What's in the Trend: Understanding the Dynamics of Corporate Credit Ratings When Ratings Are Sticky*

Argyro Panaretou† Grzegorz Pawlina‡ Qifan Zhai§

February 7, 2022

Abstract

We propose a model of credit rating migrations that allows for the possibility of rating stickiness. The model aims to capture the mechanism of the rating process, based on the notion of hysteresis in real option models, that underlies the behavior of rating agencies and is used to explain the observed gradual deterioration in ratings. The paper contributes to the ongoing debate whether the downward trend in credit ratings results from the deteriorating credit quality or tightening rating standards. It is shown that corporate credit quality actually slightly improves over time and that there is asymmetry in rating migrations as upgrades become increasingly more difficult, while downgrade standards remaining unchanged.

*We thank Alcino Azevedo, Shantanu Banerjee, Dimitris Kanelis, Andrianos Tsekrekos, seminar participants at Aston Business School as well as participants of the Lancaster-Manchester Accounting and Finance workshop for helpful comments and suggestions. Any remaining errors are ours.

†a.panaretou@lancaster.ac.uk, Department of Accounting and Finance, Lancaster University Management School, Lancaster, LA1 4YX, UK.

‡Corresponding author: g.pawlina@lancaster.ac.uk, Department of Accounting and Finance, Lancaster University Management School, Lancaster, LA1 4YX, UK.

§scutb16@scut.edu.cn, South China University of Technology, 381 Wushan Road, Tianhe District, Guangzhou, 510641, China.

1 Introduction

Credit rating, as a crucial indicator of a firm's risk level, assists investors in decision making and may act as a tool helping enforce market discipline (Kolaric et al., 2021). It also gradually becomes the standard for the national capital assessment following the Basel Capital Accord. The average credit rating in the US market has deteriorated in the past few decades, with rating downgrades dominating the migration trend. It has also been shown that rating migrations tend to move in the same direction, and the consecutive downgrades are more frequent than upgrades (Altman and Kao, 1992; Lando and Skødeberg, 2002; Du, 2003). It is therefore not surprising that the observed rating levels exhibit downward momentum and that the average rating dropped by three notches between 1985 and 2009 (Baghai et al., 2014).

The phenomenon of rating deterioration has led to the debate whether it is a result of a decreasing credit quality or the tightening of rating standards (Blume et al., 1998; Jorion et al., 2009; Alp, 2013). A rating event involves two main participants, which are firms to be evaluated and rating agencies. Possible causes of the deterioration may therefore come from either side and encompass the quality of the borrower and the stringency of the rating standard. A number of studies, such as Blume et al. (1998); Jorion et al. (2009), and Alp (2013), focus on the possible cause of the deterioration by matching ratings to firm characteristics. They typically rely on empirical models attempting to control for all relevant covariates while leaving the effect of the (time-varying) rating standards to year dummies. The general conclusion of this strand of literature is that the continuous tightening of rating standards indeed contributes to the observed rating deterioration. While offering valuable insights, this line of research cannot put a definitive end to the debate as it does not address the question of the evolution of the quality of firms' borrowing.

This paper attempts to fill the void in the literature by focusing on the mechanism of rating migrations. The migrations are informative because they consider the most recent updates of rating agencies, and hence involve the trigger and motivation of agencies' decisions. To directly assess the interaction between credit quality and varying standards, it is necessary to filter out the effect of each component. This cannot be done without considering in the empirical design the mechanism of how ratings are assigned and updated. The empirical work in this area largely agrees on two attributes of credit rating, namely, the high stability (persistence) and the slow respondence. The prior results indicate that ratings demonstrate considerably higher persistence compared with the underlying credit quality (Kealhofer et al., 1998; Carey and Hrycay, 2001; Löffler, 2004). In other words, it is not unusual that the credit quality changes but the ratings do not. Moreover, credit ratings are also characterized by their slow respondence. For example, in the run-up to the 1997 Asian crisis, rating agencies made very few downgrades despite credit rating being designed as a forward-looking indicator. The above evidence therefore indicates that credit ratings are generally sticky.

This research proposes a model of rating process that allows for sticky ratings and explores the impact of the stickiness on rating migrations.[1] Differing from existing empirical studies (Blume et al., 1998; Jorion et al., 2009; Alp, 2013), the model embeds the stickiness in the empirical design and the estimation. A typical assumption in the current empirical research is that firms' credit quality matches its rating. According to such a view, rating agencies apply immediate rating updates to eliminate any deviations outside of the current rating, resulting in ratings being accurate at all times. However, this research challenges the assumption of accurate and contemporaneous ratings. We argue that credit ratings demonstrate high persistence and slow respondence, which can make them a *lagged* indicator of credit quality. These features of credit rating therefore introduce a certain degree of tolerance for deviations of credit quality outside of the ratings' nominal range (in the same way in which lump sum cost affect optimal control policies). The tolerance may be small, but it represents a distinct

[1] Here, we define stickiness as the deviation of credit quality from the nominal range of its rating. Suppose the nominal Z-score range of rating AA is 5 to 10. The stickiness refers to the situation when credit quality moves to 4 or 11 but the rating still remains of AA.

mechanism of rating and leads to a specific definition of stickiness. The immediate migration assumed in most other studies interprets credit rating as a continuous spectrum with a single boundary between each two adjacent categories. However, the presented model allowing for stickiness allows adjacent rating categories to overlap to a certain extent, similar to the hysteresis effect in real options literature.[2] The overlapping area belongs to both categories and, hence, reflects the credit quality entering next rating range but the deviation not being sufficient to trigger rating migration. Compared to the standard view (which essentially implies zero lump sum cost of rating adjustment), the proposed model allows for adjacent rating categories to partially overlap. Entering the overlapping area does not lead to rating changes and rating migration is only triggered only when the deviation is large enough.

Our first aim is to empirically demonstrate the existence of stickiness in credit ratings, which is the basis of the proposed mechanism. This demonstration is important for a number of reasons. First, the proposed mechanism must fit the observed phenomenon in market, which is the high persistence and slow respondence. Second, whether or not stickiness exists determines whether it is appropriate to use the standard estimation models such as linear regression and multinomial logit/probit. Extant empirical studies also question the reliability of credit rating as a contemporaneous measure of credit quality (Altman, 1998; Becker and Milbourn, 2008; Hull et al., 2004; Norden and Weber, 2004). Our findings support the existence of rating stickiness and indicate that allowing for stickiness in the estimation procedure partially absorbs the decreasing trend of yearly intercepts as well as that the coefficients of explanatory variables interacting with migration dummies are statistically significant. We also decompose rating stickiness into that resulting from the delay in time (t-stickiness) from the one refecting inertia in the state space (z-stickiness). The t-stickiness measures agencies' delay in implementing rating migrations, that is, the length of period between credit quality moving outside of the nominal range and the actual migration. In

[2]For example, suppose ratings A and B are next to each other. The standard view assumes their nominal Z-score ranges are 10 to 6, and 5 to 1, respectively. However, the proposed model that allows for the stickiness allows for the ranges of ratings A and B to be, e.g., 10 to 5 and 6 to 1, respectively.

contrast, z-stickiness measures agencies' tolerance of deviations of credit quality from the target range. We find that the effect of z-stickiness dominates in most specifications.

The presented results directly address the debate regarding the relationship between potential credit quality deterioration and rating standards. The proposed framework allows for isolating the effect of rating standards so we are able to uncover the variation of credit quality during the sample period. Although the observed rating levels deteriorate during 1985 to 2015, the credit quality actually improves with an increment of 0.573 measured using the Z-score. Moreover, the findings support the view that more stringent standards contribute to the credit rating deterioration reported before (Blume et al., 1998; Jorion et al., 2009; Alp, 2013). The novel feature of the proposed model is that it shows the asymmetry between upgrade and downgrade decisions. In particular, the rating standards for upgrades strengthen over time whereas they remain largely unchanged for downgrades. Moreover, the frequency and magnitude of upgrades fall significantly, which is not the case for downgrades. This asymmetry contributes to the observed deterioration of ratings in aggregate.

Finally, our study contributes to the extant literature on determinants of credit ratings (Horrigan, 1966; Pogue and Soldofsky, 1969; Kamstra et al., 2001). Although rating agencies tend to apply complex and proprietary approaches (such as grid method) to assign ratings, academic interest in independently estimating determinants of ratings has been consistently high.[3] In this paper, we offer a novel perspective on the dynamics of the rating process.

The remainder of the paper is organized as follows. Section 2 summarizes existing literature. Section 3 explains the model and the estimation. Section 4 describes the data. Section 5 presents main findings, whereas section 6 concludes.

[3]Moody's (2018) illustrates the application of grid method in rating assignment. They firstly identify grid factors, which are dimensions agencies consider to evaluate credit quality (e.g. financial policy, leverage and coverage), and estimate a quality score for each factor. Then, calculate the weighted average value, denoted by x, as the indicator of the overall credit quality. Finally, transforms the numeric quality score x into ratings (e.g. assign Aaa if $x < 1.5$).

2 Literature

Empirical work has provided ample evidence on the persistence of credit ratings (Kealhofer et al., 1998; Carey and Hrycay, 2001; Löffler, 2004). The persistence may correspond to the situation where rating updates fall behind the changes in credit quality and, hence, detrimentally affect the accuracy of ratings. Ellis (1997), using survey data of 200 CFOs and 400 institutional fixed income investors, reports that 70% of the interviewees believe that ratings should improve to reflect previous positive changes in credit quality. Kealhofer et al. (1998) develop a synthetic rating measuring contemporaneous credit quality based on Merton (1974). Compared with this synthetic rating, the actual agency rating demonstrates a much higher probability of staying in the same rating category (around 90%) as in the preceding period. Carey and Hrycay (2001) divide issuers into rating categories following banks' internal rating methodology. The rating grade reflects borrowers' contemporaneous credit quality as their method is based on frequent information updates normally required for monitoring and risk management activities of the banks. Calculations based on the proposed contemporaneous ratings indicate that only 40%-50% observations have their rating unchanged over one-year horizon, while for agency ratings this proportion ranges from 80% to 90%, indicating that agency ratings are on average twice as persistent as the proposed contemporaneous ("current-condition") ratings.

The high persistence of credit ratings implies a certain degree of tolerance for the deviation of actual quality from the nominal rating range and empirical evidence demonstrates the correlation between the magnitude of this deviation and the probability of rating migrations (Altman and Rijken, 2004; Mora, 2006; Posch, 2011). These empirical results indicate that rating migrations are triggered when the borrowers' actual credit quality exceeds the nominal quality of their current ratings by 1.25 notches. Mora (2006) provides more direct evidence about the rating migration mechanism and concludes that ratings change when the divergence between actual quality and the assigned rating is sufficiently large. Posch (2011) further measures the amount of tolerance (inertia) by extending the model with frictions to allow for non-constant thresholds and shows that default probability has to change by at least two notches before rating agencies react.[4]

The stickiness framework proposed in this paper is also theoretically supported by the structure of agency rating market. Cheng and Neamtiu (2009) emphasize the lack of timeliness and increasing regulatory pressure in agency ratings, which implies that the accuracy of the ratings may deserve further investigation. The extant literature provides some insights into the origins of stickiness. In general, agencies have incentives to make credit rating sticky considering their profitability and reputation. Löffler (2005) documents that agencies attempt to avoid rating reversal after a migration, which contributes to rating stickiness.[5] Moreover, Jeon and Lovo (2013) introduce the notion of 'reputation build-up', which postulates that frequent rating adjustments harm the profitability of the agencies by weakening their reputation to potential issuers. More precisely, Bolton et al. (2012) discuss the "rating shopping" phenomenon, according to which agencies attract business by enhancing the stability as issuers can shop in the market for the best ratings they can receive. The issuer-paid pattern indeed results in extra cautiousness for agencies in terms of rating updates, which is potentially detrimental to rating accuracy (Xia, 2014). Xia (2014) find that introducing investor-paid rating agencies (e.g., Egan-Jones Rating Company) can improve the accuracy and timeliness relative the traditional issuer-paid ratings model.

Finally, a number of further studies investigate the changes in rating standards (Blume et al., 1998; Jorion et al., 2009; Alp, 2013; Baghai et al., 2014). Blume et al. (1998) attempts to explain the reported decline in credit quality using accounting ratios and market information. They find that the credit rating deterioration is not fully explained by changes in credit

[4]Default probability is the indicator of credit quality in Posch (2011), and it is the basis of rating assignment. For example, if the default probability p is within the first notch ($0 < p \leqslant 10\%$), this observation will be assigned the best rating AAA.

