Time to Change.
Rating Changes and Policy Implications.

Peter N. Posch*

Draft Version: April 2, 2006

Abstract Rating agencies are often subject to the criticism of being slow in adjusting their rating to current conditions. This paper examines the timeliness of rating changes and identifies factors which result in 'stickiness' of rating actions. Stickiness is characterized by not adjusting the rating even when a market-based estimate of default probability changes. Introducing an extended econometric model of friction the migration policy is modelled in terms of thresholds which have to be crossed by default probability estimates before an up- or downgrade occurs. Default probability estimates have to change by around two notches before the rating agency reacts. The timeliness differs across the rating spectrum and over the years. During periods with high defaults and for low credit quality firms agencies tend to rate more timely.

Keywords: Rating Stickiness, Migration Policy, Friction Model

JEL-Classification: C33, G20, G33

* Department of Finance, University of Ulm, 89069 Ulm, Germany. eMail: pposch@uni-ulm.de

Acknowledgement I gratefully received valuable comments of Gunter Löffler, Wolfgang Bühler, Rüdiger Kiesel, André Lucas, Stefan Pichler, Richard Stehle the participants of the conferences at FRU, Copenhagen, DGF, Augsburg, the 10th Symposium, Karlsruhe and the WHU, Vallendar. I thank Moody's Investors Service and Moody's KMV for providing the data. The usual disclaimer applies.

1 Introduction

Credit rating agencies like Moody's or Standard and Poor's claim that the stability of ratings is in favor of their clients (e.g. Fons (2001)). Rare changes could show their ability to forecast long term credit risk. But an extreme example like Enron shows the dark side of stability. During the year 2001 the market-based measure of the expected default frequency published by Moody's KMV increased from February to November by more than 2,700% without the rating to adjust. Agencies are thus often criticised to react too slowly compared to market-based measures of default risk. A survey conducted by the Association for Financial Professionals (AFP) among practitioners revealed that about one quarter of the respondent observed a downgrade more than six months after a deterioration in their company's financials. An upgrade took even longer, see AFP (2002) for details.

This essay examines to what extent a market-based measure of risk has to change to observe a rating change and what fastens and what slows this process. The required extent is given in threshold values. The knowledge of these values helps anticipating a rating change and contribute to the understanding of the rating agency's philosophy. An investment manager, for example, could easily monitor the short-term credit risk of his investments and if it changes beyond the threshold he can sell before the announcement. The identification of factors which influence the timeliness of a rating action furthermore provides a close insight to the agency's behaviour and their policy of rating migration. This migration policy is estimated in units of rating steps (notches). For example if the two-year change of a market-based measure exceeds -2 notches then this issuer is upgraded. If this change is 1.6 notches a downgrade is expected and if the change is between these numbers the rating does not react but is subject to stickiness.

In the literature justifications given by the rating agencies for the relative slowness of their actions are discussed under the term 'rating reversal avoidance'. When the rating of a company is changed the agency wants to be quite sure that this adjustment is stable and has not to be reversed shortly after. It is a political decision of the rating agency to change relatively seldom and this migration policy dampens the volatility of the ratings. But if investors rely on private or additional information contained in the agency's rating and find that their revelation lags the publicly available information the agency loses credibility. And in the bottom line all ratings are based on reputation (see e.g. Covitz and Harrison (2003)). Each time a rating change is expected the agency has to balance the gains of a stable rating with the losses from being ex post incredible. Thus a rating action might not be taken even if a change in the markets valuation of credit quality implies so. Such a migration policy would be sticky since the rating does not adjust, so the observed migrations differ from the expected ones. Thus rating stability and reputation incentives result in stylized facts such as the rating drift, i.e. it is more probable to see a subsequent rating change in the same than in the opposite direction.

There are at least two strands of literature related with the questions posed in this paper. The first one is the theoretical modelling of rating agencies and their role in the market. In this context the aspects of reputation and reliability of an intermediary are the most important for this paper. Benabou and Laroque (1992) and Morris (2001) model the interaction between an sender of information and a receiver to whom the reliability of the sender is unknown. Spence (1973) introduces asymmetric information between firm insider and market participants and examines the ability of different types of firm to signal their true quality to the market. Boot et al. (2006) show that rating agencies can coordinate between multiple equilibria and how the credit watch procedure applied by the agencies induces a monitoring. Bannier and Tyrell (2005) model an utility function of the rating agencies which includes reputation, competition and feedback effects. They find that public rating announcements and private information complement each other.

The second strand of literature analyses the empirical behaviour of rating agencies and market information. Löffler (2005) uses simulations and Altman and Rijken (2004) use actual rating data in an empirical approach to ascribe the stylized facts to the migration policy. The former author and Fledelius et al. (2004) coin the term 'reversal aversion' for the mentioned stylized fact. Altman and Rijken (2004) compare a constructed 'agency'-rating with a benchmark market-based rating using the Z-score variables proposed by Altman (1968). Through simulations they conclude that a rating change is triggered if the actual through-the-cycle credit quality exceeds 1.25 notch steps of the average credit quality for a given rating class. Delianedis and Geske (1998) compare expected default frequencies calculated by the Merton (1974) and the Geske (1977) model to Moody's and S&P ratings. They observe that the former measures of default probability contain early information about subsequent rating changes, i.e. these measures increase(decrease) long before an rating down-(up-)grade. They conclude that news affecting an issuer's credit quality is not instantaneously translated to a rating change. Perraudin and Taylor (2004) analyse the consistency of agency ratings and ratings implied by bond market yields. They find that some high credit quality bonds are rated inconsistent with their pricing. Most of these occurrence disappear after controlling for tax, liquidity and risk premiums and, moreover, the differences in rating and credit spreads become smaller over time.

Kavvathas (2000) examines to which extend the rating changes depend on the current business cycle. His methodology of the ordered probit analysis is also used by Nickell et al. (2000) to investigate sector and cyclical effects and by Blume et al. (1998) and Jorion et al. (2004) to answer the question whether the rating agency's policy has changed during the 1990s. In contrast to this branch of literature, which uses static models with constant cut points, I use a dynamic model to identify of the driving factors of the rating stickiness. This leads to a detailed insight into the rating methodology itself. The very general methodology used here to identify the stickiness of rating adjustments is an extension of models of friction, which where initially proposed by Rosett (1959). Applications of these kind of model include the estimation of central bank intervention by Almekinders and Eijffinger (1996) or Asano (2002) who examines the effects of costly reversible investment. This paper extends the methodology of frictional models with respect to the use in panels and variation in the coefficient estimates for the negative and positive parts of the model.

The remainder is organized as follows. Section 2 describes the data set. Section 3 introduces the friction models, while section 4 introduces the variables used in the analysis. Section 5 gives the results of the empirical analysis. Section 6 summarizes and draws a conclusion.

2 Data Set

The data set I use contains monthly information on Moody's long-term ratings and a market-based default risk measure by Moody's KMV ('MKMV', hereafter). It consists of 4,023 US and Non-US anonymized corporate issuers and covers the period from January 1980 to April 2005. I start my analysis in January 1983 and exclude rating changes to a default category. In April 1982 Moody's refined its rating system and added new rating modifiers (the 'alphanumerical' ratings). Due to this refinement the rating changes without any change of the underlying credit worthiness of the issuer. A rating change to the default category and changes within default cannot be considered as a political migration as this decision is forced exogenously rather than influenced by a rating committee.[1] Figure 1 shows the evolution of issuers over time.

The variables used in the model are based on MKMV's expected default frequency (EDF, hereafter) and Moody's issuer rating. The EDF gives a cardinal estimate of the one-year probability of default. It is based on balance sheet data and market measures of the value and

[1]Sensitivity analysis including defaults show no qualitative change in the results which could be due to the relatively small proportion of defaulters in the data set.

volatility of a firm's equity. In the approach of Merton (1974) the probability of default (PD, hereafter) is derived by applying the standard normal cumulative distribution function to the modelled distance to default. MKMV, in contrast, calculates the EDF by calibrating the distance to default to empirical (historical) default rates. This calibration results in a minimal EDF of 0.02% and a maximal EDF of 20%. Throughout this paper I use the logarithms of the EDF instead of the levels. Whenever certain EDF values are quoted they refer, however, to the level. A logarithmic transformation, in contrast to the application of the normal distribution, preserves the information of the tails, i.e. it can discriminate among issuers with very high or very low PD.

The response of ratings to changes in the EDF should be different across the spectrum. Consider e.g. a change from 0.04% to 0.14%. This should have an different impact on the rating response than a change from 0.3% to 0.4%. The level's change is equal for both cases, but the percentage change is 250% in the former and only 33.3% in the latter case. This effect is controlled for by using percentage changes, i.e. changes of the log-EDF, instead of changes in the levels. Figure 2 shows the mean and median of the EDF over time for all issuers and issuers with investment-grade rating.

The issuer rating is converted to a numerical scale with Aaa $\sim$ 1 to C $\sim$ 21 (RAT, hereafter). For similarity of scaling I transform the ordinal ratings to their PD using 'idealized PD' published by Moody's (see Yoshizawa (2003)). This variable is denoted by PDRAT, hereafter. The transformation can generally be done in different ways leading to different implications. Using the idealized PD sharpens the through-the-cycle approach of the rating agencies as the resulting variable is constant over time for each rating class. The mapping of can further be done using the (ex post) yearly transition matrices. This approach incorporates information on macroeconomic changes over the time and thus reduces the agency-nature of the ratings. Between these two procedures is the mapping using a transition matrix over the whole period of the data set. Such a transition matrix can either be obtained from publications of the major credit agencies or be calculated within the data set. In this paper I model as close to the rating agencies predictions as possible, which is to use the transformation by the idealized PD. Table 1 shows descriptive statistics for each rating category.[2]

For a sub sample of this data set starting in September 1991, outlook and watchlist data is available. A rating outlook is defined by Moody's to be 'an opinion of the likely direction of a

[2]Note that the mean EDF is u-shaped. For the results in section 5 I conducted robustness checks excluding the Aaa rating level. As the results remain unchanged I use the full dataset.

rating over the medium term'. An outlook can be either positive, negative, stable or developing. A watchlist is used to 'indicate that a rating is under review for possible changes in the short-term'.[3] That review can be either for upgrade, downgrade or with uncertain direction. I estimate adjusted ratings (ADJRAT, hereafter) which incorporate outlook and watchlist information by adjusting the rating level. Following Hamilton and Cantor (2004) a rating review adjusts the rating for two notches and an outlook for one notch. These adjustments are symmetric for positive and negative conditions. For example an issuer with a rating of B2 and a positive outlook has an adjusted rating of B1, while a negative review would results in an adjusted rating of Caa. Overall there are 3,462 issuer with available outlook information.

3 Methodology

The rating policy of volatility aversion and the methodology of attempting to look through the cycle are major reasons for observing stickiness in rating changes. The econometric approach used here estimates the observed rating changes as a function of the observed change of the current creditworthiness measured by the EDF. The model is estimated by three equations each describing a different rating behaviour. When the change of the EDF and the following rating change is moving in the same direction, i.e. an upgrade is observed in response to a decrease of the EDF or an increase in the EDF is followed by a downgrade, we are outside the frictional part of the model. But if the rating remains unchanged even if the EDF changes the rating is subject to stickiness. The model quantifies to what extend a latent credit process has to change to observe a change of the dependent variable, i.e. the rating.

Let RAT^*_{it} be the expected but unobserved (latent) rating for company i at time t. The operator Δ_n denotes the n-th difference of a variable x, i.e. $\Delta_n x := x_t - x_{t-n}$, where the subscript is suppressed when denoting a non-specific difference and n is measures in month. The latent rating change $\Delta RAT^*_{i,t}$ is modelled by a vector $\mathbb{X} = (x_1\ x_2\ ...\ x_k)$ of k exogenous variables as shown by equation (1), excluding the constant.[4]

$$\Delta RAT^*_{it} = \sum_{i=1}^{k} \beta_i x_{it} + \varepsilon_{it} = \beta^{'}\mathbb{X} + \varepsilon_{it} \tag{1}$$

$\varepsilon_{i,t}$ is a serially uncorrelated, homoskedastic error term with mean zero and time-constant variance. To control for issuer-specific heterogeneity I use the estimator for cluster-robust stan-

[3] See http://www.moodys.com at 'Rating Definitions'.

[4] In this model the inclusion of a constant would shift the threshold parameters α_i. To see that elongate the curves of the areas (L) and (U) in figure 3 toward the vertical axis.

dard deviations proposed by Huber (1967) and White (1980) (HWS, hereafter). The HWS is modified to deal with clustering around issuers, see e.g. Wooldridge (2002) for details.

