficient to cause a significant enhancement in rating stability.

The investor’s opinion on the timeliness of agency ratings could be partly explained by a conservative migration policy. The other part can be ascribed to the “through-the-cycle” rating methodology, which ignores the temporary decline / rise in credit quality, when the long term credit quality perspective changes. Further study is required to find out whether the “through-the-cycle” methodology corrects for an overreaction effect or represents a real change in short-term credit quality.

Literature

- Altman, E.I., 1968, "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy", *Journal of Finance* 35, 1001 - 1016
- Altman E.I. and D. L. Kao, 1992, "The implications of corporate bond ratings drift ", *Financial Analysts Journal*, June, 64 – 75
- Altman E.I., 1998, "The importance and subtlety of credit rating migration", *Journal of Banking & Finance* 22, 1231 – 1247
- Altman E.I. and G. Bana, 2003, "Default and returns on high yield bonds: the year 2002 in review and the market outlook", report NYU Salomon Center, Stern School of Business
- Altman E.I. and H.A. Rijken, 2003, "Benchmarking the credit agency ratings with credit scoring models", working paper
- Association for Financial Professionals, 2002, "Ratings Agencies Survey: accuracy, timeliness, and regulation", November, www.afponline.org
- Bangia A. F. X. Diebold, A Kronimus, C. Schagen and T. Schuermann, 2002, "Ratings migration and the business cycle with application to credit portfolio stress testing", *Journal of Banking & Finance* 26, 445 – 474
- Baker H.K. and S.A. Mansi, 2001, "Assessing credit rating agencies by bond issuers and institutional investors", working paper
- Blume M.E., F. Lim and A.C. MacKinlay, 1998, "The declining credit quality of U.S. Corporate Debt: Myth or Reality ?", *Journal of Finance* 53, 1389 - 1413
- Carey M. and M. Hrycay, 2001, "Parameterizing credit risk models with rating data", *Journal of Banking and Finance* 25, 197 - 270
- Ederington, L.H. , 1985, "Classification models and bond ratings", *The Financial Review* 20 (4), 237 – 262
- Ellis D., 1998, "Different sides of the same story: investors' and issuers' views of rating agencies", *The Journal of Fixed Income* 9, 35 – 45
- Elton E.J., M.J. Gruber, D. Agrawal and C. Mann, 2001, "Explaining the rate spread on corporate bonds", *The Journal of Finance* 56, 247 - 277
- Hu, Y, R. Kiesel and W. Perraudin, 2002, "The estimation of migration matrices for sovereign credit ratings", *Journal of Banking & Finance* 26, 1383 – 1406
- Kaplan, R.S. and G. Urwitz, 1979, "Statistical models of bond ratings: a methodology inquiry", *Journal of Business* 52, 231 - 261
- Kamstra M., P. Kennedy and T. Suan, 2001, "Combining bond rating forecasts using logit", *The Financial Review* 36, 75 - 96
- Lando D. and T.M. Skodeberg, 2002, "Analyzing rating migrations and rating drift with continuous observations", *Journal of Banking & Finance* 26, 423 – 444
- Löffler G., 2002, "An anatomy of rating through-the-cycle", to be published in *the Journal of Banking and Finance*
- McQuown J.A. , 1993, "A comment on market vs. accounting based measures of default risk", published by KMV (now Moody's)
- Moody's Investor Services, 2002a, "Understanding Moody's corporate bond ratings and rating process", special comment series

- Moody's Investor Services, 2002b, "Default & Recovery rates of corporate bond issuers", special comment series
- Nickel P., W. Perraudin and S. Varotto, 2000, "Stability of rating migrations", *Journal of Banking & Finance* 24, 203 – 227
- Saunders A. and L. Allen, 2002, "Credit risk measurement: new approaches to Value at Risk and other paradigms", John Wiley & Sons, 2nd edition, ISBN 047121910X
- Standard and Poor's, 2002, Special Report 2002
- Standard and Poor's, 2003, "Corporate Ratings criteria", www.standardandpoors.com
- Shumway T. , 2001, "Forecasting bankruptcy more accurately: a simple hazard model", *Journal of Business* 74 (1), 101 – 124
- Treacy W.F. and M. Carey, 2000, "Credit rating systems at large US banks", *Journal of Banking & Finance* 24, 167 - 201
- Zhou, C., 2001, "Credit rating and corporate defaults", *The Journal of Fixed Income* 11, 30 – 40

Figure I Concept of the benchmark study

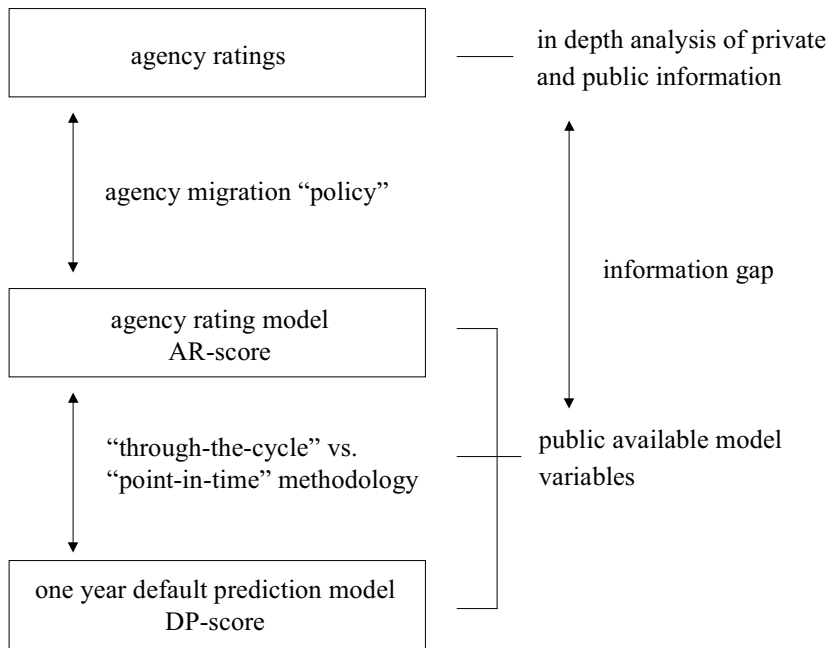


Table I Descriptive statistics of the credit scoring model-variables

The table presents the descriptive panel data statistics of the credit scoring model-variables for all 13718 observations in the period 1981 - 2001, including the 271 defaulting observations of firms less than one year before default (1Y before D). Out of these 271 defaulting observations 167 observations have sufficient data on the model-variables after the default event (D). The numbers in the first row refer to COMPUSTAT data codes.