[5]It is worth clarifying that persistence is not equivalent to stickiness as the former refers to the observed fact that rating does not change over a period of time regardless of the (lack of) changes of fundamentals. Stickiness refers to the situation in which the rating does not change following a shift in fundamentals.

quality and that it is at least partly caused by the increasingly stringent rating standards. Alp (2013) quantifies this effect, showing that the tightened standard leads to on average 1.5-notch drop in ratings between 2002 and 2007. Baghai et al. (2014) find the drop to be 3 notches between 1985 to 2009. Alp (2013) finds that the tightening standards pattern applies to investment grade bonds but the speculative grade bonds may be subject to a loosening standard. This observation is consistent with Jorion et al. (2009), who study the mechanism behind the tightening of standards. They find that the perceived tightening of rating standards for the investment grade issuers may reflect changes in accounting quality. After controlling for those changes, the tightening pattern in rating standards disappears. Baghai et al. (2014) study the impact of tightening rating standards on firms' behavior and show that firms affected more by the tightening standards (measured by the difference between the actual rating and the predicted rating assuming constant standards) issue less debt, hold more cash, experience lower growth and are less likely to access debt markets.

3 The model

Rating agencies need to strike a delicate balance when choosing migration frequency since sudden and frequent updates harm reputation among issuers (Jeon and Lovo, 2013), while significantly delayed migrations lead to the criticism from investors (Ellis, 1997). The optimal policy for agencies is to wait until the deviation of credit quality from the target for a given rating becomes sufficiently large and only then update the rating. This policy results in the stickiness in credit ratings. The model proposed in this paper extends the standard way of modelling ratings in two aspects. First, it relaxes the assumption that a single boundary separates two adjacent rating categories and is the same for migrations in either direction. The model allows for an overlap between two neighboring ratings, with the overlapping area reflecting agencies' tolerance of credit quality deviation. Obviously, the existence of such areas represents the stickiness in ratings (as in models with hysteresis). Figure 1 demonstrates the mechanism of rating migrations under the this framework. There are three rating categories, namely A, B, and C, in which A indicates the best credit quality and C indicates the worst. The nominal quality range of rating A is the area from line $A2$ and above; the quality range for rating B is the area between lines $B1$ and $B2$; and the range for rating C is the area below line $C1$. The area between lines $B1$ and $A2$ is the overlapping range of ratings A and B. It reflects the stickiness by allowing credit quality deviation from either upper (rating A) or lower (rating B) position without invoking migrations.

[Please insert Figure 1 about here.]

Second, the standard approach typically assumes a perfect match between credit quality and rating. Under this assumption, agencies immediately eliminate any deviation in credit quality by updating ratings. Empirically, however, such a setting is inconsistent with the documented persistence and slow respondence of ratings. The proposed model deals with this challenge by allowing for the rating migration mechanism that is driven by both the current rating as well as the deviation from the target value (range). Migrations are triggered by credit quality hitting either the upper or the lower boundary but, unlike in standard models, the upper threshold of rating B is not necessary the lower threshold of rating A (suppose ratings A and B are adjacent, and A indicates better quality). As in Figure 1, the green line presents the path of the firm's credit quality movements and demonstrates the mechanism of rating migrations. A migration is triggered by credit quality crossing the boundaries of its nominal range. For instance, points $q2$, $q3$, and $q4$ indicate different credit quality levels, depicting the process of a downgrade migration. Credit quality enters the range of rating C at $q1$, but this movement will not cause downgrade since the quality has not reached the lower threshold of rating B (line $B2$). When credit quality drops from $q2$ to $q3$, it moves outside of rating B's range and this magnitude of deviation exceeds the agency's tolerance. A downgrade decision is made at $q3$ but implemented at $q4$ to fit the slow-respondence feature.

Inversely, points $q9$, $q10$, and $q11$ describe a rating upgrade process. Credit quality crosses $B1$, the upper boundary of rating B, to $q10$, and the rating upgrade is observed next period at $q11$.

In general, the proposed model includes three groups of observations, namely those exceeding the upper threshold, the lower threshold, and the observations which have their rating correctly matched to the credit quality. The estimation of the model requires the accurate identification of observation in each of the three groups. The identification methodology concentrates on rating migrations. Rating upgrade (downgrade) is triggered by credit quality breaching the upper (lower) threshold of its previous rating range. Hence, the observation before an upgrade (downgrade) is informative for the upper (lower) threshold. For example, there is an upgrade in Figure 1 depicted by points $q9$, $q10$, and $q11$. The upgrade happens at $q11$, which suggests that the credit quality at $q10$ exceeds the upper threshold (i.e. $Z_{q10} \geqslant B1$).[6] Hence, observation $q10$ falls into the group of upper threshold. Similarly, $q3$ can be an example of lower threshold identification through downgrade (i.e $Z_{q10} \leqslant B2$).

In what follows, we use the Z-score to represent the credit quality (Z_{it}):

$$Z_{it} = \beta X_{it} + \varepsilon_{it}, \tag{1}$$

where β is the coefficient set, matrix X_{it} contains firm characteristics as covariates, and ε_{it} represents the normally-distributed error term. The Z-score serves as a linking function which transforms firm characteristics to ratings through categorization. The data contains five rating levels with rating 5 the best quality and rating 1 the worst.[7] R^*_{it} in equation (2) represents observations that are not thresholds. In other words, the credit quality and rating levels are correctly matched for these observations. Each rating level R_i is quantified by the

[6]The upgrade action happens one period after the breach because of the time agencies need to collect and interpret information (exemplified by t-stickiness). This setting fits the slow-respondence feature of credit rating.

[7]Details are given in Section 4.

nominal Z-score range with upper and lower boundaries U_i and L_i, respectively.[8] Hence,

$$R_{it}^{*} = \begin{cases} 5 & if\ Z_{it} \in [L_5, \infty) \\ 4 & if\ Z_{it} \in [L_4, U_4] \\ 3 & if\ Z_{it} \in [L_3, U_3] \\ 2 & if\ Z_{it} \in [L_2, U_2] \\ 1 & if\ Z_{it} \in (-\infty, U_1] \end{cases} \tag{2}$$

the credit quality, denoted by Z_{it} is within the nominal range for each rating level. This part is exactly the same as the ordered probit model because of the same underlying assumption of an accurate match between credit quality and rating.

$$R_{it}^{u} = \begin{cases} 4 & if\ Z_{it} \geqslant U_4 \\ 3 & if\ Z_{it} \geqslant U_3 \\ 2 & if\ Z_{it} \geqslant U_2 \\ 1 & if\ Z_{it} \geqslant U_1 \end{cases} \tag{3}$$

R_{it}^{u} in equation (3) is the rating level of an observation in upper threshold group. As stated before, credit quality and ratings are not matched in this group because quality exceeds upper threshold ($Z_{it} \geqslant U_{R_{it}}$) and is the trigger of upgrade.

[8] Obviously, rating level 5 indicates the best quality and hence no further upgrade available. Its upper threshold is infinity. On the side of the rating spectrum, the worst rating level 1 has minus infinity as the lower threshold.

$$R_{it}^{l} = \begin{cases} 5 & if\, Z_{it} \leqslant L_5 \\ 4 & if\, Z_{it} \leqslant L_4 \\ 3 & if\, Z_{it} \leqslant L_3 \\ 2 & if\, Z_{it} \leqslant L_2 \end{cases} \tag{4}$$

Lastly, R_{it}^{l} in equation (4) means the rating for lower threshold observations, and the trigger of downgrade requires $Z_{it} \leqslant L_{R_{it}}$. We estimate the three parts jointly through the maximum likelihood method (see the Appendix).

4 Data

The sample contains 1,488 US firms from 1985 to 2014, which leads to 20,557 observations overall. The $S\&P$ ratings are obtained from the Compustat Ratings File.[9] This sample excludes observations with negative or zero total assets, financial firms (SIC code 6000-6999), and quasi-governmental enterprises (SIC 9000 and above). Missing explanatory values reduce the sample to 20,557 firm-year observations from 1,488 unique firms for the full sample analysis. We merge ratings based on the original S&P categories: our rating A includes S&P ratings from AAA to AA; rating B includes S&P ratings from $AA-$ to A; rating C includes S&P ratings from $A-$ to BBB; rating D includes S&P ratings from $BBB-$ to $BB-$; and rating E includes S&P ratings $CCC+$ and below.[10] Therefore, this sample includes the full spectrum of S&P rating categories with ratings from A to C corresponding to the investment grade and the remaining being speculative. Table 1 presents the distribution of ratings in the sample.

[9] While our choice of the rating firm is a function of data availability, ratings of major agencies are generally seen as equally informative (Livingston et al., 2010).

[10] The merging strategy used differs from that in other studies (Blume et al., 1998; Jorion et al., 2009; Alp, 2013; Baghai et al., 2014) in terms of the width of each category and the coverage of ratings. Our categorization strategy identifies upper and lower thresholds based on rating migrations. We merge credit ratings to ensure that there are sufficient number of rating migrations for each category in each year.

[Please insert Table 1 about here.]

There are 849 and 1,215 observations that belong to the upper threshold and lower threshold group, respectively. Table 2 provides details about the threshold categories.

[Please insert Table 2 about here.]

The selection of explanatory variables follows existing literature (Blume et al., 1998; Jorion et al., 2009; Alp, 2013). $Intcov$ measures interest coverage calculated by $ebitda$ divided by interest expense ($ebitda/xint$). Variables $k1$ to $k4$ indicate different levels of $Intcov$. This is done to capture the non-linearity of interest coverage effect of a credit rating, following Blume et al. (1998). $k1$ indicates $Intcov$ range from 0 to 5 (e.g. an observation with $Intcov = 3$ will have $k1 = 3$, $k2 = 0$, $k3 = 0$, $k4 = 0$). $k2$ indicates $Intcov$ range from 5 to 10 (e.g. an observation with $Intcov = 7$ will have $k1 = 5$, $k2 = 2$, $k3 = 0$, $k4 = 0$). $k3$ indicates a range from 10 to 20, and $k4$ indicates $Intcov$ above 20. Vol is the volatility of profit, calculated as the standard deviation of the last 5 years of $ebitda/sale$. $Tlev$ refers to total leverage measured by total debt divided by total asset ($\frac{dlc+dltt}{at}$). $Rent$ is the rent expense divided by total asset ($xrent/at$). Tan refers tangibility, measured by property, plant and equipment divided by total asset (ppe/at). Dni is a dummy variable which equals to one when net income is negative. $Ddiv$ is a dummy variable which equals to one when a firm pays dividend in a given period. Rd is the research and development expense divided by total asset (xrd/at). Rd is set to zero when the expense is missing. Mtb is the market to book ratio measured by total asset minus book value of equity plus market value of equity and then divided by book value of assets ($(at - bv + mv)/at$). The market value of equity is the product of year-end price and number of shares outstanding ($prcc_f * csho$). Book value of equity is shareholders' equity minus preferred stock liquidating value plus deferred taxes and investment tax credit ($seq - pstkl + txditc$). The deferred tax credit $txditc$ is set to zero if missing. Equity (seq) will be replaced by either common equity plus preferred stock at par value ($ceq + pstk$) or total asset minus total liability ($at - lt$) if missing. Preferred stock liquidating value $prstkl$ will be replaced by either redemption value $pstkrv$ or par value $pstk$ if missing. Firm size ($Size$) is the logarithm of total asset. *Beta* and *Rmse* measure systematic risk and idiosyncratic risk, respectively.[11] They are estimated in market model regressions of a firm's daily stock returns on the CRSP value-weighted index return. The regressions are adjusted for nonsynchronous trading effects using the Dimson (1979) procedure with one leading and one lagging value of the market return. One firm-year observations of *Beta* and *Rmse* are computed from one regression using firm-specific daily stock returns from one calendar year. All continuous variables are winsorized at top and bottom 1 percentile. Table 3 reports the descriptive statistics of these variables.

[Please insert Table 3 about here.]

5 Empirical results

This section presents the paper's main empirical findings. Firstly, we test for the existence of stickiness and explore its effect on credit ratings. Subsequently, we demonstrate that credit quality actually improves over time although the rating levels deteriorate. Lastly, we document the asymmetry in rating standard stringency, which contributes to explaining the perceived deterioration of credit ratings.