The observed change in the rating ΔRAT_{it} is modelled as a function of the expected rating change according to $\Delta RAT_{it} = \xi(\Delta RAT^*_{it})$, where the function $\xi(\cdot)$ maps the unobserved latent variable RAT^* to the observed variable RAT. In the case of the friction models discussed here this function is given by equation (2). This model was originally proposed by Rosett (1959) as an generalization of Tobin (1958). It was improved by Dagenais (1969) and Dagenais (1975) and is here further extended to deal with panel-issues and non-constant thresholds.

$$\Delta RAT_t = \begin{cases} \Delta RAT^*_t - \alpha_1 + \gamma_1 & , \Delta RAT^*_t < \alpha_1 \text{ (L)-Area} \\ 0 & , \alpha_1 \leq \Delta RAT^*_t \leq \alpha_2 \text{ (F)-Area} \\ \Delta RAT^*_t - \alpha_2 + \gamma_2 & , \alpha_2 < \Delta RAT^*_t \text{ (U)-Area} \end{cases} \tag{2}$$

The reputation loss aversion determines the length of the frictional part. This part is estimated by the upgrade threshold $\alpha_1 < 0$ and the downgrade threshold $\alpha_2 > 0$ in equation 2.[5] The parameter $\gamma_1 \leq 0$ and $\gamma_2 \geq 0$ allow for jumps at the threshold level, i.e. when the threshold is crossed there is an immediate (constant) adjustment in addition to the (variable) change of observed rating. The cost for the rating change announcement itself, i.e. the publication of the new rating on the internet, press-notes etc., is proxied by the parameters γ_i.

Each line of equation (2) corresponds to an area of figure 3, where the left graph shows an idealized graphical representation of the friction model without fixed costs ($\gamma_1 = \gamma_2 = 0$), while the right graph shows the addition of vertical friction. Note that the figures show a linearised graph. With rating classes the graph would be stepwise. Thus an underlying assumption is that the ratings can be linearised. This assumption leads to an over prediction of the thresholds. Evidence from simulation analysis show, however, that this effect is only marginal.

The likelihood function for the friction model is derived in the appendix. The areas refer to the position of the derived function, where L denotes the lower part, F the friction area and U the upper part. The slope coefficients in figure 3 are equal for both the lower and the upper part. I relax this restriction by generalising the equations to include different slope coefficients. Furthermore I allow for variable threshold parameters. Instead of estimating α_j as constants I include variables $x_k^{\alpha_j}$. This results in distributions of thresholds and in the following analysis the mean of this distribution is reported whenever the threshold model includes variables.

[5] A negative change in the rating or the EDF is associated with an increased credit quality.

Threshold conversion

The threshold-parameters α_i are estimated in units of the dependent variable, in this case in rating notches. Consider e.g. the model $\Delta RAT^* = \beta_1 \cdot \Delta EDF + \varepsilon$ and suppose that the coefficient estimator for the only exogenous variable is equal to 30, i.e. $\widehat{\beta_1} = 30$ and the threshold-parameters are $\widehat{\alpha_1} = -0.5$ and $\widehat{\alpha_2} = 0.4$. Then the model tells that if ΔRAT^* is between -0.5 and 0.4 we are in the frictional part (area F). A more commonly statement would sound like 'for EDF changes between -1.6pp and 1.3pp we do not observe any rating change'. Converting these thresholds from units of ΔRAT^* to units of the exogenous variable is straightforward. What equation (2) tells is 'if ΔRAT^* is smaller than $\widehat{\alpha_2} = 0.4$ but bigger than $\widehat{\alpha_1} = -0.5$ we observe friction'. This is equivalent to $\alpha_1^{\Delta EDF} = -0.016 < \Delta EDF < \alpha_2^{\Delta EDF} = 0.013$ as obtained by solving equation (1) for ΔEDF. Extending the results to cases with more than two exogenous variables is possible using $E[\Delta RAT^*|\mathbb{X}, \Delta RAT^* > |\alpha_i|] = \sum_j \widehat{\beta_k} \cdot E[x_k]$. Solving this equation for any variable x_l gives the threshold parameters in units of x_l: $\alpha_j^{x_l} := (\alpha_j - \sum_{k \neq l} \widehat{\beta_k} \cdot E[x_k])/(\widehat{\beta_l})$ for $l \neq k$, $j = 1, 2$. Note that this procedure is appropriate only in absence of interaction terms. When interaction terms are included the correlation of the interacting variables would have to be accounted for.

Marginal Effects

The marginal effects in this model are estimated similar to the Tobin case. In a standard least square regression the marginal effect is given by the coefficient estimator. Here the expected value of the latent variable $E[\Delta RAT|\mathbb{X}, \beta, \alpha_i, \gamma_i]$ depends on the threshold parameters and thus on the probability of being in one of the parts of figure 3. For example consider the area L, a marginal effect for this part accounts for the probability of leaving that area L. The marginal effect is obtained by taken the first derivatives of $E[\Delta RAT|\mathbb{X}]$ in equation (5) with respect to β', see e.g. Greene (2003), p. 765 for a general result and the appendix for a deviation of this equation.

$$
\begin{aligned}
E[\Delta RAT|\mathbb{X}, \alpha_i, \gamma_i] &= \Phi\left(\frac{\beta'\mathbb{X} - \alpha_2}{\sigma}\right)(\beta'\mathbb{X} - \alpha_2) + \phi\left(\frac{\beta'\mathbb{X} - \alpha_2}{\sigma}\right) \cdot \sigma \\
&= -\Phi\left(\frac{-\beta'\mathbb{X} - \alpha_1}{\sigma}\right)(-\beta'\mathbb{X} - \alpha_1) - \phi\left(\frac{-\beta'\mathbb{X} - \alpha_1}{\sigma}\right) \cdot \sigma
\end{aligned}
\tag{3}
$$

Comparison to response models

A widely used methodology to explain ratings is the ordered probit model (see e.g. Kavvathas (2000), Nickell et al. (2000) or Blume et al. (1998)). This model is an extension of the probit model to the case of more than two discrete answers. An ordered probit model assumes the error term to be standard normally distributed, i.e. $\varepsilon \sim N(0, 1)$. Since the friction model estimates the variance of the error term as a parameter one has to standardize the estimators of the friction model to compare them to the ordered probit case. Doing so reveals no significant difference of the slope coefficients of the two models.

The ordered probit estimates a cutoff for each class of the dependent variable. The friction model, in contrast, estimates only the thresholds for entering and leaving the sticky part of the model (recall figure 3). The thresholds α_i from the friction model and the corresponding cutoffs from the ordered probit model are within a 99.9% confidence band. The similarity to a two-limit probit model can further be shown analytically (see Maddala (1982) for details)

Despite of these similarities the friction model contains several advantages over the ordered probit model. First, the specification of the latent process is not restricted to be equal for upgrades and downgrades. Second, the specification of the upgrade and downgrade threshold is not restricted to a constant.

The restriction of the cutoff to only two thresholds is only a shortcoming at first sight. The cutoff from the ordered probit model can be obtained by extending the threshold specification of the friction model, i.e. including class dummies in the specification of the threshold process. A direct consequence from this generalization is, for example, the specification of each rating *level*. That way the model can be used to make the thresholds dependent on the rating level at the time of the rating change (see section 5 for results).

4 Variable Selection

The friction model estimates a relationship between the rating as Moody's prediction of the long-term credit quality and the EDF as a market-based short-term measurement of the probability of default, where the former is supposed to be sticky with respect to the latter. To shed more light on the factors which trigger a rating migration I extract the information contained in the rating and the EDF. In fact the friction approach needs the specification of two models: The model of the latent rating change ΔRAT^*, which is equivalent to specifying the 'true', but unobservable, credit quality process and the specification of the threshold variables, i.e. the specification of the α_i-models. Each variable's influence can thus be twofold: it can model the impact on the credit process. Here it influences the height of an rating adjustment. Furthermore it can specify the threshold level, i.e. it influences the timeliness of a rating change.

The theoretical literature dealing with the probability of migrating from one rating to another assume these probabilities to be Markovian. Empirical work (e.g. Lando and Skodeberg (2002)) shows that this probability depends on the path of the rating. Abstracting from cyclical effects there is evidence that the time spent in a given rating class as well as the direction from which the current rating was achieved, i.e. downgrade or upgrade, influences the probability of further rating migration. The first observation of this behaviour is given by Altman and Kao (1992a), Altman and Kao (1992b) and Lucas and Lonski (1992). These authors coin the term 'rating drift' and although the exact definition differ within the literature this stylized fact is typically based on proportions of downgrades to upgrades (or vice versa) within or across rating classes. Recent papers which examine whether the history of rating migration influences the current transition probability include Kavvathas (2000), Lando and Skodeberg (2002) and Fledelius et al. (2004). These authors find similar results, namely that downgrades are autocorrelated, i.e. a downgrade is more likely to be followed by another downgrade than an upgrade. For upgrades this serial dependence is less pronounced, i.e. statistically not as significant as the downgrade momentum.

Instead of including the number of month spent in a certain rating class as a duration measure I use the difference of the PD implied by the rating and the current EDF (($PDRAT1 - EDF$), hereafter). This variable further controls for the difference of the market's and the agency's perception of the credit quality. Since KMV considers its EDF to give a 1-year PD I use the one year horizon denoted by $PDRAT1$. Consider a rating which has not changed over a long period of time, i.e. the duration is high. If the current EDF is far away from the implied one-year probability, the variable $PDRAT1 - EDF$ has high values. The sign is negative if the EDF implies a higher PD than the rating agency, and positive in the opposite case. The probability of a rating change should be lower in this case if the agency follows a through-the-cycle approach than if it would follow a point-in-time approach. The rationale is that the agency's belief about the issuers willingness to increase its creditworthiness within a short period of time is strengthened by past experience captured in the duration. This behaviour is conditional on the current rating level since a long duration in the lower non-investment grades should result in a higher transition probability than within the higher investment grades. Furthermore this measure controls for the rating drift over time.

One problem when estimating changes of agency ratings is the methodology of bucketing. Issuers with different current conditions of their credit quality are grouped in the same bucket. Since this assembly changes over the time, the distance of each issuers current point-in-time credit quality to the 'typical' quality in his current rating class accounts for the relativeness of a rating system. This variable is denoted by $BMEDIAN - EDF$, where $BMEDIAN$ refers to the 50% quantile (median) of the EDF distribution for the corresponding rating class at that time. The variable controls for the effect of rating reversal avoidance since low values indicate that the issuer's current credit quality is in line with the other issuers rated equally well, while high values indicate a difference between the market measure of credit risk of this issuer with the typical company rated this grade. The sign of the variable gives the distance of distortion from the median.

To control for the precision of the market's view on the company's credit quality I include a moving volatility of the change in the EDF, denoted by $VOLA_{iT} = \sigma(\Delta EDF_{it}|t \leq T)$. The higher the volatility the less confident the market. The expected sign of the coefficient in the threshold specification is negative. For an desired upgrade this is because high volatility is opposed to the higher stability implied by a higher rating grade. The same arguments justifies a quick downgrade.

I use a dummy variable for investment grade rated issuers ($INVEST$). This dummy is equal to one if the rating is between Aaa and Baa3 and zero for the speculative grades (Ba1 to Caa3). I control for recession periods according to the National Bureau of Economic Research (NBER), where the first recession during this data set is from July 1990 to March 1991 and the second recession period lasts from March 2001 to November 2001.[6] Both recession periods are accumulated in the dummy variable $RECESSION$.

5 Results

In all analyses I use non-overlapping time periods and control the results for the investment-grade boundary, which yields estimates for each subset and the whole data set. Beginning with the most simple model I demonstrate the improvement of releasing the slope coefficient's restriction. Then I present the stability of the model over time and different specification. Before extending the specification I conduct a simulation to assess the bias due to the rating agency's methodology. The effect of adding outlook information finalises the discussion.

5.1 Basic Model

In the initial setting I model yearly rating changes ($\Delta_{12}RAT$) on yearly percentage changes of the expected default frequency ($\Delta_{12}EDF$). Table 2 shows the results, where panel (A) gives

[6] See http://www.nber.org/cycles.html.

results for the friction model with the restriction of equal slope coefficients while panel (B) relaxes this constraint.

The generalized friction model allowing jumps at the time of crossing the threshold is rejected in favor of the standard friction model. The jump parameters γ_i are not significant and very low. All other coefficients are highly significant. The restricted estimates of panel (A) are between the unrestricted of panel (B), e.g. the coefficient of $\Delta_{12}EDF$ is 0.663 for the restricted setting and splits to a value of 0.401 for upgrades and 0.818 for downgrades when releasing the constraint. These coefficients specify the change of the expected rating ΔRAT^* in response to an observed percentage change of the EDF - given ΔRAT^* is in the appropriate region, i.e. $|\Delta RAT^*| > \alpha_i$. The inverse of this coefficients gives the percentage EDF change needed for an one-notch change of the expected rating. When $\Delta_{12}EDF = 1/0.663 = \pm 1.508$ then the expected rating changes by $\pm$ one notch. This is equivalent to a percentage change in the EDF of $\exp(1.508) - 1 = 351\%$. Although this figure seems quite high at first, it is not uncommonly observed in this data set and the initial example of Enron is only one case in which even higher percentage changes where needed to trigger a rating change. For downgrades this number is within the upper 95% of the distribution. The 99% quantile is at a percentage change of 1,275%, while for upgrades the 1% quantile value is at a percentage change of 99%. For the results of panel (B) this picture sharpens. While changes of the EDF exceeding the one-notch threshold are observed often for downgrades, the upgrade part is observed seldomly. One technical reason is the lower proportion of upgrades compared to downgrades and no-rating changes. In this setting there are 8.08% upgrades, 66.22% no rating changes and 14.53% downgrades observed. An economic reason lies in the variables used. A one-year EDF change might be a reasonable variable to predict downgrades, but a too short time period for predicting any upgrade.