agency rating	N	WK/TA (4 - 5)/6	RE/TA -ln(1-36/6)	EBIT/TA -ln(1-178/6)	ME/TL 1+ln(ME/181)	Size ln(181/Mkt)	Age
mean statistic per rating class							
AAA	317	0.15	0.64	0.17	1.98	-6.53	9.14
AA	1198	0.12	0.52	0.14	1.67	-7.51	9.07
A	2885	0.14	0.39	0.12	1.35	-7.97	8.72
BBB	2603	0.14	0.27	0.10	1.05	-8.48	7.77
BB	2396	0.18	0.11	0.09	0.82	-9.37	6.01
B	2323	0.22	-0.03	0.04	0.56	-10.05	5.29
CCC/CC	168	0.08	-0.21	-0.05	-0.66	-10.08	5.45
NR	1828	0.26	0.21	0.08	1.28	-10.55	9.06
mean of observations one year preceding default (1Y before D) and at default (D)							
1Y before D	271	0.11	-0.17	-0.02	-0.83	-10.02	5.33
D	167	-0.13	-0.45	-0.07	-2.44	-9.86	5.88
statistics for all 13718 panel observations (excluding D ratings)							
mean	13718	0.17	0.23	0.092	1.07	-8.95	7.53
median	13718	0.15	0.20	0.092	1.08	-9.05	10
St. dev.	13718	0.19	0.35	0.087	1.11	1.66	3.37
min	13718	-0.51	-1.22	-0.33	-3.78	-15.04	1
max	13718	0.73	1.56	0.40	4.16	-3.18	10
kurtosis	13718	3.36	5.68	7.13	3.83	2.89	2.04
skewness	13718	0.42	0.03	-0.42	-0.22	0.20	-0.86

Table II Parameter estimates of the default-prediction model and the agency-rating model

The dependent binary variable in the logit regression model estimation equals 0 for the observations one year preceding the default event and 1 otherwise. The dependent variable in the ordered logit regression model estimation is the numerical agency-rating scale: CCC/CC = 1, B- = 2, AA = 15, AA+/AAA = 16. The standard errors are Huber-White standard errors to control for firm dependence. The pseudo R² is a measure for the goodness of the fit.

	default-prediction model (logit model)		Agency-rating model (ordered logit model)	
	1981 - 1999	2000 - 2001	1981 - 1999	2000 - 2001
const	7.06 (9.35)	0.68 (0.76)	-	-
WK/TA	0.41 (0.75)	-1.37 (2.11)	-1.80 (6.27)	-1.71 (4.39)
RE/TA	0.20 (0.58)	0.89 (1.94)	3.33 (16.48)	2.84 (10.59)
EBIT/TA	3.78 (3.79)	8.36 (5.76)	4.70 (8.70)	8.11 (9.80)
ME/TL	1.28 (12.83)	1.01 (9.73)	0.85 (15.89)	0.52 (11.15)
Size Liabilities	0.47 (6.40)	-0.24 (2.84)	0.95 (17.59)	0.93 (14.78)
Age	0.20 (6.83)	0.14 (3.63)	0.082 (6.76)	0.082 (4.02)
	boundaries B _k			
AAA/ AA+	-	-	-0.24	0.26
AA	-	-	-1.61	-0.69
AA-	-	-	-2.25	-1.51
A+	-	-	-3.06	-2.53
A	-	-	-4.14	-3.66
A-	-	-	-4.73	-4.31
BBB+	-	-	-5.30	-5.04
BBB	-	-	-5.99	-5.79
BBB-	-	-	-6.55	-6.53
BB+	-	-	-7.01	-7.00
BB	-	-	-7.64	-7.73
BB-	-	-	-8.56	-8.78
B+	-	-	-10.31	-10.15
B	-	-	-11.65	-11.52
B-	-	-	-13.05	-12.93
CCC/CC	-	-	- ∞	- ∞
pseudo R ²	0.355	0.374	0.214	0.231
N observations	11990	1728	10345	1545
N obs. 1 year preceding default	150	121	-	-

Figure II Average CM-scores for 16 agency-rating classes

Average DP-scores and average AR-scores for all panel observations in a particular agency-rating class y .