5.1 Stickiness of credit ratings

We begin by estimating both ordered probit and the proposed extended model (adjusted ordered probit) given by equations (2), (3), and (4). The comparison of the results of the

[11] *Beta* is the coefficient in the market model estimation, and *Rmse* is the standard error.

two models demonstrates the difference in coefficients obtained with and without considering stickiness and, hence, explores the effect of stickiness. Table 4 presents the main estimation results. The ordered probit regression ignores stickiness, as in Models 1 and 5, and provides the basis of comparison in the analysis. The adjusted ordered probit regression, as in Models 2 and 6, takes into account both time delay (t-stickiness) and credit quality tolerance (z-stickiness). Moreover, we use lagged explanatory variables in the ordered probit regression, as in Models 3 and 7, to evaluate the effect of t-stickiness. We isolate z-stickiness in the adjusted ordered probit model by using contemporaneous values of explanatory variables, as in Models 4 and 8. The Fama-MacBeth models takes the decades average (85-94, 95-04, and 05-14) of coefficients for adjusted ordered regressions controlling for all stickiness based on every two years of data. Values in parentheses present the standard error for each estimation. Standard errors for adjusted ordered models are calculated using bootstrapping method, and those for Fama-MacBeth models use Delta method. Panel A presents estimation for explanatory variables, Panel B provides the estimation of cut-offs, and Panel C lists year dummy intercepts.

Model 1 in Table 4 summarizes the estimation of ordered probit, which aims to replicate the key findings in existing literature without considering stickiness. The estimated effects of covariates are consistent with prior studies (Blume et al., 1998; Jorion et al., 2009; Alp, 2013). Consistent with Blume et al. (1998), firms with better ability to pay back debt receive better ratings, as indicated by positive signs of $Intcov$, Tan, $Ddiv$, and $Size$. Interest coverage is a direct measure of the ability to service credit and tangible assets capture collateral, which both reduce debtholders' risk. The ability to pay dividends indicates a strong financial position and profitability. Moreover, potential growth opportunities also improve credit ratings consistent with Alp (2013). (The letter is captured by the positive coefficients of Mtb and Rd.) Lastly, risk factors lead to a more conservative assessment from rating agencies. Cash flow uncertainty, denoted by Vol, and systematic risk level, represented by $Beta$ generally reduce credit quality and drive down the ratings. Credit risk measured by total leverage $Tlev$ exhibits a similar effect.

[Please insert Table 4 about here.]

Model 2 in Table 4 allows for stickiness and is estimated using the adjusted ordered probit methodology described by equations (2), (3), and (4). As explained before, this estimation method explicitly allows for sticky ratings both in the time-series dimension (t-stickiness) as well as across the state space (z-stickiness). Allowing for stickiness does not result in substantive qualitative changes in regression coefficients relative to Model 1. Subsequently, we decompose the overall rating stickiness into t-stickiness and z-stickiness in Models 3 and 4, respectively. The t-stickiness assumes zero tolerance of agencies regarding credit quality deviation from its nominal range but assumes a delay in taking action by one period. Under this assumption, current ratings match therefore lagged firm features and standard ordered probit model is appropriate with all covariates lagged by one period as in Model 3. Model 4 allows for z-stickiness only and therefore assumes agencies allow some deviation in credit quality but react immediately when the deviation exceeds a critical threshold.[12] The coefficients of control variables in Model 3 and Model 4 do not substantially deviate from those in Model 2.

To address concerns expressed in Blume et al. (1998) regarding the non-linear relationship between interest coverage and ratings, we follow their method and decompose the $Intcov$ into four components according to its magnitude as discussed in subsection 4. We replicate Models 1 to 4 with $Intcov$ replaced by the piece-wise coverage variables in Models 5 to 8. This decomposition demonstrates that the correlation between components of interest coverage and ratings weakens with the magnitude of interest coverage. As in Model 5, one unit of increment in the low range of interest coverage, the $k1$, brings 0.306 extra credit

[12] For example, in the downgrade situation depicted by points $q2$, $q3$, and $q4$ in Figure 1, credit quality breaches the threshold at $q3$. Under z-stickiness, rating changes at $q3$ to reflect the immediate reaction. In comparison, under t- and z-stickiness combined, the rating update happens at $q4$ to reflect the slow-respondence.

quality measured by Z-score, but this effect shrinks to 0.003 as the range shifts to $k4$.[13] This effect is consistent with prior studies (Blume et al., 1998; Alp, 2013).

The main novel set of results based on the adjusted ordered probit model is on the evolution of the stringency of rating standards over time. Prior empirical studies (Blume et al., 1998; Alp, 2013) interpret the declining year dummy intercepts (as in Panel B of Table 4) as indicators of rating standard becoming more strict. Those studies find that the intercepts indeed move downward and, hence, conclude that ratings become more conservative. (As the decreasing year dummy intercepts implies that a firm whose all characteristics remain unchanged receives a lower rating by just stepping into the next year.[14]) Figure 2 plots the year intercepts of Models 5 to 8 (Panel B, Table 4). The downward-sloping intercepts of the (standard) ordered probit model confirm the pattern in Blume et al. (1998). However, this trend all but disappears once rating stickiness is controlled for the adjusted model. In addition, we also present results when t- and z-stickiness are separately allowed for and obtain a trend in intercepts that lies between those corresponding to models with no and full (i.e., both z- and t-) stickiness.

[Please insert Figure 2 about here.]

The disappearing downward pattern in year intercepts suggests that year dummy variables may contain some incremental information to that included in rating standards. For example, Du (2003) challenges the conclusion of tightening rating standards in Blume et al. (1998) and provides an alternative explanation. Du (2003) interprets the decreasing year dummy coefficients as the outcome of a situation in which new bonds are mainly issued by low quality firms. From the perspective of stickiness, changes in year dummies reflect both the effect of the changing standard as well as any delay inherent in the rating mechanism.

[13] As already discussed, $k1$ measures the *Intcov* range from 0 to 5, and $k4$ measures the range above 20.

[14] A lower index refers to a worse rating. This paper therefore uses the same convention as in (Blume et al., 1998; Jorion et al., 2009; Alp, 2013), which implies that better rating corresponds to a higher value of the rating index.

Estimation based the ordered probit model ignores agencies' flexibility in timing the rating migrations. Compared to the presented adjusted model that allows for stickiness, the estimation of ordered probit model effectively forces the threshold observations, either upper or lower, to be within their current rating, while the credit quality associated with these observations actually moved into the range of an adjacent rating. For example, dot $q10$ in Figure 1 is an upper threshold observation whose credit quality enters range of rating A but its current rating is still B since migration happens in the next period. Static estimation will force $q10$ to fit the range of rating B, which leads to the negative year intercepts. In a similar way, the fitting of the ignored lower threshold observations would lead to positive year intercepts. The downward trend implies asymmetric effect of the upper and lower threshold groups and the influence of the upper threshold observations on estimation results seems greater.

We further demonstrate the impact of stickiness through rating migrations (Table 5). The first three columns present the true migration observations from perspectives of an upgrade, no migration and a downgrade. The next three columns summarize the predicted migrations without considering stickiness, while the last three columns contain analogous information when stickiness is allowed for. The inferred migrations for predictions without and with stickiness are based on the same credit quality, and the only difference is the model specification.[15] As shown in the columns that correspond to cases without stickiness, this model tends to overestimate the frequency of both upgrades and downgrades. The total number of migrations predicted is almost three times the number of the actual rating adjustments. In comparison, predictions based on the framework with stickiness are closer to the actual values. Moreover, using the framework with stickiness may uncover the asymmetry between the behavior of the upper and the lower threshold since upgrades are less frequent during the sample period.

[15] Credit quality is the predicted Z-score based on Model 6 in Table 4.

[Please insert Table 5 about here.]

5.2 Credit quality

The effect of stickiness considered in the presented model absorbs effects of the rating migration process and leaves the coefficients directly determined by credit quality. To explore in more detail the latter, we further decompose the evolution of credit quality during the sample period and find that the credit quality of firms actually improves. Table 6 summarizes the changes in the predicted average level of credit quality between the last six-year period (2009-2014) average level and the first six-year period (1985-1990). The contribution of each variable is the product of increment in period average of the specific explanatory variable and the corresponding regression coefficient. Column 1 reports the contribution of variables to the credit quality variation. Overall, credit quality, measured by Z-score, increases by 0.573, while firm size contributes the most (0.312). Further, we separate the full sample by rating levels and firm size, and find that investment grade rated and large firms experience more credit quality improvement compared to speculative rated and small firms.

[Please insert Table 6 about here.]

Since credit quality improves during this period, tightening rating standards remain the most plausible explanation of the negative trend in average ratings. Unlike previous empirical studies (Blume et al., 1998; Jorion et al., 2009; Alp, 2013), the presented model allows for the quantifying effects of changing standards separately for upgrade and downgrade decisions (by isolating upper and lower thresholds from their average effect). Figure 3 illustrates the upper threshold movements from 1985 to 2014 for every two years. Upper thresholds represent cut-off levels for ratings, for which every crossing by credit quality causes an upgrade to the adjacent rating above. As shown in Figure 3, upper thresholds generally increase for each rating category, which suggests that upgrade migration becomes increasingly more difficult. For example, the threshold $U4$ has experienced a dramatic surge during this period, which implies that upgrade to the best rating category (level 5) necessitates a very significant improvement in credit quality.[16]

[Please insert Figure 3 about here.]

Figure 4 plots the lower cut-off variation in the same format. Lower thresholds exhibit substantial variation around their original levels but no visible trend is present. The latter finding implies no significant tightening of the rating standards for downgrades. Overall, the presented results suggest that the standard stringency increases for upgrade decisions but remains constant for downgrades. This type of asymmetry can therefore be concluded to be a key contributor to the deterioration in credit ratings.

[Please insert Figure 4 about here.]

The documented asymmetry may result from a partial release of information, as suggested in prior empirical studies (Altman and Kao, 1992; Lando and Skødeberg, 2002). Altman and Kao (1992) create a measure defined by the frequency of subsequent migrations in the same direction (e.g., an upgrade followed by a further upgrade) divided by the frequency of subsequent migrations in the opposite directions (an upgrade followed by downgrade). This statistic is larger than 1, which means that rating migrations tend to be followed by migrations in the same direction. The correlation between consecutive downgrades is higher than the one between consecutive upgrades. Lando and Skødeberg (2002) confirm

[16]Threshold $U4$ refers to the upper cut-off for rating level 4. The firm is upgraded to rating level 5 once its credit quality crosses this boundary, and level 5 is the higher rating.

this migration correlation pattern using a semi-parametric regression based on continuous observations. Löffler (2005) interprets this phenomenon by agencies responding partially to a piece of information so that they "dole out the bad news in small doses rather than savaging the bond issuer - who is, after all, their customer - all in one go" (The Economist, 1997). This finding is also consistent with the partial adjustment pattern documented by Altman and Rijken (2004). From that perspective, rating agencies respond to a piece of information in more than one period and, therefore, break down a big rating migration decision into several small migrations. Such a pattern naturally contributes to the stickiness of credit ratings.

Furthermore, we study the determinants of rating migrations to investigate the role of the asymmetry in credit rating deterioration. Table 7 reports the coefficients of migration determinants under the ordered probit model.[17] The dependent variable is rating migrations: it takes the value 1 if downgrade happens at that firm-year observation; 2 if rating remains; and 3 if upgrade happens. The columns 1 and 5 present the estimation results and the remaining columns report marginal effects. Most of the year intercepts in Model 1 demonstrates negative signs, which suggests that the probability of being a downgrade observation increases with time, holding all other control variables the same. This is consistent with the observed asymmetry that upgrades become increasingly difficult. Marginal effects reported in columns 2 to 4 confirm the asymmetry. $ME1$ reports the probability of being a downgrade observation and most of the year dummies exhibit positive contribution. At the same time, year dummy impact in $ME3$ has a negative sign, which means the probability to be an upgrade observation decreases. This pattern does not change after controlling for industry fixed effects as in Model 2. This evidence supports the earlier conclusion that there is asymmetry between upgrade and downgrade decision standards, which helps explain the observed credit rating deterioration.

[17] The adjusted ordered probit model is not appropriate in migration studies since it ignores stickiness embedded in decisions to upgrade or downgrade.

[Please insert Table 7 about here.]

5.3 Robustness analysis

To test the robustness of earlier results, we perform a number of additional tests. To support the earlier result regarding the existence of rating stickiness, we provide complementary evidence from the interaction analysis and lag rating analysis. Table 8 summarizes the interaction analysis, in which the specification follows Model 5 in Table 4 but with added interaction terms. Stickiness implies that agencies update ratings only after credit quality breaches a certain threshold. One way to test this conjecture is to demonstrate that the relationship between covariates and ratings at the breaching point the period right before a rating migration differs from the one in the other periods. In line with expectations, all interactions terms in Panel A are highly significant, suggesting that the threshold observations statistically differ from normal observations. Panel B controls for lagged credit rating in interaction terms and the highly significant effects provide further support.