The threshold parameter vary between -4.5 to -3.6 notches for upgrades and 3.0 to 3.6 for downgrades, depending on the specification and sample used. The thresholds in panel (B) are smaller in absolute values compared to panel (A). The threshold parameters implicate that a decline of more than 3 notches of the latent credit quality triggers a downgrade. For upgrades this unobserved credit quality has to raise by more than 3.6 notches to observe an upgrade. Concluding the initial model is sufficient to capture downgrade stickiness, but does not seem to be adequate for upgrade stickiness, which is partly due to the lower proportion of upgrades in the data set.

Relaxing the restriction of equal slope coefficients significantly improves the model's fit.

To assess this improvement I use a likelihood ratio (LR) test. The test statistic is given by $\lambda = 2(lle - llr)$ where lle is the log likelihood of the full model and llr the likelihood of the restricted model. The test statistic is approximately χ^2 distributed with degrees of freedom equal to the number of restrictions imposed (see e.g. Greene (2003) for details). For all three samples (full, investment-grade and non-investment grade) the p-value is zero and thus the hypothesis of jointly equal coefficients has to be rejected.[7] Note that this test is not appropriate if the two models are non nested as in the cases of the friction models with and without the restriction $\gamma_1 = \gamma_2 = 0$. Comparing such non nested models one would have to use a test like the one proposed by Vuong (1989).

Investment and non-investment grade rated issuers differ significantly in both the estimated coefficients and thresholds. To test this hypothesis I used a approach similar to the Chow (1960) test for structural differences. The Chow test assesses the probability of equal coefficients for two (or more) groups by applying a standard F-test to the joint hypothesis of equal coefficients for these groups. This formulation of the Chow test applies only to regressions where the sampling distribution is known such as the linear regression. The approach itself, however, is more general and can be applied to the friction model using a χ^2 test instead.[8] Doing so I estimate the basic model on the full sample interacting the explanatory variable with an indicator on each of the two groups (investment-grade and non-investment grade) and including dummies for both in the threshold specification. The hypothesis of equal estimates for both groups has to be rejected. The χ^2 statistic is 630 (p-value: 0.00) with three degrees-of freedom (one for the slope coefficients and two for the thresholds) for the model in panel A of table 2 and 606 (p-value: 0.00) with an additional degree of freedom for the more general model of panel B.

5.1.1 Sensitivity of Thresholds

The threshold estimated in the basic model refer to a time period of more than twenty years, over 4,000 different issuer and to the whole rating spectrum. They thus mark a mean threshold which on average triggers a rating event for any rating class, company and point in time. The sensitivity of the thresholds to these factors contribute to the understanding of the rating policy. Since the names of the issuers are not available in this dataset the influence of company characteristics is limited to the realisation of the included variables. It is probable that information such as

[7] The critical value of the χ^2 distribution with one degree of freedom are 3.84 for 95% confidence and 6.63 for 99% confidence level.

[8] Recall that the χ^2 distribution is the limiting distribution of the F statistic. A weaker alternative to the test described here would be the inclusion of a dummy variable for the investment grade and applying the LR test.

branches, age of the company or management structure influence the thresholds. Examining the other factors, however, is possible and appropriate before extending the thresholds specification.

Rating Classes

I include dummy variables for each rating class level in the threshold specification. Since these dummies cover the whole rating spectrum I exclude the constant. The coefficient on each dummy variable thus gives the threshold for issuers rated with the corresponding grade. Figure 6 shows the result together with an 95% confidence interval. Each threshold coefficient is highly significant for both the upgrade and downgrade specification.

For upgrades the threshold behaviour can be divided into two parts. In the first part the thresholds become higher with decreasing credit quality, i.e. an upgrade becomes more likely. That behaviour is counter intuitive at first sight, but a possible explanation is the increase of observed upgrades. An issuer rated Aa3 can upgrade one notch, but downgrade twenty notches. Thus on average one would observe less upgrades for this rating grade than for a lower grade. The second part of the upgrade behaviour starts at rating class Ba. From that class onwards an upgrades becomes less likely, the thresholds increase. The results indicate that there is a critical rating level for upgrades. If this level is crossed upgrades become slower compared to the average timeliness. Using only letter grade ratings, i.e. without the alphanumerical modifiers, sharpens this observation.

To gain additional insight to the hypothesis of a n-shaped upgrade threshold behaviour I apply a fractional polynomial regression to the estimated thresholds. Fitting a polynomial has furthermore the advantage of getting an analytical function of the threshold with the rating class as argument. Linear regression are not appropriate to investigate the observed non-linear behaviour. A fractional polynomial is a generalization of a ordinary polynomial of the form $y = \sum_{i=1}^{m} \beta_i X^{p_i} + \beta_0$, where m specifies the degree of the fractional polynomial and p_i gives the i^{th} power including fractional powers. The best fit with an adjusted R^2 of 59% is obtained for the following polynom (t-statistics in parenthesis):

$$FP(Upgrade) = 0.8604(3.4) \cdot RAT^3 - 1.4952(-4.0) \cdot RAT^3 \cdot ln(RAT) - 3.5202(-20.1)$$

That polynom has its maximum at a rating level of B3. Figure 7 shows the observed data together with the fitted polynom and its 95% confidence bands.

For Downgrades the thresholds are decreasing almost monotonically with decreasing rating classes. The worse a rating the faster a downgrade. Using letter grade ratings indicates that this decrease is again divided into two parts. One with a flatter decrease up to approximately rating class Ba and the other with a stepper decrease.

Testing this hypothesis it is sufficient to apply a linear regression to the estimated downgrade thresholds. Splitting the rating spectrum at the Ba3 level shows a ten times higher decrease for the lower part of the rating spectrum. The slope coefficient is -0.0332 with a t-statistic of -4.8 for the rating classes down to Ba3 and an slope coefficient of -0.3111 with a t-statistic of -11.3 for classes worse than Ba3. The R^2 for the former regression is at 69% and at 95% for the latter. It can thus be concluded that the threshold behaviour differs across the rating spectrum. This difference can be divided into two parts where the rating class Ba3 provides a boundary.

Threshold over different years

To examine the variation of thresholds over time I include dummy variables for each year from 1984 to 2004.[9] The coefficients on these dummies give the variation from the estimated constant threshold for the corresponding year. Whenever the coefficient is significant the corresponding year differed from the estimated mean threshold. Figure 4 shows these coefficients with 95% confidence intervals. This figure implies that the migration policy could be subject to cyclical behavior.

The timeliness of upgrades is more stable over the years than the downgrade migration policy. From 1989 to 1991 upgrades took place more slowly than on average, the threshold are significantly larger during these years. The same is true for the years 2001 to 2003. For both periods, however, the increase is less than one notch. For downgrades the deviation from the mean threshold is more pronounced. In 13 years the dummy is significant. From 1989 to 1993 companies are downgraded faster than average. This decrease of the downgrade threshold accelerates until 1991 with a decrease of -1.4 notches on the average threshold and then moves back to he average level. From 1994 to 1998 the downgrade threshold is at average, slows in 1999 to fastens with the 2000 period onwards again. The last year 2004 shows an tendency to the average again.

Using polynomial regression like before is not appropriate in this setting for two reasons. Firstly since the hypothesis to test is the cyclicality of the thresholds the polynomial would have to be of high dimension in order to fit the observed curves. Second and more important the technique is biased due to boundary effects. What we observe here is just a small section as we cannot model the thresholds before 1984 and after 2004. A more appropriate technique would

[9] The calculation of one-year changes leaves the year 1983 with no observations and the data set ends 2005, so this year is not completed.

be the application of trigonometric regression (see e.g. Lau and Studden (1985) or Eubank and Speckman (1990)). For these type of model the number of observations is too limited. Nevertheless additional insights can be found by comparing these thresholds with the realized default rates per year. During periods with many defaults downgrades should be observed more quickly and upgrades should take place only seldom. Figure 5 plots the default rate over the years. This implies a negative relationship as expected. The correlation between downgrade thresholds and realized default rates is -0.42 (p-value: 0.06) and for upgrades -0.61 (p-value: 0.00). Concluding we see graphical evidence of cyclicality of the threshold over the years which is most pronounced for downgrades. Both thresholds negatively correlate with the default rate. In periods with high frequency of defaults firms are upgraded more slowly and downgraded more aggressively.

A closer look at the post-millennium period shows a structural differences to the period prior to that date. Altman and Rijken (2004) finds that the period from the years 2000 onwards is structurally different from the former one. According to these authors the 'too-big-to-fail' protection is lost in the post-2000 period, resulting in high default rates for large sized issuers. Rationale for this observation include the market's response to the 9/11 attacks (see e.g. Altman and Arman (2002)) and the rating agency's announcement of migration policy changes in response to the Enron aftermath in 2002 (see e.g. The Economist, May 16th 2002).[10]

The announcement of the agencies to rate more timely reduced the downgrade stickiness by almost one notch, while the effect on upgrades is vice versa. The downgrade thresholds decrease by 0.462 notches for the years 2000/01 and 0.863 notches for the years from 2002 onwards. For upgrades the effect is not as pronounced. During the 2000/01 years there is no significant difference in upgrade thresholds compared to the pre-2000 period. For the post-2002 period, however, the upgrade threshold widens by 0.368 notches, i.e. the upgrade stickiness increases. The observation might sound counter intuitive at first, but can be explained by an incentive to avoid too good ratings. The agencies react faster to an deterioration of the perceived credit quality of an issuer and audit the melioration of credit quality more careful. Concluding for the years 2002 onwards companies are downgraded faster and upgraded slower, for the years 2000/01 there is a not as strong pronounced effect in the same direction.

[10] Altman and Bana (2004) report a default rate of 12.8% for the year 2002 and count 90 firms with outstanding liabilites of more than one billion US-$ defaulting in the time period from January 2001 to June 2003. With the end of 2003 the default rates are on a historical small level and further declining, as citeAltman and Fanjul (2004) points out.

Threshold over different time horizons

While the thresholds are relatively stable over the years, the question arises whether the specification of the time horizon is robust. When the EDF measures the current-condition of the credit quality, but the ratings react only to persistent changes, the changes in the latent rating might not go back long enough to capture the triggering current-condition change. Connected with this question is the robustness of the dependent variable specification. In the initial setting I used the one-year rating change, but how do thresholds change with a longer/shorter horizon? Since it can take some months from the rating decision to actually changing the rating, the current EDF changes could not capture the decision. If so either longer time horizon for the EDF change or EDF changes some months ago would be the appropriate candidate for measuring the rating trigger.

Longer changes of the EDF do not lower the thresholds set by the agency and thus do not fasten a rating change. To model this effect I start with a rating change of six months explained by a six months change of the EDF, which is increased successively by six months. Then I increase the dependent variable by six month which is again explained by increasing changes of the EDF. The final specification tested is the 30 month rating change by the 60 month change of the EDF. The thresholds are very stable for each specification of the independent variable. In fact, the volatility of all parameters is very low (between 0.01 and 0.05). Thus a longer time horizon of the latent credit quality does not alter the results. The rating changes within one period can be explained with latent changes during the same period. The same results is obtained when holding the time horizon of the changes constant but adding successively lags to the independent variable, i.e. using the one-year EDF change from six months ago. Figure 8 shows the results for the thresholds.

The main finding is that the longer the time horizon the smaller the thresholds which have to be exceeded to trigger an observed rating change. The decrease of the threshold is most pronounced when switching from a six-month rating change to a one-year rating change. Continuing to increase the time length declines the changes of the thresholds. In fact the two-year rating change and the 2.5-year rating change specification are both within a 95% confidence band. An economic reason would be the through-the-cycle methodology. During two years most of the temporary conditions should have changed and this is reflected both in a (permanent) shift of the EDF and a change of the rating. Switching form one-year rating changes to two-year rating changes almost doubles the mean of these variables. Concluding the two-year model containes valuable information not captured by the one-year specification.

5.1.2 Quantification of the Methodology Bias

All models specified here use variables based on the EDF as latent credit quality process. But the EDF is a noisy measure of the point-in-time credit quality. It is driven by the market's evaluation of an issuers asset value. Likewise any other noisy measures this leads to a decline in the precision of the estimates. Furthermore the EDF does not reflect the through-the-cycle methodology of an agency rating nor the effect of bucketing, where the credit rating agencies put issuer with a wide-range of different EDF into one rating bucket. To quantify the effect of these characteristics I use a common filter technique to obtain the trend of the EDF. This trend is free of temporary fluctuations and forward-looking. It thus comes close to the ability of looking through the cycle. With this EDF-trend I then simulate a rating system to capture the effect of bucketing. These effects sum up to about one rating notch.