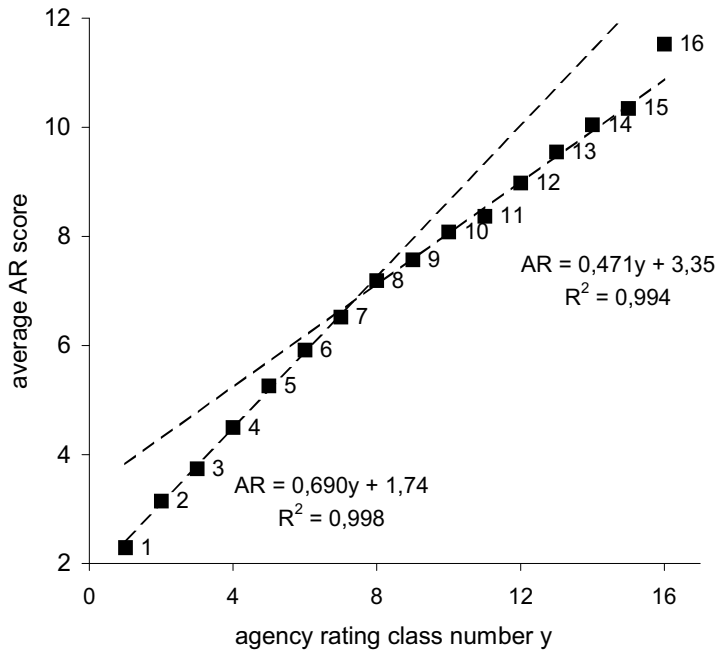
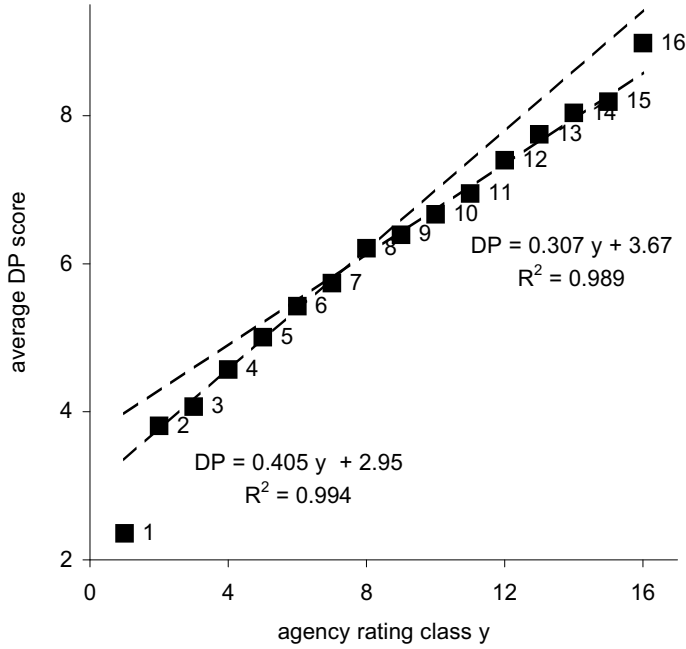


Table III Descriptive statistics of CM-scores for 16 rating classes in two time periods

The DP-scores and AR-scores are computed with the estimated parameters for the period 1981 – 1999. The AR-scores are scaled to the lower boundary of the B- rating class. The table gives the mean, median and standard deviation statistics of the CM-scores for two periods 1981 – 1990 and 1991 – 2001.

		mean		median		standard dev.		N	
		81 - 90	91 - 01	81 - 90	91 - 01	81 - 90	91 - 01	81 - 90	91 - 01
DP-score									
16	AAA/ AA+	8.88	9.09	8.92	9.17	1.39	1.68	236	206
15	AA	7.88	8.69	8.02	8.80	1.19	1.48	379	229
14	AA-	7.80	8.26	7.87	8.47	1.13	1.24	222	243
13	A+	7.61	7.91	7.74	8.00	1.06	1.23	409	359
12	A	7.20	7.59	7.28	7.68	1.04	1.21	637	676
11	A-	6.86	7.02	7.03	7.06	1.08	1.18	320	484
10	BBB+	6.66	6.67	6.74	6.75	0.96	1.13	302	541
9	BBB	6.34	6.43	6.36	6.45	1.09	1.06	405	580
8	BBB-	6.27	6.18	6.29	6.21	1.19	1.21	236	539
7	BB+	5.77	5.72	5.80	5.78	0.98	1.36	181	397
6	BB	5.75	5.30	5.75	5.32	1.24	1.50	229	547
5	BB-	5.46	4.81	5.39	4.82	1.25	1.52	324	718
4	B+	4.86	4.38	4.87	4.41	1.33	1.69	571	844
3	B	4.46	3.88	4.56	3.79	1.59	1.82	202	409
2	B-	3.96	3.72	4.11	3.69	1.50	1.89	109	188
1	CCC/CC	3.32	1.90	3.63	2.00	1.52	2.22	54	114
AR-score									
16	AAA/ AA+	11.51	11.55	11.98	11.66	1.92	2.42	236	206
15	AA	9.96	11.01	10.31	11.07	1.68	2.40	379	229
14	AA-	9.91	10.18	10.06	10.39	1.40	1.83	222	243
13	A+	9.39	9.73	9.52	9.65	1.35	1.54	409	359
12	A	8.84	9.11	8.85	9.13	1.21	1.55	637	676
11	A-	8.51	8.27	8.52	8.31	1.46	1.55	320	484
10	BBB+	8.36	7.92	8.33	7.93	1.06	1.35	302	541
9	BBB	7.74	7.45	7.71	7.50	1.26	1.36	405	580
8	BBB-	7.58	7.02	7.73	7.02	1.25	1.30	236	539
7	BB+	6.77	6.41	6.83	6.41	1.26	1.37	181	397
6	BB	6.51	5.67	6.65	5.71	1.33	1.45	229	547
5	BB-	5.93	4.95	5.86	4.97	1.31	1.37	324	718
4	B+	5.02	4.15	4.92	4.22	1.23	1.63	571	844
3	B	4.45	3.39	4.50	3.37	1.68	1.77	202	409
2	B-	3.89	2.72	3.80	2.78	1.74	1.73	109	188
1	CCC/CC	3.26	1.82	3.59	1.91	2.37	2.54	54	114

Table IV Relative CM-score conditional on a rating migration event

The table presents the ΔCM_t -scores and the agency-rating migration ΔN_t conditional on a rating migration event in period (-1,0) for $t \in (-3,3)$. For the definition of $\Delta CM(+)$ -score, $\Delta CM(-)$ -score, $\Delta N(+)$ and $\Delta N(-)$ see text. All ΔCM -scores and ΔN -values and are given in terms of notch steps and their standard deviations are given in brackets.