[Please insert Table 8 about here.]

Furthermore, we argue that past and current locations on credit quality spectrum jointly determine the rating migration behavior and the observed ratings. (In other words, rating in last period has to be a significant determinant in the model when stickiness is present.) Therefore, we use the lagged rating analysis to demonstrate this effect (Table 9). The lagged credit rating variable, denoted by $LagY$, is highly significant in all models, consistent with our expectations.

[Please insert Table 9 about here.]

To investigate whether the previous results might have been by shifts in relationships of firm characteristics rather than rating stickiness, we apply Fama-MacBeth estimation for every ten years data as in Models 9 to 11 in Table 4. We firstly estimate regressions using adjusted ordered probit on every two-year period data, controlling for both t- and z-stickiness. Then, we take the average of regression coefficients for every decade (85-94, 95-04, and 05-14) to calculate the Fama-MacBeth coefficients. The comparison between these three models demonstrates that the relations between credit rating and firm characteristics do not shift during our sample period.

[Please insert Table 10 here.]

Finally, we re-run our previous tests for investment grade and speculative grade subsamples. Table 10 replicates Table 4 within subsamples of investment grade observations and speculative grade observations. "Investment grade" refers to BBB- or above under S&P credit ratings. The subsample includes ratings A, B, and C in our merged ratings, which in total contains 10,480 observations. In comparison, observations with S&P ratings below BBB- fall into the speculative grade, which covers ratings D and E in the merged category and in total include 10,077 observations. Figures 5 and 6 plot the year intercepts for the investment grade and speculative grade subsamples, respectively. Our main results in Table 4 indicate that the downward trend of year intercepts in the ordered probit model disappears after controlling for stickiness. This pattern also exists in the investment grade subsample but is less clear for speculative grade observations.

[Please insert Figures 5 and 6 about here.]

6 Conclusions

The presented model of credit ratings that explicitly allows for ratings stickiness facilitates the understanding of the observed downward trend in ratings. By explicitly recognizing the possibility of inertia in the rating mechanism, which is essentially equivalent to the presence of lumpy costs, the model enables us to separate the effects of the evolution of the average credit quality from potential changes in rating standards. The paper's findings shed therefore light on the debate whether the downward trend in credit ratings is due to the deteriorating average credit quality or the tightening of rating standards. The results confirm the existence of stickiness in credit ratings and demonstrate its significant impact on the dynamics of ratings. After allowing for stickiness, we find that the downward trend of year dummy intercepts all but disappears and firms' credit quality actually improves during the sample period. We also document asymmetry in rating migrations. Upgrades become increasingly less frequent while downgrade frequency appears to remain unchanged. This pattern offsets the documented improvement in credit quality and leads to the perceived (i.e., inferred from ratings) deterioration of credit quality. Finally, our study contributes to the literature on the determinants of credit ratings and, more generally, to research on the effects of lumpy adjustment costs on decision making and economic inference.

A Appendix

A.1 Likelihood function

The estimation approach for adjusted ordered probit differs from that for the standard ordered probit model in two aspects: 1) the categorization of observations and 2) separation of cut-offs, which are reflected in the likelihood function.

Case 1: Upper Threshold When a firm experiences rating upgrade in the next period, its observation in current period belongs to the upper threshold group. Its credit quality now exceeds the upper cut-off of its current rating category, which leads to a rating upgrade in the next period. Hence, observations in this group contain the information about upper cut-offs. More specifically, upper cut-offs are the lowest possible level of the credit quality exhibited by observations in this group. Hence, the probability of this situation occurring is

$$p_1 = 1 - \Phi(u_r - x \cdot \beta, 0, 1). \tag{A.1}$$

where $\Phi(\cdot, 0, 1)$ denotes the standard normal CDF, x refers to the explanatory variable set, β is the coefficient vector to be estimated, r refers to the credit rating category, and u_r (l_r) is the upper (lower) cut-off of rating r.

Case 2: Ordinary Observation When a firm's credit rating remains the same in the next period, its observation in current period is in the ordinary group. The credit quality of observations in this group matches their actual ratings. The credit quality is therefore correctly bounded by the upper and lower cut-offs of their observed credit ratings in this group. Hence, the probability of this situation occurring is

$$p_2 = \Phi(u_r - x \cdot \beta, 0, 1) - \Phi(l_r - x \cdot \beta, 0, 1). \tag{A.2}$$

Case 3: Lower Threshold The lower cut-off group contains observations which are downgraded in the next period. These observations are therefore informative about lower cut-offs of credit ratings as credit quality is below the lower cut-offs of their rating category for observations in this group. The relevant probability is therefore

$$p_3 = \Phi(l_r - x \cdot \beta, 0, 1). \tag{A.3}$$

Overall, the loglikelihood function is

$$loglike = log(p_1) + log(p_2) + log(p_3). \tag{A.4}$$

The calculation of standard errors is based on the following Hessian matrix:

$$H = \begin{bmatrix} \frac{\partial^2 loglike}{\partial\beta\partial\beta} & \frac{\partial^2 loglike}{\partial\beta\partial u} & \frac{\partial^2 loglike}{\partial\beta\partial l} \\ \frac{\partial^2 loglike}{\partial u\partial\beta} & \frac{\partial^2 loglike}{\partial u\partial u} & \frac{\partial^2 loglike}{\partial u\partial l} \\ \frac{\partial^2 loglike}{\partial l\partial\beta} & \frac{\partial^2 loglike}{\partial l\partial u} & \frac{\partial^2 loglike}{\partial l\partial l} \end{bmatrix}.$$

The covariance matrix is hence $Var(\hat{\theta_{ML}}) = [-H(\hat{\theta_{ML}})]^{-1}$, where θ_{ML} is the vector that contains all variables to be estimated and includes β, u_r and l_r.

References

Alp, A. (2013), 'Structural shifts in credit rating standards', *Journal of Finance* **68**(6), 2435–2470.

Altman, E. I. (1998), 'The importance and subtlety of credit rating migration', *Journal of Banking and Finance* **22**(10), 1231–1247.

Altman, E. I. and Kao, D. L. (1992), 'Rating drift in high-yield bonds', *Journal of Fixed Income* **1**(4), 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.

Baghai, R. P., Servaes, H. and Tamayo, A. (2014), 'Have rating agencies become more conservative? implications for capital structure and debt pricing', *Journal of Finance* **69**(5), 1961–2005.

Becker, B. and Milbourn, T. (2008), *Reputation and competition: Evidence from the credit rating industry*, Harvard Business School.

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.

Bolton, P., Freixas, X. and Shapiro, J. (2012), 'The credit ratings game', *Journal of Finance* **67**(1), 85–111.

Carey, M. and Hrycay, M. (2001), 'Parameterizing credit risk models with rating data', *Journal of Banking and Finance* **25**(1), 197–270.

Cheng, M. and Neamtiu, M. (2009), 'An empirical analysis of changes in credit rating properties: Timeliness, accuracy and volatility', *Journal of Accounting & Economics* **47**(12), 108–130.

Dimson, E. (1979), 'Risk measurement when shares are subject to infrequent trading', *Journal of Financial Economics* **7**(2), 197–226.

Du, Y. (2003), 'Predicting credit rating and credit rating changes: A new approach', *Queens School of Business, Queens University* pp. 241–252.

Ellis, D. (1997), 'Different sides of the same story: Investors? And issuers? Views of rating agencies', *Working paper* .

Horrigan, J. O. (1966), 'The determination of long-term credit standing with financial ratios', *Journal of Accounting Research* **4**, 44–62.

Hull, J., Predescu, M. and White, A. (2004), 'The relationship between credit default swap spreads, bond yields, and credit rating announcements', *Journal of Banking and Finance* **28**(11), 2789–2811.

Jeon, D.-S. and Lovo, S. (2013), 'Credit rating industry: A helicopter tour of stylized facts and recent theories', *International Journal of Industrial Organization* **31**(5), 643–651.

Jorion, P., Shi, C. and Zhang, S. (2009), 'Tightening credit standards: The role of accounting quality', *Review of Accounting Studies* **14**(1), 123–160.

Kamstra, M., Kennedy, P. and Suan, T.-K. (2001), 'Combining bond rating forecasts using logit', *Financial Review* **36**(2), 75–96.

Kealhofer, S., Kwok, S. and Weng, W. (1998), 'Uses and abuses of bond default rates', *KMV Corporation* .

Kolaric, S., Kiesel, F. and Ongena, S. (2021), 'Market discipline through credit ratings and too-big-to-fail in banking', *Journal of Money, Credit and Banking* **53**(2-3), 367–400.

Lando, D. and Skødeberg, T. M. (2002), 'Analyzing rating transitions and rating drift with continuous observations', *Journal of Banking and Finance* **26**(2), 423–444.

Livingston, M., Wei, J. and Zhou, L. (2010), 'Moody's and S&P ratings: Are they equivalent? conservative ratings and split rated bond yields', *Journal of Money, Credit and Banking* **42**(7), 1267–1293.

Löffler, G. (2004), 'An anatomy of rating through the cycle', *Journal of Banking and Finance* **28**(3), 695–720.

Löffler, G. (2005), 'Avoiding the rating bounce: Why rating agencies are slow to react to new information', *Journal of Economic Behavior and Organization* **56**(3), 365–381.

Merton, R. C. (1974), 'On the pricing of corporate debt: The risk structure of interest rates*', *Journal of Finance* **29**(2), 449–470.

Moody's (2018), 'Packaging manufacturers: Metal, glass, and plastic containers'. [Online; posted 4-May-2018].
URL: *https://www.moodys.com/researchdocumentcontentpage.aspx?docid=PBC_1120393*

Mora, N. (2006), 'Sovereign credit ratings: Guilty beyond reasonable doubt?', *Journal of Banking & Finance* **30**(7), 2041–2062.

Norden, L. and Weber, M. (2004), 'Informational efficiency of credit default swap and stock markets: The impact of credit rating announcements', *Journal of Banking and Finance* **28**(11), 2813–2843.

Pogue, T. F. and Soldofsky, R. M. (1969), 'What's in a bond rating', *Journal of Financial and Quantitative Analysis* **4**(2), 201–228.

Posch, P. N. (2011), 'Time to change. rating changes and policy implications', *Journal of Economic Behavior and Organization* **80**(3), 641–656.

The Economist (1997), 'Risks beyond measure', *The Economist* **December 13**, 70–71.

Xia, H. (2014), 'Can investor-paid credit rating agencies improve the information quality of issuer-paid rating agencies?', *Journal of Financial Economics* **111**(2), 450–468.

Table 1: Number of Companies by Year and S&P Rating Category

Year	A	B	C	D	E	Total	A	B	C	D	E
1985	75	164	142	154	159	694	10.8%	23.6%	20.5%	22.2%	22.9%
1986	73	154	159	194	247	827	8.8%	18.6%	19.2%	23.5%	29.9%
1987	69	137	170	230	284	890	7.8%	15.4%	19.1%	25.8%	31.9%
1988	66	135	170	200	242	813	8.1%	16.6%	20.9%	24.6%	29.8%
1989	62	135	173	187	208	765	8.1%	17.6%	22.6%	24.4%	27.2%
1990	57	137	164	177	154	689	8.3%	19.9%	23.8%	25.7%	22.4%
1991	58	149	165	161	116	649	8.9%	23.0%	25.4%	24.8%	17.9%
1992	52	152	182	196	121	703	7.4%	21.6%	25.9%	27.9%	17.2%
1993	50	159	190	222	123	744	6.7%	21.4%	25.5%	29.8%	16.5%
1994	43	146	191	230	113	723	5.9%	20.2%	26.4%	31.8%	15.6%
1995	41	158	182	216	113	710	5.8%	22.3%	25.6%	30.4%	15.9%
1996	41	152	194	234	115	736	5.6%	20.7%	26.4%	31.8%	15.6%
1997	41	147	221	238	103	750	5.5%	19.6%	29.5%	31.7%	13.7%
1998	34	160	224	265	110	793	4.3%	20.2%	28.2%	33.4%	13.9%
1999	29	143	222	260	94	748	3.9%	19.1%	29.7%	34.8%	12.6%
2000	25	130	214	233	107	709	3.5%	18.3%	30.2%	32.9%	15.1%
2001	20	119	207	235	104	685	2.9%	17.4%	30.2%	34.3%	15.2%
2002	24	113	211	256	98	702	3.4%	16.1%	30.1%	36.5%	14.0%
2003	21	108	217	253	92	691	3.0%	15.6%	31.4%	36.6%	13.3%
2004	21	101	217	259	103	701	3.0%	14.4%	31.0%	36.9%	14.7%
2005	25	94	229	231	99	678	3.7%	13.9%	33.8%	34.1%	14.6%
2006	25	89	217	239	92	662	3.8%	13.4%	32.8%	36.1%	13.9%
2007	25	90	183	226	98	622	4.0%	14.5%	29.4%	36.3%	15.8%
2008	20	87	183	204	112	606	3.3%	14.4%	30.2%	33.7%	18.5%
2009	19	82	179	186	113	579	3.3%	14.2%	30.9%	32.1%	19.5%
2010	17	77	185	183	96	558	3.0%	13.8%	33.2%	32.8%	17.2%
2011	16	74	187	192	82	551	2.9%	13.4%	33.9%	34.8%	14.9%
2012	16	70	197	177	85	545	2.9%	12.8%	36.1%	32.5%	15.6%
2013	17	73	198	165	70	523	3.3%	14.0%	37.9%	31.5%	13.4%
2014	17	73	200	155	66	511	3.3%	14.3%	39.1%	30.3%	12.9%
Total	1099	3608	5773	6358	3719	20557	5.3%	17.6%	28.1%	30.9%	18.1%

The distribution of ratings over time. The sample contains 1,488 firms from 1985 to 2014, with 20,557 observations overall. The ratings are obtained from the Compustat Ratings File. We merge ratings based on the original S&P categories: A includes S&P ratings from AAA to AA; rating B includes S&P ratings from $AA-$ to A; rating C includes S&P ratings from $A-$ to BBB; rating D includes S&P ratings from $BBB-$ to $BB-$; and rating E includes S&P ratings $CCC+$ and below.