Leser (1961) proposes a technique to decompose any time series x in a cyclical component x^{cycle} and a trend component x^{trend}. This method is later improved by and named after Hodrick and Prescott (1997): the HP-filter. In this case the original series EDF is decomposed into a trend component called $HPEDF$ and a cyclical component EDF^{cycle}, which is not further used here, by solving the optimization problem given below.

$$\begin{aligned}\min_{EDF_{it}^{trend}} \quad & \sum_{t=1}^{T}((EDF_{it} - EDF_{it}^{trend}))^2 \\ + \quad & \lambda\Big((EDF_{i,t+1}^{trend} - EDF_{it}^{trend}) - (EDF_{it}^{trend} - EDF_{i,t-1}^{trend})\Big)^2\end{aligned}$$

Here λ is a penalty parameter which is set a priori. This parameter penalizes the roughness of the original series. Higher values of λ lead to a smoother cyclical component.[11] I set λ to the commonly used value for monthly data of 14,400. Higher values of λ slightly increase the correlation of the trend with the rating, but do not significantly alter the results. The overall correlation of Moody's issuer rating with the trend of the EDF is 70%, while using the EDF the correlation is at 66%. The filter is estimated per issuer on the logarithms of the EDF. Figure 9 gives a graphical representation of the filter's use for an arbitrary issuer.

Based on the trend $HPEDF$ I construct a benchmark rating system ($HPRAT$, hereafter). In every year I obtain the percentage of observation in each rating category. Using the density

[11] Note that in the limit $\lambda \to 0$ the trend component approaches the original series, while in the limit $\lambda \to \infty$ this component approaches linearity. For a more detailed discussion see e.g. Schlicht (2004)

of $HPEDF$ new rating classes are assigned according to the observed frequency. For example, in 1983 there are 2.6% Aaa rated company-month. The lowest 2.6% of the $HPEDF$ values are assign to the rating class $HPRAT = 1$. The constructed rating depends only on the number of observation and the trend of the EDF. The mean of the construced rating and the original rating do not differ significantly, but the former's volatility is lower. Note that this constructed rating is a pure through-the-cycle rating with respect to the EDF by construction. However, ratings are still assigned to buckets of issuers. It is thus not free of bucketing and effects caused by this method.

Figure 10 shows an example of Moody's rating, the constructed rating and the HPEDF for an arbitrary issuer and table 3 gives the results for the $HPRAT$ as dependent variable for the one-year and two-year specification for the initial model with variable slope coefficients.

The thresholds using the simulated rating vary between 0.5 and 1 notches for both upgrades and downgrades. They are higher for upgrades of speculative grade companies and lower for downgrades in this subset. Switching the specification from one-year to two-year changes has only marginal effect of about 0.1 notches. Overall when interpreting the results of the initial model the through-the-cycle methodology and the noise due to bucketing account to one notch.

5.2 Quantifying Factors of Stickiness

While the previous analyses revealed the robustness of specification of the latent credit quality process and quantified the impact of the agency's methodology, this section is dedicated to the identification and quantification of the migration policy. I refine the models specification with the additional variables described in section 4 to explain their impact on the stickiness.With this specification I relax the restricting assumption of constant thresholds. To compare the results with the former specifications I present the mean-slope coefficient of the latent model and the mean-threshold for each of the following estimations.

Starting with the results for the latent model the mean slope coefficients in table 4 do not differ much from the ones estimated without interaction terms (see table 2). However, the total effect of the previous models now subdivides into several factors. The first of these factors, the difference of the rating's implied PD to the EDF has an positive effect for upgrades. When the current rating's PD is higher than the EDF the slope becomes steeper and a upgrade is triggered by less intensive changes in the EDF. The volatility of the EDF changes is negative for the one-year setting and positive for the two-year setting. The switching sign of these coefficients implies that the uncertainty of the market's valuation, proxied by the volatility of the EDF, changes as longer time periods are considered.

Since the coefficient is not very high compared to the other interaction terms, the impact is not so strong. Still, a higher volatility results in flatter upgrade responses over the one-year horizon and steeper responses in the two-year setting. The last interaction term is the difference of the bucket median and the EDF. Here the effect is negative, i.e. if the current EDF level is below the bucket's median the upgrade response is flatter. The more an issuer moves away from the buckets median, the flatter the rating response. One reason for this observation is the avoidance of rating reversals. If an issuers expected rating differs only by a small proportion from the current rating, a rating change should be unlikely and if triggered the new rating will not differ much from the old one. For the downgrade specification all implications are the same since the coefficients are in the same direction. The interpretation of the slope coefficients assumes that the expected value of the latent process is in the respective area (recall figure 3).

For the threshold specification a recession period widens the upgrade thresholds, i.e. higher changes of the latent credit risk are needed to trigger an upgrade and downgrade thresholds become smaller. The recession increases the upgrade threshold by 0.375 notches and decreases the downgrade threshold by 0.394 notches for the one-year specification. The effect of the recession on the two-year changes is a bit less pronounced.

Turning to the volatility of the EDF. The more volatile the EDF changes the more likely is a downgrade as the coefficient of this variable is negative. This effect is stronger for speculative-grade issuers. For upgrades the effect is not so pronounced, so a small EDF volatility does not affect the speed of an upgrade.

The $PDRAT - EDF$ variable gives a mixed picture. The coefficients are negative for upgrades, i.e. if the rating's implied PD is above the current EDF the upgrade probability decreases and vice versa. But the signs for downgrades are not consistent for investment-grade issuers. This could be due to effects captured by the $BMEDIAN - EDF$ variable, which also changes the sign for this sub sample. The latter variable controls for the incentive of rating reversal aversion as one principal of the migration policy. The effect of this variable is most pronounced for upgrades of the investment-grade sample. It is particularly important to avoid rating reversals here as indicated by the positive coefficient. A high distortion from the typical issuer rated this grade changes the probability of triggering a rating change opposite to the direction of this distortion. If an issuer has a much lower current-condition risk than the mean of his bucket an upgrade is likely and vice versa for the opposite case.

The constant threshold gives the mean value which has to be exceeded to observe a rating change. Additionally to this constant, the issuer-specific characteristics of the other variables can lead to an in- or decrease of this constant. Overall the constant is between 3.4 and 4 notches for the upgrades in the one-year specification and about 0.5 notches below that for downgrades. For the two-year panel the constant varies between 2.2 and 3 notches for upgrades and 1.6 to 2.2 notches for downgrades. The mean thresholds are within these numbers. The thresholds are thus not symmetric for upgrades and downgrades. When a rating agency upgrades a company the rating reversal avoidance is more severe. The better credit quality reflected by an EDF decrease is more fragile, in terms of avoiding reversals, than the decline of credit quality reflected by an increase in the EDF. As the agency attempts to provide a through-the-cycle rating this short-term credit quality increase may not be enough to alter the business cycle the company is in. Depending on the position of the business cycle and macroeconomic conditions, a one-two year increase of the EDF might be the beginning of some trouble. For this reason the downgrade threshold is lower than the upgrade threshold as being temporarily better does not imply being good, but being temporarily worse is a smoking gun.

Comparison with other Studies

The methodology used here is new in this branch of literature, but some of the results can be compared to existing empirical findings. Altman and Rijken (2004) use a constructed agency-model which they fit via simulations to the observed rating behaviour. Their point-in-time credit quality is modelled with mainly book-value based variables as proposed by Altman (1968) in his Z-score model. The resulting estimate of the credit quality is not as volatile as the stock-price driven EDF. Thus the threshold level obtained by these variables are smaller than my threshold levels by construction. In fact, the authors find that a rating change is triggered if the prediction of the point-in-time rating differs from the current agency rating by more than 1.25 notches. The following migration is considered to 'close 75% of the gap between the actual agency-rating level and the predicted rating level' (Altman and Rijken (2004)). In my model the analogy to this difference is the $PDRAT - EDF$ variable, which is significant in each setting. As discussed above, however, the influence on the thresholds is different for upgrades and downgrades. The level of adjustment when a threshold is crossed is given by the level of the triggering EDF change and the estimated coefficient, as described for the initial model above.

Löffler (2005) examines the avoidance of rating reversals as an explanation of the stylized facts. He points out that 'managed or unmanaged, ratings change when credit quality crosses a threshold; the main difference is that, under rating management thresholds are path-dependent'. To illustrate this model's prediction and the rating management implied hereby I derive the rating level. Since the model's predictions are given in rating *changes* I assume the first rating of each issuer to reflect the true credit quality. I compute the predicted changes without the presence of stickiness, i.e. the plain prediction of the upgrade and downgrade specification and constrain the predicted rating changes to the estimated thresholds, i.e. estimate the predicted sticky rating. Figure 11 shows the two resulting rating systems compared to the EDF for an arbitrary issuer. The specification used to produce this figure is given in panel (B) of table 4.

The left graph of figure 11 shows a quite volatile rating, which heavily relates on the movements of the EDF. While in the right graph the predicted sticky rating is less volatile and even reacts before the actual Moody's rating. These pictures are consistent with the discussed migration policy employed by rating agencies. Instead of following a volatile but timely market risk measure, they react only if the changes exceed thresholds, i.e. if the incentive of providing a stable rating is overpowered by the expected reputational losses.

Investment-grade Boundary & Effects of the Initial Rating

Johnson (2004) examines rating changes around the investment-grade boundary (Baa3 or BBB-) for Standard & Poor's ratings. He concludes that downgrades from this rating are higher and more frequent than from its neighbouring grades. The investment-grade dummy and the results obtained in each sub sample provide information on the structural difference along the boundary. The investment dummy is negative for upgrades and positive for downgrades. The thresholds widen for investment-grade rated companies, which is confirmed in the sub sample settings. This observation is in-line with the higher rating stability of high-rated issuers. The reputation is based on the correct prediction of the long-term credit risk and thus a investment-grade default results in a severe loss of reputation. Since many investors are subject to investment-grade restrictions crossing the investment-boundary is associated with large transaction costs.

To gain additional insights into the dynamics which influence the thresholds levels across the investment-grade boundary I re-estimate the model with a threshold-specification of three dummy variables: one for the investment-boundary at Baa3 and two for this grade's neighbours. For downgrade thresholds the Baa3-dummy is negative, but not significant, i.e. there is a tendency for faster downgrades from this level, but not so pronounced. The upper neighbour

(Baa2) is positive with 0.597 notches, i.e. a downgrade to Baa3 is less likely than the mean constant threshold level of 2.195 notches. The lower neighbour (B1) is again positive with 0.360 notches, but not highly significant. For upgrades the dummies increase with decreasing grades. For Baa2 there is an addition of 0.311 notches, for Baa3 of 0.433 and for B1 of 0.805 notches to the mean constant threshold level of 2.993. Concluding the results by Johnson are present in this model as well. There is a tendency for different rating behaviour around the investment-boundary, which is, however, not that pronounced. Thus the migration policy of Moody's can be considered consistent across the investment-grade boundary.

Recent research by Jorion and Zhang (2005) implies that the first rating which is assigned to a bond effects subsequent rating changes. The effect of such initial ratings in the friction model is examined analogously as for rating levels. For each initial rating class a dummy is included in the threshold specification. In this data set the number of companies within each initial rating class varies quite a lot. The mean issuers across the rating grades Aaa and Caa3 is at 201 with a volatility of 149. There are no observations for issuers initially rated worse than Ca. The minimum is observed for an initial rating of Caa2 with 5 companies and the maximum for a first rating of B1 with 472 issuers. Table 5 gives additional descriptive statistics.

Figure 12 shows the thresholds for each initial rating for the basic setting. As before these figures show the deviation of the thresholds. For upgrades (left hand graph) the initial rating matters for all rating classes from Aaa to B3. Here the upgrade timeliness is reduced by approximately two notches on the average upgrade threshold of -1.97 notches. For the remaining rating classes there is no significant deviation, the initial rating does not matter here with respect to the stickiness of rating actions. For downgrades the right hand graph shows no significant deviations for all initial rating classes but Caa2, which can be considered as outlier. Concluding the initial rating does not significantly affect the downgrade but lowers the upgrade threshold. The latter finding could imply that initial ratings tend to be too good, so the agencies avoid subsequent upgrades. Evidence for this hypothesis, however, is to weak to be more than speculative.

5.3 Outlooks and Watchlists

Current research by Hamilton and Cantor (2004) concludes that the rating momentum disappears when rating changes are enriched with outlooks and watchlists. The authors point out that 'Moody's credit opinion consists of both an issuer's current credit rating and its current outlook or review status'. When quantifying the migration policy of a rating agency it is thus appropriate to control for secondary rating information.

The effects of controlling for outlooks significantly reduces the threshold levels, i.e. adding outlooks increases the likelihood of observing a rating change in response to a change in the EDF. The level of reduction is a proxy for the informational value of outlooks. The information added by publishing outlook information amounts to about one notch and is symmetric for up- and downgrades.

The following calculations are based on a sub sample of the data set (see section 2 for details). A re-estimation of the models specified so far on this sub sample showed, however, no significant changes. The results are particularly stable for inclusion of issuers without any outlook information during the time-period for which other issuers have outlooks. The previous results can thus be directly compared with the outlook results, which are reported in table 6 for the full specification.

The coefficient estimates remain almost unchanged compared to the plain ratings setting. Since the volatility of the EDF becomes insignificant for upgrades, outlook information controls for the variation of the short-term measure of credit risk, i.e. signaling an active outlook or review might be an response to lower precession of the market-based risk measure.

6 Summary and Conclusion

The aim of this paper is to contribute to the understanding of rating changes. Facing the criticism that credit agency's ratings react slowly to changes in the credit quality this slowness is analyzed in terms of threshold factors, which have to be crossed by a latent credit process to observe a rating change. The threshold parameter can be interpreted as parameters of the agency's migration policy.