agency migration event in period (-1,0)		year t						
		-3	-2	-1	0	1	2	3
ΔAP -score at the end of June in year t								
upgrade	$\Delta AR(+)$	0.45 (2.84)	0.64 (2.69)	1.03 (2.65)	0.41 (2.62)	0.39 (2.64)	0.18 (2.61)	0.16 (2.63)
downgrade	$\Delta AR(-)$	-0.10 (2.72)	-0.16 (2.59)	-0.51 (2.65)	0.29 (2.60)	0.37 (2.62)	0.65 (2.66)	0.65 (2.58)
ΔAR -score at the end of June in year t								
upgrade	$\Delta DP(+)$	0.28 (3.70)	0.77 (3.63)	1.57 (3.62)	1.33 (3.66)	1.06 (3.30)	0.68 (3.19)	0.50 (3.13)
downgrade	$\Delta DP(-)$	-0.33 (3.40)	-0.59 (3.15)	-1.40 (3.24)	-0.77 (3.19)	-0.34 (3.08)	0.27 (3.16)	0.42 (3.00)
ΔN agency-rating migration in period (t - 1,t)								
upgrade	$\Delta N(+)$	-0.03 (1.13)	-0.01 (1.05)	0.14 (0.87)	1.34 (0.70)	0.22 (0.78)	0.13 (0.93)	0.01 (1.05)
downgrade	$\Delta N(-)$	0.00 (0.81)	-0.15 (0.93)	-0.19 (0.86)	-1.63 (1.03)	-0.33 (1.23)	-0.15 (1.18)	0.05 (1.03)
number of observations								
upgrade	-	648	770	904	904	798	713	618
no migration	-	5525	6544	7862	7874	6620	5559	4720
downgrade	-	915	1093	1251	1255	993	819	682

Figures III Cumulative CM-scores conditional on a agency-rating migration event

The figure shows the cumulative $\Delta CM_t(+)$ -score and cumulative $\Delta CM_t(-)$ -scores for $t \in (-3,3)$. The expressions cum $\Delta DP(+)$ and cum $\Delta DP(-)$ refer to the cumulative ΔCM -score of the DP-model conditional on respectively an upgrade in period $(-1,0)$ and a downgrade in period $(-1,0)$. Comparable definitions hold for cum $\Delta AR(+)$ and cum $\Delta AR(-)$.

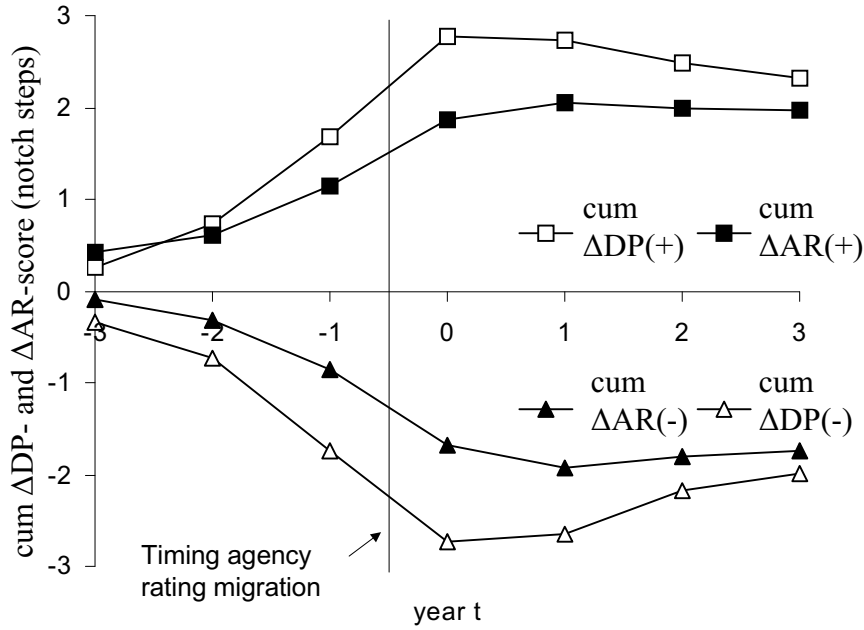


Table V Unconditional rating migration

The table presents the calculation of the unconditional rating migration of agency ratings, AR-ratings and DP-ratings. From a technical perspective the average rating migration is the difference between the average rating level of firms entering the dataset and the average rating level of firms exiting the dataset divided by the average number of years of unbroken stay in the dataset (= 6.35 years).

category	reasons for firms to enter or exit the dataset	N	weight	average agency ratings	average AR-ratings	average DP-ratings
enter	start dataset in 1981	442	26.8%	9.98	10.00	9.94
	newly rated after 1981	942	57.2%	6.05	5.82	5.65
	NR status lifted up	89	10.7%	7.29	7.13	6.29
	COMP. data becomes available ¹	176	5.4%	6.70	7.31	8.53
exit	end of the dataset in 2001	621	37.7%	7.80	7.65	7.47
	default	240	14.5%	0	0	0
	rating changed to NR status	340	20.6%	5.55	5.72	6.69
	COMP. data becomes unavailable ¹	448	27.2%	8.10	7.89	7.90
no exit no enter		8621	-	9.22	9.23	9.15
total enter		1649	100.0%	7.24	7.17	7.14
total exit (excl default)		1410	85.5%	7.35	7.26	7.42
total exit		1649	100.0%	6.28	6.20	6.34
exit - enter				-0.95	-0.96	-0.80
unconditional annual migration				-0.15	-0.15	-0.13

Table VI Panel distribution of migrations for agency ratings, AR-ratings and DP-ratings

The table presents, for agency ratings, AR-ratings and DP-ratings, the panel distributions of the rating migration M, (1) unconditionally for all observations, (2) conditionally to the migration in previous year, (3) for a sub-samples of non-investment graded firms (rating number ≤ 7) and investment graded firms (rating number ≥ 8) and (4) for two periods 1981 – 1990 and 1991 – 2001. In last 4 columns the symmetric properties of these panel distributions are assessed in terms of the average rating migration and the ratio in number of rating downgrades and number of rating upgrades (D/U ratio). The two-sided sign test determines whether the number of downgrades significantly differs from the number of upgrades.