Table 2: Number of Cutoffs Identified by Rating Category

	Upper	No migration	Lower	Total
A	0	1001	98	1099
B	23	3272	313	3608
C	159	5207	407	5773
D	329	5632	397	6358
E	338	3381	0	3719
Total	849	18493	1215	20557

Cutoff observations identified for the purpose of model estimation. The identification relies on the observed rating migration. For a firm in two consecutive years t and $t+1$, if there is an upgrade in year $t+1$ (eg. $Rating_t = C$ and $Rating_{t+1} = B$), the credit quality in t is regarded as exceeding the upper cutoff of its original rating, and hence observation in t is an upper cutoff observation. Analogously, we identify lower cutoff observations based on downgrades. The observations not subject to a rating migration are in the middle category.

Table 3: Descriptive Statistics

	Mean	Median	Std.	10th	90th
Intcov	9.848	5.350	15.221	1.536	19.548
k1	4.044	5.000	1.356	1.801	5.000
k2	1.896	0.544	2.172	0.000	5.000
k3	1.615	0.000	3.326	0.000	10.000
k4	2.491	0.000	10.667	0.000	0.628
Vol	0.035	0.020	0.050	0.006	0.071
Tlev	0.333	0.312	0.180	0.129	0.560
Rent	0.023	0.013	0.031	0.003	0.052
Tan	0.389	0.344	0.236	0.104	0.743
Dni	0.196	0.000	0.397	0.000	1.000
Ddiv	0.671	1.000	0.470	0.000	1.000
Rd	0.016	0.000	0.031	0.000	0.053
Mtb	1.496	1.275	0.682	0.933	2.346
Size	7.942	7.904	1.614	5.848	10.101
Beta	0.952	0.958	2.012	0.231	1.943
Rmse	0.068	0.020	0.167	0.010	0.083

Descriptive statistics of covariates for the whole sample. All continuous variables are winsorized at top and bottom 1 percentile. $Intcov$ measures interest coverage is calculated by $ebitda$ divided by interest expense ($ebitda/xint$). $k1$ measures the $Intcov$ range from 0 to 5 (e.g. an observation with $Intcov = 3$ will have $k1 = 3$, $k2 = 0$, $k3 = 0$, $k4 = 0$). $k2$ measures the range from 5 to 10 (e.g. an observation with $Intcov = 7$ will have $k1 = 5$, $k2 = 2$, $k3 = 0$, $k4 = 0$). $k3$ measures the range from 10 to 20, and $k4$ measures the range above 20. Vol is the volatility of current and past four year profits ($ebitda/sale$). $Tlev$ is leverage measured by debt divided by total asset ($\frac{dlc+dltt}{at}$). $Rent$ is the rent expense divided by total asset ($xrent/at$). Tan refers tangibility, measured by property, plant and equipment divided by total asset (ppe/at). Dni is the dummy variable which equals one when net income is negative. $Ddiv$ is the dummy variable which equals one when firms pay dividend. Rd is research and development expense divided by total asset (xrd/at). Rd is set to zero when the expense is missing. Mtb is the market to book ratio measured by total asset minus book equity plus market equity and then divided by book assets ($(at - bv + mv)/at$). The market value of equity is the product of year-end price and number of shares outstanding ($prcc_f * csho$).[18] Equity seq will be replaced by either common equity plus preferred stock par value ($ceq + pstk$) or total asset minus total liability ($at - lt$) if missing. Preferred stock liquidating value $prstkl$ will be replaced by either redemption value $pstkrv$ or par value $pstk$ if missing. Firm size $Size$ is the logarithm of total asset. $Beta$ and $Rmse$ measure Systematic Risk and Idiosyncratic Risk, respectively. They are estimated in market model regressions of a firms daily stock returns on the CRSP value-weighted index return. The regressions are adjusted for nonsynchronous trading effects using the Dimson (1979) procedure with one leading and one lagging value of the market return. One firm-year observations of $Beta$ and $Rmse$ are computed from a regression using firm-specific daily stock returns from one calendar year.

[18]Book equity is shareholders' equity on balance sheet minus preferred stock plus deferred taxes and investment tax credit ($seq - pstkl + txditc$). The deferred tax credit $txditc$ is set to zero if missing.

Table 4: Regressions Estimating the Determinants of Credit Ratings

Panel A: Explanatory Variable Coefficients											
	Ordered Model 1	Adj. Model 2	Ordered Model 3	Adj. Model 4	Ordered Model 5	Adj. Model 6	Ordered Model 7	Adj. Model 8	FM(85-94) Model 9	FM(95-04) Model 10	FM(05-14) Model 11
Stickiness		t, z	t	z		t, z	t	z	t, z	t, z	t, z
Intcov	0.012	-0.001	0.012	0.001							
	(0.001)	(0.001)	(0.001)	(0.001)							
k1					0.306	0.047	0.315	0.097	0.022	0.115	0.016
					(0.010)	(0.013)	(0.010)	(0.016)	(0.021)	(0.025)	(0.035)
k2					0.078	0.001	0.082	0.012	0.001	0.006	0.014
					(0.006)	(0.009)	(0.006)	(0.011)	(0.018)	(0.017)	(0.019)
k3					0.026	-0.003	0.027	0.002	0.014	-0.005	-0.003
					(0.004)	(0.005)	(0.004)	(0.006)	(0.015)	(0.010)	(0.011)
k4					0.003	-0.001	0.004	0.000	-0.006	-0.002	0.002
					(0.001)	(0.001)	(0.001)	(0.001)	(0.005)	(0.002)	(0.002)
Vol	-3.318	-2.124	-6.493	-2.404	-5.410	-1.958	-4.864	-2.056	-2.479	-1.495	-2.780
	(0.212)	(0.303)	(0.222)	(0.337)	(0.230)	(0.316)	(0.223)	(0.345)	(0.547)	(0.520)	(0.747)
Tlev	-2.038	-1.229	-1.922	-1.561	-1.162	-1.121	-0.809	-1.309	-1.381	-0.674	-1.287
	(0.058)	(0.084)	(0.058)	(0.108)	(0.065)	(0.096)	(0.064)	(0.108)	(0.150)	(0.160)	(0.228)
Rent	-2.830	-4.747	-5.824	-4.803	-7.975	-4.821	-7.253	-5.013	-6.185	-3.370	-7.623
	(0.316)	(0.407)	(0.289)	(0.475)	(0.303)	(0.431)	(0.298)	(0.475)	(0.753)	(0.673)	(1.080)
Tan	0.830	0.377	0.814	0.428	0.859	0.366	0.752	0.413	0.204	0.514	0.421
	(0.039)	(0.056)	(0.038)	(0.065)	(0.039)	(0.059)	(0.039)	(0.064)	(0.099)	(0.111)	(0.131)
Dni	-0.651	-0.514	-0.398	-0.725	-0.418	-0.488	-0.210	-0.675	-0.448	-0.489	-0.668
	(0.024)	(0.033)	(0.023)	(0.036)	(0.025)	(0.033)	(0.024)	(0.042)	(0.060)	(0.061)	(0.094)
Ddiv	1.169	0.103	1.123	0.207	0.995	0.096	1.041	0.192	0.141	0.063	0.226
	(0.022)	(0.029)	(0.021)	(0.035)	(0.022)	(0.029)	(0.022)	(0.033)	(0.057)	(0.053)	(0.068)
Rd	0.061	-1.523	0.567	-1.938	-0.341	-1.673	-0.229	-2.275	-1.659	-3.098	0.627
	(0.308)	(0.440)	(0.307)	(0.553)	(0.314)	(0.446)	(0.311)	(0.526)	(0.760)	(0.818)	(1.206)
Mtb	0.517	0.411	0.430	0.434	0.387	0.405	0.305	0.413	0.436	0.424	0.407
	(0.014)	(0.021)	(0.014)	(0.026)	(0.015)	(0.023)	(0.015)	(0.023)	(0.049)	(0.037)	(0.059)
Size	0.465	0.133	0.419	0.152	0.442	0.138	0.439	0.162	0.147	0.123	0.203
	(0.007)	(0.010)	(0.007)	(0.012)	(0.007)	(0.010)	(0.007)	(0.013)	(0.019)	(0.020)	(0.024)
Beta	-0.032	-0.019	-0.035	-0.006	-0.036	-0.019	-0.035	-0.006	-0.021	-0.022	-0.067
	(0.004)	(0.005)	(0.004)	(0.006)	(0.004)	(0.005)	(0.004)	(0.007)	(0.008)	(0.010)	(0.025)
Rmse	-0.006	0.003	-0.387	0.011	-0.196	0.026	-0.169	0.060	0.112	0.047	-0.436
	(0.057)	(0.076)	(0.055)	(0.077)	(0.057)	(0.077)	(0.056)	(0.071)	(0.115)	(0.111)	(0.213)
Pseudo R	0.334	0.089	0.323	0.139	0.371	0.090	0.353	0.143	-	-	-
Nobs.	20557	20557	20557	20557	20557	20557	20557	20557	-	-	-

Panel B: Cutoffs

	Ordered Model 1	Adj. Model 2	Ordered Model 3	Adj. Model 4	Ordered Model 5	Adj. Model 6	Ordered Model 7	Adj. Model 8	FM(85-94) Model 9	FM(95-04) Model 10	FM(05-14) Model 11
U1/C12	2.489	1.844	1.701	1.784	2.903	2.046	3.120	2.210	1.795	2.515	2.624
	(0.073)	(0.106)	(0.072)	(0.099)	(0.085)	(0.116)	(0.084)	(0.129)	(0.193)	(0.204)	(0.275)
U2/C23	4.233	2.673	3.403	2.685	4.815	2.913	4.957	3.183	2.640	3.510	3.688
	(0.076)	(0.111)	(0.075)	(0.107)	(0.089)	(0.125)	(0.088)	(0.137)	(0.218)	(0.214)	(0.297)
U3/C34	5.574	3.335	4.696	3.352	6.236	3.593	6.322	3.885	3.148	4.238	4.791
	(0.078)	(0.117)	(0.077)	(0.115)	(0.092)	(0.132)	(0.091)	(0.149)	(0.241)	(0.287)	(0.511)
U4/C45	6.983	4.247	6.058	4.309	7.686	4.510	7.725	4.846	4.256	5.965	6.758
	(0.083)	(0.156)	(0.081)	(0.154)	(0.096)	(0.160)	(0.095)	(0.171)	(0.501)	(0.384)	(0.519)
L2		-0.812		-1.302		-0.580		-0.836	-0.813	-0.193	-0.105
		(0.106)		(0.108)		(0.122)		(0.136)	(0.214)	(0.207)	(0.290)
L3		-0.274		-0.619		-0.017		-0.095	-0.300	0.479	0.631
		(0.112)		(0.111)		(0.128)		(0.146)	(0.224)	(0.225)	(0.311)
L4		0.130		-0.141		0.395		0.412	0.078	0.995	1.083
		(0.117)		(0.119)		(0.138)		(0.154)	(0.235)	(0.234)	(0.333)
L5		0.470		0.235		0.733		0.797	0.428	1.401	1.078
		(0.137)		(0.143)		(0.151)		(0.171)	(0.274)	(0.441)	(0.615)

Regression estimates based on 20,557 observations from 1,488 unique firms in the sample. Models 1, 3, 5, and 7 use ordered probit regression which assumes no stickiness. The other results are estimated by the adjusted model controlling for stickiness. Row "Stickiness" indicates different dimensions of stickiness considered in model: z-stickiness, t-stickiness or both. The Fama-MacBeth coefficients are decade average of regression coefficients of the adjusted model. Panel A presents estimation results for explanatory variables and Panel B provides the estimates of cut-offs. We report standard errors in parentheses.