The model used here to quantify the impact of different exogenous variables on the rating stickiness is a generalization of the Tobin model. This model requires the specification of the latent rating-change process, which is done using a market-based point-in-time measure of credit quality. The thresholds are specified separately for upgrades and downgrades, yielding to definable insights to both processes.

The time horizon of the model is carefully examined. Overall the thresholds become smaller with increasing time horizons. This process diminishes when observing longer than two-year changes. All results are obtained by using non-overlapping time periods. Sensitivity analysis imply robustness across time- and issuer-specifications. Extending the examination on investment and sub-investment grade shows that the agency's reaction differ for these subsets. The thresholds are between 3.6 to 4 rating notches for a one-year horizon upgrade and 3 to 3.4 for downgrades over the same period. For two-year changes an upgrade is trigger if the point-in-time measure exceeds 2.5 to 2.8 notches and 1.6 to 2.1 notches lead to an downgrade. These numbers are means over the whole panel used and further biased by the through-the-cycle components not captured by the latent credit process and furthermore influenced by the process of bucketing. To quantify these effects I conduct a simulation study based on the forward-looking trend of the point-in-time measure. It can be concluded that the bias added by these effects amount up to one notch.

To control for the effect of secondary rating information, such as outlooks and reviews, the thresholds using adjusted ratings are estimated for a subsample of the data set. Here the threshold levels decrease to 1.2-2.4 notches, depending on the direction of the expected change.

Overall the migration policy used by Moody's and examined through friction models is robust. This implies that the agency's policy is elaborated and consistent with the markets implications. Particularly for investment grade rated companies the variation of the EDF slows an upgrade but fastens an expected downgrade. The estimated thresholds can be used e.g. by managers which are subject to investment restrictions as a monitoring device. If the short term credit risk over a certain period changes toward the thresholds a rating change is probable and an portfolio adjustment before the actual announcement would be advantageous.

So is it time to change the migration policy? Well that depends on what to achieve with a credit rating. If one wants timely information based on the markets prediction he should not look at the rating alone. Monitoring market measures in addition to the credit rating helps to understand the markets reaction and enables to predict the rating changes. The rating itself hereby includes valuable information and stability is neither bad per se nor does it lead to a misjudgement of the credit quality of an issuer. Concluding the rating agencies do not react as slow as often accused of. Outlook information shows that they know what's going on.

Appendix

Derivation of the likelihood function for the friction model

Consider the friction model given by equation (4) below and the model of the latent rating change $\Delta RAT^*_{it} = \beta^{'}\mathbb{X} + \varepsilon_{it}$ with $\varepsilon_{it} \sim N(0, \sigma^2)$. Since the calculations are done in a pooled setting, the subscript i describing the issuers is suppressed.

$$\Delta RAT_t = \begin{cases} \Delta RAT^*_t - \alpha_1 + \gamma_1 & , \Delta RAT^*_t < \alpha_1 \text{ (L)-Area} \\ 0 & , \alpha_1 \leq \Delta RAT^*_t \leq \alpha_2 \text{ (F)-Area} \\ \Delta RAT^*_t - \alpha_2 + \gamma_2 & , \alpha_2 < \Delta RAT^*_t \text{ (U)-Area} \end{cases} \tag{4}$$

The likelihood function for this model consist of three part according to the areas of equation (4). The derivation of these parts for upgrades ($\Delta RAT^*_t < 0$) and downgrades ($\Delta RAT^*_t > 0$) is essentially the same. Consider $Prob(\Delta RAT < 0) = Prob(\Delta RAT_t - \Delta RAT^*_t - \alpha_1 - \gamma_1)$, where $Prob()$ denotes the probability measure of expectations given $\mathbb{X}, \alpha_i, \gamma_i$. Inserting the model of the latent rating change and solving for the error term ε_t gives the first part of the likelihood function below. The frictional part, i.e. the (F)-area, is given by solving $Prob(\Delta RAT_t = 0) = Prob(\alpha_2 \leq \Delta RAT^*_t \leq \alpha_1)$. Let $\Phi()$ denote the cumulative distribution of the standard normal function and $\phi()$ the analogous density.

$$\begin{aligned} F_t(\Delta RAT_t|\mathbb{X}, \alpha_i, \gamma_i, \beta, \sigma) \quad = \quad & \Pi_{\Delta RAT_t<0}\, \sigma^{-1}\phi\left(\frac{\Delta RAT_t + \alpha_1 - \gamma_1 - \beta^{'}\mathbb{X}}{\sigma}\right) \\ & \cdot\Pi_{\Delta RAT_t=0}\left[\Phi\left(\frac{\alpha_2 - \beta^{'}\mathbb{X}}{\sigma}\right) - \Phi\left(\frac{\alpha_1 - \beta^{'}\mathbb{X}}{\sigma}\right)\right] \\ & \cdot\Pi_{\Delta RAT_t>0}\, \sigma^{-1}\phi\left(\frac{\Delta RAT_t + \alpha_2 - \gamma_2 - \beta^{'}\mathbb{X}}{\sigma}\right) \end{aligned}$$

The Log-likelihood is given by $logL() = \sum_t ln(F_t)$, which is

$$\begin{aligned} logL() \quad = \quad & \sum_{\Delta RAT_t<0} -0.5 \cdot \log(2\pi\sigma^2) - (\Delta RAT_t + \gamma_1 + \alpha_1 - \beta^{'}\mathbb{X})^2/(2\sigma^2) \\ & + \sum_{\Delta RAT_t=0} \Phi((\alpha_2 - \mathbb{X})/\sigma) - \Phi((\alpha_1 - \beta^{'}\mathbb{X})/\sigma) \\ & + \sum_{\Delta RAT_t>0} -0.5 \cdot \log(2\pi\sigma^2) - (\Delta RAT_t + \gamma_2 + \alpha_2 - \beta^{'}\mathbb{X})^2/(2\sigma^2) \end{aligned}$$

The Rosett-Model is obtained by setting $\gamma_1 = \gamma_2 = 0$. Allowing for different slope parameters changes the ML equation to $\beta_U' x_i$ for the U-Part and $\beta_L' x_i$ for the L-Part and analogous for the F-part. The frictional (F-Part) becomes $\Pi_F \left[\Phi\left(\frac{\alpha_2 - \beta_U' x_i}{\sigma}\right) - \Phi\left(\frac{\alpha_1 - \beta_L' x_i}{\sigma}\right)\right]$ For more details of these models see also Maddala (1982), chapter 6. The deviation of the locus $E[\Delta RAT|\mathbb{X}, \alpha_i, \gamma_i]$

in equation 3 is straightforward as:

$$\begin{aligned}
E[\Delta RAT|\mathbb{X}, \alpha_i, \gamma_i] &= E[\Delta RAT^* - \alpha_1|\mathbb{X}, \Delta RAT^* < \alpha_1] \cdot P(\Delta RAT^* < \alpha_1|\mathbb{X}) \\
&\quad +0 \cdot P(\Delta RAT^* \geq \alpha_1, \Delta RAT^* \leq \alpha_2|\mathbb{X}) \\
&\quad +E[\Delta RAT^* - \alpha_2|\mathbb{X}, \Delta RAT^* > \alpha_2] \cdot P(\Delta RAT^* > \alpha_2|\mathbb{X})
\end{aligned}$$

References

AFP (2002). Rating agencies survey: Accuracy, timeliness, and regulations. report of survey results. http://www.afp.com.

Almekinders, G. J. and Eijffinger, S. C. (1996). A friction model of daily bundesbank and federal reserve intervention. *Journal of Banking & Finance*, 20:1365–1380.

Altman, E. I. (1968). Financial ratios, discriminant analysis, and the prediction of corporate bankcrupty. *Journal of Finance*, 23:589–609.

Altman, E. I. and Arman, P. (2002). Defaults & returns on high yield bonds: Analysis through 2001. *Journal of Applied Finance*, 12.

Altman, E. I. and Bana, G. (2004). Defaults and returns on high-yield bonds. *The Journal of Portfolio Management*.

Altman, E. I. and Fanjul, G. (2004). Defaults and returns on high yield bonds: First quarter 2004. NYU Law and Economics Research Paper No. 04-008.

Altman, E. I. and Kao, D. L. (1992a). The implications of corporate bond ratings drift. *Financial Analysts Journal*, pages 64–75.

Altman, E. I. and Kao, D. L. (1992b). Rating drift in high yield bonds. *The Journal of Fixed Income*, (1):15–20.

Altman, E. I. and Rijken, H. A. (2004). How rating agencies achieve rating stability. *Journal of Banking and Finance*, 28(11):2679–2714.

Asano, H. (2002). Costly reversible investments with fixed costs: An empirical study. *Journal of Business & Economic Statistics*, 20(2):227–240.

Bannier, C. E. and Tyrell, M. (2005). Modelling the role of credit rating agencies - do they spark off a virtuous circle? Working Paper, Dept. of Finance, Goethe University Frankfurt.

Benabou, R. and Laroque, G. (1992). Using privileged information to manipulate markets: Insiders, gurus and credibility. *Quarterly Journal of Economics*, 107(3):921–958.

Blume, M. E., Lim, F., and MacKinlay, A. C. (1998). The declining credit quality of US corporate debt: Myth or reality? *Journal of Finance*, 53(4):1389–1413.

Boot, A. W. A., Milbourn, T. T., and Schmeits, A. (2006). Credit ratings as coordination mechanisms. *Review of Financial Studies*, page forthcoming.

Chow, G. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, 28:591–605.

Covitz, D. M. and Harrison, P. (2003). Testing conflicts of interest at bond ratings agencies with market anticipation: Evidence that reputation incentives dominate. Federal Reserve Board Working Paper.

Dagenais, M. G. (1969). A threshold regression model. *Econometrica*, 37(2):193–203.

Dagenais, M. G. (1975). Appliction of a threshold regression model to household purchases of automobiles. *The Review of Economics and Statistics*, 57(3):275–285.

Delianedis, G. and Geske, R. (1998). Credit risk and risk neutral default probabilities: Information about migrations and defaults. Working Paper, Anderson School, UCLA, Los Angeles.

Eubank, R. A. and Speckman, P. (1990). Curve fitting by polynomial-trigonometric regression. *Biometrika*, 77(1):1–9.

Fledelius, P., Lando, D., and Nielsen, J. P. (2004). Non-parametric analysis of rating transition and default data. Working Paper.

Fons, J. (2001). Understanding moody's corporate bond ratings and rating process. Special Comment.

Geske, R. (1977). The valuation of corporate liabilities as compound options. *Journal of Financial and Quantitative Analysis*, pages 541–552.

Greene, W. H. (2003). *Econometric Analysis*. Prentice Hall.

Hamilton, D. T. and Cantor, R. (2004). Rating transitions and defaults conditional on watchlist, outlook and rating history.

Hodrick, R. J. and Prescott, E. C. (1997). Postwar u.s. business cycles: An empirical investigation. *Journal of Money, Credit, and Banking*, 29(1):1–16.

Huber, P. J. (1967). The behavior of maximum likelihood estimates under nonstandard conditions. In *Proceedings of the Firth Berkeley Symposium on Mathematical Statistics and Probability*, pages 221–223. University of California Press.

Johnson, R. (2004). Rating agency actions around the investment-grade boundary. *The Journal of Fixed Income*. Federal Reserve Bank of Kansas City RWP 03-01.

Jorion, P., Shi, C., and Zhang, S. (2004). Tightening credit standards: Fact of fiction? Working Paper, http://www.gsm.uci.edu/ jorion/.

Jorion, P. and Zhang, G. (2005). Non-linear effects of bond ratings changes. Working Paper, http://www.gsm.uci.edu/ jorion/.

Kavvathas, D. (2000). Estimating credit rating transition probabilites for corporate bonds. SSRN Working Paper No. 252517.

Lando, D. and Skodeberg, T. (2002). Analyzing rating transitions and rating drift with continuous observations. *Journal of Banking and Finance*, 26:423–444.

Lau, T.-S. and Studden, W. J. (1985). Optimal designs for trigonometric and polynomial regression using canonical moments. *The Annals of Statistics*, 13(1):383–394.

Leser, C. E. V. (1961). A simple method of trend construction. *Journal of the Royal Statistical Society. Series B (Methodological)*, 23:91–107.

Löffler, G. (2005). Avoiding the rating bounce: Why rating agencies are slow to react to new information. *Journal of Behavior & Organization*, 56:365–381.

Lucas, D. and Lonski, J. (1992). Changes in corporate credit quality 1970-1990. *The Journal of Fixed Income*, (1):7–14.

Maddala, G. S. (1982). *Limited-Dependent and Qualitative Variables in Econometrics*. Econometric Society Monographs.

Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance*, 28:449–470.

Morris, S. (2001). Political correctness. *Journal of Political Economy*, 109:231–265.

Nickell, P., Perraudin, W., and Varotto, S. (2000). Stability of rating transitions. *Journal of Banking and Finance*, 24(1):203–227.

Perraudin, W. and Taylor, A. P. (2004). On the consistency of ratings and bond market yields. *Journal of Banking & Finance*, 28:2769–2788.

Rosett, R. N. (1959). A statistical model of friction in economics. *Econometrica*, 27:263–267.

Schlicht, E. (2004). Estimation the smoothing parameter in the so-called hodrick-prescott filter. Dep. of Economics, University of Munich, Discussion paper 2004-2.