	N	distribution rating migration steps M								D/U ratio	sign test z -value	rating migration	
		< -3	-3	-2	-1	0	1	2	> 2			mean	sd
agency rating													
all	10257	1.5%	1.6%	3.6%	7.9%	76.6%	6.5%	1.8%	0.5%	1.65	-12.7	-0.15	0.97
downgrade previous year	1076	3.2%	3.8%	6.4%	13.8%	68.4%	3.1%	0.9%	0.5%	6.08	4.8	-0.47	1.23
no change previous year	6738	1.4%	1.4%	3.5%	7.9%	76.8%	6.8%	1.8%	0.4%	1.57	-9.4	-0.14	0.96
upgrade previous year	806	0.6%	0.9%	0.9%	4.2%	78.9%	10.5%	3.4%	0.6%	0.46	-13.4	0.08	0.78
investment graded	6304	0.9%	0.9%	3.4%	7.8%	80.4%	5.4%	1.0%	0.2%	1.95	-11.7	-0.14	0.84
non-investment graded	3953	2.5%	2.7%	3.9%	8.0%	70.7%	8.4%	2.9%	0.9%	1.39	-6.3	-0.17	1.14
1981 – 1990	4336	1.2%	1.5%	4.1%	6.7%	78.2%	5.7%	2.0%	0.7%	1.61	-7.6	-0.13	0.95
1991 – 2001	5921	1.7%	1.7%	3.2%	8.8%	75.5%	7.2%	1.6%	0.4%	1.67	-10.7	-0.17	0.99
AR-rating													
all	10257	1.4%	1.6%	4.5%	20.0%	51.0%	17.9%	3.0%	0.5%	1.29	-10.2	-0.15	1.12
downgrade previous year	1854	1.1%	1.3%	5.2%	23.7%	46.9%	17.3%	3.9%	0.6%	1.44	-5.8	-0.16	1.09
no change previous year	4588	0.9%	0.9%	2.9%	18.7%	58.7%	15.7%	1.9%	0.3%	1.31	-6.3	-0.11	0.95
upgrade previous year	2178	2.5%	2.7%	6.3%	19.2%	43.4%	21.7%	3.7%	0.6%	1.18	-4.5	-0.21	1.30
investment graded	6290	1.0%	1.5%	4.1%	19.6%	55.9%	16.0%	1.7%	0.2%	1.46	-10.9	-0.18	1.03
non-investment graded	3967	2.1%	1.8%	5.2%	20.7%	43.2%	21.0%	5.1%	1.0%	1.10	-3.2	-0.12	1.25
1981 – 1990	4336	1.0%	1.6%	4.5%	18.7%	52.0%	18.6%	3.0%	0.6%	1.16	-4.2	-0.11	1.06
1991 – 2001	5921	1.7%	1.7%	4.5%	21.0%	50.3%	17.4%	3.0%	0.4%	1.39	-9.8	-0.19	1.16
DP-rating													
all	10257	2.7%	2.8%	6.9%	19.9%	39.3%	18.6%	6.6%	3.1%	1.14	-6.1	-0.13	1.51
downgrade previous year	2615	3.4%	3.4%	8.0%	19.5%	34.8%	20.4%	7.5%	3.1%	1.11	-2.5	-0.16	1.59
no change previous year	3531	1.2%	1.7%	5.2%	19.2%	47.9%	18.4%	4.8%	1.5%	1.10	-2.4	-0.07	1.19
upgrade previous year	2474	2.5%	2.6%	7.2%	21.3%	36.8%	18.6%	7.0%	4.0%	1.14	-3.2	-0.09	1.52
investment graded	6206	3.2%	3.1%	7.2%	20.0%	42.1%	17.8%	5.0%	1.6%	1.37	-11.1	-0.27	1.45
non-investment graded	4051	2.0%	2.4%	6.4%	19.8%	35.1%	19.9%	9.0%	5.4%	0.89	3.6	0.08	1.55
1981 – 1990	4336	2.6%	2.5%	5.9%	17.8%	39.6%	19.8%	7.9%	4.0%	0.91	2.1	0.00	1.52
1991 – 2001	5921	2.8%	3.0%	7.6%	21.5%	39.2%	17.8%	5.6%	2.4%	1.35	-9.9	-0.22	1.48

Table VII Impact migration policy parameters on rating dynamics properties

The table presents the influence of the threshold parameter TH and the adjustment factor AF on the rating-migration probability and rating drift properties of AR(TH,AF)-ratings. The average rating migration, standard deviation of rating migrations and rating-migration probability are given unconditionally for all observations ("all") and conditionally on the migration in previous year.

rating	all	rating migration in previous year				
		< -1	-1	0	1	> 1
average rating migration						
agency rating	-0.09	-0.30	-0.30	-0.08	0.06	0.18
AR-rating, TH = 0, AF = 1	-0.07	0.09	-0.06	-0.07	-0.15	-0.06
AR-rating, TH = 1.25, AF = 1	-0.06	-0.10	-0.08	-0.07	0.00	-0.01
AR-rating, TH = 1.25, AF = 0.83	-0.06	-0.25	-0.15	-0.07	0.14	0.19
AR-rating, TH = 1.25, AF = 0.66	-0.06	-0.38	-0.31	-0.05	0.21	0.33
AR-rating, TH = 1.25, AF = 0.50	-0.06	-0.59	-0.33	-0.05	0.29	0.50
standard deviation rating migration						
agency rating	0.82	1.16	0.92	0.79	0.70	0.85
AR-rating, TH = 0, AF = 1	0.93	1.21	0.99	0.82	0.98	1.28
AR-rating, TH = 1.25, AF = 1	0.78	0.84	0.70	0.78	0.70	0.86
AR-rating, TH = 1.25, AF = 0.83	0.78	0.84	0.70	0.78	0.70	0.86
AR-rating, TH = 1.25, AF = 0.66	0.68	0.83	0.71	0.65	0.69	0.80
AR-rating, TH = 1.25, AF = 0.50	0.60	0.80	0.63	0.55	0.63	0.75
rating-migration probability						
agency rating	8416	4.4%	7.4%	78.7%	7.1%	2.4%
AR-rating, TH = 0, AF = 1	8416	5.2%	19.0%	53.8%	18.3%	3.6%
AR-rating, TH = 1.25, AF = 1	8416	6.3%	5.5%	79.4%	4.6%	4.2%
AR-rating, TH = 1.25, AF = 0.83	8416	5.2%	6.9%	79.4%	5.3%	3.3%
AR-rating, TH = 1.25, AF = 0.66	8416	3.8%	8.4%	79.5%	6.0%	2.3%
AR-rating, TH = 1.25, AF = 0.50	8416	2.7%	9.1%	80.5%	6.5%	1.2%