Table 5: Predicted Rating Migrations

	Actual			W/o Stickiness			With Stickiness		
	Up.	Stay	Down.	Up.	Stay	Down.	Up.	Stay	Down.
1986	18	755	54	81	673	73	0	822	5
1987	23	823	44	104	717	69	0	839	51
1988	30	744	39	93	639	81	0	765	48
1989	30	704	31	70	618	77	0	719	46
1990	21	636	32	120	531	38	0	652	37
1991	25	598	26	100	477	72	0	634	15
1992	28	648	27	166	480	57	1	683	19
1993	41	674	29	89	519	136	0	728	16
1994	22	678	23	92	548	83	1	704	18
1995	24	658	28	71	560	79	0	696	14
1996	31	678	27	128	566	42	0	723	13
1997	35	695	20	37	546	167	0	733	17
1998	35	720	38	90	574	129	0	765	28
1999	23	688	37	48	523	177	0	711	37
2000	14	647	48	37	482	190	0	652	57
2001	19	610	56	66	523	96	0	639	46
2002	9	645	48	100	528	74	0	647	55
2003	13	625	53	123	516	52	0	667	24
2004	18	651	32	63	556	82	0	678	23
2005	18	623	37	32	542	104	0	664	14
2006	19	603	40	47	554	61	0	645	17
2007	27	563	32	35	468	119	0	598	24
2008	10	559	37	83	456	67	0	574	32
2009	13	533	33	251	303	25	0	561	18
2010	20	516	22	42	427	89	0	552	6
2011	26	510	15	34	451	66	0	548	3
2012	16	512	17	140	378	27	0	538	7
2013	19	497	7	34	416	73	0	521	2
2014	16	484	11	13	373	125	0	507	4
Total	643	18277	943	2389	14944	2530	2	19165	696

Predicted rating migrations. The first three columns report the true number of upgrade observations, unchanged observations, and downgrade observations. The next three columns reports the predicted migration volume assuming no stickiness. The last three columns present the predicted migration numbers with stickiness. The credit quality is calculated using the regression results in Model 6 in Table 4.

Table 6: Determinants of Deviation in Credit Quality

	Full	Invest.	Specu.	Diff.	Large	Small	Diff.
k1	0.016	0.022	0.008	0.015	0.018	0.013	0.005
	(0.001)	(0.001)	(0.000)		(0.001)	(0.001)	
k2	0.002	0.001	0.002	-0.001	0.002	0.001	0.000
	(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	
k3	0.026	0.014	0.047	-0.032	0.028	0.024	0.004
	(0.001)	(0.001)	(0.002)		(0.002)	(0.002)	
k4	-0.027	-0.013	-0.053	0.040	-0.025	-0.029	0.003
	(0.002)	(0.002)	(0.004)		(0.002)	(0.003)	
Vol	0.020	0.036	0.001	0.035	0.027	0.014	0.013
	(0.003)	(0.005)	(0.002)		(0.004)	(0.004)	
Tlev	0.091	0.146	0.015	0.131	0.105	0.077	0.028
	(0.006)	(0.008)	(0.006)		(0.008)	(0.008)	
Rent	0.040	0.034	0.054	-0.021	0.038	0.041	-0.002
	(0.004)	(0.006)	(0.004)		(0.006)	(0.005)	
Tan	-0.017	-0.009	-0.029	0.021	-0.018	-0.016	-0.002
	(0.001)	(0.001)	(0.002)		(0.002)	(0.002)	
Dni	0.026	0.039	0.012	0.027	0.027	0.025	0.003
	(0.004)	(0.006)	(0.003)		(0.005)	(0.005)	
Ddiv	0.007	0.014	0.000	0.014	0.009	0.004	0.005
	(0.001)	(0.002)	(0.001)		(0.002)	(0.002)	
Rd	0.005	0.008	0.000	0.007	0.003	0.008	-0.005
	(0.001)	(0.001)	(0.002)		(0.002)	(0.002)	
Mtb	0.087	0.040	0.173	-0.133	0.095	0.080	0.014
	(0.006)	(0.006)	(0.011)		(0.008)	(0.008)	
Size	0.312	0.334	0.295	0.039	0.319	0.304	0.015
	(0.005)	(0.005)	(0.007)		(0.007)	(0.007)	
Beta	-0.007	-0.011	0.000	-0.011	-0.008	-0.006	-0.003
	(0.001)	(0.001)	(0.001)		(0.001)	(0.001)	
Rmse	-0.007	-0.011	-0.001	-0.010	-0.008	-0.007	-0.001
	(0.000)	(0.001)	(0.000)		(0.001)	(0.001)	
Total	0.573	0.646	0.525	0.121	0.612	0.535	0.077

Variation of predicted credit quality between the last six-year period (2009-2014) average level and the first six-year period (1985-1990) average level. The Z-score following Model 9 of Table 4 serves the proxy of credit quality: Z-score$= 0.023 * k1 + 0.001 * k2 + 0.014 * k3 - 0.006 * k4 - 2.479 * Vol - 1.381 * Tlev - 6.185 * Rent + 0.204 * tan - 0.448 * Dni + 0.141 * Ddiv - 1.659 * Rd + 0.436 * Mtb + 0.147 * Size - 0.021 * Beta + 0.112 * Rmse$. The contribution of each variable is the product of an increment in period average and the corresponding coefficient.[19] The standard errors in parentheses are calculated using Delta method.

[19]The average Vol in period 2009 to 2017 is 0.033, and the average level in period 1985 to 1990 is 0.041. Then, the contribution of Vol to credit quality variation is 0.020 (which is $(0.033 - 0.041) * (-2.479)$).

Table 7: Regressions of Migration Determinants

	Model 1				Model 2			
	Oprobit	ME1	ME2	ME3	Oprobit	ME1	ME2	ME3
k1	-0.042	0.004	-0.001	-0.003	-0.031	0.003	-0.001	-0.002
	(0.014)	(0.001)	(0.000)	(0.001)	(0.015)	(0.001)	(0.000)	(0.001)
k2	-0.015	0.001	0.000	-0.001	-0.011	0.001	0.000	-0.001
	(0.010)	(0.001)	(0.000)	(0.001)	(0.011)	(0.001)	(0.000)	(0.001)
k3	-0.006	0.001	0.000	0.000	-0.005	0.000	0.000	0.000
	(0.006)	(0.001)	(0.000)	(0.000)	(0.007)	(0.001)	(0.000)	(0.000)
k4	-0.002	0.000	0.000	0.000	-0.003	0.000	0.000	0.000
	(0.001)	(0.000)	(0.000)	(0.000)	(0.002)	(0.000)	(0.000)	(0.000)
Vol	-0.038	0.003	-0.001	-0.002	-0.479	0.043	-0.011	-0.032
	(0.271)	(0.023)	(0.006)	(0.016)	(0.320)	(0.028)	(0.007)	(0.021)
Tlev	-0.728	0.060	-0.016	-0.044	-0.838	0.074	-0.019	-0.056
	(0.089)	(0.007)	(0.003)	(0.005)	(0.100)	(0.009)	(0.003)	(0.007)
Rrent	-1.451	0.121	-0.033	-0.088	-0.492	0.044	-0.011	-0.033
	(0.407)	(0.034)	(0.010)	(0.025)	(0.670)	(0.060)	(0.015)	(0.045)
Tan	0.107	-0.009	0.002	0.006	-0.024	0.002	-0.001	-0.002
	(0.057)	(0.005)	(0.001)	(0.003)	(0.108)	(0.010)	(0.002)	(0.007)
Dni	-0.511	0.056	-0.033	-0.023	-0.512	0.037	0.002	-0.039
	(0.034)	(0.005)	(0.004)	(0.001)	(0.036)	(0.002)	(0.002)	(0.003)
Ddiv	-0.275	0.021	-0.003	-0.018	-0.270	0.018	0.007	-0.024
	(0.031)	(0.002)	(0.001)	(0.002)	(0.037)	(0.002)	(0.003)	(0.004)
Rd	-1.224	0.102	-0.027	-0.074	-2.436	0.218	-0.054	-0.165
	(0.444)	(0.037)	(0.011)	(0.027)	(0.757)	(0.069)	(0.018)	(0.052)
Mtb	0.235	-0.020	0.005	0.014	0.290	-0.025	0.006	0.019
	(0.022)	(0.002)	(0.001)	(0.001)	(0.025)	(0.002)	(0.001)	(0.002)
Size	-0.030	0.002	-0.001	-0.002	-0.052	0.005	-0.001	-0.003
	(0.010)	(0.001)	(0.000)	(0.001)	(0.013)	(0.001)	(0.000)	(0.001)
Beta	0.007	-0.001	0.000	0.000	0.006	-0.001	0.000	0.000
	(0.006)	(0.001)	(0.000)	(0.000)	(0.006)	(0.001)	(0.000)	(0.000)
Rmse	0.142	-0.012	0.003	0.009	0.139	-0.012	0.003	0.009
	(0.080)	(0.007)	(0.002)	(0.005)	(0.088)	(0.008)	(0.002)	(0.006)
Cut1	-2.506				-2.880			
	(0.124)				(0.255)			
Cut2	1.206				0.899			
	(0.122)				(0.253)			
Pseudo R	0.0511				0.0708			
Ind. FE	No				Yes			

Determinants of rating migrations under ordered probit model with year dummies. The dependent variable is 1 if a downgrade happens at that firm-year observation; it is 2 if the rating remains; and it is 3 if an upgrade happens. The columns 1 and 5 present the estimation results and the remaining columns provide marginal effect. The marginal effect for continuous variables refers to the slope of a specific probability[20] at the mean level of that variable. For dummy variables, marginal effect measures the probability deviation due to changes of that variable from 0 to 1. Model 2 controls for industry fixed effects.

[20]For "ME1" columns, the probability refers to the probability of being a downgrade observation. For "ME2" columns, the probability refers to the probability of being an observation with remaining ratings. For "ME3" columns, the probability refers to the probability of being an upgrade observation.