Spence, A. M. (1973). Job market signaling. *Quarterly Journal of Economics*, 87(3):355–374.

Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica*, 26:24–36.

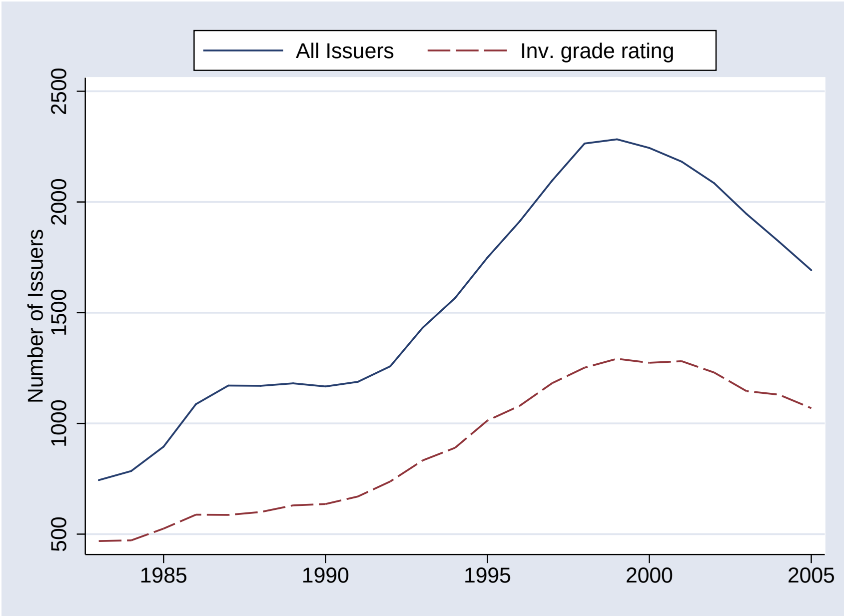
Vuong, Q. H. (1989). Likelihood ratio test for model selection and nonnested hypothesis. *Econometrica*, 57:307–333.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimation and a direct test for heteroskedasticity. *Econometrica*, 48:817–830.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. The MIT Press.

Yoshizawa, Y. (July 29, 2003). Moody's approach to rating synthetic CDO's. Moody's Investor Service.

Figure 1: Evolution of number of issuers over time.



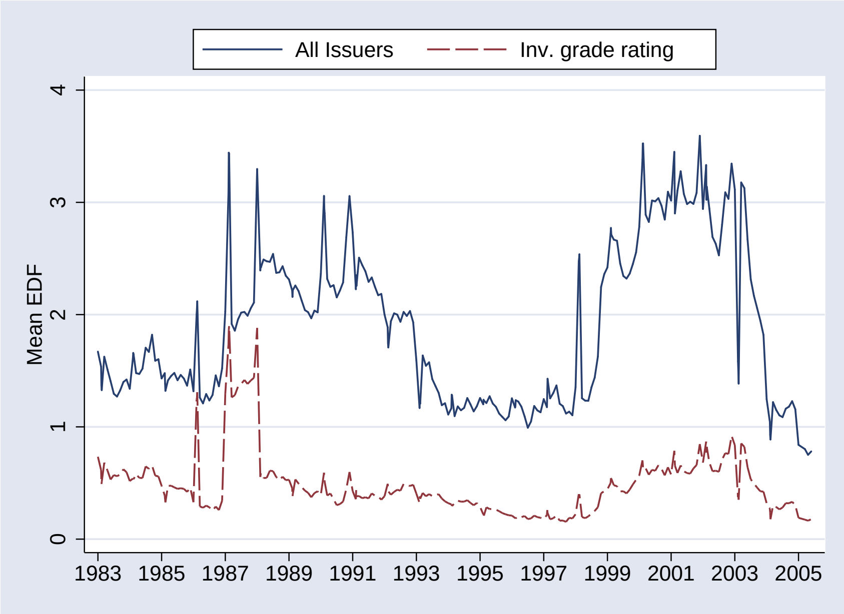
This graph shows the number of issuers and the number of issuers with investment-grade rating of the data set used. Overall there are 4,023 US and Non-US issuer with a total of 411,109 observations present.

Table 1: Descriptive Statistics of MKMVs EDF per Rating Class.

			EDF				
Moody's Rating	Numerical Value	Obs.	Mean	Median	Std. dev.	5% Quantile	95% Quantile
Aaa	1	7,077	0.82	0.07	3.50	0.01	1.02
Aa1	2	4,417	0.51	0.08	2.59	0.01	0.73
Aa2	3	12,099	0.35	0.13	1.34	0.01	0.93
Aa3	4	18,279	0.34	0.10	1.57	0.01	0.84
A1	5	25,542	0.34	0.12	1.36	0.02	1.00
A2	6	41,261	0.35	0.18	0.93	0.02	1.03
A3	7	33,018	0.37	0.18	0.89	0.02	1.19
Baa1	8	28,258	0.51	0.23	1.27	0.02	1.60
Baa2	9	37,283	0.61	0.31	1.10	0.03	2.04
Baa3	10	26,276	0.78	0.38	1.58	0.04	2.58
Ba1	11	22,989	1.02	0.50	1.84	0.06	3.51
Ba2	12	23,587	1.63	0.88	2.46	0.09	5.43
Ba3	13	30,271	1.92	1.05	2.82	0.13	6.76
B1	14	28,283	3.25	1.57	4.42	0.16	14.01
B2	15	17,733	5.07	2.65	5.75	0.28	20.00
B3	16	14,994	8.26	5.16	7.39	0.46	20.00
Caa1	17	3,885	10.66	8.81	7.90	0.57	20.00
Caa2	18	6,550	13.75	20.00	7.62	1.01	20.00
Caa3	19	1,861	16.00	20.00	6.58	1.75	20.00
Ca	20	3,429	15.18	20.00	7.31	1.15	20.00
C	21	705	17.69	20.00	5.68	0.93	20.00

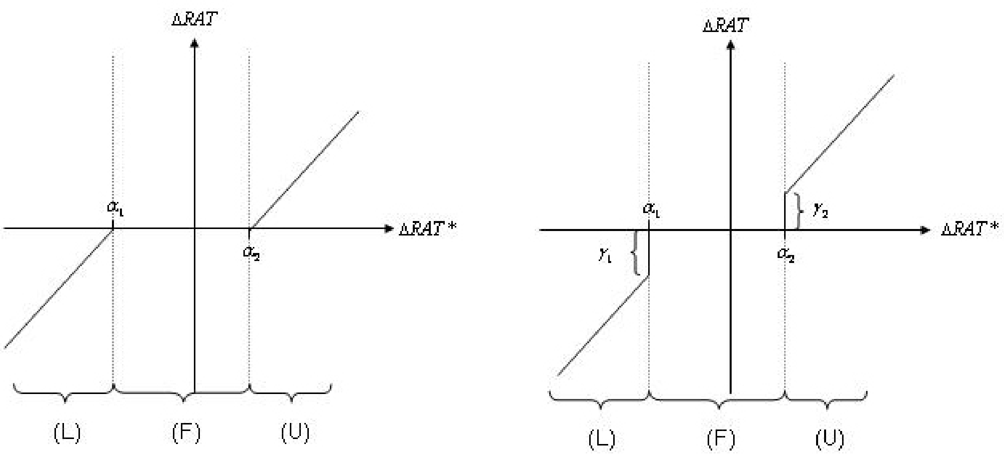
This tables shows descriptive statistics for Moody KMV's Expected Default Frequency (EDF) by each class of Moody's issuer rating. The second column shows the numerical value of the rating class.

Figure 2: Evolution of the Mean and Median of MKMV's EDF over time



These graphs show the evolution of the mean (left side) and median (right side) of Moody-KMV's Expected Default Frequency (EDF) over time.

Figure 3: How do the friction models look like?



These graphs show the friction model of Rosett (1959) (left side) and the extension by Dagenais (1975) (right side). See section 3 for details and the appendix for a derivation of the likelihood.

Table 2: Initial model of one-year rating changes.

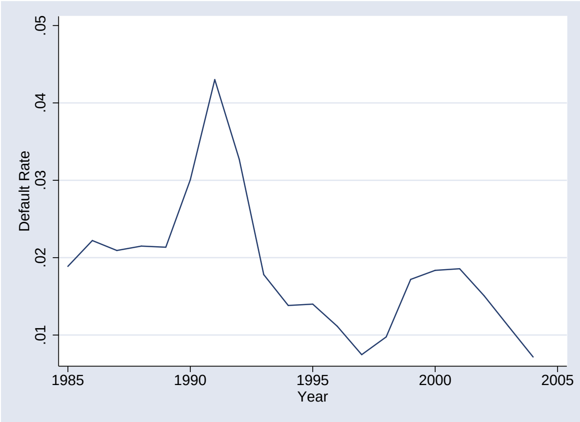
	Panel (A)			Panel (B)		
	Full	Invest. Grade	Spec.	Full	Invest. Grade	Spec. Grade
Dependent Variable			$\Delta_{12}RAT$			
			Upgrade Specification			
$\Delta_{12}LNEDF$	0.663	0.291	0.790	0.401	0.169	0.576
	(24.8)	(10.3)	(23.4)	(12.6)	(4.3)	(12.2)
α_1	-4.544	-4.106	-4.270	-3.947	-3.909	-3.663
	(67.6)	-(59.6)	-(46.8)	(45.5)	(41.8)	(30.7)
γ_1	0.000	0.000	0.000	0.000	0.000	0.000
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
			Downgrade Specification			
$\Delta_{12}LNEDF$	0.663	0.291	0.790	0.818	0.368	0.907
	(24.8)	(10.3)	(23.4)	(24.5)	(10.5)	(22.0)
α_2	3.381	3.212	3.022	3.627	3.296	3.265
	(61.6)	(54.9)	(44.3)	(52.1)	(45.7)	(38.3)
γ_2	0.000	0.000	0.000	0.000	0.000	0.000
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Wald	616.08	106.19	546.14	158.42	18.31	149.25
Log-Likelihood	-50,265	-23,751	-20,124	-50,165	-23,738	-20,091
			Number of Observations			
Total	51,772	32,167	19,605	51,772	32,167	19,605
Upgrade (L)	4,181	2,467	1,714	4,181	2,467	1,714
Frictional (F)	40,068	26,175	13,893	40,068	26,175	13,893
Downgrade (U)	7,523	3,525	3,998	7,523	3,525	3,998

The table presents the results for the regression of yearly rating changes on yearly percentage changes of the EDF. The first panel (A) imposes the restriction of equal coefficients for the upgrade- and downgrade specification, while panel (B) releases this constraint. t-statistics are given in parenthesis and calculated using a generalised Hubert-White sandwich estimator for cluster robust standard errors (see section 2) with clustering among issuers.

Figure 4: How do the Thresholds change over the years?

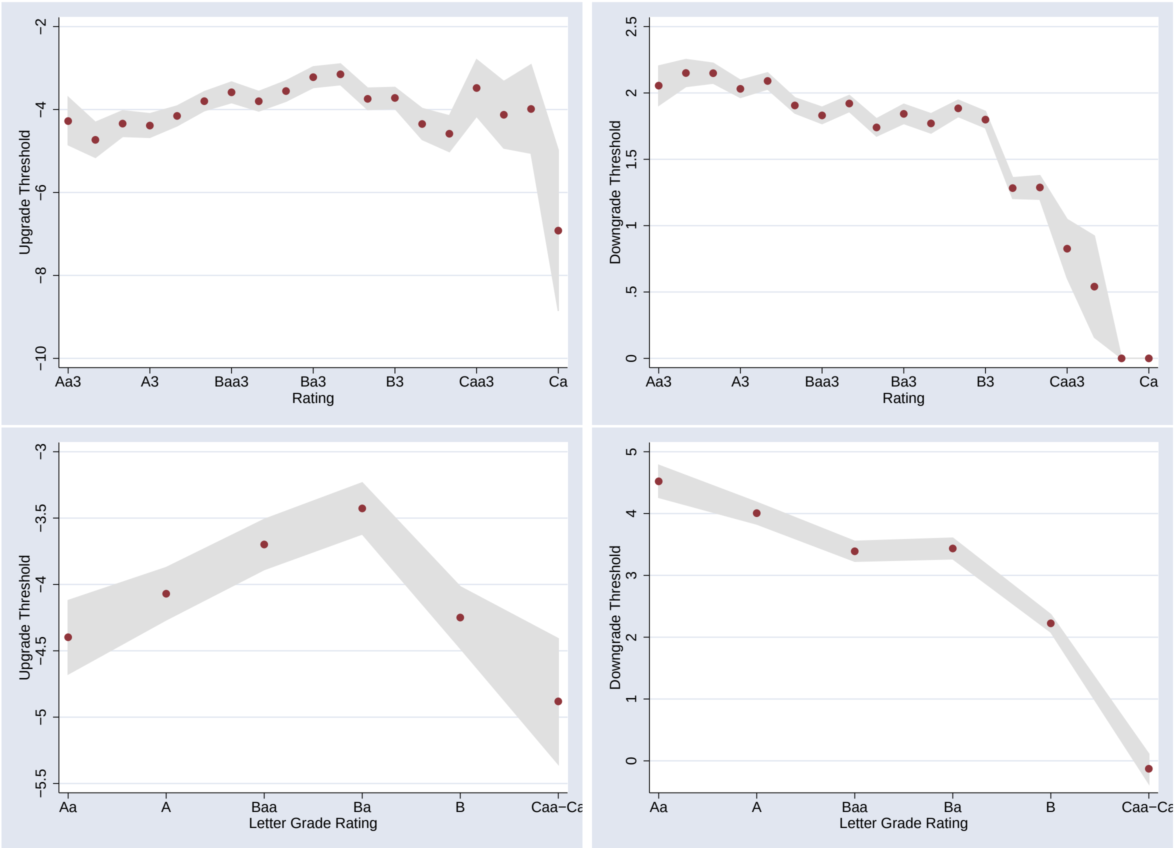
This graph shows the difference in the threshold parameter for different years. The lefthand graph shows the upgrades threshold while the righthand gives the downgrades threshold using the initial setting of table 2. The shaded area represents an 95% confidence interval. Whenever zero is within the 95% area this years difference is not significantly different from the overall threshold.