Table VIII One-year rating migration matrix

The elements in the migration matrix (y_{t-1}, y_t) represent the one-year probability of a migration from rating class y_{t-1} to rating class y_t .

		agency rating						
	AAA	AA	A	BBB	BB	B	CCC	default
AAA	0.92	0.07	0.01	0.00	0.00	0.00	0.00	0.00
AA	0.01	0.92	0.07	0.01	0.00	0.00	0.00	0.00
A	0.00	0.02	0.91	0.07	0.00	0.00	0.00	0.00
BBB	0.00	0.00	0.05	0.88	0.06	0.01	0.00	0.00
BB	0.00	0.00	0.00	0.06	0.84	0.07	0.00	0.02
B	0.00	0.00	0.00	0.00	0.06	0.82	0.04	0.08
CCC/CC	0.00	0.00	0.00	0.00	0.01	0.14	0.48	0.38

		AR(1.25,0.66)-rating						
	AAA	AA	A	BBB	BB	B	CCC	default
AAA	0.92	0.07	0.01	0.00	0.00	0.00	0.00	0.00
AA	0.03	0.88	0.09	0.00	0.00	0.00	0.00	0.00
A	0.00	0.02	0.92	0.06	0.00	0.00	0.00	0.00
BBB	0.00	0.00	0.04	0.89	0.05	0.00	0.00	0.01
BB	0.00	0.00	0.00	0.05	0.85	0.08	0.00	0.02
B	0.00	0.00	0.00	0.00	0.06	0.84	0.03	0.07
CCC/CC	0.00	0.00	0.00	0.00	0.00	0.16	0.50	0.34

		AR-rating						
	AAA	AA	A	BBB	BB	B	CCC	default
AAA	0.82	0.17	0.01	0.00	0.00	0.00	0.00	0.00
AA	0.04	0.81	0.15	0.00	0.00	0.00	0.00	0.00
A	0.00	0.05	0.84	0.11	0.00	0.00	0.00	0.00
BBB	0.00	0.00	0.10	0.77	0.12	0.01	0.00	0.00
BB	0.00	0.00	0.00	0.13	0.71	0.14	0.00	0.02
B	0.00	0.00	0.00	0.00	0.13	0.75	0.04	0.08
CCC/CC	0.00	0.00	0.00	0.00	0.00	0.25	0.38	0.38

		DP-rating						
	AAA	AA	A	BBB	BB	B	CCC	default
AAA	0.78	0.21	0.00	0.00	0.00	0.00	0.00	0.00
AA	0.05	0.74	0.20	0.01	0.00	0.00	0.00	0.00
A	0.00	0.08	0.76	0.14	0.02	0.00	0.00	0.00
BBB	0.00	0.00	0.16	0.64	0.17	0.02	0.00	0.00
BB	0.00	0.00	0.01	0.22	0.58	0.18	0.00	0.01
B	0.00	0.00	0.00	0.01	0.18	0.69	0.05	0.07
CCC/CC	0.00	0.00	0.00	0.00	0.01	0.26	0.22	0.51

Table IX Rating migration and rating-migration probability in a one-year migration matrix

The table presents the average rating migration RM and rating-migration probability RMP in the one-year migration matrix (for definitions of RM and RMP see section 5.3) for agency ratings, AR(TH,AF)-ratings and DP-ratings. Migrations to default are excluded from the computation of these figures.

rating	AAA	AA	A	BBB	BB	B	CCC	all
rating migration RM								
agency rating	-0.086	-0.073	-0.060	-0.024	-0.012	0.030	0.247	-0.029
AR(1.25,0.66)-rating	-0.088	-0.062	-0.048	-0.016	-0.029	0.033	0.244	-0.026
AR(1.25,1)-rating	-0.110	-0.067	-0.062	-0.016	-0.020	0.046	0.276	-0.027
AR-rating	-0.193	-0.118	-0.070	-0.036	-0.013	0.101	0.398	-0.031
DP-rating	-0.219	-0.164	-0.114	-0.047	0.053	0.160	0.565	-0.026
rating-migration probability RMP								
agency rating	8.0%	8.1%	9.0%	11.5%	14.3%	10.5%	23.5%	10.9%
AR(1.25,0.66)-rating	7.8%	12.1%	8.2%	10.2%	13.7%	9.0%	24.4%	10.4%
AR(1.25,1)-rating	8.5%	13.4%	10.6%	13.7%	16.2%	10.7%	27.6%	12.8%
AR-rating	17.9%	19.2%	15.7%	22.3%	27.9%	18.5%	39.8%	20.8%
DP-rating	21.6%	26.1%	24.0%	36.0%	40.8%	25.9%	55.1%	30.9%