Table 8: Interaction Analysis

Panel A: Interaction without Lagged Rating	Model 1	Model 2	Model 3	Model 4		Model 5		
Dependent	y	y	lag y	y		y		
Interac.					M		U	D
Inter M					-0.235			
					(0.267)			
Inter U	1.314	1.522	1.128				-2.126	
	(0.045)	(0.049)	(0.044)				(0.488)	
Inter D	2.087	2.487	1.774					2.352
	(0.056)	(0.060)	(0.055)					(0.347)
k1	0.312	0.363	4.253	0.308	0.059	0.319	0.123	-0.108
	(0.010)	(0.012)	(0.095)	(0.011)	(0.031)	(0.011)	(0.058)	(0.040)
k2	0.075	0.135	6.166	0.073	0.045	0.075	0.022	0.000
	(0.007)	(0.007)	(0.100)	(0.007)	(0.022)	(0.007)	(0.038)	(0.028)
k3	0.025	0.024	7.578	0.023	0.022	0.024	-0.014	0.022
	(0.004)	(0.004)	(0.103)	(0.004)	(0.014)	(0.004)	(0.024)	(0.018)
k4	0.003	0.005	9.019	0.002	0.008	0.003	0.024	-0.008
	(0.001)	(0.001)	(0.107)	(0.001)	(0.004)	(0.001)	(0.007)	(0.005)
Vol	-5.332	-3.959	0.313	-5.536	3.201	-5.739	2.055	3.572
	(0.233)	(0.275)	(0.010)	(0.250)	(0.672)	(0.252)	(1.278)	(0.876)
Tlev	-1.392	-1.866	0.080	-1.413	1.698	-1.464	1.985	0.004
	(0.066)	(0.079)	(0.006)	(0.070)	(0.209)	(0.071)	(0.387)	(0.275)
Rrent	-9.271	-8.289	0.026	-9.111	4.489	-9.475	0.443	0.399
	(0.310)	(0.529)	(0.004)	(0.325)	(0.962)	(0.329)	(2.085)	(1.204)
Tan	0.859	0.961	0.003	0.842	-0.211	0.873	-0.249	-0.197
	(0.040)	(0.079)	(0.001)	(0.041)	(0.131)	(0.042)	(0.220)	(0.173)
Dni	-0.490	-0.392	-4.841	-0.496	0.531	-0.515	-0.097	0.143
	(0.025)	(0.028)	(0.225)	(0.027)	(0.074)	(0.027)	(0.164)	(0.091)
Ddiv	0.981	0.830	-1.004	0.955	0.304	0.995	0.148	-0.292
	(0.022)	(0.027)	(0.065)	(0.023)	(0.068)	(0.024)	(0.115)	(0.093)
Rxrd	-1.637	-0.043	-8.082	-1.648	0.921	-1.715	0.265	-0.612
	(0.319)	(0.569)	(0.302)	(0.336)	(1.029)	(0.338)	(1.790)	(1.337)
Mtb	0.461	0.480	0.770	0.456	-0.465	0.472	-0.165	-0.038
	(0.015)	(0.018)	(0.039)	(0.016)	(0.049)	(0.016)	(0.074)	(0.073)
Size	0.459	0.629	-0.259	0.454	-0.050	0.469	-0.015	-0.113
	(0.007)	(0.010)	(0.025)	(0.007)	(0.022)	(0.008)	(0.039)	(0.029)
Beta	-0.038	-0.026	1.026	-0.035	-0.011	-0.037	-0.011	-0.051
	(0.005)	(0.005)	(0.022)	(0.005)	(0.017)	(0.005)	(0.030)	(0.025)
Rmse	-0.171	-0.120	-1.087	-0.141	-0.212	-0.149	0.233	-0.389
	(0.058)	(0.067)	(0.314)	(0.060)	(0.202)	(0.060)	(0.384)	(0.259)
Cut1	4.293	6.461	0.362	2.957		3.035		
	(0.097)	(0.204)	(0.015)	(0.091)		(0.091)		
Cut2	6.314	8.920	0.449	4.907		5.085		
	(0.102)	(0.209)	(0.007)	(0.094)		(0.095)		
Cut3	7.798	10.738	-0.037	6.342		6.573		
	(0.105)	(0.213)	(0.004)	(0.097)		(0.098)		
Cut4	9.300	12.553	-0.208	7.803		8.080		
	(0.110)	(0.216)	(0.057)	(0.101)		(0.103)		
Pseudo R	0.395	0.491	0.371	0.378		0.400		
Ind. FE	No	Yes	No	No		No		
Year FE	Yes	Yes	Yes	Yes		Yes		
Panel B: Interaction with Lagged Rating								
	Model 1	Model 2		Model 3		Model 4		

Dependent	y	y	y			y	
Inter M				-0.414			
				(0.406)			
Inter U	-0.840	-1.048				-1.727	
	(0.067)	(0.070)				(0.736)	
Inter D	0.589	0.777					1.117
	(0.051)	(0.053)					(0.525)
lag y	2.936	2.867	2.920	0.342	2.947	-0.035	0.019
	(0.024)	(0.026)	(0.025)	(0.053)	(0.025)	(0.111)	(0.076)
k1	0.085	0.135	0.080	0.049	0.083	0.145	-0.013
	(0.015)	(0.017)	(0.016)	(0.048)	(0.016)	(0.089)	(0.061)
k2	0.010	0.042	0.009	-0.005	0.009	0.000	-0.009
	(0.010)	(0.011)	(0.010)	(0.033)	(0.010)	(0.058)	(0.042)
k3	0.006	0.008	0.003	0.030	0.003	0.022	0.028
	(0.006)	(0.006)	(0.006)	(0.022)	(0.006)	(0.037)	(0.028)
k4	-0.001	-0.001	-0.001	0.003	-0.001	0.006	-0.002
	(0.001)	(0.002)	(0.001)	(0.006)	(0.001)	(0.011)	(0.007)
Vol	-2.568	-2.701	-2.570	0.907	-2.627	-3.325	1.798
	(0.330)	(0.385)	(0.350)	(1.003)	(0.353)	(1.902)	(1.296)
Tlev	-1.554	-1.952	-1.629	1.722	-1.665	1.574	0.769
	(0.097)	(0.113)	(0.102)	(0.315)	(0.103)	(0.584)	(0.411)
Rrent	-5.766	-4.855	-5.878	5.761	-6.048	2.276	1.951
	(0.452)	(0.751)	(0.473)	(1.444)	(0.478)	(3.054)	(1.821)
Tan	0.513	0.517	0.513	-0.388	0.525	-0.098	-0.350
	(0.058)	(0.114)	(0.061)	(0.200)	(0.062)	(0.326)	(0.261)
Dni	-0.705	-0.680	-0.711	0.287	-0.728	-0.603	0.274
	(0.037)	(0.039)	(0.039)	(0.110)	(0.039)	(0.224)	(0.135)
Ddiv	0.192	0.165	0.191	-0.088	0.201	0.218	-0.222
	(0.033)	(0.039)	(0.035)	(0.110)	(0.035)	(0.181)	(0.145)
Rxrd	-1.973	-2.690	-1.884	-0.661	-1.944	-0.433	-1.689
	(0.466)	(0.807)	(0.490)	(1.526)	(0.494)	(2.710)	(1.936)
Mtb	0.415	0.484	0.423	-0.395	0.432	-0.156	-0.128
	(0.022)	(0.026)	(0.023)	(0.075)	(0.023)	(0.112)	(0.108)
Size	0.169	0.250	0.170	-0.095	0.173	0.059	-0.067
	(0.011)	(0.015)	(0.011)	(0.036)	(0.011)	(0.062)	(0.047)
Beta	-0.012	-0.009	-0.011	0.014	-0.012	-0.026	-0.008
	(0.007)	(0.007)	(0.007)	(0.027)	(0.007)	(0.047)	(0.038)
Rmse	0.041	0.051	0.069	-0.336	0.068	0.324	-0.548
	(0.086)	(0.099)	(0.090)	(0.308)	(0.091)	(0.573)	(0.395)
Cut1	4.976	5.592	4.935		4.975		
	(0.128)	(0.282)	(0.135)		(0.136)		
Cut2	8.798	9.658	8.717		8.838		
	(0.140)	(0.291)	(0.146)		(0.148)		
Cut3	12.365	13.386	12.254		12.409		
	(0.154)	(0.300)	(0.159)		(0.161)		
Cut4	15.968	17.124	15.836		16.020		
	(0.172)	(0.312)	(0.176)		(0.179)		
Pseudo R	0.791	0.804	0.789			0.793	
Ind. FE	No	Yes	No			No	
Year FE	Yes	Yes	Yes			Yes	

Interaction analysis of the stickiness in credit ratings using ordered probit model. The specification generally follows Model 5 in Table 4 but interaction terms added. The dependent variable is the observed credit rating or a lagged ratings, as described. "Inter M" refers to the dummy variable with value 1 indicating rating migration (either upgrade or downgrade). Similarly, "Inter U" and "Inter D" refer to dummies specifically indicating an upgrade and a downgrade, respectively. Model 4 includes terms interacting with migration dummies, and model 5 further replaces the migration interaction terms with upgrade and downgrade interaction terms. Panel A report the results without lagged rating interactions and Panel B controls these interaction terms.

Table 9: Lag Rating Analysis

	Model 1	Model 2	Model 3	Model 4
Lag Y	2.916	2.900	2.905	2.870
	(0.023)	(0.024)	(0.023)	(0.024)
Intcov	-0.001	0.002		
	(0.001)	(0.001)		
k1			0.045	0.087
			(0.015)	(0.015)
k2			-0.002	0.014
			(0.009)	(0.010)
k3			0.000	0.008
			(0.006)	(0.006)
k4			-0.001	0.000
			(0.001)	(0.001)
Vol	-2.901	-2.683	-2.750	-2.378
	(0.327)	(0.331)	(0.332)	(0.336)
Tlev	-1.495	-1.613	-1.394	-1.360
	(0.088)	(0.090)	(0.096)	(0.098)
Rent	-4.410	-5.023	-4.460	-5.245
	(0.437)	(0.445)	(0.439)	(0.449)
Tan	0.605	0.471	0.600	0.462
	(0.057)	(0.058)	(0.057)	(0.059)
Dni	-0.660	-0.704	-0.637	-0.663
	(0.035)	(0.036)	(0.036)	(0.037)
Ddiv	0.189	0.196	0.183	0.185
	(0.033)	(0.033)	(0.033)	(0.033)
Rd	-0.489	-1.588	-0.555	-1.876
	(0.457)	(0.469)	(0.458)	(0.472)
Mtb	0.371	0.392	0.363	0.368
	(0.021)	(0.021)	(0.021)	(0.022)
Size	0.079	0.145	0.079	0.153
	(0.009)	(0.011)	(0.009)	(0.011)
Beta	-0.010	-0.008	-0.010	-0.008
	(0.007)	(0.007)	(0.007)	(0.007)
Rmse	0.035	0.006	0.060	0.049
	(0.086)	(0.087)	(0.087)	(0.088)
Cut1	4.538	4.919	4.709	5.273
	(0.098)	(0.113)	(0.111)	(0.128)
Cut2	8.135	8.573	8.321	8.959
	(0.108)	(0.125)	(0.123)	(0.141)
Cut3	11.524	12.029	11.706	12.416
	(0.124)	(0.141)	(0.137)	(0.155)
Cut4	14.951	15.547	15.120	15.919
	(0.145)	(0.162)	(0.155)	(0.174)
Pseudo R	0.769	0.774	0.769	0.775
Year Dummy	No	Yes	No	Yes

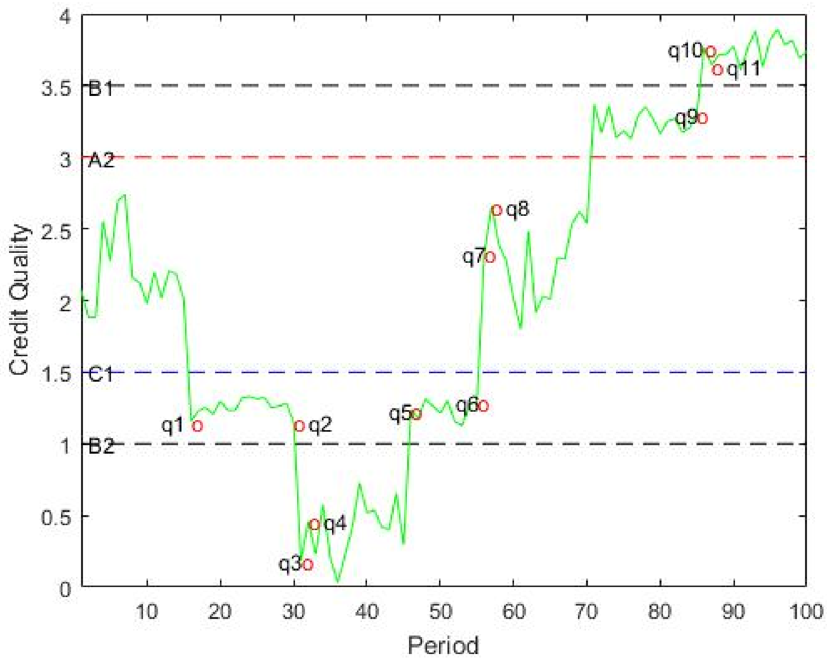
Ordered probit regression results controlling for lagged rating. *LagY* refers to the rating lagged by one year, and all other variables are defined as in Table 3. There are 19,069 observations after creating the lagged rating variable. Standard errors are in the parentheses.