Figure 5: Default rates per year



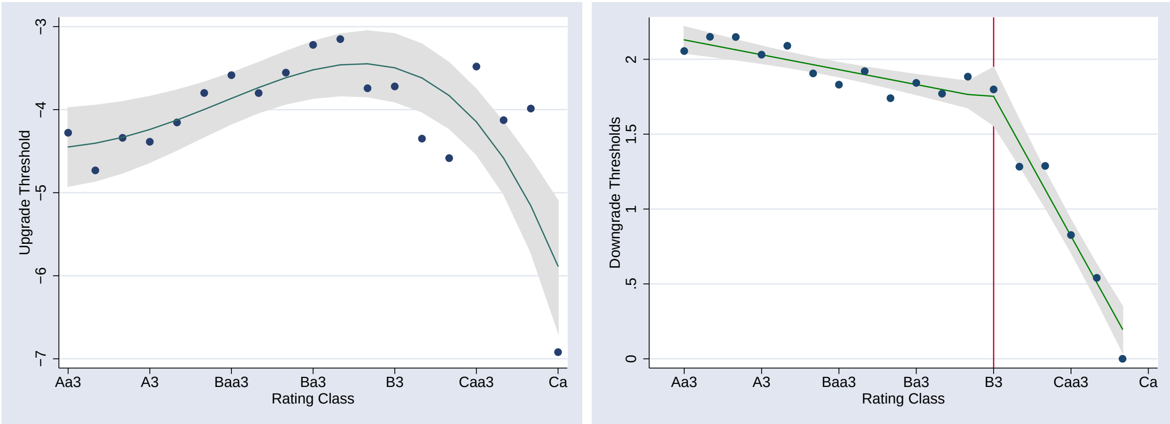
This graph shows the realized default rate per year. Migrations to default and within default are excluded in the analysis (see section 2 for details). The default rate is calculated as the number of defaulted company during that year over the number of total companies in that year.

Figure 6: Thresholds per alphanumerical and letter grade Rating



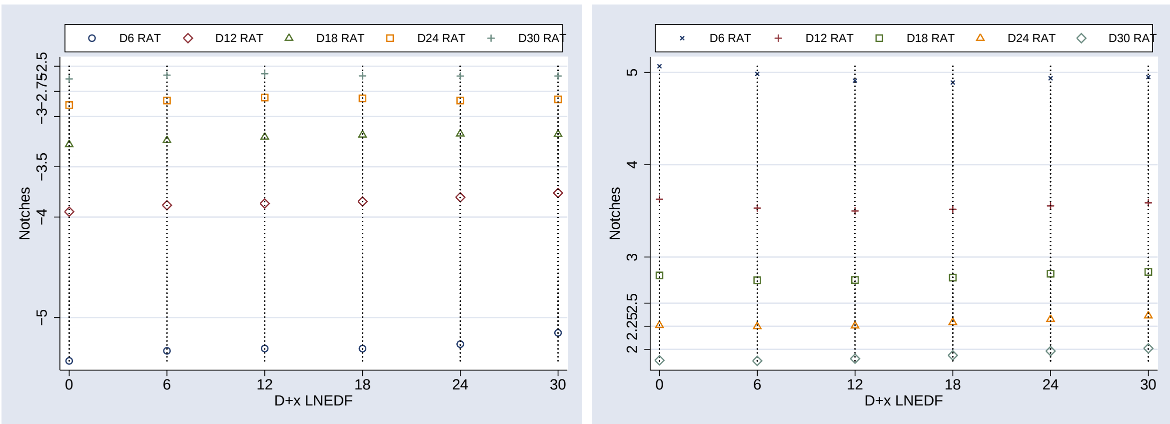
This graph shows the threshold parameter for upgrades (lefthand graph) and downgrades (righthand graph) with an 95% interval. The upper graphs show the thresholds for each alphanumerical rating class while the lower graphs restrict the ratings to the letter grades. Each set is estimated seperately using the initial setting of table 2

Figure 7: Describing the Thresholds per Rating Class by a Polynomial



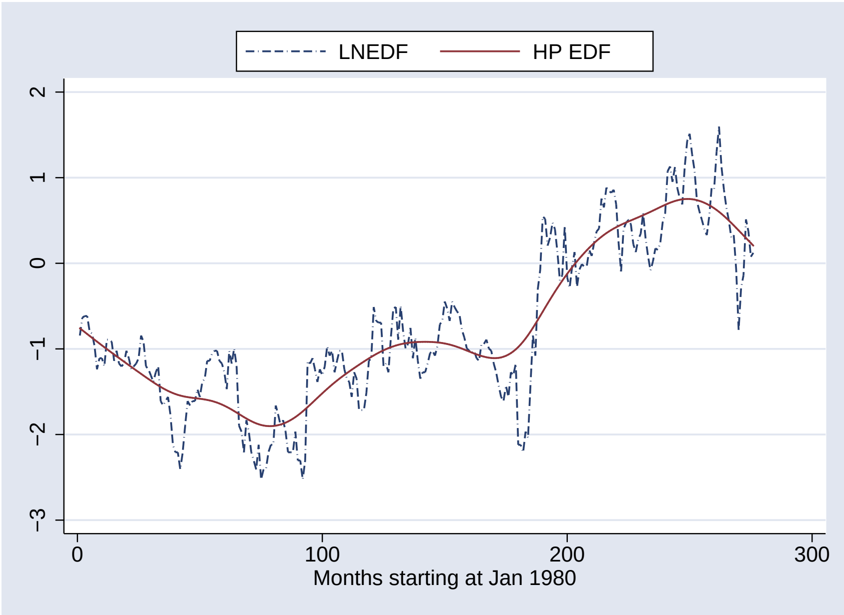
These graphs show a polynomial fit for the upgrade (lefthand graph) and downgrade (righthand graph) thresholds per rating class with an 95% confidence band. See figure 6 for details.

Figure 8: How do the Thresholds depend on the horizon?



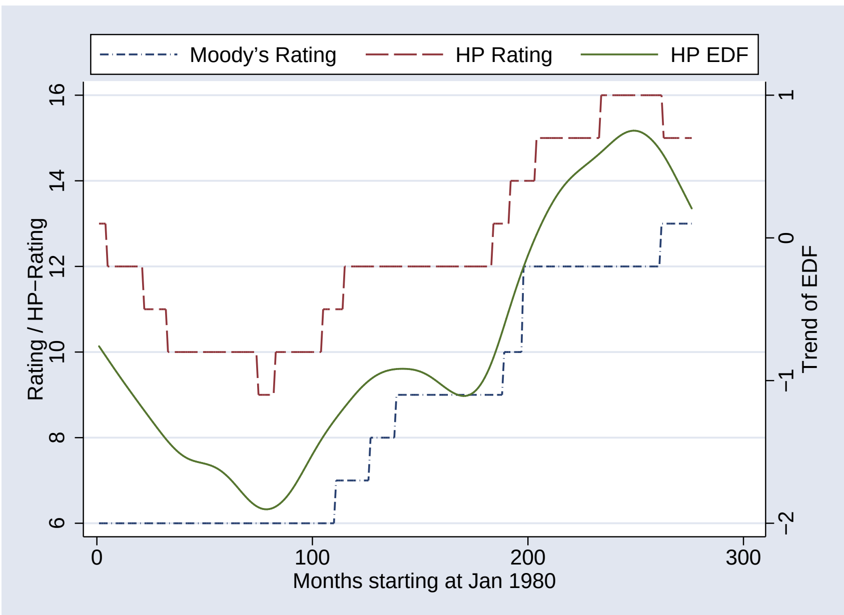
This graphs shows the threshold parameter for upgrades (left side) and downgrades (right side) for different parameter specifications. On the x-axis the time-horizon for the independent variable $\Delta_{y+x} LNEDF$ increases, while each row shows a different specification of the dependent variable $\Delta_y RAT$.

Figure 9: How does the Trend of the EDF look like?



This graphs represents the use of the Hodrick/Prescott filter for an arbitrary issuer, which is used to compute the trend of the EDF for an arbitrary issuer. See section 5.1.2 for details.

Figure 10: Example of the simulated rating for an arbitrary issuer.



This graphs compares the constructed rating ($HPRAT$) based on the Hodrick/Prescott trend of the EDF ($HPEDF$) and Moody's issuer rating (RAT).

Table 3: Impact of the through-the-cycle & bucketing methodology.

	Panel (A) - One-Year Changes			Panel (B) - Two-Year Changes		
	Full	Invest. Grade	Spec. Grade	Full	Invest. Grade	Spec. Grade
Dependent Variable		$\Delta_{12}HPRAT$			$\Delta_{24}HPRAT$	
			Upgrade Specification			
$\Delta_{12/24}LNEDF$	1.301	1.276	1.359	1.734	1.718	1.758
	(82.3)	(66.4)	(46.5)	(98.1)	(81.6)	(52.0)
α_1	-0.782	-0.651	-1.040	-0.671	-0.522	-0.963
	(38.9)	(29.7)	(25.3)	(28.8)	(20.3)	(20.1)
γ_1	0.000	0.000	0.000	0.000	0.000	0.000
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
			Downgrade Specification			
$\Delta_{12/24}LNEDF$	1.296	1.377	1.203	1.639	1.754	1.499
	(89.1)	(71.0)	(52.8)	(114.9)	(92.8)	(63.4)
α_1	0.834	0.909	0.729	0.596	0.667	0.498
	(40.5)	(35.0)	(21.0)	(27.0)	(23.9)	(12.9)
γ_1	0.000	0.000	0.000	0.000	0.000	0.000
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Wald	6,775	4,411	2,163	9,633	6,659	2,707
Log-Likelihood	-71,370	-44,181	-25,106	-66,391	-42,001	-20,722
			Number of Observations			
Total	51,772	32,167	19,605	45,579	29,211	16,368
Upgrade (L)	4,181	2,467	1,714	6,463	4,144	2,319
Frictional (F)	40,068	26,175	13,893	27,525	19,275	8,250
Downgrade (U)	7,523	3,525	3,998	11,591	5,792	5,799

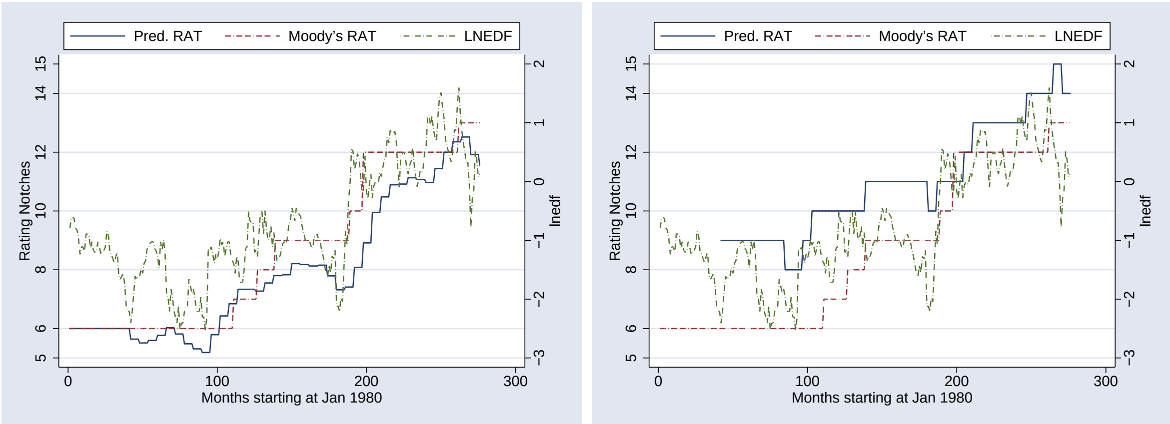
The table presents the results for the regression using a simulated rating based on the Hodrick/Prescott filter of the EDF. This estimates the bias of the friction model due to the effect of through-the-cycle rating and the method of bucketing. *t*-statistics are given in parenthesis and calculated using a generalised Hubert-White sandwich estimator for standard errors (see section 2).

Table 4: Factors which influence the stickiness and adjustment level of rating actions.