Endnotes

- ¹ The critique on rating agencies is mainly focused on the timeliness properties of agency ratings, not on the accuracy level itself. The AFP survey reveals that 83% of the investors believe that most of the time, agency ratings reflect accurately the issuers' creditworthiness.
- ² In their disclosure on corporate ratings criteria, Standard & Poor's presents their view on the interpretation of their credit ratings (Standard & Poor's, 2003, Corporate Ratings Criteria, page 41): *"Standard & Poor's credit ratings are meant to be forward looking; that is, their time horizon extends as far as is analytically foreseeable. Accordingly, the anticipated ups and downs of business cycles - whether industry specific or related to the general economy - should be factored in the credit rating all along. This approach is in keeping with Standard's and Poor's belief that the value of its rating products is greatest when it's rating does not fluctuate with near term performance. Ratings should never be a mere snapshot of the present situation. There are two models for how cyclicity is incorporated in credit ratings. Sometimes, ratings are held constant throughout the cycle. Alternatively, the rating does vary- but within a narrow band"*.
- ³ The designation "migration policy" in this paper does not refer to an official migration policy of Standard & Poor's, Moody's and Fitch.
- ⁴ In the remainder of this paper, agency ratings refer to the corporate issuer credit ratings of agencies.
- ⁵ In Altman and Rijken (2003) the default prediction performance of DP-scores and agency ratings are compared with each other. On a one-year time horizon the DP-scores show a better default prediction performance than the agency ratings. On a longer time horizon the agency ratings perform slightly better.
- ⁶ In Altman and Rijken (2003) a long term default-prediction model is estimated by modeling the default rates for time horizons up to six years. The benchmark results for this credit-scoring model do not differ substantially from the DP-model, which models the one-year default rates. Therefore we have ignored in this paper the influence of the time horizon on the "point-in-time" credit quality estimates.
- ⁷ Rating numbers 1 – 16 are assigned to respectively the agency-rating classes CCC, B-, B, B+, BB-, BB, BB+, BBB-, BBB, BBB+, A-, A, A+, AA-, AA, and AA+/AAA. In order to have a reasonable number of observations in each rating class the agency rating classes C, CC, CCC-, CCC and CCC+ are combined to CCC/CC and the agency ratings classes AA+ and AAA are combined to AA+/AAA.
- ⁸ Most bond classification models are based on an ordered probit model. In order to be consistent with the default-prediction model we have chosen for an ordered logit model.
- ⁹ Thereby ignoring the influence of the time horizon in the "point-in-time" credit quality estimate.
- ¹⁰ The Age-variable is set to 10 for observations with Age-values above 10 and set to 10 for observations of firms already rated at the start of the dataset in 1981.
- ¹¹ The distribution of the ME/BL-variable is extreme positively skewed. To a lesser extent the distributions of the RE/TA-variable and EBIT/TA-variable are negatively skewed. The information content in the fat tails is relatively low. For example, the difference between a ME/BL-value of 50 and 25 is far less informative than a difference between a ME/BL-value of 2 and 1, which might distinguish a healthy firm from a firm approaching default. The effectiveness of the ME/BL-variable in the logit regression estimation can be improved by a log-transformation of the ME/BL-variable: $\rightarrow 1 + \ln(\text{ME/BL})$. This stretches the informative part at the lower side of the ME/BL-scale and compresses the non-informative part at the upper side of the ME/BL-scale. For the same reason the RE/TA- and EBIT/TA-variables are log-transformed: $-\ln(1 - \text{RE/TA})$ and $-\ln(1 - \text{EBIT/TA})$. The log-transformation reduces the skewness in the distributions of these variables. The average value of these distributions is hardly affected as the log-transformation is centered around one for the ME/BL-variable and centered around zero for the EBIT/TA- and RE/TA-variables.

-
- ¹² The sales-to-asset ratio is not included in the default-prediction model. This variable has little additional value in a default-prediction model, when estimated for a sample of firms covering a wide range of industries.
- ¹³ In a report on their rating methodology, Standard & Poor's (2003) describes a set of 8 key ratios. These ratios include two interest coverage ratios, two cash flow ratios, two earnings profitability ratios and two leverage ratios. In numerous empirical studies on credit-scoring models, different sets of variables are proposed to proxy for these four groups of credit-risk fundamentals. In general interest coverage ratios and cash flow ratios appear to add surprisingly little in explaining default. Presumably, the strong correlation with earnings profitability and leverage prevents a significant marginal contribution. Moreover interest coverage ratios suffer often from ambiguity problems, because both the denominator values (interest) and numerator values (EBIT) are centered close to zero. Therefore only the profitability and leverage ratios are included in the benchmark credit scoring models.
- ¹⁴ Often included in default-prediction models is the size of total market equity. However, this variable strongly correlates with the ME/BL-variable. Instead, we include the size of total liabilities. Apart from this technical reason, the size of total liabilities is more directly related to the "too-big-to-fail" protection. One explanation for the "too-big-to-fail" protection is that credit holders might shy away from the large consequences of a default or bankruptcy and hope problems will be solved by time. The potential loss of large loans, potential damage to bank reputations and the number of credit holders involved, might slow down the decision process, thereby leaving time for large companies with large loans to restructure themselves.
- ¹⁵ A strong negative relationship exists between Age and the default rate, for Age-values below 10. An exception forms the low default rate in the first year after being firstly rated (Age = 1). This would plea for a dummy variable. However, in a multivariate logit model estimate, the parameter of this dummy variable is not statistically significant. Probably the lower default rate in the first year is captured by the healthier financial ratios. New ratings often coincide with bond issues, which enhances, at least temporarily, the financial condition of the issuing firms.
- ¹⁶ Most of the literature on credit-scoring models has been written in the seventies and the eighties. The research on credit-scoring models has recently gained renewed interest for two reasons. First, the record high default rates in the years 2001 - 2002 (see Moody's, 2003, Altman and Bana, 2003) stimulate a further improvement and refinement of these models. Second, the expected implementation of the Basle II accord in 2007 has triggered efforts to upgrade internal rating systems of banking institutions.
- ¹⁷ The empirical analysis is conducted using data on the corporate issuer credit ratings of Standard & Poor's, so in the remainder of this paper agency ratings refer strictly to the ratings of Standard & Poor's. However, we are not aware of a reason why the empirical results and the conclusions presented here for ratings of Standard & Poor's should not apply for the ratings of Moody's and Fitch.
- ¹⁸ Apart from minor deviations, the distribution in S&P ratings is not affected by this data reduction and selection of public firms. However, the percentage of defaulting observations at the beginning of the nineties shrinks, while the percentage of defaulting observations in the years 2000 and 2001 increases. Presumably the credit quality of public firms is less affected by the recession in the beginning of the nineties than the credit quality of private firms. In the years 2000 and 2001 the opposite happened.
- ¹⁹ The reason to include panel observations with a S&P non-rated status in the estimation of the DP-model is to maximize the number of observations in the estimate. When defaulting, the rating status of firms with a NR status changes to a D status.
- ²⁰ For companies with a December fiscal year-end the accounting information refers to the previous fiscal year. For about 30% of the companies the fiscal year end differs from December. In order to keep the accounting information up-to-date the income statement data is averaged for the four fiscal quarters