Table 10: Subsample Analysis

Coefficients				
	Investment		Speculative	
	Model 1	Model 2	Model 3	Model 4
k1	0.398	0.100	0.292	0.059
	(0.028)	(0.045)	(0.015)	(0.020)
k2	0.097	-0.011	0.018	0.004
	(0.010)	(0.017)	(0.015)	(0.020)
k3	0.070	0.016	-0.024	-0.021
	(0.005)	(0.009)	(0.011)	(0.014)
k4	0.006	0.001	0.010	-0.003
	(0.001)	(0.002)	(0.003)	(0.004)
Vol	-10.011	-2.740	-1.958	-1.332
	(0.714)	(1.147)	(0.272)	(0.364)
Tlev	-0.541	-1.284	-0.888	-1.035
	(0.149)	(0.253)	(0.092)	(0.123)
Rrent	-14.343	-7.472	-4.149	-3.421
	(0.759)	(1.213)	(0.412)	(0.549)
Tan	0.869	0.342	0.097	-0.001
	(0.066)	(0.110)	(0.073)	(0.098)
Dni	-0.426	-0.476	-0.390	-0.509
	(0.060)	(0.103)	(0.035)	(0.046)
Ddiv	0.830	0.213	0.656	0.030
	(0.060)	(0.096)	(0.034)	(0.045)
Rxrd	1.404	-0.125	-5.208	-3.397
	(0.511)	(0.856)	(0.531)	(0.714)
Mtb	0.328	0.354	0.324	0.448
	(0.022)	(0.037)	(0.033)	(0.044)
Size	0.366	0.123	0.349	0.142
	(0.011)	(0.019)	(0.014)	(0.019)
Beta	-0.075	-0.010	-0.011	-0.022
	(0.011)	(0.019)	(0.006)	(0.008)
Rmse	0.245	0.573	0.035	0.059
	(0.143)	(0.225)	(0.077)	(0.105)
U1/Cut1	5.303	3.615	2.945	2.201
	(0.190)	(0.305)	(0.149)	(0.198)
U2/Cut2	6.984	4.602		
	(0.196)	(0.327)		
L1		0.443		-0.439
		(0.310)		(0.205)
L2		0.881		
		(0.325)		
Pseudo R	0.2705	0.0828	0.3294	0.1204
Obs.	10480	10480	10077	10077

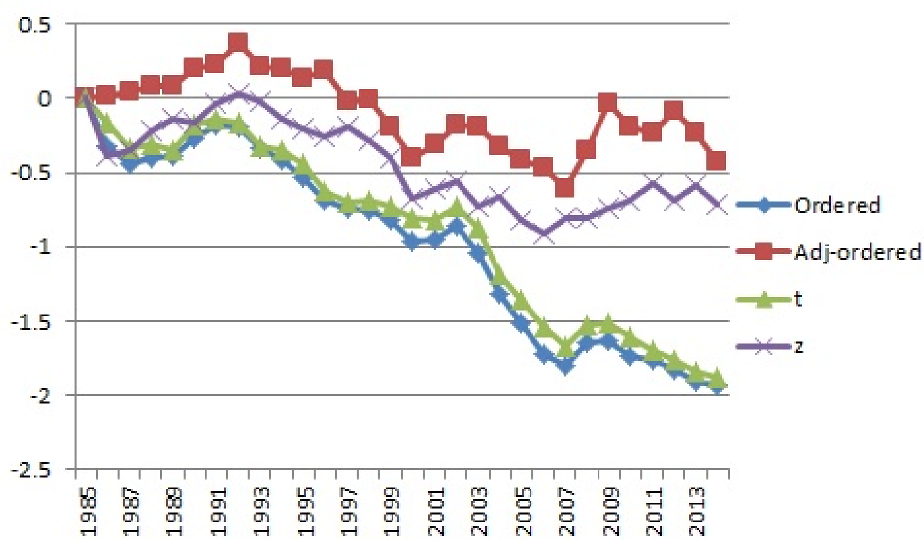
A comparison of models of Table 4 between estimations of ordered probit and adjusted ordered probit within subsamples of investment grade observations and speculative observations. Investment grade refers to BBB- or above under S&P credit rating framework. It includes ratings A, B, and C, in the data and contains 10,480 observations. Observations with S&P ratings below BBB- falls into the speculative grade, which covers ratings D and E in the data and includes 10,077 observations.

Figure 1: Credit Rating Mechanism



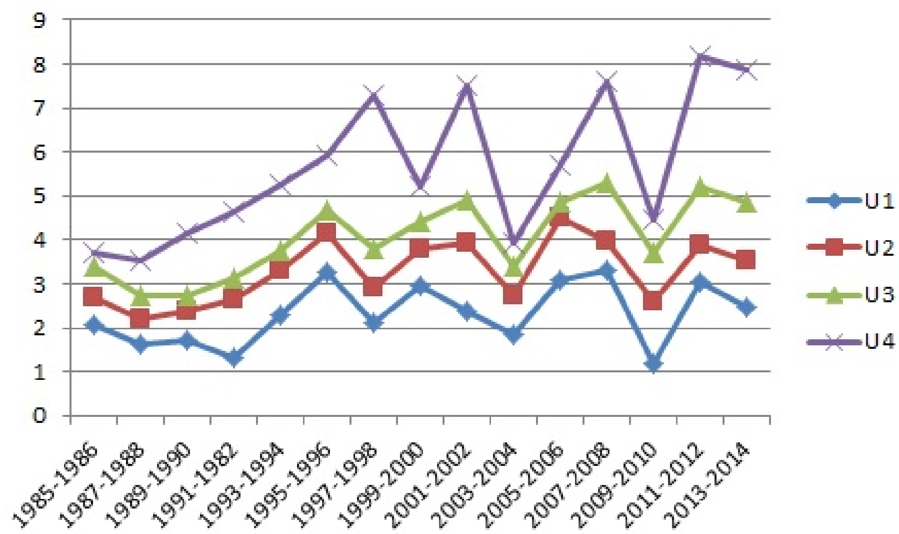
The mechanism of rating migrations under stickiness framework. This example contains 100 observations from one firm within consecutive 100 periods. The vertical axis indicates credit quality of the firm, and the horizontal axis represents period. There are three rating categories, A, B, and C, in which A indicates the best credit quality and C indicates the worst. The nominal quality range of rating A is the area from line $A2$ and above; the quality range for rating B is the area between lines $B1$ and $B2$; and the range for rating C is the area below line $C1$. The green line presents the path of the firm's credit quality movements, and demonstrates the mechanism of rating migration. A migration is triggered by credit quality crossing the boundaries of its nominal range. For instance, points $q2$, $q3$, and $q4$ depict the process of a downgrade migration. When credit quality drops from $q2$ to $q3$, it moves outside of rating B's range and this magnitude of deviation exceeds agency's tolerance. A downgrade decision is made at $q3$ but implemented at $q4$ to fit the slow-respondence feature. Points $q9$, $q10$, and $q11$ describe a rating upgrade process. Credit quality crosses $B1$, the upper boundary of rating B, to $q10$, and the rating upgrade is observed next period at $q11$.

Figure 2: Plot of the Estimates of the Year Dummy Intercepts



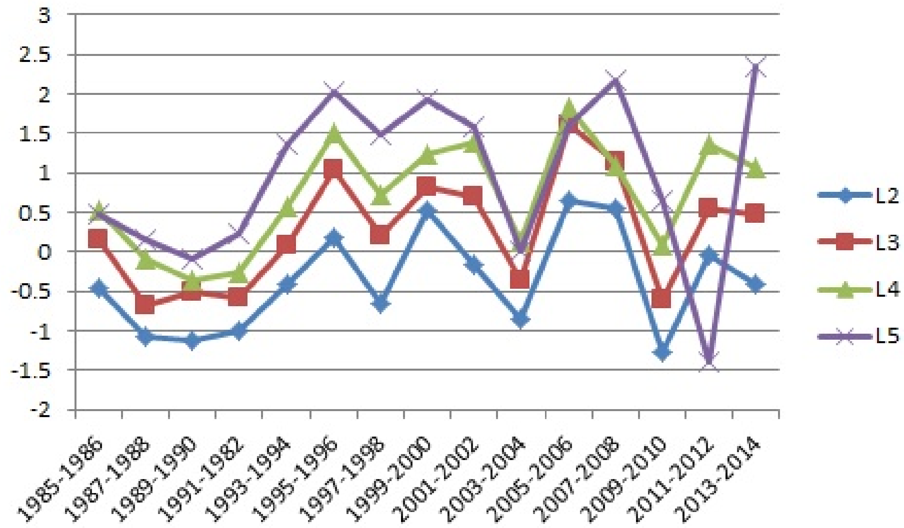
Year dummy intercepts estimated from ordered probit and adjusted ordered probit models in Table 4, based on a panel dataset containing 20,557 firm-year observations from 1985 to 2014. Rating stickiness is further separated into the t-dimensional and z-dimensional ones. The t stickiness refers to the delay of rating adjustment in time series. It is measured by the year intercepts from ordered probit model with lagged rating being the dependent variable. The z stickiness measures the tolerance of credit quality deviation, and the year intercepts are from adjusted ordered probit with different categorization method. The difference in categorization refers to neglecting the time-series delay of rating adjustment. More precise, our main adjusted ordered probit model assumes the rating migration at time t is caused by the breaching of rating threshold at time $t-1$. However, the z stickiness model assumes the rating migration at time t happens because of the breaching at the time t as well. Hence, it is an assumption of immediate adjustment.

Figure 3: Plot of the Estimates of the Upper Cut-off for Rating Categories



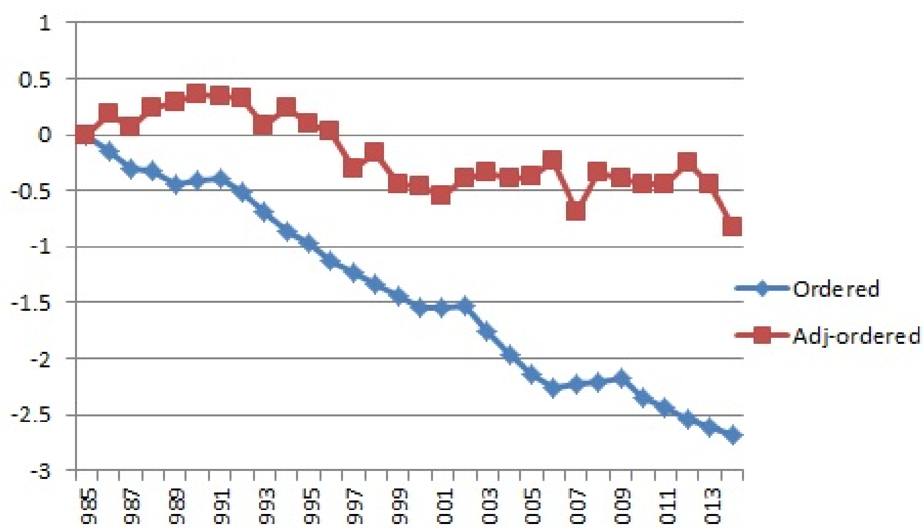
Evolution of upper cut-offs between adjacent ratings. The plots are based on the estimates of upper cutoff from adjusted ordered probit model for every two-year period. Overall, there are five rating categories, but the upper cutoff for the highest rated category is infinity (no further upgrade available).

Figure 4: Plot of the Estimates of the Lower Cut-off for Rating Categories



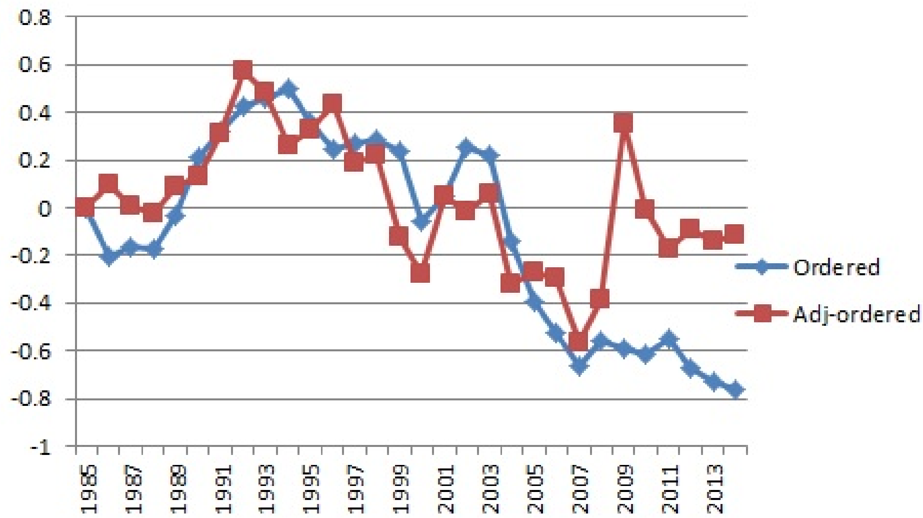
Evolution of upper cut-offs between adjacent ratings. The plots are based on the estimates of lower cutoff from adjusted ordered probit model for every two-year period. Overall, there are five rating categories, but the lower cutoff for the worst rated category is negative infinity (no further downgrade available).

Figure 5: Investment grade: Plot of the Estimates of the Year Dummy Intercepts



Year dummy intercepts based on the ordered probit and adjusted ordered probit models of investment grade subsample in Table 10.

Figure 6: Speculative: Plot of the Estimates of the Year Dummy Intercepts



Year dummy intercepts based on the ordered probit and adjusted ordered probit models of speculative grade subsample in Table 10.