	Panel (A) - One-Year Changes			Panel (B) - Two-Year Changes		
	Full	Invest. Grade	Spec. Grade	Full	Invest. Grade	Spec. Grade
Dependent Variable	$\Delta_{12}RAT$			$\Delta_{24}RAT$		
			Upgrade Slope Coefficients			
$\Delta_{12/24}LNEDF$	0.352	0.196	0.510	0.315	0.223	0.471
	(9.9)	(3.8)	(8.5)	(11.2)	(5.6)	(10.0)
· PDRAT-EDF	0.042	-0.193	0.011	0.044	-0.086	0.007
	(3.5)	(0.7)	(0.9)	(4.9)	(0.5)	(0.8)
· VOLAEDF	-0.012	-0.029	0.008	0.022	-0.004	0.045
	(1.7)	(3.5)	(0.7)	(2.3)	(0.5)	(3.1)
· BMEDIAN-EDF	-0.053	0.189	-0.027	-0.063	0.068	-0.025
	(3.4)	(0.7)	(1.5)	(4.6)	(0.4)	(1.6)
Mean Slope	0.390	0.34	0.460	0.387	0.320	0.491
	(2.7)	(6.3)	(2.2)	(1.8)	(3.2)	(1.7)
			Upgrade Threshold			
VOLAEDF	0.036	0.031	0.013	0.019	0.006	-0.005
	(2.4)	(1.2)	(0.8)	(0.9)	(0.2)	(0.2)
PDRAT-EDF	-0.072	-1.825	-0.069	-0.136	-2.276	-0.134
	(4.1)	(6.6)	(4.3)	(5.1)	(7.4)	(5.1)
BMEDIAN-EDF	0.094	1.907	0.085	0.097	2.320	0.089
	(4.6)	(6.8)	(4.3)	(3.7)	(7.5)	(3.3)
INVEST	-0.323			-0.289		
	(4.3)			(3.2)		
RECESSION	-0.375	-0.574	-0.328	-0.253	-0.234	-0.679
	(3.5)	(4.0)	(2.1)	(2.9)	(2.4)	(4.6)
Cons.	-3.666	-4.073	-3.427	-2.588	-3.035	-2.215
	(36.0)	(39.0)	(28.1)	(26.0)	(31.1)	(18.7)
Mean	-3.933	-3.894	-3.621	-2.873	-2.788	-2.569
	(13.7)	(12.8)	(12.0)	(7.5)	(9.6)	(4.3)
			Downgrade Slope Coefficients			
$\Delta_{12/24}LNEDF$	0.571	0.297	0.678	0.519	0.278	0.663
	(17.4)	(6.2)	(14.1)	(18.8)	(7.4)	(16.3)
· PDRAT-EDF	0.036	-0.085	0.011	0.022	0.107	-0.004
	(6.0)	(0.3)	(2.1)	(5.3)	(0.6)	(0.9)
· VOLAEDF	-0.004	-0.009	0.006	0.018	0.001	0.025
	(0.7)	(1.6)	(0.6)	(1.9)	(0.1)	(2.4)
· BMEDIAN-EDF	-0.070	0.095	-0.036	-0.054	-0.117	-0.017
	(7.7)	(0.4)	(3.6)	(7.5)	(0.6)	(2.2)
Mean Slope	0.621	0.573	0.696	0.572	0.529	0.639
	(4.0)	(11.8)	(3.1)	(3.4)	(5.8)	(2.8)
			Downgrade Threshold			
VOLAEDF	-0.065	-0.069	-0.106	-0.113	-0.143	-0.161
	(5.2)	(3.1)	(8.0)	(5.8)	(4.1)	(8.4)
PDRAT-EDF	-0.097	0.777	-0.128	-0.101	0.357	-0.136
	(11.3)	(2.9)	(14.1)	(10.6)	(1.2)	(13.2)
BMEDIAN-EDF	0.110	-0.594	0.129	0.037	-0.237	0.068
	(8.0)	(2.1)	(8.6)	(2.2)	(0.8)	(4.1)
INVEST	0.807			0.967		
	(11.9)			(11.9)		
RECESSION	-0.394	-0.327	-0.324	-0.237	0.046	-0.720
	(4.5)	(3.3)	(2.4)	(3.5)	(0.6)	(7.6)
Cons.	3.079	3.450	3.535	1.658	2.221	2.246
	(43.6)	(43.8)	(42.0)	(22.4)	(28.7)	(27.8)
Mean	3.397	3.280	3.014	2.077	2.107	1.666
	(4.6)	(7.1)	(3.5)	(2.1)	(2.6)	(4.7)
Wald	136.07	32.10	116.48	202.68	55.19	181.03
Log-Likelihood	-48,263	-23,595	-18,734	-61,823	-32,031	-20,788

The table presents the results for the regression of one-year (panel (A)) and two-year (panel (B))rating changes with the respective changes of the EDF. The specification of the latent credit process includes interaction terms with the EDF changes. The specification of the thresholds is variable. t-statistics are given in parenthesis and calculated using a generalised Hubert-White sandwich estimator for cluster robust standard errors (see section 2) with clustering among issuers. The rows 'Mean' of the threshold specification give the mean thresholds calculated as weighted sum of the threshold specifications.

Figure 11: Example of Ratings predicted by the friction model.



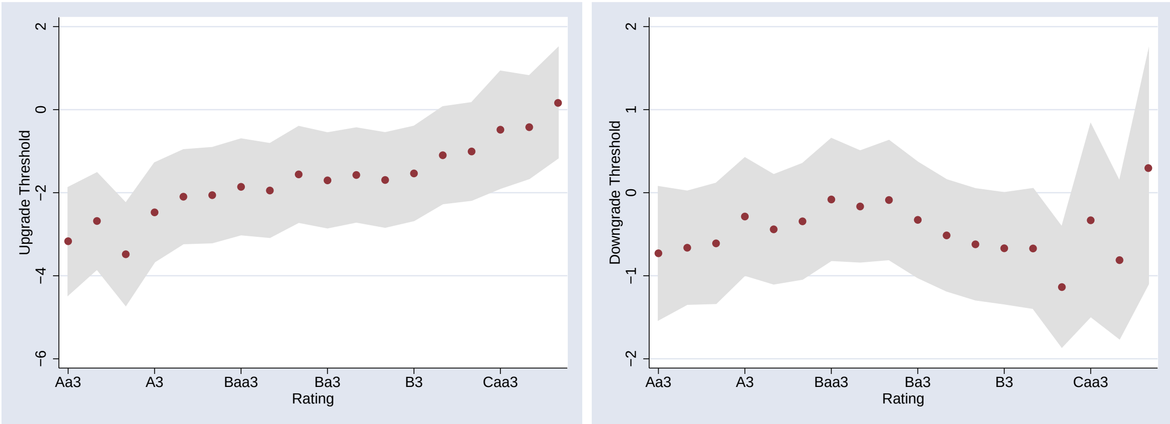
These graphs compare the Moody's rating with the models prediction without imposing stickiness (left side) and with sticky ratings (right side) for the parameter specification of table 4, panel (B).

Table 5: Descriptive Statistics for issuers with different initial rating.

Initial Rating	Obs	Mean Lifetime (Month)	Mean $\Delta_{12}RAT$	Mean $\Delta_{24}RAT$
Aaa	92	237	0.219	0.461
Aa1	36	167	0.226	0.475
Aa2	135	233	0.228	0.451
Aa3	113	147	0.218	0.448
A1	145	155	0.161	0.329
A2	426	214	0.175	0.348
A3	234	146	0.162	0.334
Baa1	213	124	0.073	0.140
Baa2	425	187	0.114	0.218
Baa3	203	123	0.095	0.162
Ba1	223	143	0.128	0.239
Ba2	393	159	0.222	0.409
Ba3	431	120	0.230	0.419
B1	471	96	0.235	0.419
B2	203	109	0.178	0.244
B3	175	78	0.307	0.458
Caa1	44	51	0.174	0.150
Caa2	40	113	0.092	0.159
Caa3	5	69	0.092	0.292
Ca	10	85	-0.215	-0.290
Mean	211	141	0.175	0.324
SD	146	51	0.062	0.117

This tables show descriptive statistics for issuers with different initial rating classes. 'Obs' gives the number of companies with an given initial rating. 'Lifetime' gives the mean lifetime of these issuers in month, while the two rightmost columns give the mean rating changes over a one-year and two-year period. 'SD' refers to the standard deviation. See figure 12 for a graphical representation of the thresholds for these classes.

Figure 12: Does the initial rating matter?



This figure shows the impact on the initial rating on the threshold as deviation from the average with a 95% confidence band. The righthand graph gives the upgrade thresholds and the lefthand graph the downgrade thresholds for companies with a given initital rating.

Table 6: Impact of Outlook & Watchlists on rating thresholds.

	Panel (A) - One-Year Changes			Panel (B) - Two-Year Changes		
	Full	Invest. Grade	Spec. Grade	Full	Invest. Grade	Spec. Grade
Dependent Variable	$\Delta_{12}OUTLKRAT$			$\Delta_{24}OUTLKRAT$		
			Upgrade Specification			
$\Delta_{12/24}LNEDF$	0.395	0.336	0.473	0.350	0.273	0.475
	(11.1)	(6.2)	(8.0)	(10.7)	(5.9)	(8.5)
· PDRAT-EDF	0.030	0.630	0.017	0.031	0.484	0.017
	(3.5)	(2.0)	(2.0)	(3.9)	(2.1)	(2.0)
· VOLAEDF	0.031	-0.001	0.045	0.067	0.013	0.071
	(2.8)	(0.1)	(3.1)	(5.0)	(0.6)	(4.5)
· BMEDIAN-EDF	-0.054	-0.668	-0.034	-0.059	-0.557	-0.039
	(4.5)	(2.1)	(2.6)	(4.7)	(2.4)	(2.7)
			Upgrade Thresholds			
VOLAEDF	-0.001	0.084	-0.031	0.008	0.104	-0.044
	(0.1)	(2.0)	(1.6)	(0.3)	(1.9)	(1.5)
PDRAT-EDF	-0.042	0.616	-0.043	-0.106	-0.174	-0.115
	(4.2)	(2.0)	(4.3)	(6.0)	(0.5)	(6.1)
BMEDIAN-EDF	0.056	-0.569	0.051	0.077	0.194	0.076
	(3.7)	(1.8)	(3.3)	(3.4)	(0.5)	(3.2)
INVEST	-0.384			-0.291		
	(5.7)			(3.2)		
RECESSION	-0.208	-0.032	-0.318	0.052	0.085	-0.266
	(2.1)	(0.2)	(2.1)	(0.5)	(0.7)	(1.7)
Cons.	-2.054	-2.440	-1.916	-1.617	-1.922	-1.334
	(27.5)	(30.2)	(21.9)	(16.6)	(21.4)	(11.2)
Mean	-2.353	-2.498	-2.129	-1.893	-1.880	-1.746
	(12.1)	(10.2)	(7.3)	(5.1)	(6.7)	(2.6)
			Downgrade Specification			
$\Delta_{12/24}LNEDF$	0.556	0.361	0.604	0.525	0.317	0.642
	(14.4)	(6.2)	(10.8)	(14.8)	(6.3)	(12.3)
· PDRAT-EDF	0.026	-0.038	0.014	0.016	0.064	0.002
	(4.6)	(0.1)	(2.6)	(3.4)	(0.3)	(0.6)
· VOLAEDF	0.007	-0.007	0.015	0.051	0.089	0.036
	(0.7)	(0.2)	(1.3)	(3.8)	(1.7)	(2.9)
· BMEDIAN-EDF	-0.057	0.026	-0.031	-0.044	-0.085	-0.020
	(5.8)	(0.1)	(2.9)	(5.6)	(0.4)	(2.3)
			Downgrade Thresholds			
VOLAEDF	-0.025	0.099	-0.076	-0.035	0.139	-0.115
	(1.7)	(1.8)	(5.0)	(1.4)	(1.9)	(4.7)
PDRAT-EDF	-0.064	-0.173	-0.075	-0.078	-0.645	-0.093
	(7.5)	(0.5)	(9.0)	(7.5)	(1.7)	(9.0)
BMEDIAN-EDF	0.109	0.426	0.106	0.039	0.795	0.047
	(7.2)	(1.3)	(6.9)	(2.1)	(2.1)	
INVEST	0.735			1.109		
	(9.9)			(11.1)		
RECESSION	0.071	-0.013	0.295	0.204	0.381	-0.269
	(0.7)	(0.1)	(1.8)	(2.2)	(3.3)	(2.1)
Cons.	1.854	2.270	2.205	0.662	1.313	1.263
	(26.1)	(30.0)	(25.7)	(7.7)	(15.9)	(13.3)
Mean	2.166	2.251	1.827	1.223	1.400	0.807
	(3.8)	(6.2)	(3.1)	(1.6)	(3.4)	(1.1)
Wald	244.18	45.25	186.92	359.91	104.01	320.23
Log-Likelihood	-29,667	-14,863	-12,408	-69,416	-37,702	-23,343

The table presents the results for the regression of one-year (panel (A)) and two-year (panel (B)) rating changes with addition of outlook and watchlist information the respective changes of the EDF. The specification of the latent credit process includes interaction terms with the EDF changes. The specification of the thresholds is variable. The rows 'Mean' of the threshold specification give the mean thresholds. t-statistics are given in parenthesis and calculated using a generalised Hubert-White sandwich estimator for standard errors (see section 2).