-
- ending in previous calendar year and the balance sheet data is taken from the latest ending fiscal quarter in previous calendar year.
- ²¹ The surviving observations represent observations of firms at the end of June in year X, which have a stock exchange listing at the end of June in year X+1 as well. This imposes a forward-looking survivorship bias. Robustness tests show that this bias does not affect the parameter estimates of the benchmark models significantly.
- ²² The defaulting observations represent observations of firms at the end of June in year X, which default in between the end of June in year X and the end of June in year X+1.
- ²³ As the raw COMPUSTAT data produces some extreme values for the model-variables, which mostly contain no relevant information, the 0.5% highest values and the 0.5% lowest values are truncated for each model-variable. These values are replaced by values ranked at respectively 99.5% and 0.5%. Since the default event is an extreme event we have checked whether this truncation procedure affects the distribution of the model-variables for defaulting observations. For defaulting observations the maximum and minimum values of the model-variables fall well outside the interval of one standard deviation from the mean of the model-variables.
- ²⁴ A logit model, excluding the ME/TL-variable, is less effective in explaining the default rates than a logit- model, including only the ME/BL-variable. When including the ME/BL-variable, the weight of the EBIT/TA-variable and the RE/TA-variable is strongly reduced.
- ²⁵ By exception the estimated parameter of the WK/TA-variable increases to a significant value of 0.94.
- ²⁶ For six industry sectors, defined by the first digit of the SIC code, the logit model has been estimated separately. The sign of the estimated parameters does not change; the magnitude of the parameters varies within a factor of 2 among these six industry sectors. The parameter of the WK/TA-variable is an exception to this finding. It appears to be significantly positive for the infrastructure services sector.
- ²⁷ Due to space considerations these results are not presented in this paper. Results are available on request.
- ²⁸ An exception to this finding, again, is the estimated parameter of the WK/TA-variable. An estimation of the DP-model with the bankruptcy dataset, gives the WK/TA-variable relatively more weight in the DP-score.
- ²⁹ The WK/TA-variable forms an exception to these findings. After correcting for profitability and (long-term) financial leverage, firms with a higher WK/TA-variable are rated lower by the rating agency, which shows up mostly in the infrastructure services industry and manufacturing industry.
- ³⁰ The only major difference is the absence of a significant model-parameter for the Age-variable for non-investment graded firms.
- ³¹ The distinction between the non-investment grades (notch number 2 – 7) and investment grades (notch number 8 – 15) is determined by eye. The breaking point could equally well be chosen one notch below or above.
- ³² The $\gamma(y,t)$ is computed as follows. For each year t the average $CM(y,t)$ scores are computed for 16 rating classes y. For $y = [2,...,7]$ $\gamma(y,t)$ equals to the slope of the regression equation $CM = \alpha + \gamma y$, with $y \in [2,...,7]$. For $y = [8,...,15]$ $\gamma(y,t)$ equals to the slope of the regression equation $CM = \alpha + \gamma y$, with $y \in [8,...,15]$. $\gamma(1,t)$ equals $CM(2,t) - CM(1,t)$. $\gamma(16,t)$ equals $CM(16,t) - CM(15,t)$. In order to reduce noise the average γ is computed of the γ in current and last two previous years: $\gamma(y,t) \rightarrow (\gamma(y,t) + \gamma(y,t-1) + \gamma(y,t-2)) / 3$. Exceptions are made for $t = 1982$: $\gamma(y,t) \rightarrow (\gamma(y,t) + \gamma(y,t-1)) / 2$ and $t = 1981$: $\gamma(y,t)$ is not replaced.
- ³³ The robustness test results for the two periods are not shown here but are available on request.
- ³⁴ The $\Delta CM_{1(+)}$ -score is slightly higher for non-investment graded firms. After an upgrade the $\Delta CM(+)$ -score remains relatively high for non-investment graded firms. After a downgrade the $\Delta CM(-)$ -score tends to be lower for non-investment graded firms.

-
- ³⁵ We have not explored the reasons for the availability or missing of COMPUSTAT data and the reasons for a switch to and from a non-rated status. Likely reasons are mergers, acquisitions etc.
- ³⁶ The minimum threshold level is 0.5 notch steps, which is the average threshold level implied by the numerical agency-rating scale.
- ³⁷ Including default migrations does not affect the conclusions of the simulation experiment.
- ³⁸ Only conditional on $M_{t-1} < -1$ and $M_{t-1} = 1$ slight differences in the average rating migration are observed.
- ³⁹ The insensitivity of rating drift properties to the threshold level TH dispels a concern that the absence of rating drift in CM-ratings is due to a countervailing effect of continuously reverting noisy CM-scores. If this effect would exist than the average rating conditional on M_{t-1} should depend on the threshold level, which is not the case.
- ⁴⁰ On a major rating level the rating numbers 1 – 7 refer respectively to the following agency rating-major level rating classes CCC/CC, B, BB, BBB, A, AA and AAA. The rating class CCC/CC is a combination of the rating classes C, CC and CCC.
- ⁴¹ $N(y)$ equals to $\Sigma M(y,k)$.